

DISSERTATION

INVESTIGATING THE RELATIONSHIP BETWEEN COVER CROP SPECIES DIVERSITY,
COMPOSITION AND FUNCTION OF THE SOIL MICROBIOME.

Submitted by

Valerie Seitz

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2023

Doctoral Committee:

Advisor: Jessica Prenni

Kelly Wrighton
Meagan Schipanski
Marc Nishimura

Copyright by Valerie Seitz 2023

All Rights Reserved

ABSTRACT

INVESTIGATING THE RELATIONSHIP BETWEEN COVER CROP SPECIES DIVERSITY, COMPOSITION AND FUNCTION OF THE SOIL MICROBIOME.

Cropping diversification, such as cover cropping, can contribute to sustainable agriculture by enhancing soil health and promoting ecosystem services through interactions with the soil microbial community. One important mechanism through which cover crops impact soil health is via root exudation, the release of organic compounds from plant roots into the soil region surrounding the roots, the rhizosphere. Root exudation varies among cover crop species, growth stages, and edaphic and environmental conditions resulting in a myriad of effects on the rhizosphere. Plant-derived inputs, like root exudates, modulate the soil microbial community, influencing microbial biomass, community structure, and catalyzing biogeochemistry. As a result, cover crops are linked to microbial changes that impact soil nutrient cycling and organic matter decomposition leading to a legacy impact on primary crop yield and health. Understanding the intricate relationship between cover crop root exudation composition and the soil microbiome is crucial for optimizing cover crop selection, management practices, and harnessing cover crops for precision microbiome management in agroecosystems. My dissertation demonstrates that cover crop root exudation differs considerably across cover crop species, and cultivars within species, and reveals cover crop metabolic impacts on soil microbial composition and function, which play a large role in the generation and maintenance of healthy soils to support our agricultural needs.

ACKNOWLEDGEMENTS

I want to start by expressing my heartfelt gratitude to the many incredible people and sources of support that have been instrumental in making this PhD a reality. This journey has been an adventure filled one with a ton of learning and personal growth, and I'm incredibly thankful for the guidance, inspiration, and encouragement that has come my way over the years. It's thanks to the contributions of amazing colleagues, friends, and family that I've been able to reach this milestone.

First, thank you to **Dr. Jessica Prenni** for letting me join your lab all those years ago as an undergraduate, I can't think of a better mentor to have throughout the phases of earning this degree. You have been instrumental in pushing me to become the best version of myself. Thank you for sharing your wisdom, allowing me the independence to grow (and admittedly make a few mistakes) as a young scientist, and poising me with more confidence to pursue tough challenges.

I would also like to thank **Dr. Jacqueline Chaparro** for your mentorship and training during my time in the Prenni Lab. Thank you for constant support through all the trials and tribulations of data analysis. I appreciate your time, advice, and support and I know I am a much better scientist because of your guidance.

I am grateful to the rest of my committee, **Dr. Meagan Schipanski, Dr. Kelly Wrighton, and Dr. Marc Nishimura. Dr. Nishimura**, thank you for your constant support and always asking the big picture questions. Thank you, **Dr. Schipanski**, for your guidance with practical applications of these experiments and results, and for training me in field sampling. Thank you, **Dr. Kelly Wrighton**, for your constant support, wisdom, words of encouragement, time and training me in all things microbiome data analysis. I feel so lucky to have such a supportive committee.

Thank you to all members of the Prenni and Wrighton lab, past and present, that have provided so much support and guidance. Thank you to those that have helped me with my projects – all the big and small things have added up and I am so grateful for the support I received. Thank you, **Beth Mitchel**, for your help processing my many, many samples; thank you **Brandon Sandoval** for helping on field sampling days; and for everyone else whose unwavering support I always felt. A special thank you to **Dr. Kayla Borton** and **Dr. Bridget McGivern**, of the Wrighton Lab, for your many hours of training, guidance and sharing your knowledge (and code!) whenever I needed assistance.

Thank you to my friends and family who have provided immeasurable support. I couldn't have done it with my community! To my best friends, **Katie Feighner, Hailey Davis, Siena Tornillo, Jasmine Coffin, Darian Sheldon, Emily McMillan, Phoebe Millsap, Joey Stewart & the whole crew** who kept my head above water, made me laugh, who always listened when I needed a friend, and who were *always* down for a margarita (or two!), a concert, or anything really. I could not have done this without my amazing, supportive friends.

Thank you to my incredibly supportive mother, **Teri Lindstrom**, who I can always rely on for love, encouragement, advice, and support. This achievement is not only mine but also a reflection of your sacrifices, love, and guidance. I am also grateful for my in-laws, **Jodee and Tim Seitz and the whole Seitz crew** who have become such amazing supporters and I can always count on them for a good laugh.

Finally, to my loving husband, **Nate Seitz**, who fills my life with love, laughter, unbounding support (for more houseplants, of course), and the occasional (absolutely absurd, borderline psychotic) outdoor adventure that's sure to fulfil my adrenaline-seeking side. I am so much happier with you, and our animals, in my life.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS	iii
CHAPTER 1: ROOTING FOR SUCCESS: HARNESSING COVER CROP ROOT EXUDATION TO CULTIVATE THRIVING SOIL MICROBIOMES FOR SUSTAINABLE AGRICULTURE	1
1.1 Introduction.....	1
1.2 Influence of Cover Crop Root Exudates on the Bacterial and Archaeal Microbiome	3
1.2.1 Bacteria.....	3
1.2.2 Archaea	8
1.3 Functional Influence of Cover Crop Root Exudates on the Eukaryotic Microbiome	9
1.3.1 Fungi.....	9
1.3.2 Nematodes.....	11
1.4 Conclusions.....	12
CHAPTER 2: COVER CROP CULTIVAR, SPECIES, AND FUNCTIONAL DIVERSITY IS REFLECTED IN VARIABLE ROOT EXUDATION COMPOSITION	14
2.1 Introduction.....	14
2.2 Results & Discussion	16
2.2.1 Root exudation rate is not explained by cover crop root morphology alone.	16
2.2.1 Root exudate quality and quantity is influenced by plant genotype.	19
2.2.3 Identification of two soyasaponins, a soyasapogenol aglycone base, and sixty-nine soyasaponin molecular family members highlight legume cover crop functional capacity for soyasaponin synthesis.	26
2.2.4 Soyasaponins exclusively enriched in legume cover crop species.	30
2.2 Conclusions.....	32
2.3 Materials and Methods	33
2.3.1 Cover Crop Root Exudate Collections	33
2.3.2 Root Scanning, Analysis, and Drying	35
2.3.3. Targeted UPLC-MS/MS for Phytohormone Analysis of Root Exudates	35
2.3.4 Non-targeted GC-MS Analysis of Root Exudates	36
2.3.5 Non-targeted FastDDA LC-MS/MS Analysis	37
2.3.6 Metabolomics Data Analysis	38
2.3.7 Molecular Networking and Spectral Library Search	39

2.3.8 Compound Superclass Annotations	40
2.3.9 Statistics	40
CHAPTER 3: VARIATION IN ROOT EXUDATE COMPOSITION INFLUENCES SOIL MICROBIOME MEMBERSHIP AND FUNCTION	42
3.1 Introduction.....	42
3.2 Results & Discussion	43
3.2.1 Sorghum genetics influence root exudation patterns	43
3.2.2 Laboratory microcosms reveal that exudate treatment structures agricultural soil microbial communities.....	46
3.2.3 Exometabolites hint at potential microbial metabolisms stimulated by exudates.....	48
3.2.4 Curation of a sorghum exudate responsive microbial genome catalog	51
3.2.5 Bacterial salicylic acid production is assigned to a potentially-new species of Pseudomonas	53
3.2.6 Diverse benzoic acid production pathways encoded by multiple MAGs.....	55
3.2.7 Bacterial indoleacetic acid production encoded by diverse MAGs	56
3.3 Conclusions.....	57
3.4 Materials & Methods	58
3.4.1 Sorghum Root Exudate Collections.....	58
3.4.2 Sample Derivatization for Sorghum Root Exudates	59
3.4.3 Non-Targeted and Targeted GC-MS Analysis for Sorghum Root Exudates.....	60
3.4.4 Synthetic Exudate Preparation	60
3.4.5 Soil Samples	61
3.4.6 Microcosm Experimental Set-Up	61
3.4.7 Sample Preparation and Extraction for Exometabolomics Analysis	62
3.4.8 Targeted UPLC-MS/MS for Phytohormone Analysis of Exometabolites	62
3.4.9 Non-targeted GC-MS Analysis of Exometabolites	63
3.4.10 GC-MS and LC-MS Data Analysis.....	64
3.4.11 Soil DNA Extraction and Library Preparation	65
3.4.12 Metagenomics Analysis.....	65
3.4.13 Statistics.....	66
CHAPTER 4: GENOME-RESOLVED MULTI-OMICS REVEAL COVER CROP ROOT EXUDATES DIFFERENTIALLY INFLUENCE SOIL MICROBIAL FUNCTION AND PHYTOHORMONE PRODUCTION IN AGRICULTURAL SOIL MICROBIOMES	68
4.1 Introduction.....	68
4.2. Results & Discussion	72

4.2.1. Agricultural Exudate-Responsive Metagenomic (ARM) Database Expands Genomic Knowledge in Colorado Soils	72
4.2.2. Microbial Community Function, but Not Membership, is Altered by Exudate Addition	75
4.2.3. Root Exudates from Cereal Rye and Hairy Vetch Had a More Pronounced Impact on Restructuring Microbial Metabolic Outputs.	77
4.2.4. Specific Microbial Phytohormone Metabolisms are Captured in Laboratory Microcosms.....	79
4.2.3. Microbial Production of ACC-Deaminase in Response to Addition of ACC.....	80
4.2.4. Production of IAA & IAA-like Indole Derivatives is Stimulated by Rapeseed, Hairy Vetch and Sorghum Treatments, but not Cereal Rye.	81
4.2.5. Cereal Rye Exudates Stimulate Potentially New Bacterial Species with Gibberellic Acid 4 Biosynthetic Capacity.	86
4.3 Conclusions.....	89
4.4 Materials & Methods.....	90
4.4.1 Cover Crop Root Exudate Collections.....	90
4.4.2 Cover Crop Root Exudate Microcosm Experimental Setup	91
4.4.2 Metabolomics: Targeted UPLC-MS/MS for Phytohormone Analysis.....	92
4.4.4 Metabolomics: Nontargeted LC-MS and fastDDA Analysis	93
4.4.5 Metabolomics: Data Analysis	94
4.4.6 Metagenomics: Sample Preparation.....	96
4.4.7 Metagenomics: Sample Preparation.....	96
4.4.8 Metatranscriptomics: Sample Preparation.....	97
4.4.9 Metatranscriptomics: Data Analysis.....	98
4.4.10 Statistics.....	98
REFERENCES	100
APPENDICES.....	127
Appendix A: Chapter 2 Supplementary Files	127
Appendix B: Chapter 3 Supplementary Files	128
Appendix C: Chapter 4 Supplementary Files	129

CHAPTER 1: ROOTING FOR SUCCESS: HARNESSING COVER CROP ROOT EXUDATION TO CULTIVATE THRIVING SOIL MICROBIOMES FOR SUSTAINABLE AGRICULTURE

1.1 Introduction

Current sustainable agricultural practices, like cover cropping, use secondary crops to support the growth of a primary crop and stabilize the soil microbiome. Cover cropping is defined as the use of a secondary, unharvested crop in rotation or in coordination with a primary cash crop. Various plant species serve as cover crops and often multiple species are used at once to maximize desired functions. Specifically, cover crops have been shown to increase functions such as soil aggregation, water infiltration, weed suppression, nutrient cycling, and organic matter levels and in many cases, increase primary crop yields¹⁻⁷. Specialized cover crop ecosystem services, like the implementation of cover cropping for weed- and pest-suppressive effects, are known to vary as these desirable functional traits are typically expressed in a species- or genus-dependent manner. Thus, the selection of cover crops is primarily driven by desired primary crop outcomes through cover crop functional trait selection that is closely aligned with farmer objectives. However, many of these specialized traits have yet to be quantified in cover crops which ultimately hinders the potential impact cover crop selection can have on precision agriculture.

Cover crops elicit some of these ecosystem services through their physical and chemical interactions with the soil microbiome. These agricultural soil microbial communities are composed primarily of bacteria, fungi, archaea, nematodes, and viruses and harbor an incredible amount of taxonomic and functional diversity that directly respond to cover crops. While these soil microorganisms directly impact cover crop efficiency, their interactions with cover crops also play a vital role in the sustained functioning of agricultural ecosystems through their ability to aid in nutrient cycling, organic matter decomposition, and pest and disease management, leading to the

beneficial 'legacy effect' model that many cover cropping systems have reported⁸⁻¹¹. However, primary crops can only benefit from cover cropping (i.e., legacy effects) if the desired effects persist until primary crop germination, which means cover crops have the potential to boost and activate selected soil microbes that could improve these ecosystems function, yet this remains a relatively understudied area in agricultural research.

One way in which cover crops can interact with soil microbes, and modulate these microbial interactions and legacy effects, is through the release of plant-derived carbon and nitrogen containing compounds, called root exudates. Root exudates, a term encompassing plant root secretion of organic compounds, are a critical chemical cue that plants use to alter soil microbial community composition and function. Root exudate compounds range from simple primary metabolites to much larger secondary metabolites, proteins, and mucilage¹²⁻¹⁴. Plants use root exudates as currency in exchange for microbial assistance in accessing soil nutrients, resisting pathogens, and increasing nutrient and water availability¹⁵. However, the composition of root exudates is not static. Factors such as plant species and genotype, growth stage, and edaphic conditions lead to variations root exudate quantity and composition influencing microbial associations that crops may support¹⁶⁻¹⁸. In particular, crop genotype is a target for leveraging specialized root exudation as even genetic variants of the same species have shown dissimilarity in root exudate composition resulting in different rhizosphere establishment, highlighting the strict genetic regulation of root exudates and the potential for root exudates to be a target for precision microbiome management¹⁸⁻²¹.

Despite the growing appreciation for the impacts of cover crops on soil microbial communities, mechanistic research using modern microbiome technologies to link plant root exudate chemistry, soil microbial community metabolism, and soil health responses remains limited. As a result, there is a current critical need to identify how cover crop selection shapes the chemical landscape of the soil and its impacts on the soil microbial community composition and

function to enable the advancement of precision agricultural management strategies that aim to shape favorable microbiomes for primary crops.

Here, I review the current state of cover cropping research that investigates the potential for root exudates to influence the bacterial, archaeal, and eukaryotic establishment and function in agricultural systems and the legacy impacts as a result of these interactions. A primary focus of this review is to highlight current progress and emerging strategies for exudate-focused intentional microbiome manipulation to enhance primary crop performance and sustainability.

1.2 Influence of Cover Crop Root Exudates on the Bacterial and Archaeal Microbiome

1.2.1 Bacteria

Soil bacteria make up the vast majority of below-ground biomass and diversity closely followed by fungi, protists, viruses, and other eukaryotes²². At the local scale, bacteria from bulk soil associate with plants through chemotaxis towards plant roots and root exudates and densely inhabit the soil adjacent to the roots (i.e. the rhizosphere). Small root exudates from cover crops, like sugars, organic, and amino acids are the most common primary metabolites released via root exudation from the root tips and hairs of plants. Primary root exudates may function either directly to mobilize nutrients through adjusting soil pH, or by stimulation and feeding rhizosphere bacteria (rhizobacteria) which in turn mobilize nutrients into plant available forms. Aside from growth promotion and rhizosphere establishment, sugar assimilation in rhizobacteria promotes nutrient cycling. Nitrogen (N), phosphorus (P), and potassium (K) are considered the most limiting minerals to crop growth and productivity due to the limited quantity of naturally-occurring plant bioavailable forms in many soils²³. Root exudates therefore provide the necessary signals required to stimulate and metabolically compensate beneficial rhizobacteria like phosphorus-solubilizing bacteria (PSB), potassium solubilizing bacteria (KSB), heterotrophic bacteria involved in nitrogen mineralization processes, and nitrogen fixing bacteria (rhizobia) which all function to convert minerals into plant bioavailable forms (NH_4^+ , NO_3^- , K^+ , H_2PO_4^- , HPO_4^{2-}), reducing the

need for external inputs in cropping systems^{7,24} (**Figure 1A**). Additionally, PSB, KSB, and siderophore-releasing bacteria benefit from metabolizing root exudate sugars resulting in organic acid by-products that lower soil pH which further aids in the solubilization of plant bioavailable P, K, and iron²⁵⁻²⁷.

Variation in root architecture is an additional factor known to influence the spatial distribution and abundance of roots and root exudates, determining the extent and effect of root-soil interactions modulating the quantity and composition of exudates released into the rhizosphere²⁸. For example, cover crops *Vicia faba*, *Fagopyrum esculentum*, and *Lupinus sp.* grown under P deficiency respond through increasing root biomass and exudation of primary root exudates like organic acids, carboxylates, protons, and enzymes necessary for P acquisition, enhancing P availability²⁹⁻³².

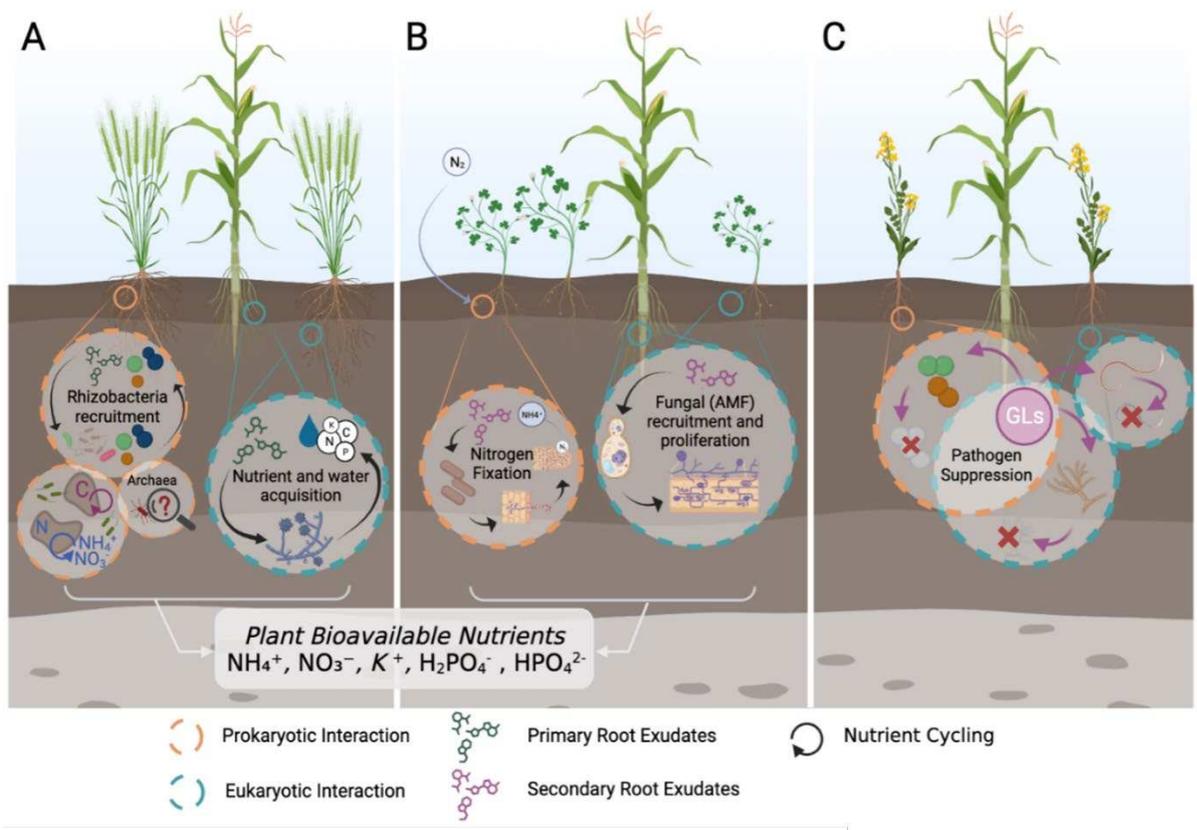


Figure 1. Rhizosphere intercropping interactions. The varying interactions between commonly used cover crop functional groups (a. grasses; b. legumes; c. brassicas) and primary crop (i.e., corn) grown in close proximity (intercropping) where root exudates and root-associated microorganisms influence soil biogeochemistry, crop function, and crop disease suppression. Circles depict functional changes in the rhizosphere from root exudation, circle colors indicate prokaryotic (orange) or eukaryotic (blue) interactions. A) Influence of grass cover crop species on the rhizosphere. Eukaryotic recruitment via general/primary root exudates improve nutrient and water acquisition. Rhizobacteria recruitment from primary root exudates, nutrient cycling (C, Carbon; N, Nitrogen) from beneficial bacteria, and the potential role of archaea in the rhizosphere all of which contributing to increase nutrient availability for plants. B) Influence of legume cover crops on the rhizosphere through exuded flavonoid compounds that recruit N cycling rhizobia, which form root nodules, and fix atmospheric N. Other specialized flavonoid exudates recruit AMF and contribute to increased nutrient availability to plants. C) Influence of brassica cover crops on the rhizosphere. Glucosinolates (GLs) from brassica root exudates control disease pressure against pathogenic bacteria and mycotoxigenic fungi, as well as parasitic nematodes. This figure was generated in BioRender.

Secondary root exudates are specialized metabolites that can improve plant fitness and influence soil N cycle dynamics, improving N use efficiency. Secondary plant metabolites commonly found in soils include flavonoids, phenolics, polyketides, terpenoids, alkaloids, and many others in varied concentrations across species³³⁻³⁶. Using secondary root exudates, plants have the ability to regulate N transformations carried out by fungal and bacterial populations. For example, leguminous cover crops species can enhance N availability by releasing species specific flavonoids which target compatible nitrogenase-producing rhizobial bacteria that convert N₂ gas into NH₃ within special legume root nodule structures (**Figure 1B**)³⁷. An additional impact root exudates may have on nitrogen cycling is with nitrifying rhizobacteria, which convert NH₄⁺ to NO₃⁻ to be assimilated by plants. This also causes NO₃⁻ leaching and the loss of the greenhouse gas, nitrous oxide (N₂O). Recently, biological nitrification inhibition by root exudation has been characterized as a way to control soil N cycle and reduce environmental pollution. For example, the cover crops *Sorghum bicolor* and *Brachiaria humidicola* produce secondary metabolites sorgoleone, sakuranetin, methyl 3-(4-hydroxyphenyl) propionate, brachialactone which can inhibit ammonia monooxygenase, the enzyme responsible for the first and rate-limiting step of nitrification (NH₃ oxidation to NH₂OH). Sorgoleone, sakuranetin, and brachialactone also inhibit

hydroxylamine oxidoreductase, which catalyzes the second step of nitrification (oxidation of NH_2OH to NO_2^-)³⁸⁻⁴⁰. Therefore, endogenously supporting soil N cycling using specialized cover crops with specific root exudate patterns can greatly reduce the pressure to use N-based fertilizers alleviating harmful effects from N overapplication such as eutrophication of water systems^{7,41}.

Antimicrobial nitrogen- and sulfur-containing secondary metabolites called glucosinolates (GLs) are another method some cover crops use as a functional strategy to improve soil microbial community structure. GLs are found commonly in brassicas cover crops like *Brassica rapa* and *Brassica napus* which can be planted to reduce reliance on pesticides in agroecosystems and reduce pest pressure on primary crop, leading to improved yield and quality^{42,43}. GLs are hydrolyzed by myrosinase into specialized thiocyanates, isothiocyanates, nitriles among other hydrolysis products. Hydrolyzed GLs can inhibit the growth of soil-borne pathogenic bacteria and parasitic fungi, protecting primary crops from these disease and providing the potential for endogenous microbial pathogen suppression in sustainable cropping systems without the need for traditional pesticide applications^{41,44} (**Figure 1C**). However, reports on the impacts of brassica biofumigation off-target effects of GLs are varied, making GLs a complex biofumigation strategy what warrants more attention⁴⁵⁻⁴⁷.

Benzoxazinoids (BXs) are allelochemicals exuded from the roots of a number of cover crop species (e.g., *Poaceae* species) into the rhizosphere, where they can have multiple functions such as a plant defense system reducing the growth of neighboring plants and altering root-associated fungal and bacterial infections. Because of this, *Poaceae* crops (e.g., cereal rye, maize, wheat) offer a competitive advantage by limiting the growth of weeds. In maize, the BX, 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA), acts as a chemoattractant for plant growth promoting (PGP) *Pseudomonas putida* KT2440, a competitive root colonizer of the maize rhizosphere that promotes colonization by other PGP bacteria and elicits systemic defense

priming⁴⁸⁻⁵¹. Despite maize being a primary crop, this perspective of specialized root exudates in maize also offers a potential target in other cereal grain crops. In contrast, wheat BXs are primed for defense against mycotoxigenic fungi and cereal rye BXs illicit weed suppressive effects^{52,53}. While they are primarily known for their roles in plant defense against pests, pathogens, and allelopathy, recent research has suggested that BXs can also have beneficial effects when released into the soil as root exudates offer a new perspective for cereal grains as useful cover crops that can modulate soil microbial composition and functionality.

Generally, primary and secondary root exudates will positively influence the bacterial rhizosphere dynamically whether cover crops are intercropped or in rotation with main crops (**Figure 2**). These effects are largely through increased host-defense priming, soil enzymatic activity, microbial taxonomic diversity and function, nutrient cycling, mineral solubilization, and other plant-beneficial interactions which are estimated to enhance cover crop legacy effects on primary crop productivity^{13,54-56}. Although intensive future research is needed to untangle the proximal mechanisms governing cover crop root exudation, the bacterial rhizosphere function and composition, and the potential for meaningful legacy effects, current work in the field suggests that cover crop root exudates and their manipulation could be promising targets for enhancing bacterial function in agricultural microbiomes.

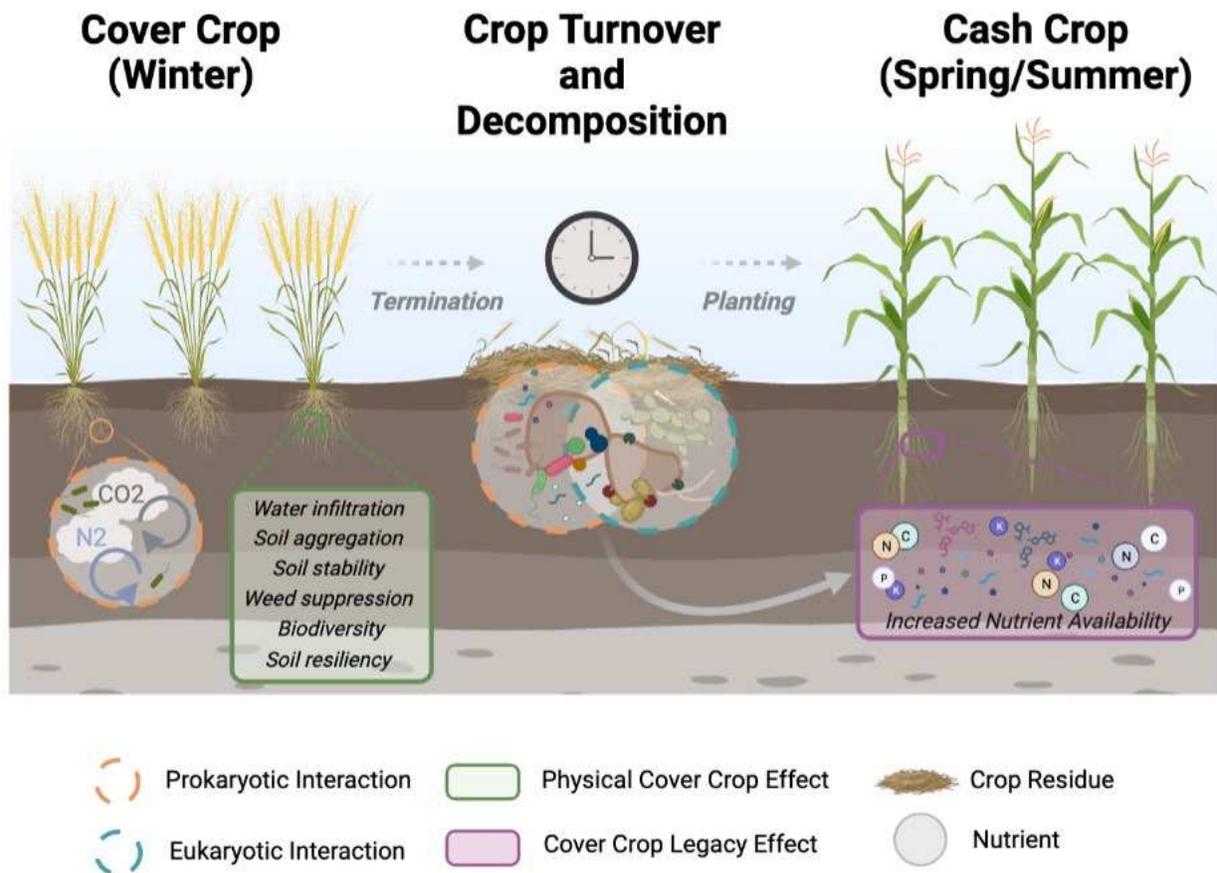


Figure 2. Influence of rotational cover crops on the soil microbiome and their legacy effects. Winter cover crops planted during fallow allow soil protection from wind and rain erosion, among other physical properties due to root presence (green box). Nutrient cycling and organic matter addition take place during periods of cover cropping (orange circle). Upon termination of the cover crop, organic matter decomposition and turnover begins from the bacteria and fungi (orange and blue circles). Legacy effects from the increase in organic input and carbon turnover positively impacts the growth of the primary crop (e.g., corn) (purple box).

1.2.2 Archaea

Archaea are distinct from bacteria in their functional potential (although some functional redundancy exists) and ecosystem preferences, as archaea are best known to thrive in earth's most extreme environments. Although underrepresented in microbial profiling of agricultural soils, with the exception of archaeal methanogenesis in rice fields, plant-associated archaea are gaining an appreciation for their functional potential as PGP organisms, nutrient cyclers, and

ability to shield plants during periods of stress^{57,58}. As obligate anaerobes, archaeal taxa in agricultural microbiomes are typically methanogens from the orders *Methanobacteriales*, *Methanomicrobiales*, *Methanosarcinales* and *Methanocellales*^{59,60}. Recent work that characterized an alpine peat bog associated archeome using metagenomics has uncovered the potential role of archaea in PGP due to the detection of auxin biosynthesis genes as a domain-wide signature⁵⁷. In the context of the rhizosphere, auxins are critical signaling molecules between rhizosphere microbes and plants, as these compounds can be produced by both plants and microbes and is often linked to improved plant health through elongation of plant roots, faster nodulation by rhizobia, and improved plant productivity. These results assign an emerging role for archaea to be significant members in PGP model^{61,62}. Previous work has also shown that within a chemically rich and highly competitive rhizosphere, slower-growing archaea are out-competed by faster-growing fungi and bacteria, potentially contributing to their lower abundance in rhizosphere despite their similar carbon utilization⁶³. Despite an appreciation for archaeal abundance and functional potential in the rhizosphere of some crop plants, little work has supported the mechanistic response of archaea to root exudation in a targeted manner, leaving open the possibility for archaea to become an untapped resource in precision microbiome management.

1.3 Functional Influence of Cover Crop Root Exudates on the Eukaryotic Microbiome

1.3.1 Fungi

The agricultural rhizosphere harbors various types of fungi, including arbuscular mycorrhizal fungi (AMF), saprophytic fungi, endophytic fungi, and pathogenic fungi influencing nutrient uptake, decomposition, plant health, and disease dynamics. Fungi are highly efficient at colonizing plants and roots due to their unique physiological and morphological characteristics. Because fungi are heterotrophs and rely on root exudates, soil organic matter, and other external energy sources for survival, they have developed specialized structures like hyphae and mycelium that enable

them to penetrate and explore the plant tissues and large soil volumes, allowing for nutrient acquisition and symbiotic interactions. Root associated fungi, called mycorrhizae, have adapted two strategies for plant root symbiosis. The fungal mycelium of ectomycorrhiza surround the roots forming a mantle and Hartig net, providing extensive surface area for nutrient exchange. In contrast, endomycorrhiza, most observed as arbuscular mycorrhiza fungi (AMF), form a mutualistic symbiotic association with plant roots, where the fungal hyphae penetrate the root cells, forming arbuscules and vesicles, facilitating nutrient exchange and enhancing plant nutrient uptake.

Cover crops represent a key opportunity to enhance soil health through fungal recruitment by root exudation. While mycorrhiza do metabolize primary root exudates, in general they receive a relatively constant source of glucose from their root-associated hyphae while external hyphae and mycelium scavenge for nutrients through the release of extracellular hydrolytic enzymes like cellulase and lignin-degrading enzymes⁶⁴. Therefore, secondary metabolites are the prime target of interest for cover crop fungal associations.

Secondary root exudates play a larger role in influencing fungal diversity and function in the rhizosphere. For example, leguminous cover crops like alfalfa, hairy vetch, and clover release higher quantities of flavonoids such as quercetin and quercitrin as root exudates which are recognized by AMF as stimulatory signaling molecules that induce in hyphal growth, differentiation, and root colonization (**Figure 1B**)^{65,66}. Similarly, strigolactone root exudates released from cover crops species like alfalfa, pea, and sorghum have also been shown to induce hyphal branching of AMF and stimulation of spore germination⁶⁷⁻⁶⁹. AMF fungi, in return for plant carbon, extend their hyphae into the soil and increase the root surface area for nutrient absorption, particularly to scavenge for phosphorus^{70,71}. Mycorrhizae colonization also act to suppress pathogens in cover and primary crops (**Figure 1C**). For example, enhanced mycorrhizal populations associated with hairy vetch cover cropping enhanced subsequent mycorrhizal

abundance and reduced the incidence of fusarium wilt in watermelon crops⁷². Similar effects have been recorded in other cover crop systems⁷³⁻⁷⁶. Cover crops that associate with AMF also improve water infiltration, enhance soil structure, and protect against erosion due to the extensive hyphal networks, although these functions are less directly associated with cover crop root exudate chemistry and more likely a physical effect from plant roots (**Figure 2**). These biological functions highlight the importance of AMF colonization, which is typically disturbed by traditional tillage systems in conventional agriculture⁷⁷. Recent results further support the role root exudates and cover cropping as AMF-colonizing cover crops act to increase AMF colonization of primary crops through their legacy effects⁷⁷.

1.3.2 Nematodes

Nematodes also play a vital role in agricultural soil ecosystems. Beneficial nematode species decompose organic matter and mineralize nutrients contributing to plant and soil fertility. Some nematodes feed on bacteria, fungi, and other soil organisms, regulating their populations and maintaining a balanced soil ecosystem. Plant parasitic nematodes, however, can cause significant damage to crops by feeding on various plant tissues as they puncture plant cells, withdraw nutrients, and cause stunted growth, wilting, and ultimately, crop yield losses. Consequently, nematicides are utilized but their use may affect non-target organisms and reduce the efficiency of cropping systems. Thus, many secondary metabolites from cover crops have been investigated for their nematocidal or nematode inhibitory effects⁷⁸. For instance, cereal rye BXs, primarily DIBOA (2,4-dihydroxy-(2H)-1,4-benzoxazin-3(4H)-one), DIMBOA, and their degradation products BOA (benzoxazolin-2(3H)-one) and MBOA (6-methoxy-benzoxazolin-2(3H)-one) exhibit nematocidal effects in vitro and thus represent a target for biocidal root exudates⁷⁹. Alfalfa root exudates containing medicagenic acid, a triterpene glycoside, can control for parasitic nematodes such as *Xiphinema index*, the root-knot nematode *Meloidogyne incognita* and the potato cyst parasite, *Globodera rostochiensis*⁸⁰. Finally, brassica glucosinolates are anti-

herbivory against nematodes as the hydrolyzed glucosinolates products can be lethal to root knot-nematodes and cyst nematodes in potato crop populations^{42,81} (**Figure 1C**). Thus, cover crop root exudates have the potential to play a significant role in managing nematodes by acting as deterrents or through inhibitory effects on nematode development and reproduction.

1.4 Conclusions

The intentional manipulation of the rhizosphere microbiome is limited in practice and has historically involved bacterial or fungal soil inoculation, such as inoculating seeds with PGP microorganisms, to encourage proliferation without the need for external chemicals⁸²⁻⁸⁴. The primary pitfall to microbiome-focused strategies is the lack of inoculant resilience to competition or lack of proliferation due to unfavorable soil conditions which decreases their impact. Alternatively, successful experiments using the complementary approach of altering root exudate chemistry to induce changes in the microbiome are sparse but suggest great potential. This approach alters the expression of transporter proteins or metabolic pathways to regulate the release of specific root exudates to encourage consistent and favorable rhizosphere colonization from bulk soil⁸⁵. Current work has investigated how regulating the quantity of primary root exudates, like citrate, malate, and γ -amino butyric acid, can alter the soil microbiome of transgenic wheat and rice roots, but root tissue type and soil still played a significant role⁸⁵. Despite these efforts, an outstanding question remains on how and if important secondary metabolites, such as beneficial rhizobia inducing flavonoids or pathogen suppressing glucosinolate expression can be altered in a productive and profitable way. Future work is therefore a necessity required to enable the development of efficient and precise cover crop management strategies in regenerative agricultural systems that utilize root exudates to intentionally shape favorable agricultural soil microbiomes.

In this work, I addressed this knowledge gap using multi-omic approaches within laboratory systems to interrogate the microbial responses to variable cover crop root exudate stimulation. Specifically, the three objectives of my dissertation research were to:

1. Determine the variability in primary and secondary root exudate composition across 19 commonly used cover crop species (Chapter 2).
2. Utilize laboratory-scale soil microcosms to measure soil microbial responses to root exudate treatments and curate an agriculturally relevant exudate responsive metagenomics database (Chapter 3).
3. Examine differential multi-omic responses of soil microorganisms under stimulation by four cover crop root exudate treatments to identify mechanisms with which cover crops benefit soil biogeochemical cycles, microbial inhabitants, and primary crop productivity. A major community outcome of this work is the availability of the exudate-responsive metagenomic, metatranscriptomic, metaproteomic, and metabolome repositories (Chapter 4).

Taken together, this work supports the development of efficient and precise cover crop management strategies in regenerative agricultural systems that utilize root exudates to intentionally shape favorable agricultural soil microbiomes.

CHAPTER 2: COVER CROP CULTIVAR, SPECIES, AND FUNCTIONAL DIVERSITY IS REFLECTED IN VARIABLE ROOT EXUDATION COMPOSITION¹

2.1 Introduction

Current agricultural practices such as monoculturing, soil tillage, and application of synthetic inputs can negatively impact soil health and the surrounding environment, degrading the potential of arable farmlands to produce crops for an increasing population⁸⁶⁻⁸⁸. Regenerative agricultural practices offer a practical solution for maintaining crop productivity without compromising future needs. Such agricultural practices work to reduce reliance on external inputs and improve soil health by increasing biodiversity and soil cover while reducing soil disturbance, amongst other practices. While there is no individual agricultural innovation that meets every sustainability requirement, cover cropping represents a versatile management strategy that is highly adaptable to soil type, primary crop species, and desired soil health outcomes. Integrating cover crops has been reported to improve soil health parameters and provide an alternative approach to maintaining primary crop productivity while promoting soil health through increased above- and below-ground biodiversity^{5,89,90}.

Cover cropping is defined as the use of a secondary, unharvested crop in rotation or in coordination with a primary cash crop. Various plant species serve as cover crops and often multiple species are used at once to maximize desired functions^{91,92}. Specifically, cover crops have been shown to increase functions such as soil aggregation, water infiltration, weed suppression, nutrient cycling, and organic matter levels^{1,3,93,94}. Additionally, cover crops contribute to sustainable pest control, acting as biofumigants and weed suppressors which is reported to reduce the need for herbicides and pesticides⁹². Cover crop ecosystem services, like the

¹ This chapter was reproduced verbatim from “Seitz, et al. Cover Crop Cultivar, Species, and Functional Diversity is Reflected in Variable Root Exudation Composition. *Journal of Agriculture and Food Chemistry* (2023)”. The text benefitted from writing and editing contributions from other contributing authors and reviewers selected by the publisher. The ordering of the materials in this dissertation are consistent with the content available online but have been renumbered to reflect incorporation into this dissertation.

implementation of cover cropping for weed- and pest-suppressive effects, are known to vary as desirable functional traits are typically expressed in a species- or genus-dependent manner. Thus, the selection of cover crops is primarily driven by desired primary crop outcomes through cover crop functional trait selection that is closely aligned with farmer objectives.

Several cover crop functions, such as pest suppression and nutrient cycling, are mediated by soil microbial communities. One of the primary mechanisms by which cover crops exert functional changes to soil biological activity is through stimulation from root exudates, which are low molecular weight compounds (e.g., organic acids, sugars, phenolics, amino acids, and other secondary metabolites) that are secreted by plant roots into soils¹⁵. Within the soil region adjacent to roots (the rhizosphere), root exudates can be involved in multiple plant-beneficial processes including modification of soil properties, biological nitrogen fixation, resistance to pests, and recruitment of beneficial rhizosphere soil microbes. Importantly, multiple studies have shown that root exudate composition will vary between crop species and with plant growth, which can likely support different microbial associations^{85,95}. Thus, as a result of above- and below-ground phenotypic traits and the influence on microbial associations, cover crop species have been loosely classified into various plant functional types (e.g., grasses, legumes, brassicas). For example, crops species classified as legumes can enhance soil nitrogen availability by targeting free-living rhizobia (nitrogen fixing taxa), while crop species within the brassica functional group produce glucosinolates that control soil pest and disease pressures⁴⁴. More recently, the use of diverse mixtures of cover crop species from multiple functional groups has gained popularity as this is thought to provide several soil health benefits simultaneously⁵. Despite these observations, a detailed understanding of the root exudate chemical profiles representative of each cover crop species, and select cultivars, remains lacking, leaving a major gap in the knowledge necessary to enable cover cropping as a precision agricultural management strategy.

Here, we investigated the chemical heterogeneity of root exudates across 16 commonly used cover crop species as well as 3 distinct cultivars within a hairy vetch species⁹⁶. We hypothesized that cover crop species would exhibit significant variation in root exudate composition and these differences would be attributable to diverse morphological and functional traits linked to plant fitness. We collected pure root exudates from hydroponically grown cover crops and profiled the chemical heterogeneity in cover crop metabolisms to characterize root exudate profiles. To our knowledge, this is the first comprehensive characterization of root exudates from commonly used cover crops in U.S. agriculture. The results of this study will directly contribute to the development of efficient and precise management strategies in cover cropping agricultural systems while paving the way for manipulation of root exudate composition to intentionally shape favorable agricultural soil microbiomes.

2.2 Results & Discussion

2.2.1 Root exudation rate is not explained by cover crop root morphology alone.

The cover crop species chosen for this work represent the most commonly used species as reported in the 2017 USDA Sustainable Agriculture Research and Education program and the Conservation Technology Information Center cover crop survey that included responses from more than 2,000 farmers across the U.S. on their cover crop practices⁹⁷. Cover crop species represented four diverse functional groups: (i) cool season grasses, (ii) brassicas, (iii) legumes, and (iv) non-legume summer annuals. Additionally, to evaluate the genetic diversity typically found in cover crop seed sources, three cultivars of hairy vetch were included (**Figure 3A**). This selection of cover crops is representative of both fibrous and taproot root systems as some measures of cover crop effectiveness, such as soil physical properties, can be attributed to root morphology. For example, fine-root species have been shown to increase soil aggregate stability, overall porosity, and protect against erosion^{98,99}.

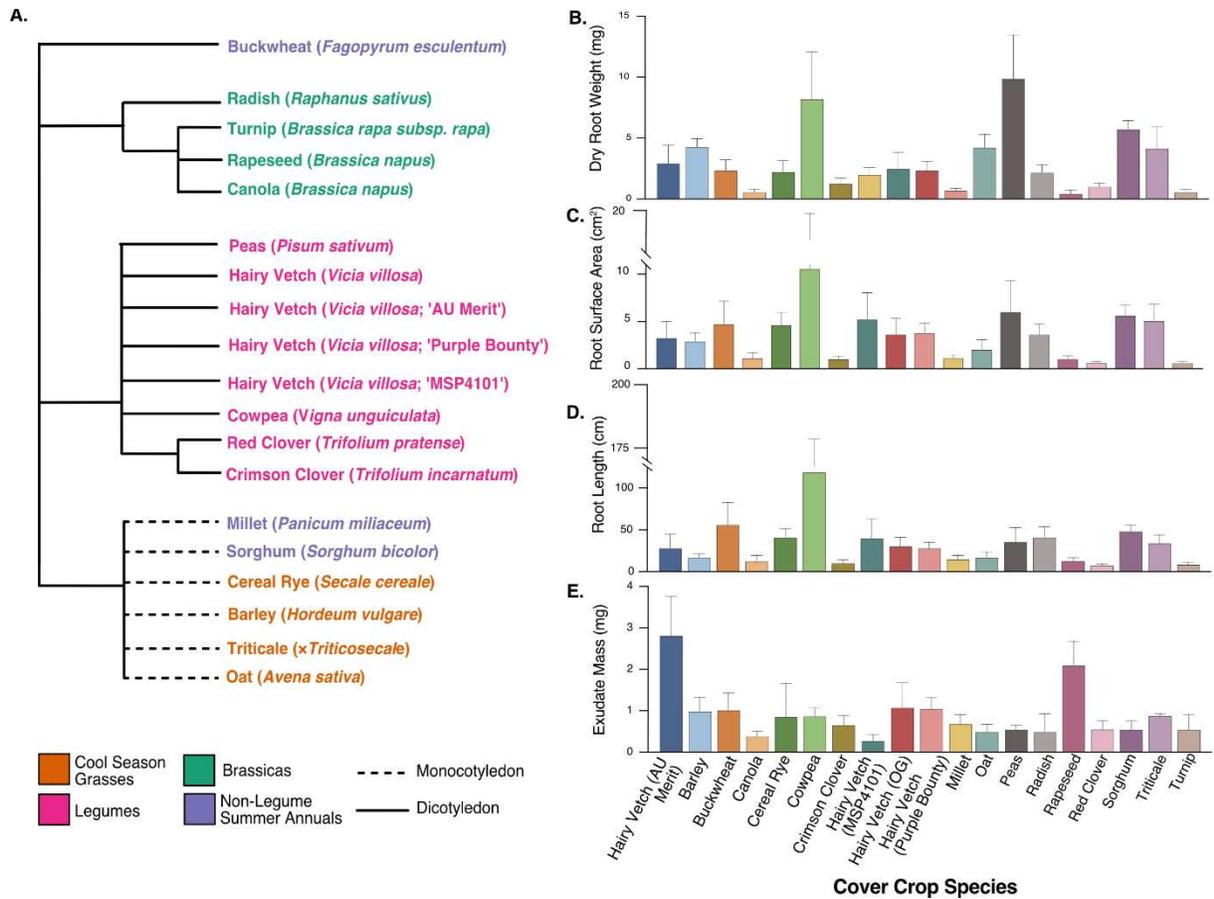


Figure 3. Cover crop phylogeny and root trait analyses. Measurement of three root morphological traits (dry root weight (DRW) mg; root length (cm); and root surface area (cm²)) using WinRHIZO™. A) Cover crop species were phylogenetically clustered using NCBI Taxonomy Browser with NCBI Taxonomy ID (**Appendix A Table S2**) to highlight functional group taxonomic distributions. Functional group annotations are colored in purple (non-legume summer annuals), green (brassicas), pink (legumes) and orange (cool season grasses). Dashed lines indicated monocotyledon species and solid lines indicate dicotyledon species. B-E). Measurement of three root morphological traits colored by cover crop species (**Appendix A File S1**). B) Dry root weight (DRW) (mg) per species; C) root length (cm) per species; and D) root surface area (cm²) per species using WinRHIZO™. E) Root exudation rate (mg/day) per species. Error bars represent 1 standard deviation away from the mean.

Despite these observations, the impact of root morphology on exudate composition and exudation release rate in common cover crop species has yet to be investigated. As expected, cover crop species exhibited high variation between root biomass (DRW), root length, root surface area, and root exudation rate (**Figures 3B-E**). As such, we evaluated the effects of species, DRW, root length, and root surface area on root exudation rate across hydroponically grown cover crop species normalized to a growth timepoint of two weeks. Using a linear model with root exudation

rate as the response variable and species as the predictor variable, we first tested the relationship between cover crop species and root exudation rate and observed that root exudation rate was significantly different across cover crop species (ANOVA $p < 0.001$; **Appendix A Table S3**). Root length, DRW, and root surface area, were also significantly influenced by cover crop species (ANOVA $p < 0.001$; **Appendix A Table S4-6**). These findings align with others who have reported on the phylogenetic control between crops species and specific functional root traits¹⁰⁰. In the context of cover crops, understanding species-specific root exudation rates (both in quantity and composition) may confer functional ecosystem advantages to the development of a productive rhizosphere microbial community.

To understand potential drivers of rhizosphere development, whether morphological, chemical or both, we explored interactions between root morphology, genotype, and exudation rate. This was tested using a linear model where species, root length, DRW, and root surface area were included as predictor variables. The results revealed no significant association to explain the variances in root exudation rate per day ($p = 0.925$), suggesting that the root morphological traits we measured were not sufficient to explain root exudation rates in these cover crop species, a result that also aligns with previous studies¹⁰¹. However, because the range of root morphologies observed may interact broadly with soils, or exhibit adaptive changes to hydroponics systems, it seems unlikely that root morphology in cover crop species is entirely inconsequential for plant growth and root exudation rate outcomes. The absence of evidence for associations between root exudation rate and root morphological traits in our experiment may be a function of the limited set of characteristics measured. For example, previous studies have suggested that measurement of root diameter and root tissue density are more predictive of exudation rate¹⁰². Alternatively, it could be that, as suggested by our data, plant genotype represents a stronger factor in root exudation rate than root morphology.

2.2.1 Root exudate quality and quantity is influenced by plant genotype.

Plants release varying quantities and types of root exudates according to plant genotype, environment, and developmental stages^{14,16} which systematically alters rhizosphere community structure and biogeochemical cycles that ultimately aid in plant fitness. Thus, the dynamic chemical composition of root exudates released from cover crops represents a primary target for enhancing agricultural rhizosphere function and composition to meet desired plant and soil health outcomes. Using a controlled hydroponic plant growth system, we collected pure root exudates from 16 commonly used cover crop species in U.S. agriculture⁹⁶, as well as 3 distinct cultivars within the hairy vetch species. Collecting root exudates from hydroponic systems ensures the collection of significant quantities of purified root exudates without interference from soil sorption, soil matrix effects, organic matter breakdown, or microbial transformation which presents substantial impediments to collecting root exudates from soil-grown plants¹⁰³. Hydroponic growth systems are easy to implement and maintain, allowing for the inclusion of 19 cover crops which represents, to our knowledge, the broadest survey of cover crop root exudation profiling to date.

Previous studies assessing root exudate variability across plant species normalize to a plant growth stage (i.e., seedling). Due to the array of cover crop species diversity in our study, plants reached seedling growth at marginally different rates, and thus using this strategy would introduce variability in root exudation composition as plants are known to change exudation strategies depending on growth stage¹⁷. To account for this, we chose to normalize to a two-week growth timepoint which most closely reflects the seedling stage normalized across all species. After growth for two weeks, we collected the soluble fraction of root exudates from each cover crop species and profiled a broad range of exudate compounds using a combination liquid- and gas chromatography-mass spectrometry platforms (LC-MS/MS and GC-MS). Specifically, non-targeted GC-MS and LC-MS/MS analyses were performed to capture primary and secondary metabolites, respectively, and a targeted LC-MS/MS analysis was performed to detect low

abundance phytohormones. However, it is important to note that there will still be a fraction of root exudates that remain uncharacterized (i.e., unknown metabolites) or that were undetected due to analytical choices (i.e., large polymers or proteins).

Here, we report the detection of 164 annotated metabolites. Of which, a total of 14 compounds were identified at a level 1 annotation, 56 compounds at the level 2 annotation, and 94 compounds at a level 3 annotation according to the Metabolomics Standards Initiative¹⁰⁴. To broadly evaluate cover crop root exudate composition, we taxonomically classified root exudates using the Global Natural Products Social (GNPS) analysis tool, Molecular Network Enhancer (MolNetEnhancer)¹⁰⁵, and ClassyFire¹⁰⁶ into the following superclasses: lipids and lipid-like molecules (49), phenylpropanoids and polyketides (39), organic acids and derivatives (36), organic compounds (15), organic oxygen compounds (12), organoheterocyclic compounds (7), benzenoids (5), homogenous non-metal compounds (1), and organic nitrogen compounds (1) (**Figure 4A**). Our results demonstrate that at a two-week growth timepoint, cover crop species exude a wide array of compounds ranging from sugars, amino and organic acids, phytohormones, terpenes, lipids, and many other secondary plant metabolites (**Appendix A Figure S1**). Furthermore, the high chemical diversity observed between species aligns with previously reported data for similar crop species grown in both hydroponics¹⁴ and field settings^{100,107}. To explore the chemical diversity among species, we used the number of annotated compounds within each species to define chemical richness used here as a proxy for chemical diversity. An ordinary one-way ANOVA demonstrated that the chemical richness significantly depended on species (**Figure 4B**; ANOVA, $p < 0.0001$; **Appendix A Table S7**). Chemical richness was highest in all hairy vetch cultivars (OG, Purple Bounty, AU Merit and MSP4101), cowpea, and peas ($p < 0.0001$) and lowest in triticale, radish, canola, and millet ($p < 0.0001$).

We next wanted to determine if chemical richness significantly differed across cover crop functional type, as cover crops are typically chosen based on functional capacity within the soil,

and secondarily by species. Using a linear model, species were categorized into either legumes, cool season grasses, brassicas, or non-legume summer annual functional groups (**Figure 3A**). We found that the number and types of detected metabolites was significantly different between the legumes and all other functional groups (**Figure 4C and Appendix A S2 ANOVA, $p < 0.0001$; Appendix A Table S8**). Legumes are broadly considered one of the most effective cover crop groups due to species diversity within the group, versatile plant secondary metabolite production, and capacity for symbiosis between nitrogen fixing rhizobia and legume roots³⁵. In our data, legume cover crops released the widest range of compounds with 91.5% of the detected metabolites represented in at least one legume species.

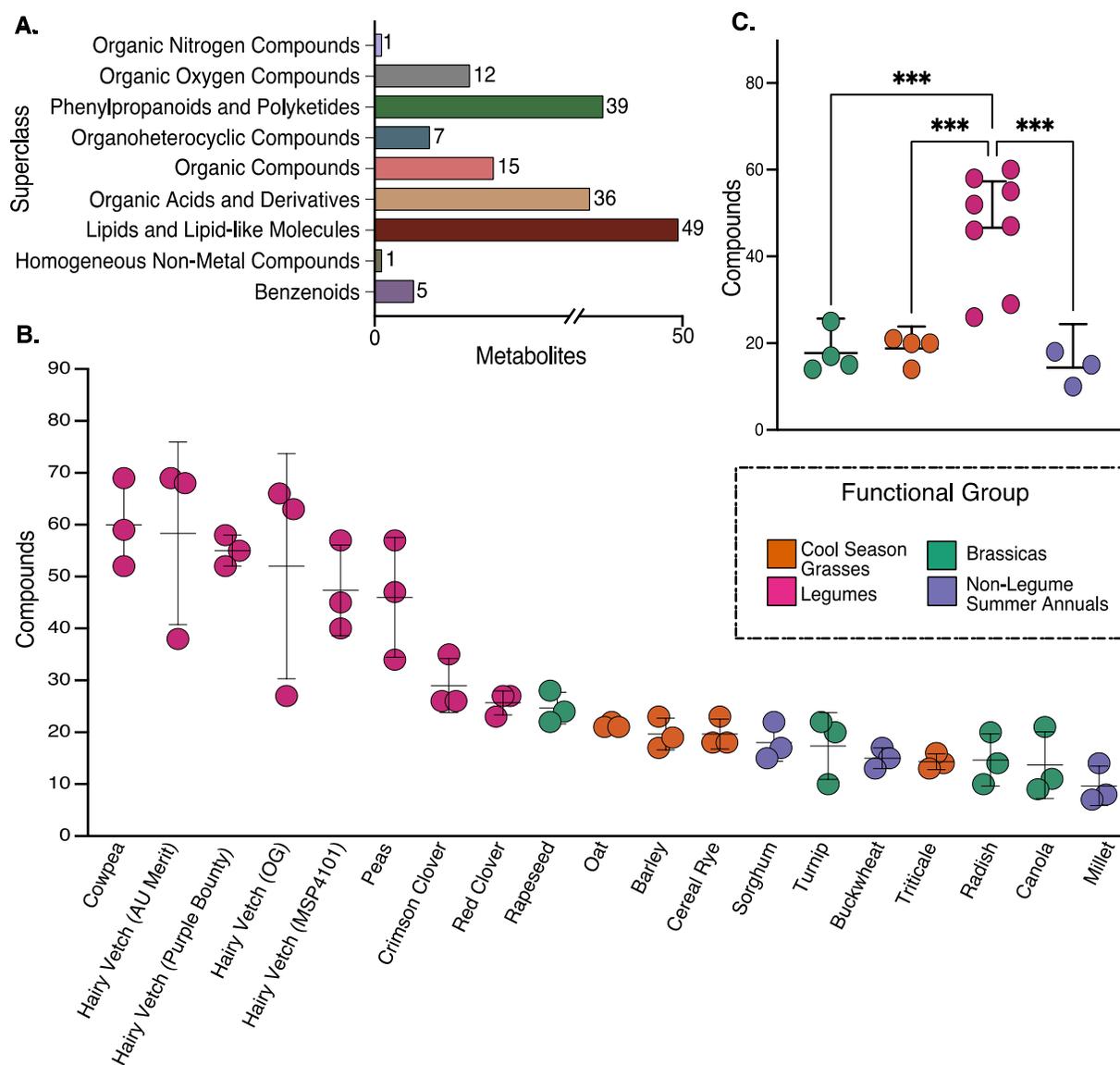


Figure 4: Chemical composition and richness analysis of root exudates. A) Root exudates were classified using MolNetEnhancer¹⁰⁵ (for non-targeted LC-MS/MS) and ClassyFire¹⁰⁶ (for GC-MS and targeted LC-MS/MS) at the superclass level. From 164 features, nine superclasses were identified: lipids and lipid-like molecules (49), phenylpropanoids and polyketides (39), organic acids and derivatives (36), organic compounds (15), organic oxygen compounds (12), organoheterocyclic compounds (7) benzenoids (5), homogenous non-metal compounds (1), and organic nitrogen compounds (1). Bars represent the number of compounds within each superclass. B) Chemical richness of root exudates by species ordered from highest to least average compound counts per species colored by functional group. Each circle represents a replicate sample. Error bars represent one standard deviation from the mean chemical richness count. Chemical richness was determined using z-score normalized peak areas. Normalized peak areas above the mean (greater than 0) were included as a count for that metabolite. Metabolites with a normalized peak area below the mean (less than 0) were assigned a zero count. The summation of normalized peak areas above the mean (greater than 0) resulted in the chemical richness count. C) Chemical richness of root exudates by cover crop functional group. Cover

crop species within each functional type were grouped together (legumes (7), cool season grasses (4), brassicas (4), non-legume summer annuals (3)). Circles represent the average number of detected metabolites per species within functional group. *** is equivalent to $p < 0.001$. Error bars represent 95% CI from the mean.

Legume root exudate heterogeneity is not unexpected, as regulation of root exudation is an imperative fitness response in controlling for favorable plant-microbiome interactions. For example, a range of flavonoid root exudates are known to play an essential role in plant fitness³⁷. Namely, flavonoid biosynthesis in legumes is fundamental for establishing the symbiosis (through induction of rhizobial nod genes) between legumes and their nitrogen-fixing symbionts, the rhizobia. We detected the presence of four flavonoids known to participate in nodule induction in legume species: kaempferol-7-O-hexoside, apigenin, quercetin-3,4'-O-di-beta-glucoside, and 3-genistein-8-C-glucoside^{108,109}. In addition, we detected 23 compounds classified as flavonoids and 5 compounds classified as isoflavonoids (subclass level) reflecting the diversity of flavonoid synthesis for root exudation across legume species.

The pattern of chemical richness across species and functional group was also reflected in the multivariate modeling of metabolite composition, as revealed by an unsupervised principal components analysis (PCA) (**Appendix A Figure S3**), and a supervised partial least squares discriminate analysis (PLS-DA) (**Figure 5**). PCA analysis visualized differences in the root exudate profiles broadly across species (**Appendix A Figure S3**) with trends of chemical compositional differences between species shown in the PLS-DA loadings (**Appendix A Figure S4 and S5**). Variable importance in projection (VIP) scores were derived from the PLS-DA model to estimate the importance of each variable in the projection used in model. 54 metabolites had $VIP > 1.0$, 13 metabolites showed $VIP > 1.5$, but only two metabolites had $VIP > 2.0$ (soyasaponin Bb ($VIP = 2.1049$) and soyasaponin Bb-like molecule (loss of H_2O) ($VIP = 2.12$)) (**Appendix A Figure S6 and Table S9**). Results of the PLS-DA and VIP scores show that broad differences between groups are primarily accounted for by a relatively small subset of metabolites.

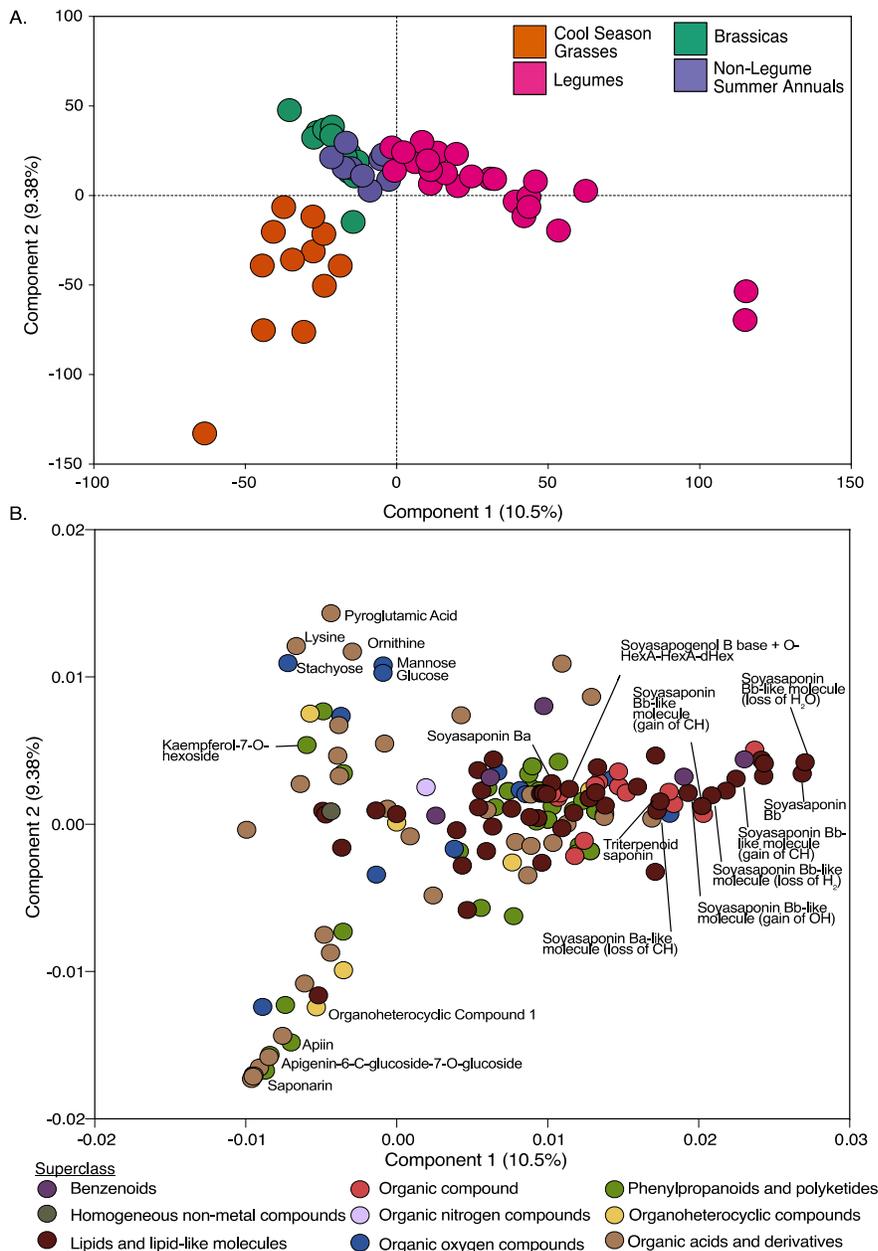


Figure 5. Multivariate analysis of cover crop root exudation by functional type. A) Two-dimensional scores plot of PLS-DA ($R^2X = 0.105$, $R^2Y = 0.0938$, $Q^2 = 0.662$) of root exudates by functional cover crop group. Each colored dot represents a species replicate within a functional group: legume (pink), cool season grasses (orange), brassicas (green), non-legume summer annuals (purple). B) PLS-DA metabolite loadings colored by superclass of exudate compounds. Purple = benzenoids, dark green = homogenous non-metal compounds, dark red = lipids and lipid-like molecules, red = organic compound, light purple = organic nitrogen compound, blue = organic oxygen compounds, light green = phenylpropanoids and polyketides, yellow = organoheterocyclic compounds, tan = organic acids and derivatives. Metabolites mentioned within the text are called out on the PLSDA loadings.

Legume and cool season grasses each clustered into functional group metabolite profiles, driven primarily by the presence of compounds from the lipids and lipid-like molecules superclass within the legume functional group and phenylpropanoids and polyketides in the cool season grasses. PLS-DA model VIP scores revealed differences separating the legumes from other functional groups is driven by the exudation of sixteen compounds from the superclass lipids and lipid-like molecules, six organic compounds, two benzenoids (4-aminobenzoic acid and benzoic acid), pinitol, and aminomalonic acid (**Appendix A Figure S6**; VIP > 1.2). Separation of cool season grass cover crop species from other functional groups was found to be primarily due to the exudation of one organoheterocyclic compound (VIP = 1.09) and phenylpropanoids and polyketides from the flavonoid glycosides subclass (apiin, saponarin, apigenin-6-C-glucoside-7-O-glucoside), however the later compounds were not found to be significant (VIP < 1.0).

Additionally, within the four hairy vetch cultivars (AU Merit, OG, Purple Bounty, MSP4101) that were included to investigate cultivar-dependent root exudation, differences in the root exudate profiles were appreciable using an additional PLS-DA model from the four hairy vetch cultivars (**Appendix A Figure S7A**). PLS-DA loadings and VIP scoring highlight the production of diverse secondary metabolites in the hairy vetch cultivars. Specifically, Purple Bounty enriches for three soyasaponin Bb isomers and other unknown metabolites belonging to the phenylpropanoids and polyketides superclass while AU Merit and MSP4101 enrich for root exudates in the superclasses of organic oxygen compounds and organic acids and derivatives (**Appendix A Figure S7B and S7C**). In particular, AU Merit enriches for the organic acids asparagine, salicylic acid, and aminomalonic acid (VIP > 1.5). These data suggest hairy vetch root exudation strategies are tightly genetically controlled, varying across cultivars, highlighting the impact that plant genotype may have on the soil through root deposition of secondary metabolites.

Non-legume summer annuals and brassicas, on the other hand, exhibit similar root exudate profiles containing compounds from the superclasses organic acids and derivatives, organic oxygen compounds, and phenylpropanoids and polyketides (**Appendix A Figure S5**). Still, brassica species root exudate profiles were significantly enriched for citric acid, pyroglutamic acid, ornithine, lysine, and stachyose as compared with the other functional groups (**Figure 5B and Appendix A S6; VIP > 1.5**). These compounds are primarily regarded for their use in microbial regulation or as carbon source; however, brassicas are non-mycorrhizal and are also known to exude organic acids as one strategy to enhance soil phosphorus availability¹¹⁰. Although brassicas are a functional group known to exhibit pest and weed-suppressive effects on the rhizosphere through the release of glucosinolates, we did not expect to see those compounds here in the absence of both allelopathy and microbial interference due to the nature of hydroponics. Non-legume summer annuals, however, were not found to have any significantly abundant root exudates. This observation is likely because non-legume summer annuals were grouped by planting season and not entirely on shared similarities in soil functional utility. In fact, sorghum and millet, two members of the non-legume summer annuals group are more phylogenetically related to cool season grasses while buckwheat is considered a pseudocereal (**Figure 3A**). Therefore, it is an unsurprising result that non-legume summer annuals lacked functional-group specificity and did not form a distinct root exudate profile.

2.2.3 Identification of two soyasaponins, a soyasapogenol aglycone base, and sixty-nine soyasaponin molecular family members highlight legume cover crop functional capacity for soyasaponin synthesis.

Feature-based molecular networking (FBMN) was then used to visualize and annotate the chemical complexity of root exudates analyzed via non-targeted LC-MS/MS using GNPS. GNPS generated twenty-five annotated molecular families using 372 MS/MS spectra (**Appendix A Figure S8**) resulting in 91 compound annotations at various annotation levels. The size and

robustness of the soyasaponin molecular family suggested importance and led us to further explore these compounds using GNPS tools.

Soyasaponins are triterpenoid saponins commonly found in legume plant tissues consisting of a soyasapogenol aglycone and oligosaccharide moieties¹¹¹. They can be further classified into four groups according to the aglycone backbone, sugar moieties, and functional group attachments. Group A, B, and E soyasaponins are derivatives from glycosides of soyasapogenols A–E, respectively while the fourth group, DDMP (2,3-dihydro-2,5-dihydroxy-6-methyl-4H-pyran-4-one), is a derivative from the glycoside of soyasapogenol B, containing DDMP residues at C22. Consumption of soyasaponins has been previously investigated in human health for biological activity as anti-inflammatory, antimutagenic, and anticarcinogenic compounds^{112,113}. However, little is known about the effect of soyasaponins from root tissues and root exudates, as soyasaponins were only recently discovered to be released as root exudates in soybeans¹¹⁴. Most work has investigated the role of saponins as specialized plant metabolites highlighting diverse biological functions in the rhizosphere such as acting as antibacterial, antifungal, allelopathic agents^{115,116} or beneficial rhizosphere-shaping metabolites as shown in tomatoes and oat^{117,118}. However, work investigating microbial responses to triterpenoid saponins, like soyasaponins, in the rhizosphere is minimal. Excitingly, soyasaponins were found to be microbially degraded in a soybean rhizosphere experiment which influenced rhizosphere assembly¹¹⁹. However, further work is necessary to determine 1) which legume species are capable of soyasaponin exudation; 2) how soyasaponins are utilized in the rhizosphere; and 3) which microorganisms are capable of degradation. We used FBMN to show quantitative evidence that soyasaponins, and putatively annotated isomers of soyasaponins in the molecular family, are uniquely enriched in all legume cover crops, highlighting species which have yet to be profiled for exudation of soyasaponins in current literature.

The soyasaponin molecular family, generated from 72 MS/MS spectra, resulted in the identification of two group B soyasaponins (soyasaponins Ba and Bb), one soyasapogenol base aglycone (Bbase+O-HexA-HexA-dHex), eight putatively annotated soyasaponin-like molecules, and sixty-one unannotated members of the soyasaponin molecular family (unknowns or superclass level annotations) (**Figure 6**). Using GNPS enabled confident identification of soyasaponin Ba and Bb as well as its aglycone base, soyasapogenol B (level 2 annotation). To aid in identification of unknown molecules within the soyasaponin molecular family, manual MS/MS spectrum and mass comparisons were used as a method to identify isomers of soyasaponins Ba and Bb. We identified eight structurally related, but unannotated soyasaponins containing small (< 30Da) mass shifts resulting from the loss or gain of H₂O, O⁻, H₂, CH₂ or CH₂O⁻ from soyasaponin Ba or Bb. The remaining unknown compounds are shown as such or with their chemical superclass annotation (**Figure 6; Appendix A Table S10**).

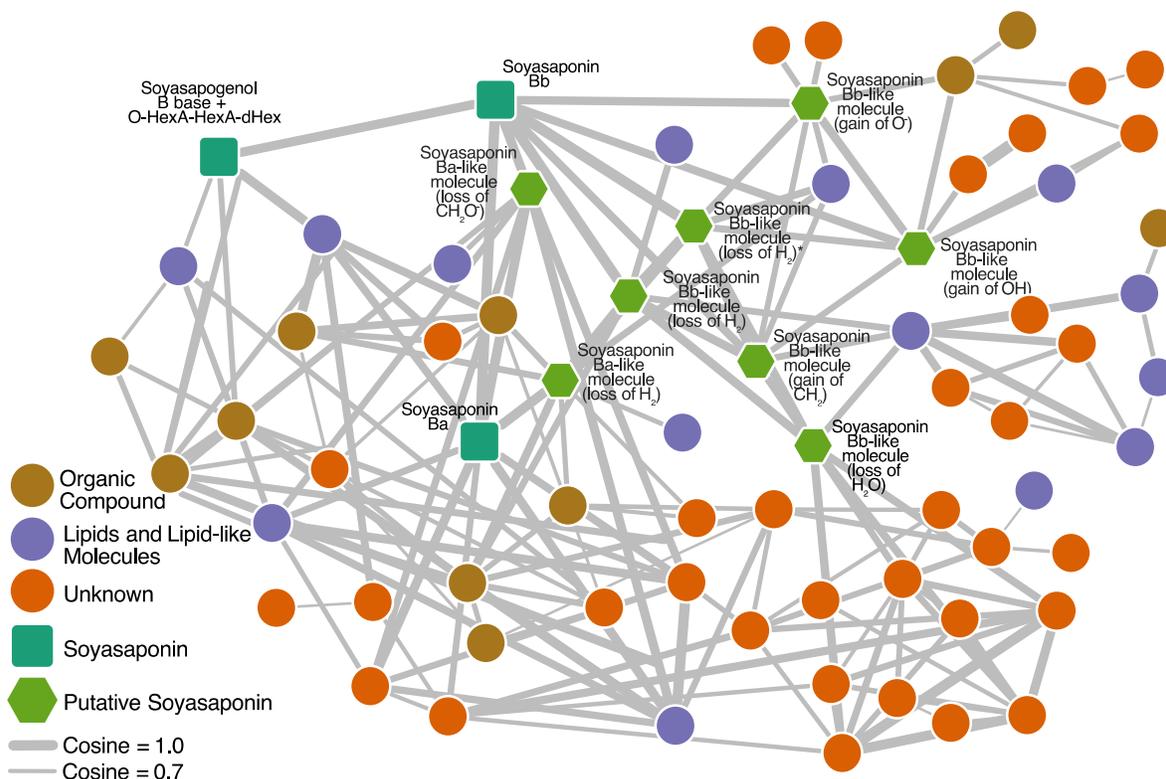


Figure 6. Soyasaponin molecular family and soyasaponin proportions in legumes. FBMN was utilized to generate individual molecular families from annotated compounds. Soyasaponin molecular networking resulted in the annotations of 2 soyasaponins (teal), 1 soyasapogenol aglycone (teal), 8 putatively annotated soyasaponin-like molecules (light green), and 61 unannotated members of the soyasaponin molecular family (purple, orange and brown circles). Node shape represents level of molecule annotation (circles = unknown member of soyasaponin molecular family with level 3 or lower annotation; hexagon = putative annotation assignment based on common functional group gains/losses as visualized using the *Edge Annotation* feature in GNPS; square = level 2 compound assignment). Node color represents compound annotation. Edge size represents cosine similarity score (0 to 1 where 1 indicates an identical spectrum and 0 indicates no similarity; here edge scores range from 0.7 (thinnest line; FBMN spectral similarity cutoff) to 1.0 (thickest line)) between two MS/MS spectra. The full-sized images of the entire network to zoom in on the molecular networks/families can be found as supporting information (**Appendix A Figure S8**). Because there were two soyasaponin Bb-like molecule (loss of H₂) putative annotations, an asterisk (*) denotes which annotation corresponds to that node in subsequent text and figures. The full-sized image of the soyasaponin molecular family, with every node charted as a proportion of soyasaponins found in each species, can be found as **Appendix A Figure S9**.

2.2.4 Soyasaponins exclusively enriched in legume cover crop species.

Legume cover crop species exhibited root exudate profiles that clearly distinguish them from other functional groups and potentially point to selective advantages that prime legumes for specific functional potential in the soil. Of the sixteen significantly abundant compounds belonging to the lipids and lipid-like molecules superclass within legumes (**Appendix A Figure S6**; VIP > 1.5), eleven were structurally related or identified as group B soyasaponins and were exclusively found in legume species (**Figure 7A**; VIP > 1.5). This finding aligns with others who have demonstrated soyasaponin secretion dynamics in soybeans over time and concluded the highest release of group B soyasaponins occurs during the early growth stages in soybean¹¹⁹. Of the eleven soyasaponins annotated, seven were significantly (VIP > 1.5) enriched in the legume functional group: soyasaponin Bb, soyasaponin Ba-like molecule (loss of CH₂O⁻), soyasaponin Bb-like molecule (gain of CH₂), soyasaponin Bb-like molecule (gain of OH), two soyasaponin Bb-like molecules (loss of H₂), soyasaponin Bb-like molecule (loss of H₂O) (**Figure 7B**).

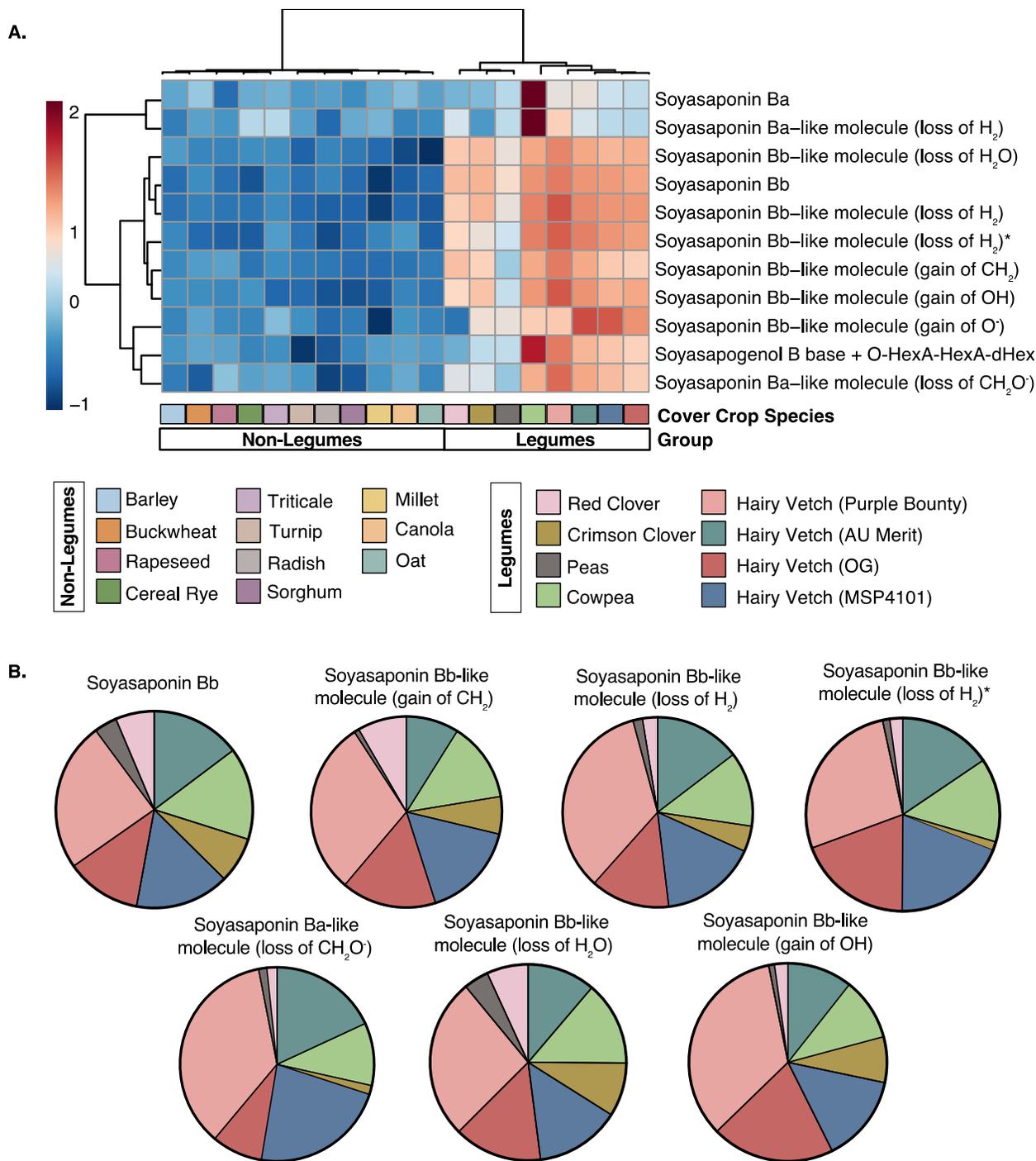


Figure 7. Heatmap of eleven soyasaponins annotated from non-targeted LC-MS/MS data. A) Heatmap showing the relative abundances (represented as a \log_{10} transformation across samples) of 11 root exudates annotated as soyasaponins averaged across each cover crops species. Ward's clustering method was applied with a Euclidean distance measure for both features and samples. Cover crop species are indicated by colors in the legend and organized within the legend as shown in the heatmap from left to right. B) Individual pie graphs of seven significantly abundant soyasaponins in legumes show the proportion of the molecule found in

each legume species or cultivar. Because there were two soyasaponin Bb-like molecule (loss of H₂) putative annotations, an asterisk (*) denotes which annotation corresponds to the second isomer of soyasaponin Bb-like molecule (loss of H₂).

Cowpeas were unique among the legume functional group in that they were significantly enriched (ANOVA, $p < 0.0001$; **Appendix A Figure S10**) for the presence of two soyasaponins (soyasaponin Ba and a soyasaponin Ba isomer (loss of H₂)). Soyasaponin Ba is different from soyasaponin Bb as it contains glucuronic acid and two sugars (galactose and glucose) bound to soyasapogenol B at position C3 while soyasaponin Bb contains glucuronic acid, galactose and rhamnose bound to soyasapogenol B at position C3. Interestingly, cowpea seeds have previously been shown to contain higher concentrations of soyasaponin Ba than other legume seeds, such as common bean, scarlet runner bean, adzuki bean, and chick pea¹²⁰. Because of this distinctiveness to cowpea, soyasaponin Ba and a soyasaponin Ba isomer (loss of H₂) were not found to be significantly important as a legume functional group soyasaponin (VIP < 1.0). Although direct biochemical evidence is lacking to document root exudation in cowpea until now, it is possible the difference in cowpea exudation of soyasaponin Ba could play a role in a selective advantage of this species for optimal rhizosphere assembly as has been suggested for field-grown soybean with soyasaponin Bb¹¹⁹ and for other saponins in tomato and oat^{117,118}. More research is needed to determine the influence of soyasaponins across all legume cover crops and their potential functional role as a plant-specialized metabolite in rhizosphere priming.

2.2 Conclusions

While it is known that cover crops can influence soil health based on traditional physical measures, their influence on the composition, diversity, and function of the soil microbiome is less understood. There is currently very little information on how variation in root exudation between cover crop species, and cultivars within species, can drive functional changes in the soil microbiome and in turn how these changes influence soil health. The results of this study address

a necessary first step to understand variation in cover crop metabolism and allocation of carbon resources through analysis of root exudate composition. We demonstrate that root exudate profiles are heterogeneous across surveyed cover crop species and cultivars (**Figure 5, Appendix A S1 and S2**). The results suggest that this chemical diversity is under phylogenetic control and representative of functional utility as legumes, hairy vetch cultivars within legumes, cool season grasses, and brassicas exhibited functional class separation in their metabolic profiles (**Figure 5 and Appendix A S7**). Notably, legumes stand out as a unique group of plants capable of producing distinct chemical environments rich with complex secondary metabolites, such as a diverse array of triterpenoid saponins, isoflavonoids, and flavonoids (**Figure 6, 7 and Appendix A S6**). Brassicas and cool season grass species were also found to exhibit distinct exudate profiles, albeit with fewer functional group specific compounds (**Appendix A Figure S6**). In light of the metabolic evidence from this study, the extent to which species-specific root exudation from cover crops, or combinations of cover crops, will influence the functional relationship between these crops and the agricultural rhizosphere warrants further investigation. Collectively, our findings pave the way for the development of efficient and precise cover crop management strategies in regenerative agricultural systems enabling the possibility for root exudates to be a source for intentionally shaping favorable agricultural soil microbiomes.

2.3 Materials and Methods

2.3.1 Cover Crop Root Exudate Collections

Mature seeds (see **Appendix A File S1** for seed information) from 19 cover crops were first surface sterilized by placing seeds in 15 mL conical tubes with either (A) 2 mL of Clorox bleach solution (3% sodium hypochlorite. 1:1 bleach to water) + 1 drop of Tween-20 to reduce surface tension and improve serialization or (B) 2 mL of Clorox bleach solution (3% sodium hypochlorite. 1:1 bleach to water) (see **Appendix A File S1** for sterilization optimization) in a sterile tissue culture hood. Sterilization method A or B was chosen based on optimization of

germination rate for each cover crop species (unpublished data). Seeds suspended in sterilization solution were shaken for 1 min. In a sterile tissue culture hood, bleach was removed using sterile techniques and the seeds were rinsed five times with sterile deionized (DI) water. Seeds were either (A) plated to agar plates with MS basal salt mixture (2.16 mg Murashige and Skoog (MS) media in 500 mL sterile DI water; MP Biomedicals, Santa Ana, CA) germinated in the dark for 3 days, and then transferred to sterile tubes with 3 mL liquid MS media after 3 days; or (B) placed directly in sterile growth vessel with 3 mL MS media. Growth vessels were either glass tubes which were used for monocotyledon species or Magenta boxes which were used for dicotyledon species based on preliminary experiments to optimize growth of each seedling (unpublished data). All seeds were incubated in a growth chamber with photoperiod 16 h light/8 h night at 25°C for 14 days. The 14-day-old seedlings did not introduce any microbes to the system as no bacterial growth was observed on the MS agar plates or growth containers. Germination method was optimized for each species (see **File Appendix A S1** for germination rate and method for each species). 18 seedlings of each cover crop species were grown for 14-days before root exudate collections. Root exudates were collected by first rinsing the roots in sterile DI water 3 times and transferring to a new, sterile vessel filled with 5mL sterile DI water for a 24-h root exudate collection period. The root exudate suspensions containing the root exudates were then filtered through a 0.2µm filter membrane to remove root detritus. 6 root exudate suspensions were pooled for 3 biological replicates of each species to achieve enough pure root exudates for analysis (n=18/6 for 3 biological reps). Root exudate suspensions were frozen at -20°C in 200 mL sterile glass jars. Frozen root exudate samples were lyophilized completely (~72 h) before resuspension in 1.5 mL sterile HPLC-grade water (Thermo Fisher Scientific, Waltham, MA). Resuspended exudates were vortexed thoroughly to remove all residue from the bottle (1min of vortexing along all edges of the bottle). Samples were then aliquoted into 2 mL glass autosampler vials (three 500 mL fractions for each of the three analytical platforms used) and dried completely under N₂(g). Total root exudate weights for each bioreplicate can be found in **Appendix A File S1**. For each

analysis, a volume equivalent to 0.25 mg of root exudates was transferred to a new, pre-weighed 2 mL glass autosampler vial and dried under N₂(g).

2.3.2 Root Scanning, Analysis, and Drying

After exudate collections, roots were carefully separated and evenly spread out into deionized water on a transparent tray (300 x 200 mm) and scanned using an Epson Perfection V750 Pro Photo Scanner (Epson, Los Alamitos, CA). After image acquisition, roots were analyzed using the 2013 Arabidopsis WinRHIZO™ Pro software (Regent Instrument Inc, QC, Canada). Total root length and surface area were analyzed via the use of WinRHIZO™ Pro analysis software and can be found in **Appendix A File S1**. Roots were then blotted dry and gently placed on a tray with paper towels, dried at 70°C for 72 h in an oven, and weighed (see **Appendix A File S1** for DRW).

2.3.3. Targeted UPLC-MS/MS for Phytohormone Analysis of Root Exudates

Targeted detection of phytohormones was performed as previously described¹²¹. Briefly, 0.25 mg subsamples were extracted in 75 µL of a spiked methanol solution containing 80% methanol (MeOH) with 65.2 ng/mL ABA-d6, 62.5 ng/mL salicylic acid-d6, and 90.0 ng/mL jasmonic acid-d5 (Sigma Aldrich, St. Louis, MO). After solvent addition, samples were placed on a shaker plate for 1 h at the highest speed setting, centrifuged at 3500 x g at 4°C for 5 min, and transferred to glass inserts. A final centrifuge step at 3500 x g for 15 min at 4°C was completed to ensure any precipitate was in the bulb of the vial insert. Five microliters of root exudate samples were injected onto a Perkin Elmer UPLC MS/MS system, equipped with a PerkinElmer QSight LX50 Solvent Delivery Module (PerkinElmer, Waltham, MA). An ACQUITY UPLC T3 column (1 × 100 mm, 1.8 µM; Waters Corporation, Milford, MA) was used for chromatographic separation. Mobile phase A consisted of LC-MS grade water with 0.1% formic acid and mobile phase B consisted of 100% LC-MS grade acetonitrile (Thermo Fisher Scientific, Waltham, MA). The elution gradient was initially set at 0.1% B for 1 min, which was increased to 55.0% B at 12 min and

further increased to 97.0% B at 15 min, then decreased to 0.1% B at 15.5 min. The column was re-equilibrated for 4.5 min for a total run time of 20 min. The flow rate was set to 200 μ L/min and the column temperature was maintained at 45°C. Samples were held at 4°C in the autosampler. Detection was performed on a PerkinElmer QSight™ 220 triple quadrupole MS in selected reaction monitoring (SRM) mode. The transitions monitored for each phytohormone compound can be found in **Appendix A Table S1**. The MS was operated with ESI voltage 4500 V in positive mode and -3500 V in negative mode. Nebulizer gas flow was set at 350 arbitrary units and drying gas was set to 120 arbitrary units. The source temperature was 315°C and hot-surface induced desolvation temperature was set to 200°C. Pooled QCs were injected after every 6th sample to ensure proper instrument function and to detect any analytical variation.

2.3.4 Non-targeted GC-MS Analysis of Root Exudates

Sample preparation was conducted as previously described¹²². Briefly, dried samples were resuspended in 50 μ L of pyridine containing 25 mg/mL methoxyamine hydrochloride (Sigma Aldrich, St. Louis, MO), centrifuged, incubated at 60 °C for 45 min, vortexed for 30 s, sonicated for 10 mins, centrifuged briefly, and incubated a second time for 45 min at 60 °C. Samples were centrifuged for 2 mins. Then, 50 μ L of MSTFA + 1% TMCS (Thermo Fisher Scientific, Waltham, MA) was added, samples were vortexed for 30 s, centrifuged, and incubated a third time at 60°C for 35 min. Samples were cooled to room temperature, centrifuged, and 80 μ L of supernatant was transferred to glass vial inserts within glass vials. Samples were centrifuged a final time for 10 min before analysis. Metabolites were separated with a 30 m TG-5MS column (Thermo Fisher Scientific, Waltham, MA, 0.25mm i.d. 0.25 μ m film thickness) and detected using a PerkinElmer Clarus 690 GC coupled to a Clarus SQ 8S mass spectrometer (PerkinElmer). Samples (1 μ L) were injected at a 12:1 split ratio onto the column with a 1.0 ml/min helium gas flow rate. The gas chromatography inlet was held at 285°C, and the transfer line was held at 300°C, and the source temp was held at 260°C. The GC oven program started at 80°C for 30 s, followed by a ramp of

15 °C/min to 330°C, followed by an 8 min hold. Masses between 50–620 m/z were scanned at 4 scans/s under electron impact ionization. Pooled QCs were injected after every 6th sample to ensure proper instrument function and to detect any analytical variation.

2.3.5 Non-targeted FastDDA LC-MS/MS Analysis

0.25 mg subsamples were extracted in 75 µL 80% MeOH in LCMS-grade water (Thermo Fisher Scientific, Waltham, MA). After solvent addition, samples were placed on a shaker plate for 1 h at the highest speed setting, centrifuged at 3500 x g at 4°C for 5 min, and transferred to glass inserts. A final centrifuge step at 3500 x g for 15 min at 4°C was completed to ensure any precipitate was in the bulb of the vial insert. Five microliters of extract were injected onto a Waters Acquity UPLC system (Waters Corporation, Milford, MA). Separation was achieved using a Waters Premier Acquity Premier VanGuard FIT HSS T3 column (Waters Corporation, Milford, MA; 1.8 µM, 2.1 x 100 mm), using a gradient from solvent A (LC-MS Water, 0.1% formic acid) to solvent B (LC-MS Acetonitrile, 0.1% formic acid). Injections were made in 99% A, held at 99% A for 1 min, ramped to 98% B over 12 minutes, held at 98% B for 3 minutes, and then returned to starting conditions over 0.05 minutes and allowed to re-equilibrate for 3.95 minutes, with a 600 µL/min constant flow rate. The column and samples were held at 65°C and 6°C, respectively. The column eluent was infused into a Waters G2-XS Q-TOF-MS with an electrospray source in positive mode, scanning 50-1200 m/z at 0.1 seconds per scan, using Waters FastDDA (Data Dependent Acquisition Algorithm) to select up to 4 precursor ions per MS scan at a collision energy ramp of 23 V. Dynamic exclusion was set such that once a precursor was sampled, it was excluded for 6 seconds, and the full DDA cycle was limited to 0.1 s to ensure sufficiently responsive precursor selection. Waters software AutocatV1 was used to exclude previously scanned masses within the sample reps. Calibration was performed using sodium formate with 1 ppm mass accuracy. The capillary voltage was held at 700 V, source temperature at 150°C, and nitrogen desolvation temperature at 600°C with a flow rate of 1000 L/h.

2.3.6 Metabolomics Data Analysis

Non-targeted GC-MS data were processed within the R statistical software¹²³ using methods previously described¹⁸. For GC-MS samples, .cdf files were processed through the following workflow: 1) XCMS software was used for preprocessing to identify molecular features¹²⁴; 2) features were further normalized to total ion current (TIC); 3) the package RAMClustR¹²⁵ was used for clustering features into spectra and prepared for subsequent spectra identification in RAMSearch¹²⁵ via spectral searching against external databases including Golm (<http://gmd.mpimp-golm.mpg.de/>) and NIST (<http://www.nist.gov>). Prior to TIC normalization, features were normalized by linearly regressing run order versus QC feature intensities to account for instrument signal intensity drift. Z-scored GC-MS peak area can be found in **Appendix A File S2**.

For phytohormone analysis of root exudates, LC-MS/MS data were processed using Simplicity 3Q (v1.5, Perkin Elmer, Waltham, MA) bioinformatics software. Briefly, the peak area for each phytohormone compound was normalized to the corresponding internal standard peak area and quantification was assessed using a linear regression against an external calibration curve (**Appendix A Table S10**). Phytohormone concentrations can be found in **Appendix A File S2**.

For LC-MS/MS Data Dependent Acquisition (DDA) data, .mzML files were first processed in MS-DIAL¹²⁶ using the conventional DDA MS Method Type Data collection type was set to centroided with the following parameters: MS1 tolerance 0.01 Da, MS2 tolerance 0.025 Da, MS1 mass range begin: 50 Da, MS1 mass range end: 1200 Da, MS2 mass range begin: 50 Da, MS2 mass range end: 1200 Da. Peak detection parameters were as follows: minimum peak height 1000, mass slice width 0.1 Da and smoothing was set to linear weighted moving average with a smoothing scan of 1 and minimum peak width of 5. Identification was conducting using GNPS public libraries with default MS-DIAL parameters. Adducts chosen were: [M+H]⁺, [M+NH₄]⁺,

[M+Na]⁺, [M+K]⁺, [M+H-H₂O]⁺. MS2Dec was set to default parameters. Alignment parameters used a retention time of 0.05 min and an MS1 tolerance of 0.015 Da. Features were removed based on blank information at a 5-fold sample max/blank average ratio. 'Suggested (w/o MS2) metabolite features' were excluded. Raw peak area data matrix and the GNPS export feature was then used to generate the MS/MS spectral summary file containing the list of representative MS/MS spectra (the most intense MS/MS spectrum) and the feature quantification table containing compound intensity and annotation. Z-scored LC-MS/MS peak areas can be found in **Appendix A File S2**

2.3.7 Molecular Networking and Spectral Library Search

A molecular network was created with the Feature Based Molecular Networking workflow¹²⁷ on GNPS (<https://gnps.ucsd.edu>) The MS/MS spectral summary file and the quantification results file generated from MS-DIAL were imported to GNPS for FBMN analysis. The data were filtered by removing all MS/MS fragment ions within +/- 17 Da of the precursor m/z. MS/MS spectra were window filtered by choosing only the top 6 fragment ions in the +/- 50 Da window throughout the spectrum. The precursor ion mass tolerance was set to 0.01 Da and the MS/MS fragment ion tolerance to 0.01 Da. A molecular network was then created where edges were filtered to have a cosine score above 0.7 and more than 4 matched peaks. Further, edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. Finally, the maximum size of a molecular family was set to 100, and the lowest scoring edges were removed from molecular families until the molecular family size was below this threshold. The spectra in the network were then searched against public GNPS spectral libraries. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least 6 matched peaks. The individual molecular networks were downloaded from GNPS and visualized using Cytoscape software (v3.9.1). The Network Annotation

Propagation (NAP) tool was used in GNPS to propagate network results to unknown metabolites¹²⁸.

2.3.8 Compound Superclass Annotations

The GNPS tool, MolNetEnhancer¹⁰⁵, was applied to the molecular network from non-targeted LC-MS/MS data to determine chemical superclass classification. Classyfire¹⁰⁶ was used to manually assign superclass annotations to GC-MS and targeted LC-MS/MS data.

2.3.9 Statistics

Significance of association between cover crop species, DRW, root surface area, root length, and root exudation rate was modeled using the *lm* command for fitting linear models from the stats (v4.4.0) package in R (v4.2.2). The *anova* command was then used to compute analysis of variance tables for the fitted model objects. Peak area files for each analysis were combined into one file (**Appendix A File S2**) and normalized by z-scoring. Chemical richness of species and functional groups was determined using z-score normalized peak areas. Normalized peak areas above the mean (greater than 0) were included as a count for that metabolite. Metabolites with a normalized peak area below the mean (less than 0) were assigned a zero count. The summation of normalized peak areas above the mean (greater than 0) resulted in the chemical richness count and were used in subsequent analyses. To determine differences in chemical richness between species and functional groups, an ordinary one-way ANOVA was conducted in GraphPad Prism (v9.4.1) using grouped analysis with Tukey's multiple comparison testing with alpha = 0.05. Multivariate statistics were conducted in SIMCA (v17.0.1) to generate PCA and PLS-DA models. PCA and PLS-DA was performed using annotated and unannotated metabolites using z-scored and UV-scaled data. The list of scores and loadings coordinates were then plotted in GraphPad Prism. Prism was used to visualize PCA and PLS-DA scores and PLS-DA loadings. VIP scores were generated using the PLS-DA Important Features tool in MetaboAnalyst (v5.0) using annotated and unannotated z-scored metabolites and pareto-scaled data to identify

metabolites that were significantly different between functional groups. The top 40 metabolites were used to generate VIP scoring and plotted according to VIP score. Heatmaps were generated in MetaboAnalyst (v5.0) using annotated and unannotated z-scored metabolites and pareto-scaled data, standardized by autoscaling features, and clustered using a Euclidean distance measure and Ward's clustering method. ANOVA-significant compounds and cowpea soyasaponin significance analysis was generated in MetaboAnalyst using annotated and unannotated z-scored metabolites and pareto-scaled data using the One-Way ANOVA analysis option with a *p*-value cutoff of 0.05.

CHAPTER 3: VARIATION IN ROOT EXUDATE COMPOSITION INFLUENCES SOIL MICROBIOME MEMBERSHIP AND FUNCTION²

3.1 Introduction

The region where plant roots interface with soil known as the rhizosphere, is one of the most intricate and diverse microbial ecosystems on earth¹²⁹. It is widely recognized that exudation of compounds from plant roots is key to assembling these rhizosphere microbial communities¹³⁰⁻¹³², which in turn can influence plant health and soil biogeochemistry. Plant roots exude up to 40% of their photosynthetically derived carbon, with these compounds ranging from small (e.g. carbohydrates, amino acids, organic acids, phytohormones) to large (e.g. proteins, mucilage) molecules^{19,133}. These exudates provide microbial substrates, as well as signaling molecules, that stimulate microbial activity. For instance, rhizosphere enriched microorganisms can suppress plant pathogens and provide plant nutrients through nitrogen fixation and metal chelation, among other mechanisms^{130,134-136}. Moreover, some members of the rhizosphere can produce plant hormones like indole-3-acetic acid (IAA), gibberellins, cytokinins, and abscisic acid, which modulate plant physiological health^{137,138}. The importance of these various contributions highlights the clear need to understand the mechanistic link between plant root exudate chemistry and metabolic changes in the soil microbiome.

Recently there has been an increased focus on plant rhizodeposition and its impact on soil microorganisms. To enhance tractability in a complex system, many of these keystone studies selected model plants (e.g. *Brachypodium distachyon*, *Arabidopsis thaliana*) and

² This chapter was reproduced verbatim from “Seitz, et al. Variation in Root Exudate Composition Influences Soil Microbiome Membership and Function. *Applied and Environmental Microbiology* (2022)”. This manuscript was co-written with BBM, with my contributions as the following: experimental design planning, sample preparation, 16S and exometabolomics data analysis and interpretation, manuscript writing. The text benefitted from writing and editing contributions from other contributing authors and reviewers selected by the publisher. The ordering of the materials in this dissertation are consistent with the content available online but have been renumbered to reflect incorporation into this dissertation.

evaluated their impact on known plant growth promoting rhizobacteria (e.g. *Pseudomonas*)^{16,17,20,139}. However, the scaling of these concepts to more complex chemical profiles and soil microbial communities has been constrained by methodological challenges. For example, where exudates were chemically profiled, they were often limited to a small set of known metabolites¹⁴⁰⁻¹⁴³, or when microbial communities were analyzed they often relied on methods that tracked changes in membership (16S rRNA, Phospholipid Fatty Acid profiles) but not metabolic potential over the course of exudation^{144,145}. Therefore, there is a current need for studies that employ high-resolution microbial and chemical methods to begin to mechanistically describe the exudate-microbe interactions in agriculturally relevant crop species.

Here, we used a multi-omics approach to characterize how root exudation from a model crop plant can drive soil microbial community structure and function. Using three common, but phenotypically diverse, Sorghum (*Sorghum bicolor* (L.) Moench) genotypes, we determined the variation in root exudate chemical composition amongst these plant varieties. To narrow in on the microbial response to these different exudate profiles, we created plant-free soil microcosms and fed these with two sorghum-informed exudate amendments. We tracked these microcosms over 20 days using exometabolomics, 16S rRNA gene profiling, and genome-resolved metagenomics. Integrating these data, we show that different sorghum exudate treatments led to distinct microbial communities and metabolisms, including the production of three phytohormones. Harnessing this knowledge could support the growing need for sustainable agroecosystems by developing holistic agricultural management strategies that optimize the metabolic capabilities of the soil microbiome.

3.2 Results & Discussion

3.2.1 Sorghum genetics influence root exudation patterns

Sorghum is one of the most widely produced agricultural crops in the world, serving as a grain, forage, and cover crop for use as human food, livestock feed, and as a biofuel feedstock^{122,146}. Due to its diverse agronomic uses, we leveraged three sorghum genotypes known to have

distinct aboveground phenotypes that we hypothesized would contribute to distinct root exudate profiles. We selected 1) the grain sorghum, BTx623; 2) the sweet sorghum type, Leoti; and 3) the bioenergy sorghum PI 505735^{122,147,148}. We grew sorghum genotypes *in vitro* (hydroponically) for seven days, and soluble exudates were collected in water and analyzed with non-targeted gas chromatography-mass spectrometry (GC-MS).

Metabolites from sorghum seedlings spanned known root exudate metabolite classes ranging from sugars, sugar alcohols, organic acids, and amino acids (**Figure 8A, Appendix B Figure S1**). Chemical differences were observed between genotypes BTx623 and Leoti, while PI 505735 represented an intermediate root exudate chemical profile (**Figure 18A**). Among BTx623 and Leoti, statistical differences in exudate profiles were found (ANOVA, $p < 0.001$), with BTx623 exudates enriched in monosaccharides and disaccharides, while Leoti was significantly enriched in organic and amino acids (**Figure 8A, Appendix B File S1**). Given these sorghum genotypes released root exudates with distinct metabolic profiles, we next pursued how these different exudate regimes would alter microbiome membership and metabolic responses.

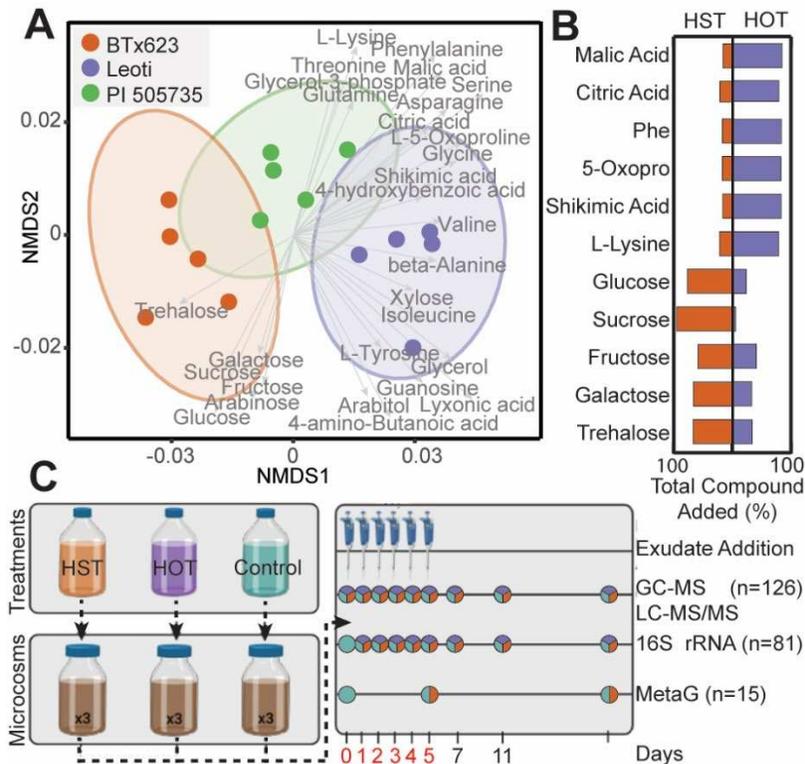


Figure 8. Root exudation metabolite profiles from three sorghum genotypes were used to design relevant amendments for soil microcosms. A) Non-metric multidimensional scaling (NMDS) of Bray-Curtis distances across metabolite abundance of Leoti (purple), BTx623 (orange), and PI 505735 (green) sorghum genotypes (stress=0.06). Ellipses denote the 95% confidence interval for each treatment. Significant metabolite vectors are shown (p-value <0.05) and labelled in gray. B) Butterfly plots show the summed total amount of each compound added across the treatments (Phe, Phenylalanine; 5-Oxopro, 5-Oxoproline); orange represents proportion added to High-Sugar Treatment (HST) and purple represents proportion added to the High Organic acid Treatment (HOT). For example, of the sucrose added, 95% went to the HST and 5% to the HOT. For actual added concentrations, see **Appendix B Table S3**. C) Microcosm schematic depicts treatment formulation from where HOT represents Leoti, HST represents BTx623, and Control was buffered media lacking an exudate treatment. Triplicate microcosms with agricultural soil were maintained for a 20-day experiment with the sampling schematic denoting when samples were obtained. Days with exudate addition are noted with pipettes and day number is colored in red. Time points with circles represent the samples taken each day, with the total number (n) of samples for each analysis listed. Colors within circles represent the type of sample (HST = orange, HOT = purple, control = teal) taken at that time point.

3.2.2 Laboratory microcosms reveal that exudate treatment structures agricultural soil microbial communities

We formulated root exudate solutions for laboratory-scale soil reactors (microcosms) using either a high sugar treatment (HST) representing BTx623 or a high organic acid treatment (HOT) representing Leoti. These treatments were run in parallel with exudate-lacking control microcosms. Both HST- and HOT-amended microcosms received the same eleven root exudate compounds (glucose, galactose, fructose, sucrose, trehalose, malic acid, lysine, phenylalanine, 5-oxoproline, citric acid, and shikimic acid, **Appendix B Table S3**), but in varying concentrations to model the relative concentrations observed in BTx623 and Leoti exudates (**Figure 8B**). The soil microcosms were constructed using soil from semi-arid agricultural plots and were amended with exudate treatments daily for the first 5 days and sampled for chemical and microbiological analyses over 20 days (**Figure 8C**).

Gene amplicon sequencing of the 16S rRNA gene was used to temporally profile the microbial diversity and membership across the three treatments (HOT, HST, exudate control, **Figure 8C**). We sequenced 81 samples, generating 2,148,831 reads with an average of more than 30,000 reads per sample (**Appendix B File S2**). After denoising, a total of 9,818 amplicon sequencing variants (ASVs) were detected, representing 43 phyla (**Appendix B File S2**). We first assessed alpha and beta diversity metrics to understand soil microbial community changes in response to exudates.

Over time, we observed distinct changes in response to amendments with either the HST or HOT (**Figure 9A**). For instance, after one day of exudate amendment, HST-amended microcosms had significant decreases in ASV richness ($p = 0.007$), Shannon's Diversity Index ($p = 0.019$), and Pielou's evenness ($p = 0.017$). In contrast, we did not detect significant changes in these metrics for HOT or control microcosms in this time period. This suggests HST enriched for select members of the microbial community. In support of this result, we observed an ASV

belonging to the genus *Pseudomonas* (**Appendix B File S2**) accounted for 59.4% of the HST community at day 1. Following 5 days of exudate addition, the HST-amended microcosms retained low diversity ($p = 0.038$), richness ($p = 0.042$), and evenness ($p = 0.017$) compared to the HOT and control microcosms (**Figure 9A**). By the final day of the experiment, these diversity metrics were no longer significantly different between HST and HOT, highlighting the importance of these amendments in structuring the microbial communities.

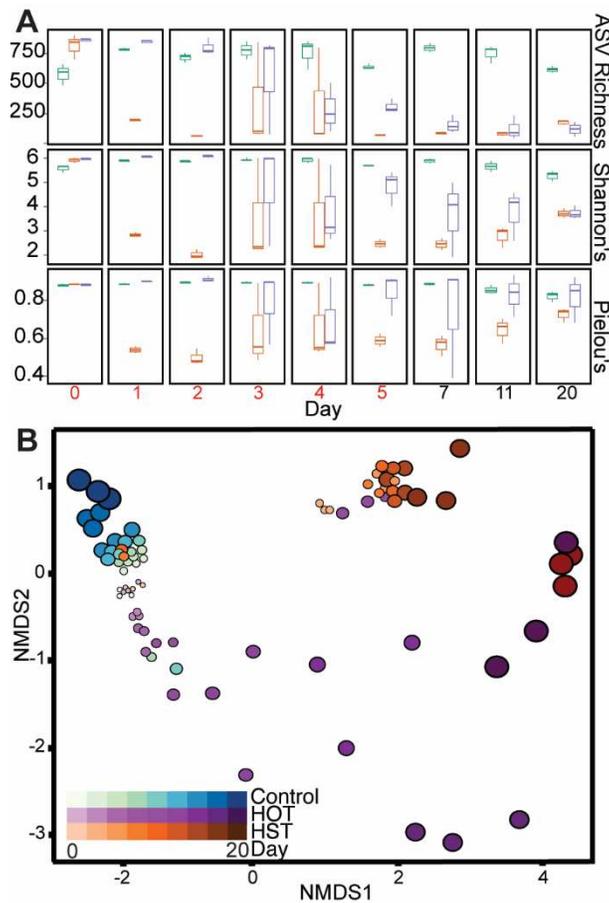


Figure 9. 16S rRNA gene diversity metrics and membership changes with exudate amendment. A) ASV richness (top), Shannon's Diversity Index (middle), and Pielou's Evenness for each day highlighting differences in treatment richness, alpha diversity, and community evenness. Coloring corresponds to treatment: control (teal), HST (orange), and HOT (purple). Days where exudates were added are colored red (days 0-5). B) Non-metric multidimensional scaling (NMDS) of Bray-Curtis distances of 16S rRNA amplicon communities showing changes in microbial community structure and membership over time with colors representing treatment and the size and darkness of circles representing time (stress=0.09).

Next, changes in the membership and structure of the microbial community across treatments was assessed using Bray Curtis dissimilarity matrix and visualized using non-metric multidimensional scaling (NMDS) (**Figure 9B**). We found that HST and HOT amendments significantly shifted microbial community composition and structure during and after periods of synthetic root exudate addition (multi-response permutation procedure, mrpp = 0.01, p-value < 0.05). A PERMANOVA analysis revealed that microbial communities were significantly altered across treatments ($p = 0.001$) and time ($p = 0.001$). Building on this, beta dispersion analyses indicated the control exudate communities were the most stable over time, while the HOT treatment had the greatest across time variation in microbial community membership (**Figure 9B**). These microbial community alpha and beta diversity metrics highlight that both treatment and time were factors shaping the soil microbiome in exudate-amended microcosms. Scaling beyond these microcosms, this research highlights the need for understanding the composition and temporal dynamics of natural root exudates.

3.2.3 Exometabolites hint at potential microbial metabolisms stimulated by exudates

In these soil reactors, we tracked the temporal dynamics of exometabolites (**Figure 8C**), which are the extracellular fraction of molecules that are inferred to be produced and/or utilized by soil microorganisms. We classified and assigned the detected exometabolites to three chemical classes: (i) central carbon metabolism, (ii) amino acids and derived compounds, and (iii) phytohormones (**Figure 10, Appendix B File S1**). The exometabolites were coordinated to the microbial communities identified by 16S rRNA gene sequencing, and thus demonstrated differences by treatment and time (**Appendix B Figure S2, 10**). Relative to the HST and controls, the HST was enriched in most of the detected exometabolites, likely reflecting a greater metabolic stimulation from the sugar rich treatment (**Appendix B Figure 10**). This finding is not completely unexpected, as prior studies have noted sugar metabolism is more efficient than amino acid metabolism in soil microorganisms¹⁴⁹.

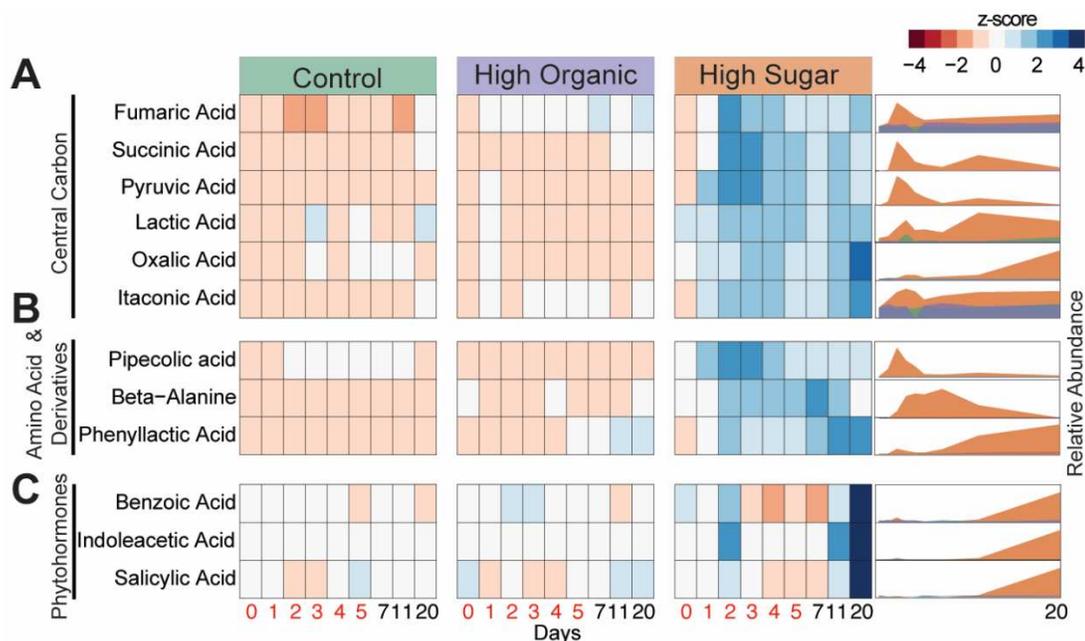


Figure 10. Exometabolite abundances across treatments. We categorized the detected exometabolites into three chemical classes: (A) central carbon metabolism, (B) amino acids and derivatives, and (iii) phytohormones. Within each, heatmaps (left) showing the relative abundances (represented as a z-score across samples) of each exometabolite detected. Time increases from left (day 0) to right (day 20) within each treatment, with days where exudate treatments were added colored red (days 0-5). Ridgeline plots (right) show the relative abundance (0 – 100%) of these exometabolites over time. Coloring corresponds to treatment: control (teal), HST (orange), and HOT (purple).

Metabolites relevant to microbial central carbon metabolism were some of the most enriched exometabolites in HST microcosms (**Figure 10A**). We detected six organic acids that changed significantly over time in HST and were significantly enriched in HST compared to the control and HOT. For example, between days 2 and 3, succinic acid and fumaric acid increased 1.5 and 0.9-fold (log₂), respectively, in the HST-amended microcosms relative to HOT and the control treatments (**Figure 10A**). We also detected increases in itaconic acid (0.7-fold) and oxalic acid (1.6-fold) over time, both of which can be derived from tricarboxylic acid (TCA) cycle intermediates cis-aconitate and oxaloacetate, respectively¹⁵⁰. Furthermore, consumption of malate and citrate was inferred in HST-amended microcosms at the same time points due to their

loss over time (**Appendix B Figure S3**). Perhaps indicating metabolite cross feeding across the community, we also observed a spike in pyruvate at days 2-3 in the HST, followed by consumption concomitant with the production of lactic acid across the experiment (**Figure 10A**). Collectively, this exometabolite data pointed to HST exudate treatments differentially altering soil carbon pools via the metabolism of the microbiome.

We also detected chemical evidence for production and consumption of amino acid and amino acid derived compounds by these soil microbial communities. Like the central carbon metabolites, these organic nitrogen exometabolites were most enriched in HST-amended microcosms relative to the other treatments, a somewhat unexpected response given the HOT amendments were initially dosed with these types of compounds (**Figure 8B**). Compared to control and HOT-amended microcosms, we observed the production of two non-proteinogenic amino acids (β -alanine and pipercolic acid) and one aromatic compound (phenyllactic acid) in HST-amended microcosms (**Figure 10B**). β -alanine, which was produced over the first 7 days but removed by day 20, is an important amino acid precursor and also necessary for biosynthesis of coenzyme A, a critical cofactor of microbial metabolism in soils¹⁵¹.

Next, we identified and tracked the production of pipercolic acid (PA), whose production peaks between days 2-4 in HST-amended microcosms. This amino acid is thought to be microbially synthesized from lysine¹⁵², a component added to both treatments (**Figure 8B**), that was consumed during this period (**Appendix B Figure S3**). PA synthesis has broader ramifications for the entire microbiome, as it is a required intermediate in secondary metabolite production of antibiotics and anthelmintics in many bacteria¹⁵². PA dynamics, which indicate production and subsequent consumption by day 20, suggest its use as a public good¹⁵³. Finally, phenyllactic acid (PLA), a phenylalanine derivative, increased 2.6-fold in the HST microcosms over time, peaking at day 11 and 20 (**Figure 10B**). Microbes that release amino acid and amino

acid derived compounds could be competitive root colonizers or act as antagonistic agents against target pathogens¹⁵⁴.

Most notably, we observed a significant increase in the abundance of three phytohormones. Salicylic acid (SA), benzoic acid (BA), and indole-3-acetic acid (IAA) significantly increased ($p < 0.01$) over the experiment in the HST relative to microcosms treated with HOT or exudate-lacking control (**Figure 10C**). In HST microcosms, BA and SA increased 1-fold (log2) while IAA increased the most with a 3-fold (log2) increase from day 0 to day 20. In plants, SA, BA, and IAA are vital phytohormones integral to physiological processes like plant defense and development¹⁵⁵. Our detection of these metabolites over time highlights that certain root exudation chemical profiles can stimulate soil microbes to produce phytohormones critical for plant growth and defense.

3.2.4 Curation of a sorghum exudate responsive microbial genome catalog

In light of the phytohormone production observed in our HST microcosms, we sought to identify microbial genomes capable of producing these compounds. Towards this goal, we constructed metagenome-assembled genomes (MAGs) representing the suite of soil microbes in the HST and control microcosms. To maximize MAG recovery, we obtained more than 365 Giga base pairs total of sequencing from control and HST microcosms, at three different timepoints spanning the experiment (control: day 0, 5, 20; HST: day 5, 20, each in triplicate; $n = 15$ metagenomes, **Figure 8C**). With this data, we reconstructed 371 MAGs that were dereplicated at 99% identity into 243 MAGs, of which 28% were high-quality¹⁵⁶ (**Figure 11A, Appendix B File S2**). The genome recovery demonstrated here exceeds those from field-based studies where assembling and binning are hampered by the vast complexity in agricultural soils¹⁵⁷, further validating the relevance of stimulated laboratory microcosms to increase the microbial genome tractability of soil systems^{158,159}.

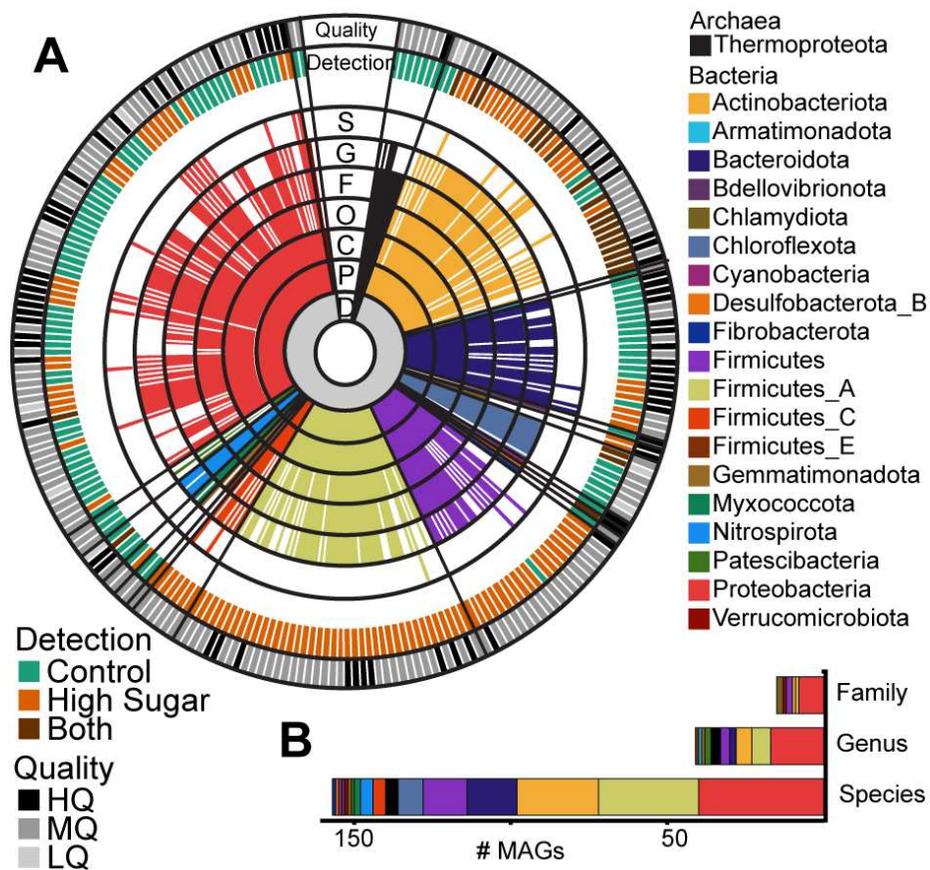


Figure 11. Taxonomy of the 243 dereplicated metagenome assembled genomes in the genome database. A) Sequential colored rings indicate the most resolved taxonomic level that could be assigned by GTDB-tk¹⁶⁰. Taxonomic level (D=Domain, P=Phylum, C=Class, O=Order, F=Family, G=Genus, S=Species) is denoted in black with a single letter abbreviation. Ring color corresponds to phylum assignment, with the color legend at right. The treatment condition corresponding to MAG detection is illustrated in the outer ring labeled “Detection”; MAGs detected only in Control metagenomes are indicated by teal, MAGs detected only in HST metagenomes are indicated by orange, and MAGs that were detected in both indicated with brown (see Materials & Methods for detection thresholds). The MAG quality is shown in the outermost ring following the MIMAG standards¹⁵⁶: high quality (HQ, >90% complete, <5% contamination), medium quality (MQ, >50% complete, <10% contamination), and low quality (LQ, here defined as >48% complete, <10% contamination). B) Stacked bar graph shows the number of dereplicated MAGs recovered that represent novel families, genera, or species according to taxonomy assignments from GTDB-tk. Coloring corresponds to MAG phylum.

Of these MAGs, 49% (n=119) were detected in HST-amended microcosm samples while 41% (n=100) were detected in the control microcosm samples (**Figure 11A, Appendix B File**

S2). Notably, only 10% (n=24) of these MAGs were shared across the two treatments, further supporting our 16S rRNA and exometabolite data showing HST amendments significantly altered the soil microbial community. Our MAG database contains representative soil microbes across 20 phyla, dominated by members of the Actinobacteriota and Proteobacteria (Figure S4), and includes MAGs belonging to 15 previously unidentified families (6%), 41 previously unidentified genera (17%), and 157 previously unidentified species (65%) (**Figure 11B**). We recovered 26 MAGs with partial or full 16S rRNA genes and could directly link 19 of these MAGs to 16S rRNA amplicon sequencing identified ASVs, providing metabolic blueprints for these taxa. These results highlight the advantage of coupling community profiling (16S rRNA gene) with high resolution metagenomics to capture the functional potential encoded within the soil microbiome.

3.2.5 Bacterial salicylic acid production is assigned to a potentially-new species of Pseudomonas

We examined which of our MAGs encoded the two genes (*pchA*, isochorismate synthase; *pchB*, isochorismate pyruvate lyase) for SA production. The first gene, *pchA*, catalyzes the conversion of chorismate to isochorismate with the second gene, *pchB*, catalyzing the subsequent conversion to salicylate (**Figure 12A**)¹⁶¹. Two of our HST-detected MAGs encoded a *pchA*, while one of these MAGs also encoded *pchB*. Notably, the MAG that encoded both *pchA* and *pchAB* likely represents a novel species within the genus *Pseudomonas_E* (L_E1_T20_B_bin.65) and had two copies of this gene set (**Figure 12B**). Consistent with our exometabolite data for SA production (**Figure 10C**), this *Pseudomonas_E* MAG was abundant in HST metagenomes at day 20 (**Figure 12E**).

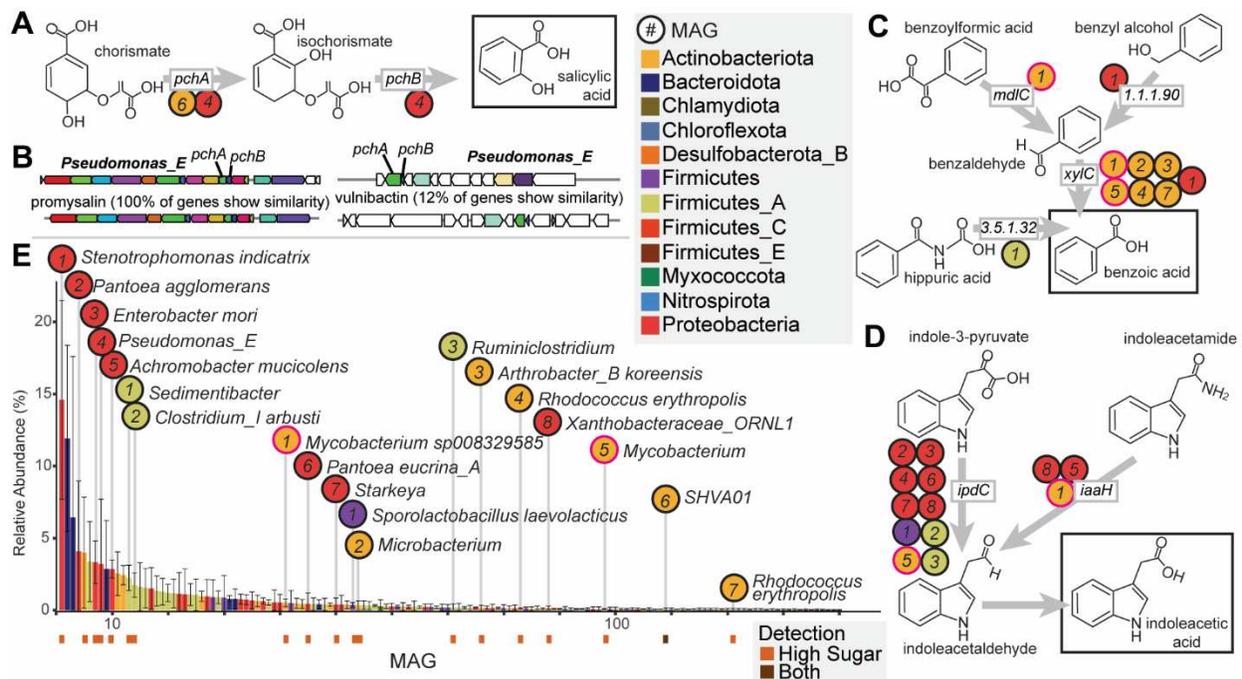


Figure 12. Diverse MAGs encode biosynthetic potential for salicylic acid, benzoic acid, and indoleacetic acid. A) Salicylic acid production from chorismate by *pchAB*. B) A *Pseudomonas* MAG encoded two biosynthetic gene clusters containing *pchAB*, including one for the antimicrobial Promysalin (left) and another for a predicted siderophore (right). The clusters on the bottom are the reference clusters from antiSMASH. C) Benzoic acid production from multiple pathways. D) Indoleacetic acid production pathways. E) Rank abundance curve of MAGs detected in HST metagenomes at day 20. Bars represent the average relative abundance (n=3), and error bars represent one standard deviation. Bars are colored by MAG phylum, and key phytohormone producing MAGs are indicated. MAG detection is denoted below abundance bars for potential phytohormone producing MAGs, with MAGs only detected in HST colored orange (n=18), and MAGs detected in both HST and control in brown (n=1). In A, C, and D, circles correspond to MAGs encoding each gene in E. The circles for the two *Mycobacterium* MAGs with BA and IAA potential are outlined in pink.

Previous research has shown that bacterial species such as *Pseudomonas*, *Bacillus*, *Mycobacteria*, and *Azospirillum* synthesize this important compound, with pseudomonads, like those enriched in our reactors, exhibiting a propensity for SA production¹⁶². It is recognized that microbially produced SA can act as a plant defense hormone providing pathogen resistance and modulating developmental signaling cascades¹⁶³⁻¹⁶⁵. Beyond phytohormone roles, bacterially produced SA can function as a siderophore for metal acquisition from soils^{162,163,166} or to support antibiotic production. Excitingly, further genomic analysis of the *Pseudomonas_E* MAG revealed

one set of *pchAB* which occurred in a predicted siderophore biosynthetic gene cluster (**Figure 12B**). Furthermore, the second *pchAB* cluster occurred in a 16kb gene cluster with 100% gene similarity to the promysalin biosynthetic gene cluster from *Pseudomonas putida* (**Figure 12B**). Promysalin is an antibiotic that contains SA and was shown to be selectively antagonistic to other closely related pseudomonads to enable rhizosphere colonization¹⁶⁷. Our data suggest the HST microcosms created conditions that stimulated the *Pseudomonas_E* MAG to produce SA, which could ultimately aid in its competition for limited resources in the complex soil microbiome. Additionally, the accumulation of SA in the enrichment leaves open the possibility that beyond enhanced microbial metabolism, this exometabolite could also be available to act as a plant defense hormone.

3.2.6 Diverse benzoic acid production pathways encoded by multiple MAGs

We next investigated BA production potential in our MAGs. Known BA production pathways derive from degradation of aromatic compounds (**Figure 12C**)¹⁶⁸. The terminal step of most of these pathways is the oxidation of benzaldehyde to BA by benzaldehyde dehydrogenase (*xyIC*). In these microcosms, seven MAGs encoded *xyIC*, including *Stenotrophomonas indicatrix*, *Arthrobacter koreensis*, two MAGs of *Rhodococcus erythropolis*, and two species of *Mycobacterium* (**Figure 12C**). These MAGs were only detected in the HST microcosms at day 20 (**Figure 12E**), matching when BA was observed in exometabolomic data (**Figure 10C**).

The *S. indicatrix* MAG encoded an aryl-alcohol dehydrogenase (EC: 1.1.1.90) adjacent to *xyIC*, which would enable reduction of benzyl alcohol to benzaldehyde (**Figure 12C**). This is also the most abundant MAG in HST metagenomes at day 20 (**Figure 12E**), indicating it may be important to BA production in the HST microcosms. Beyond *S. indicatrix*, a MAG representing *Mycobacterium* sp008329585 encoded a benzoylformate decarboxylase (*mdIC*), which decarboxylates benzoylformic acid to benzaldehyde (**Figure 12C**). Finally, an abundant MAG representing a novel species of *Sedimentibacter* encoded a gene for hippurate hydrolase (**Figure**

12E). This enzyme cleaves glycine from hippuric acid, producing BA (EC: 3.5.1.32, **Figure 12C**). Consistent with the literature, these findings position BA as a central aromatic metabolite in these microcosms.

While there are limited reports today that microbial BA production can have direct benefit to plants, it is well known that plant BA and derivatives are precursors to the production of secondary metabolites and hormones¹⁶⁹. BA is also known to improve plant stress tolerance and contribute to growth regulation¹⁶⁹. In support, exogenous application of BA has been associated with improved growth and yield in quinoa and drought-stressed soybean^{170,171}. As such, BA synthesis by the soil microbiome and its relationship with plant health is an area of research that warrants more investigation.

3.2.7 Bacterial indoleacetic acid production encoded by diverse MAGs

Tryptophan (Trp) is the primary precursor for IAA biosynthesis in microorganisms¹⁷², with five known biosynthetic pathways characterized in bacteria¹⁷³. We surveyed our MAG database for these genes and found 14 HST-detected MAGs capable of IAA production using two different pathways (**Appendix B File S3**), consistent with reports that auxin biosynthesis is widely encoded in plant^{172,174}. Nine MAGs encoded indole-pyruvate decarboxylase (*ipdC*), the key gene for converting Trp to IAA via indole-3-pyruvate (**Figure 12D**). These MAGs spanned four phyla, with six belonging to Proteobacteria, covering several of the most abundant MAGs at day 20 in the HST metagenomes (**Figure 12E**). Beyond this pathway, three MAGs encoded *iaaH* for IAA production via indoleacetamide (**Figure 12D**). One of these MAGs, a novel species of *Xanthobacteraceae*, encoded both *iaaH* and *ipdC*. This fits with previous observations of single organisms encoding redundant IAA biosynthesis pathways⁶². Of note, two Mycobacterial MAGs with IAA-production potential also had BA-production potential. One potentially novel *Mycobacterium* species encoded *ipdC* and *xyIC*, while *Mycobacterium* sp008329585 encoded

iaaH in addition to *mdlC* and *xyIC*, indicating some soil microbes can produce multiple phytohormones.

IAA is a phytohormone in the auxin family that is necessary for proper plant development and can be used to stimulate root biomass growth^{175,176}. From a microbial perspective, IAA production offers a competitive colonization strategy over non-IAA producing strains^{177,178} and it has been shown that microbially-produced IAA in rhizosphere can elicit positive plant growth promotion and disease suppression effects^{142,177-179}. However, these effects are dose dependent, as high concentrations of IAA can also inhibit plant growth¹⁸⁰⁻¹⁸². The reduced complexity of our laboratory microcosm system allowed us to demonstrate that IAA can accumulate in soil microbiomes, and its production is redundantly encoded across 4 different phyla, offering new perspectives on the microbial contributions to soil IAA pools.

It is important to note that the metagenomic results presented here only represent the genomic capabilities of a hypothesized genome, thus, additional experimental evaluation of gene expression would be required to definitively confirm specific bacterial synthesis of SA, BA, and IAA and their roles as potential PGPR species. Furthermore, gene expression could inform whether exometabolite production was a direct microbial response to exudates versus an indirect responses via the priming effect¹⁸³. Yet, we find these MAG results, coupled with exometabolite evidence for these compounds through time, to be an exciting platform for targeted studies aimed at harnessing the power of microbial metabolisms for improving agroecosystems.

3.3 Conclusions

Coupling a reduced complexity, tractable laboratory soil system with high-resolution multi-omics approaches afforded new perspectives on the complex interactions between plant root exudation and the soil microbiome. We first showed that root exudate amendments modeled after two different sorghum genotypes drastically impacted overall soil microbial community diversity and membership. Second, shifts in microbial community structure mirrored differences in

exometabolites. Finally, we linked interesting plant-relevant metabolites to bacterial genomes, highlighting new roles for specific microorganisms in phytohormone biosynthesis. We have provided a publicly available, genome-resolved microbial database for researchers interested in the sorghum rhizosphere. Our findings suggest root exudate composition influenced microbial production of possible phytohormones that could impact host plant physiology, as well as govern metal acquisition and antibiotic production from soils. Future studies can scale these hypotheses to the field to analyze the impact of plant genotype, root exudates, and associated microbial communities on overall plant performance. In summary, the results of this study represent an important step towards decoding the complex chemical language between plants and their rhizosphere microbial communities, a translation required to optimize these interactions for enhanced agricultural management and production.

3.4 Materials & Methods

3.4.1 Sorghum Root Exudate Collections

Mature seeds from three sorghum genotypes (Leoti, BTx623, and PI 505735) were first sterilized by placing seeds in sterile 50 mL conical tubes with 45 mL 95% ethanol. Tubes were vortexed, ethanol was removed and 45 mL of Captan fungicide solution (0.2g Captan fungicide in 45 mL sterile water) was added to remove any fungal components on the seeds and vortexed for 3 hours. After removal of the Captan fungicide, 45 mL of 100% Clorox bleach (8.25% sodium hypochlorite active ingredient) + 3 droplets of dish soap (Ajax Triple Action Orange) was added for the final sterilization step and seeds were shaken for 20 mins. In a sterile tissue culture hood, the bleach was removed using sterile techniques and the seeds were rinsed five times with sterile water. Seeds were blotted dry and placed on germination paper inside a clean 600 mL beaker with a solution of 1 mM CaCl₂. After 7 days of growth, seedlings were removed from the paper and pooled into a 250 mL glass bottle filled with 80 mL of ultrapure water. Root exudates were collected for each genotype in separate bottles. The bottles were covered with aluminum foil to

protect roots from the light and were placed on a rotary shaker. After 2 hours, the roots were removed, blotted dry, and weighed. The root exudate suspensions containing the root exudates were filtered through a 0.2 µm filter membrane to remove root detritus and microbial cells. Samples were frozen at -80°C. We recognize that the use of hydroponic growth chambers does not exactly replicate soil growth conditions of sorghum¹⁸⁴; however, this system provided a highly controllable, tractable, and sterile environment that eliminated confounding microbial or soil influences for more accurate downstream analytical detection. Samples were lyophilized completely before resuspension in 15 mL sterile HPLC-grade water. Resuspended exudates were vortexed thoroughly to remove all residue from the bottle, transferred to a clean 50 mL falcon tube and dried completely under nitrogen gas. 1mL LCMS-grade 80% acetonitrile was added to the falcon tube with dried exudates and vortexed until thoroughly resuspended. 0.5 mL was transferred to a clean 2 mL glass vial, and 5 µg of glucose-¹³C₆, D-arabinose-¹³C₅, sucrose-¹³C₁₂, galactose-¹³C₆, and fructose-¹³C₆ were added to each vial and dried completely under nitrogen prior to derivatization.

3.4.2 Sample Derivatization for Sorghum Root Exudates

Dried samples were resuspended in 75 µL of 0.2M of methoxyamine hydrochloride in pyridine (Sigma), incubated at 60°C for 45 min, vortexed, sonicated for 10 min, incubated a second time at 60 °C for 45 min and allowed to cool to ambient temperature. 30 µL of each sample were then transferred to a separate 2mL glass vial for MSTFA derivatization and the other 30 µL for Acetic Anhydride (AA) derivatization. 90 µL of 100% AA was added to AA vials, vortexed for 30 s, incubated at 60°C for 60 min, and cooled for 10 min at ambient temperature. 30 µL of N-methyl-N-trimethylsilyltrifluoroacetamide plus 1% trimethylchlorosilane (MSTFA + 1% TMCS, Thermo Scientific) was added to the MSTFA vials, vortexed for 30 s, incubated at 60°C for 35 min and cooled to ambient temperature before transferring to vial inserts. AA samples were dried

completely under nitrogen and resuspended in 40 μ L of 100% ethyl acetate and then loaded into glass inserts for analysis.

3.4.3 Non-Targeted and Targeted GC-MS Analysis for Sorghum Root Exudates

Metabolites were analyzed and detected using a Trace 1310 GC (Thermo) coupled to an ISQ mass spectrometer (Thermo). Samples (1 μ L) were injected into an injection port at 285°C and 1:10 split ratio. Separation was accomplished with a 30 m TG-5MS column (Thermo Scientific, 0.25 mm i.d., 0.25 μ m film thickness) and a helium gas at 1.2 mL/min flow rate. The oven temperature program started at 80°C for 30 s, ramped to 330°C at 15°C/min, and then held at the final temperature for 8 min. The transfer line and ion source were maintained at 300°C and 260°C, respectively. Masses between 50-650 m/z were scanned at 5 scans/sec after electron impact ionization. A pooled QC sample (made by combining an equal volume of all samples) was injected after every 6 samples to monitor instrument stability throughout the analysis.

3.4.4 Synthetic Exudate Preparation

Exudate treatments were formulated based on the results of the exudate profiles for the two most divergent sorghum genotypes (Leoti and BTx623). High Sugar Treatment (HST) was formulated to mimic BTx623 with higher sugar and lower organic acid composition and the High Organic acid Treatment (HOT) was formulated to mimic the Leoti exudate profiles with higher organic acid, and lower sugar composition (**Appendix B Table S3, File S5 and File S6**). The 11 synthetic root exudate compounds were chosen based on practical considerations and availability of standards. Organic acid root exudate relative quantification values were normalized to fructose peak area and scaled to the absolute quantification of fructose in that sample. Treatment calculations can be found in supplementary files S5 and S6. The exudate-lacking media control treatment (Control) was a 10mM phosphate buffer, consisting of ammonium chloride, disodium phosphate, and sodium dihydrogen at a pH of 6.5 (**Appendix B Table S1**). Exudates found in each genotype were weighed out to their respective masses (**Appendix B Table S3**) to supply

equal exudates for 6 days of exudate addition (**Fig 8C**, days 0-5) to the microcosms, homogenized, suspended in 10mL 10mM phosphate buffer, vortexed, aliquoted into 6 tubes for each day of exudate addition, and frozen at -80°C until use. Aliquots were removed 30 min before use, incubated at 24°C until thawed, and added to microcosms at the respective time point.

3.4.5 Soil Samples

The soil (microbial inoculum) was collected from agricultural fields at the Colorado State University Agricultural Research and Education Center (CSU-ARDEC) near Fort Collins, CO on October 4th, 2019. The climate at the site is semi-arid, with 408 mm mean annual precipitation and a mean annual temperature of 10.2°C (1981-2010 average, <https://usclimatedata.com/>). The soil is classified as an Aridic Haplustalf. Three 2-cm diameter soil core samples to approximately 15cm depth were collected from each of seven different plots. The soil was stored at -20°C until microcosm construction. 20g of soil from each replicate/plot was pooled and homogenized to create a representative bulk soil repository used in the following microcosms experiment.

3.4.6 Microcosm Experimental Set-Up

Microcosms were established and sampled as previously described^{158,185}. Briefly, 5 g of homogenized soil and 35mL of phosphate buffered (pH 6.5) minimal medium (**Appendix B Table S1**) was added to sterile 50 mL conical tubes to construct each microcosm. Microcosms were vortexed and allowed to settle for 5 min. Then day 0 samples were taken by removing 1 mL of soil slurry for exometabolomics analysis and 1 mL of soil slurry for DNA extraction. After this initial sampling, 2 mL of exudate treatment were added to each microcosm, vortexed, and a second 1 mL aliquot was immediately taken for exometabolomics analysis. At this point, for each microcosm the bottle caps were removed and replaced with a sterile foam stopper for the rest of the experiment to maintain oxic conditions and prevent colonization by contaminating microbes. Microcosms were incubated in an orbital shaker set at 200 rpm at 24°C for 20 days. Each exudate treatment (i.e., HST, HOT and Control) was conducted in triplicate and treatments were added (2

mL) to microcosms on days 0, 1, 2, 3, 4, 5. After day 5, no additional exudate treatments were applied but microcosms were maintained until day 20 which afforded additional samples taken at day 7, 11, and 20 (**Figure 8C**). Samples were collected at roughly the same time each day and collected with aseptic techniques to ensure no additional microbial influence was introduced. All collected samples were immediately frozen at -80°C until processing.

3.4.7 Sample Preparation and Extraction for Exometabolomics Analysis

Samples were thawed at 4°C overnight, centrifuged for 20 minutes at 18,000 x g and supernatant was transferred to a pre-weighed 1.5-dram vial and refrozen at -80°C for subsequent lyophilization. Samples were lyophilized for 24 hours until all water was sublimated. Samples were weighed to calculate total exometabolite mass and then were resuspended in 4mL sterile HPLC-grade water. A volume equivalent to 0.50 mg was transferred to a new, pre-weighed 2 mL glass vial and dried under N₂. Lastly, each sample was resuspended in 500 µL of sterile HPLC-grade water, vortexed for 1 min, and sonicated for 15 min. Two 250 µL aliquots were transferred to new 2 mL vials, respectively, and samples were dried under N₂. This yielded two 0.25 mg subsamples for analysis by GC-MS and UPLC-MS/MS as described below.

3.4.8 Targeted UPLC-MS/MS for Phytohormone Analysis of Exometabolites

0.25 mg subsamples were extracted in 75 µL of a spiked methanol solution containing 100% methanol with 65.2 ng/mL ABA-d6, 62.5 ng/mL salicylic acid-d6, and 90.0 ng/mL jasmonic acid-d5 (Sigma). After solvent addition, samples were placed on a shaker plate for 1 hour at the highest speed setting, centrifuged at 3500 x g at 4°C for 5 minutes, and transferred to glass inserts. A final centrifuge step at 3500 x g for 15 minutes at 4°C was completed to ensure any precipitate was in the bulb of the vial insert. Five microliters of exometabolite samples were injected onto a Perkin Elmer UPLC MS/MS system, equipped with a PerkinElmer QSight LX50 Solvent Delivery Module (PerkinElmer). An ACQUITYUPLC T3 column (1 × 100 mm, 1.8 µM; Waters Corporation) was used for chromatographic separation. Mobile phase A consisted of LC-

MS grade water with 0.1% formic acid and mobile phase B consisted of 100% LC-MS grade acetonitrile. The elution gradient was initially set at 0.1% B for 1 min, which was increased to 55.0% B at 12 min and further increased to 97.0% B at 15 min, then decreased to 0.1% B at 15.5 min. The column was re-equilibrated for 4.5 min for a total run time of 20 min. The flow rate was set to 200 $\mu\text{L}/\text{min}$ and the column temperature was maintained at 45°C. Samples were held at 4°C in the autosampler. Detection was performed on a Perkin Elmer QSight™ 220 triple quadrupole MS in selected reaction monitoring (SRM) mode. The transitions monitored for each phytohormone compound can be found in **Appendix B Table S2**. The MS was operated with ESI voltage 4500 V in positive mode and -3500 V in negative mode. Nebulizer gas flow was set at 350 arbitrary units and drying gas was set to 120 arbitrary units. The source temperature was 315 °C and hot-surface induced desolvation (HSID) temperature was set to 200°C.

3.4.9 Non-targeted GC-MS Analysis of Exometabolites

Sample preparation was conducted as previously described^{122,186} and as described above for sorghum root exudates. Briefly, dried samples were resuspended in 50 μL of pyridine containing 25 mg/mL methoxyamine hydrochloride (Sigma), centrifuged, incubated at 60°C for 45 minutes, vortexed for 30 seconds, sonicated for 10 mins, centrifuged briefly, and incubated a second time for 45 mins at 60°C. Samples were cooled to room temperature and centrifuged for 2 mins. Then, 50 μL of MSTFA + 1% TMCS (Thermo Fisher) was added, samples were vortexed for 30s, centrifuged, and incubated a third time at 60°C for 35 minutes. Samples were cooled to room temperature, centrifuged, and 80 μL of supernatant was transferred to glass vial inserts within glass vials. Samples were centrifuged a final time for 10 mins before analysis. Metabolites were separated with a 30 m TG-5MS column (Thermo Scientific, 0.25mm i.d. 0.25 μm film thickness) and detected using a Perkin Elmer Clarus 690 GC coupled to a Clarus SQ 8S mass spectrometer (Perkin Elmer). Samples (1 μL) were injected at a 10:1 split ratio onto the column with a 1.0 ml/min helium gas flow rate. The gas chromatography inlet was held at 285°C, and the

transfer line was held at 300°C, and the source temp was held at 260°C. The GC oven program started at 80°C for 30s, followed by a ramp of 15°C/min to 330°C, followed by an 8-min hold. Masses between 50–620 m/z were scanned at 4 scans/s under electron impact ionization. Injection of QCs were analyzed after every 6th sample to ensure proper instrument function and to detect any analytical variation.

3.4.10 GC-MS and LC-MS Data Analysis

Non-targeted GC-MS data (both sorghum exudates and exometabolites) were processed within the R statistical software¹²³ using methods previously described¹⁸⁶. For GC-MS samples, .cdf files were processed through the following workflow: 1) XCMS software was used for preprocessing to identify molecular features¹²⁴; 2) features were further normalized to total ion current (TIC); 3) the package RAMClustR¹²⁵ was used for clustering features into spectra and prepared for subsequent spectra identification in RAMSearch¹⁸⁷ using external databases including Golm (<http://gmd.mpimp-golm.mpg.de/>) and NIST (<http://www.nist.gov>). For exometabolite data, prior to TIC normalization, features were normalized by linearly regressing run order versus QC feature intensities to account for instrument signal intensity drift. For root exudate data, relative quantitation was also normalized by root weight. Targeted quantification of sugars from the GC-MS data was performed using Chromeleon 7.2 (Thermo Fisher). The integrated peak area for each sugar was normalized to its corresponding internal standard with the exceptions that arabinose 13C⁵ was used for xylose, glucose 13C⁶ was used for mannose and sucrose 13C¹² was used for trehalose and maltose. Quantification was determined using a linear regression of an 8 point standard curve for each sugar. Final concentrations were normalized to root weight. For phytohormone analysis of exometabolites, LC-MS data were processed using Simplicity 3Q (v1.5, Perkin Elmer) bioinformatics software for sample processing. Briefly, the peak area for each phytohormone compound was normalized to internal standard peak area and

quantification was assessed using a linear regression against an external calibration curve. Exudate and exometabolite data are provided in **Appendix B File S1**.

3.4.11 Soil DNA Extraction and Library Preparation

Total genomic DNA was extracted from the microcosms using the Zymo Quick-DNA Fecal/Soil Microbe Microprep kit. 16S rRNA gene amplicon sequencing was performed on the Illumina MiSeq using 251-bp paired-end reads and the Earth Microbiome Project primers 515F/806R¹⁸⁸, for an average of more than 30,000 reads per sample (see **Appendix B File S2** for individual sample data). The 16S rRNA partial gene reads were analyzed and reads were demultiplexed using QIIME2¹⁸⁹ (2019.10). Using DADA2¹⁹⁰, demultiplexed reads were denoised to produce an amplicon sequence variant (ASV) table and filtered to remove noisy sequences, chimeras and singletons. Feature classification was completed by comparing the ASV table against the trained full-length SILVA classified (silva132.250) database for taxonomic classification. The ASV table was filtered to contain ASVs that were observed in at least 2 samples and the output files were visualized in QIIME2. Samples L_E1_T5_A (HST day 5 rep A), L_E2_T11_A, and L_E2_T20_A (HOT days 11 and 20 rep A) yielded insufficient sequencing results and were excluded from subsequent analyses. The ASV feature table is provided in **Appendix B File S2**.

3.4.12 Metagenomics Analysis

Metagenomic DNA from day 0 (control) and day 5 (control and HST) metagenomes (n = 9) was sequenced at the Genomics Shared Resource at the University of Colorado Cancer Center using the NovaSeq6000 platform. Metagenomic DNA from day 20 (control and HST, n = 6) was prepared for metagenomic sequencing using the Nextera XT low input-Illumina library creation kit and samples were sequenced at the Department of Energy Joint Genome Institute on the Illumina NovaSeq 6000. FastQ files were trimmed using Sickle (v1.33)¹⁹¹. Day 0 and day 5 reads were concatenated within each timepoints/treatments (i.e. triplicate controls at day 5) for co-assembly.

Day 0 and day 5 (both control and HST) coassemblies, and the day 20 individual assemblies were assembled with IDBA-UD¹⁹². Within each assembly, scaffolds greater than 2.5kb were binned into metagenome-assembled genomes (MAGs) using MetaBAT2 (v2.12.1)¹⁹³. MAGs were assessed for completion and contamination using checkM¹⁹⁴. A MAG was retained if it was >48% complete with <10% contamination, and assigned quality following MIMAG guidelines¹⁵⁶. Using dRep¹⁹⁵, MAGs were dereplicated to 99% identity. MAG taxonomy was assigned using GTDB-tk (v1.5.0, R06-RS202), and taxonomic novelty was defined as the first unnamed taxonomy level¹⁶⁰. To obtain MAG abundance, trimmed metagenomic reads from individual samples were mapped to the dereplicated MAG set using bmap¹⁹⁶ (v38.70) at minid=95, and output as sam files which were converted to sorted bam files using samtools¹⁹⁷ (v1.9). CoverM (v0.3.2) was used to determine MAG relative abundance as described in McGivern et al ¹⁵⁸. MAGs were annotated using DRAM ¹⁹⁸. Biosynthetic gene clusters were detected using the antiSMASH webserver using default parameters (v6.0) ¹⁹⁹. See **Appendix B File S2** for MAG quality, taxonomy, and mapping; **Appendix B File S3** for DRAM annotations; and **Appendix B File S4** for raw annotations accessed using <https://doi.org/10.5281/zenodo.5639650>.

3.4.13 Statistics

Sorghum genotype ordinations were generated using metaMDS function from the vegan ²⁰⁰ package in R¹²³ (v4.0.2) and visualized with ggplot2²⁰¹. Metabolite loadings were calculated with the envfit function in vegan. Alpha diversity of microbial communities was calculated using the diversity function from the vegan package²⁰⁰ in R¹²³ (v4.0.2) using Shannon's (H), Pielou's, and richness indices. To estimate beta diversity across treatments, we utilized Bray-Curtis dissimilarity matrix visualized by non-metric multidimensional scaling (NMDS) in R with the ggplot2 package with stress of the non-parametric fit for the goodness of fit for the model. Significance of compositional differences across treatments was quantified using mrpp and the betadisper commands from the vegan package with an ANOVA model in R. Significance between

communities across time and treatment were quantified using beta-group-significance commands in QIIME2 ¹⁸⁹ (2019.10) using a PERMANOVA model. Fold changes for exometabolite dynamics were calculated and converted to log₂ abundances via the log function. 2-way ANOVA tests and pairwise comparisons were completed in GraphPad Prism (v8.2.1) using grouped analyses with Sidak's multiple comparison testing with alpha = 0.05.

CHAPTER 4: GENOME-RESOLVED MULTI-OMICS REVEAL COVER CROP ROOT EXUDATES DIFFERENTIALLY INFLUENCE SOIL MICROBIAL FUNCTION AND PHYTOHORMONE PRODUCTION IN AGRICULTURAL SOIL MICROBIOMES

4.1 Introduction

Plant roots and their root exudates alter the physical and chemical properties of the soil immediately surrounding them, representing a primary target for enhanced soil health outcomes. Addressing crop-microbiome metabolic exchanges and the impacts those interactions have on nutrient cycling is of paramount importance in agroecosystems. Enhancing rhizosphere mediated processes during traditionally non-crop phases through crop diversification via cover cropping is an attractive complementary approach to breeding efforts that also prioritize selection for functional attributes that cultivate favorable microbiomes.

Crop diversification through cover cropping is the incorporation of a secondary, unharvested crop grown to support the development of a primary cash crop species. Plants used as cover crops are cultivated in conjunction with or following primary cash crops to enhance soil health, nutrient cycling, and pest management as a function of the increased biodiversity these crops introduce to the soil. Various plant species serve as cover crops and often cropping multiple species synergistically can maximize functional outcomes^{89,91}. Cover cropping can influence aboveground plant growth responses and belowground soil characteristics to promote soil and cash crop health alike^{93,202}. For instance, aboveground, cover crops act as a physical barrier against wind and water erosion, protect against weeds, and when utilized between periods of fallow, cover crops can contribute to the maintenance of soil functionalities that bolster primary cash crop yields^{2,5,99,203}. Additionally, belowground, cover crops stimulate diverse microbial populations that catalyze beneficial biogeochemical processes^{86,204}. The impact of cover cropping belowground are influenced by both the additional plant litter and responses stimulated by cover

crop root exudation^{3,14}. Here we focus on the latter, investigating how different cover crops can stimulate specific microbial populations and processes through exudation.

Metabolites released as root exudates include diverse sugars, organic and amino acids, enzymes, and secondary plant metabolites. These compounds, primarily water-soluble and of low molecular weight, are released into the rhizosphere, altering the chemical landscape adjacent to plant roots and influencing the types and quality of microbial recruitment and associations that are beneficial for plant hosts^{19,50,205}. For example, sugars and amino acids are primarily carbon and nitrogen sources acting as growth substrates for rhizosphere microorganisms, stimulating microbial activity near roots, which provides secondary benefits of enhanced nutrient acquisition and pathogen suppression from generalist that rapidly consume primary metabolites^{7,24,206}. Secondary metabolites in exudates, like flavonoids, contribute to microbially-mediated plant defense responses and plant-microbe signaling and play a crucial role in shaping the overall ecology of the rhizosphere^{207,208}. However, harnessing exudates for targeted stimulation of the rhizosphere is complicated by the diversity in both quantity and quality of root exudate compounds resulting from different crop species, and cultivars within species, which is altered in response to growth phase, abiotic and biotic stress, and many edaphic conditions^{14,16,36,209,210}.

In the most recent national survey conducted in 2020, more than 15 million acres utilized cover cropping across 153,000 U.S. farms⁹⁶. Despite the popularity of cover cropping as a regenerative agriculture practice, the complex interface between plants and the soil environment underpins essential ecological processes for crop health, yet remains largely unresolved in cover cropping systems. There is a paucity of data on cover crop root exudation chemical diversity, which extends to an even larger knowledge gap on how cover crops can modulate soil biogeochemistry, both during cover crop cultivation and after periods of termination and extension to the cash crop (i.e., legacy effects). In fact, unlike many primary cash crops, most cover crop

species have undergone very little selection and breeding, leaving an opportunity for chemical tunability through the manipulation of root exudates in the soil landscape for desired outcomes.

We previously used metabolomics to characterize variability in root exudation across 19 commonly used cover crop species. We assessed primary and secondary root exudates, but importantly, quantified low abundance phytohormones released as exudates³⁶. The cover crop species selected here are representative of four functional classes (legumes, brassicas, cool-season grasses, non-legume summer annuals) of cover crops and our team's work demonstrated these species had distinct root exudate chemical profiles, a finding we hypothesized could yield distinct microbial outcomes. To decipher the microbial responses to amendments with different cover crop root exudate treatments, we utilized laboratory-scale soil microcosms and stimulated microbial communities with daily amendments of pure root exudates over 5 days (**Figure 13A**), tracking the responses over a 21-day experiment using metabolomics and genome-resolved metagenomics and metatranscriptomics (**Figure 13B**). Integrating genome-resolved metagenomics and metatranscriptomics, our findings demonstrate microbial membership yielded a less sensitive response than metabolomics, but functional gene expression and chemical behavior were more tightly coupled, especially for phytohormone metabolisms. Importantly, we provide a public genomic resource for exudate-stimulated microorganisms and new insights provide an important first step towards the development of precision cover cropping, supporting sustainable agriculture management practices which safeguard valuable ecosystem resources.

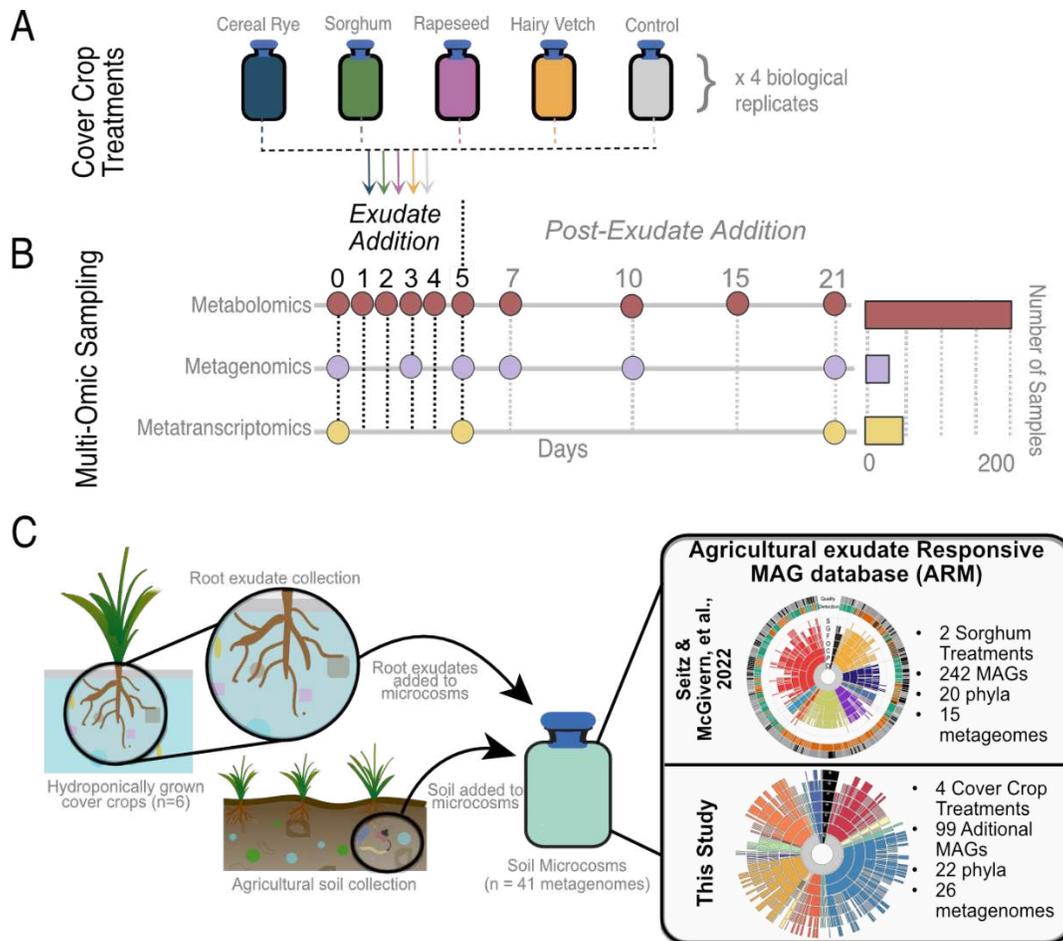


Figure 13. Curation of an exudate-responsive MAG database. A) Root exudates were collected from 4 hydroponically grown cover crops³⁶. The experimental design included soil reactors (amended with soil from the agricultural research station at Colorado State University) in biological quadruplets from 5 treatments (i) cereal rye (dark blue), (ii) sorghum (green), (iii) rapeseed (light purple), (iv) hairy vetch (orange) and a water amended soil control (grey). B) Microcosm reactors were amended with pure root exudates from treatments in (a) for 5 days (denoted as the exudate addition phase) and responses were surveyed for 21 days. Metagenomes, metatranscriptomes, and metabolomes were collected at specific timepoints, indicated by circles, to profile microbial responses. The number of samples collected for each multi-omic measurement is indicated by the bar charts on the right. C) Schematic summary of the collection of data for the MAG database which spans two studies to collectively make up the Agricultural Exudate-Responsive Metagenomic (ARM) Database.

4.2. Results & Discussion

4.2.1. *Agricultural Exudate-Responsive Metagenomic (ARM) Database Expands Genomic Knowledge in Colorado Soils*

Obtaining high-quality metagenomic data from field settings is challenging due to the dynamic and variable nature of these environments, despite the need for field-level experiments to determine the extent to which precision agriculture can be tailored. Challenges like diverse microbial communities, difficulties in sample collection and preservation, and high strain variability complicate the ways in which metagenomics can be applied at a large scale. Thus, a necessary first step toward building a metabolic blueprint of these agriculturally relevant taxa, many of which have not been genome sequenced or cultivated, is through the use of intermediate complexity laboratory-scale ecosystems (microcosms) to link the simplified chemical cover crop exudate data from findings in the more complex field trials. While not a perfect replica of natural ecosystems, here, microcosms provide controlled and replicable environment for amplification of microbial genomes to aid in genome tractability in these soils.

Our dereplicated metagenome-assembled genome (MAG) database provides, to our knowledge, the first exudate-responsive resource of the microbial community members of agricultural soils, which we call ARM (Agricultural Exudate-Responsive Metagenomic (ARM) Database). Two shotgun metagenomic sequencing datasets derived from experiments utilizing root exudate stimulated microcosms were used to taxonomically classify abundant MAGs that respond to root exudate amendments (**Figure 13C**). In both experiments, cover crops were grown hydroponically to collect the fraction of water-soluble root exudates and their metabolomes were assessed using a combination of untargeted gas- and liquid chromatography-mass spectrometry (GC-MS and LC-MS), and a targeted LC-MS/MS assay to characterize low abundance phytohormones. In experiment 1 (**Figure 13C**; denoted as “Seitz & McGivern, *et al* 2022”), root exudate treatments were formulated after a high sugar exuding sorghum genotype (*Sorghum*

bicolor BTx623) and a high organic acid sorghum genotype (*Sorghum bicolor* Leoti) in which synthetic treatments were created to amend to soil microcosms to determine how sorghum genotypes influence microbial composition and function¹⁸. In the second experiment (**Figure 13C**; denoted as “*This study*”), four cover crops were chosen based on their distinct root exudate profiles³⁶, and pure root exudates were collected hydroponically from those plants to be directly added to microcosms. These cover crop root exudate treatments represent common cover crops that exude distinct profiles during the seedling stage. The cover crops used in experiment 2 were: sorghum (*Sorghum bicolor*), hairy vetch (*Vicia villosa*), rapeseed (*Brassica napus*), and cereal rye (*Secale cereale*) (**Figure 13A**). In both experiments, root exudates were amended to microcosms during a 5-day exudate addition phase and surveyed until 21 days (post exudate-addition phase). Metagenomes were collected from experiment 1 in triplicate from days 0, 5, and 21 (n=15) and in experiment 2, metagenomes were collected from each treatment on days 0 (control only), 3, 5, 7, 10, 21 (n=26) to capture changes in community composition and functional potential occurring before, during, and after exudate addition for a total of 41 metagenomes that make up ARM (**Figure 13C**).

With a total of 41 metagenomes across two experiments, we reconstructed 342 medium- and high-quality MAGs dereplicated at 99% identity into 326 clusters, representative of distinct microorganisms (**Figure 14A**). Highlighting the genomic novelty of the soils, the ARM genomic database contains 10 MAGs which may represent previously unidentified (i.e., lacking a taxonomic assignment at this level) orders (3%), 39 MAGs that belong to previously unidentified genera (11.9%) and 170 MAGs that belong to previously unidentified species (52.1%) further emphasizing the phylogenetic novelty yet to be captured in agricultural soils (**Figure 14B**). Additionally, a large proportion of our MAG database belonged to lineages only recognized by alphanumeric identifiers (e.g., Draft Genome Sequence, DSM) in GTDB-tk at the class (n=7), order (n=46), family (n=104), and genus (n=126) levels indicating these lineages are undescribed

in soils and larger MAG catalogs. Ultimately, ARM provides new genomic information and resources for ecologically relevant taxa and is a public MAG resource that can be used to enable taxonomic analyses and metabolic reconstruction of microbial metabolisms in agricultural soils.

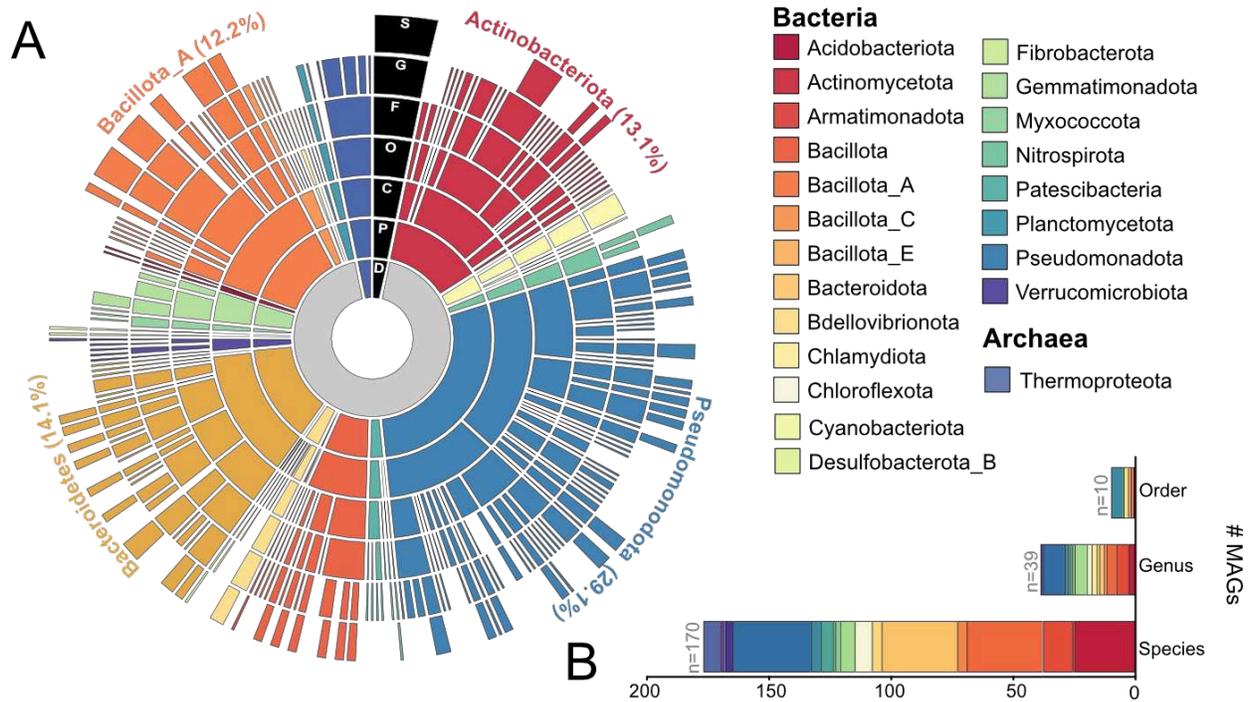


Figure 14. ARM Database taxonomy. A) Taxonomy of the 326 dereplicated MAGs in ARM. Ring color corresponds to phylum assignment with the color legend to the right. The top four most dominate phyla are written in text on the outer ring and colored according to phyla. B) Stacked bar graph shows the number of dereplicated MAGs recovered that represent novel orders, genera, or species according to taxonomy assignments from GTDB-tk. Coloring corresponds to MAG phylum.

We used ARM to inform on the community make-up across our microcosm treatments and timeseries. Based on read mapping, an average of 38% of reads that assembled mapped back to ARM, signifying that the underlying assemblies were well represented in the MAG database. Accounting for relative abundance, the top 10 most dominate MAGs across all treatments, including the control, were either from the *Sphingomicrobium*, *Arthrobacter*, or

Pseudomonas_E genera or from alphanumeric genera from the classes *Vicinamibacteria*, *Acidimicrobiia*, *Bacteroidia*, and *Gammaproteobacteria*. During the exudate addition phase (days 0-5), all treatments enriched for MAGs from the classes *Vicinamibacteria*, *Acidimicrobiia*, *Bacteroidia*, while only the exudate treatments enriched for the MAGs from the *Gammaproteobacteria* class and a MAG from the *Sphingomicrobium* genus (**Appendix C**). During the post exudate addition phase (days 7-21), the enrichment of a *Pseudomonas_E borbori* MAG was shared across all five treatments, but by day 21, the four exudate treatment microcosms were more enriched for this species than the control. By day 21, the treatment microcosms were more diversely enriched for these dominate taxa than in the control. The top 4 bacteria enriched only in the treatment microcosms were from four diverse lineages representing taxa from the *Arthrobacter* and *Sphingomicrobium* genera and the *Burkholderiales* and *Flavobacteriales* orders. *Arthrobacter* and *Burkholderiales* in particular contain members that are known rhizosphere colonizers and taxa within these lineages are typically associated with plant-growth promotion traits like nitrogen fixation and pest suppression, respectively ²¹¹⁻²¹³

4.2.2. Microbial Community Function, but Not Membership, is Altered by Exudate Addition

Next, we wanted to examine the microbial community response to exudates at the membership and functional levels. We used the ARM database to resolve metatranscriptome expression profiles across the enrichment timeseries. At the MAG and function levels, metatranscriptome expression differed significantly by time, but not by exudate treatment. This suggests the agricultural soils used as inoculum may already be adapted to exudate inputs, or that this measure of community function was not strongly affected by the amount of exudates added per gram of soil in the context of this experiment, which aimed to replicate average quantities of exudates release over a 24h period. In both the control and exudate treated microcosms, MAGs from four genera in *Nitrososphaeraceae* (phylum *Thermoproteota*, previously

Thaumarchaeota) contributed roughly 50% of exudate-addition phase metatranscriptome expression. By the end of the enrichment, all control and treatment microcosm metatranscriptomes were dominated by MAGs from the *Bacteroidota*, including unclassified genera in the *Crocinitomicaceae* and *JADKCL01* families.

As we wanted to understand the impact of exudate addition on the soil microbiomes, we focused on the metatranscriptomes from day 5, representing the exudate addition phase. We first examined MAG expression for individual exudate treatments relative to the controls. No MAGs were discriminant between control and cereal rye, or between control and rapeseed, exudate treatments. One MAG from an undescribed species in the *Actinomycetota* (*JAJTCL01 sp021323255*) was discriminant to the hairy vetch exudate microcosms, and a MAG from a novel genus in the *Patescibacteria* (family *UBA1547*) was discriminant to the sorghum exudate microcosms (Maaslin2, $q < 0.25$). Next, we wanted to know what MAGs broadly responded to exudate addition. We found 5 MAGs were enriched in the exudate microcosm metatranscriptomes (Maaslin2, $q < 0.25$). The latter group included two MAGs in the *Gammaproteobacterial* genus *Hydrogenophaga*, one MAG classified as *Nitrobacter vulgaris*, and MAGs representing novel species of *Arthrobacter* and *Paucimonas*. Collectively, this suggests exudate compounds stimulated specific microorganisms in the broader community.

Next, we wanted to examine the impact of exudate stimulation on microbial community gene expression. We aggregated genes at the annotation level to determine functions that were discriminant to control and exudate amended microcosms. We identified 114 microbial functions that were enriched in the control metatranscriptome at day 5, and 145 functions that were enriched in the exudate amended microcosms (Maaslin2, $q < 0.25$). This analysis revealed genes involved in cellular activity (ex. ribosomal proteins, RNA polymerase, anabolic pathways, etc.) were more highly expressed in exudate treatment microcosms than in control. This suggests the exudate inputs enhanced microbial metabolic activity. In support of this, genes for sugar and amino acid

transport are enriched in the control metatranscriptome, which may suggest that exudate treated microcosms are replete with these resources.

4.2.3. Root Exudates from Cereal Rye and Hairy Vetch Had a More Pronounced Impact on Restructuring Microbial Metabolic Outputs.

Our next aim was to investigate the microbial community response to exudates at the metabolite level. We combined non-targeted and targeted LC-MS/MS data which resulted in the detection of 641 molecular features to resolve community metabolic changes across the enrichment timeseries. Compared to the metagenome and metatranscriptome responses, the metabolite data showed much stronger structuring by treatment and time (PERMANOVA, $p < 0.001$). We found that cereal rye and hairy vetch microcosms were significantly different compared to soil controls at each timepoint following the first day of exudate addition (PERMANOVA, $p < 0.05$), indicating these exudates rapidly influenced microbial outputs. In the sorghum and rapeseed treatment, soil metabolic changes in response to exudate amendments were only significantly different compared to controls following day 3 and 5 of exudate addition, respectively, as well as the post-exudate addition phase, suggesting these treatments do not stimulate as strong a response as cereal rye and hairy vetch exudates, however these treatments still prompted changes (PERMANOVA, $p < 0.05$).

To broadly visualize changes in metabolisms over time and across treatments, we utilized partial least squares discriminate analysis (PLS-DA) to visualize differences in metabolomes (**Figure 15A**). The most distinct metabolic changes were observed between the metabolomes of cereal rye and hairy vetch treatments relative to the control during periods of exudate addition. Then, in the post-exudate addition phase (days 10, 15, and 21), the metabolomes clustered more similarly to the starting inoculum, suggesting a functional convergence following the lack of stimulation by exudates, although these metabolomes were still significantly different from the

controls (PERMANOVA, $p < 0.05$). This may suggest a shorter-lived legacy effect following the microbial changes stimulated by exudate treatments.

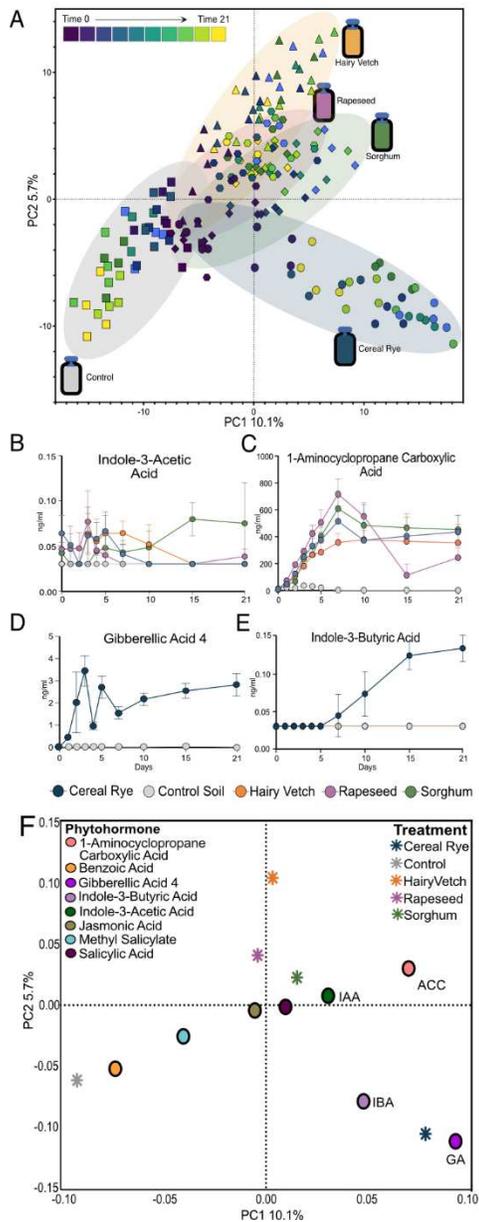


Figure 15. Cover crop treatment influences microbial metabolome and phytohormone biosynthesis across time. A) Two-dimensional scores plot of PLS-DA ($R^2_X = 0.101$, $R^2_Y = 0.0597$, $Q^2 = 0.784$, PERMANOVA $p < 0.001$) between all treatment metabolomes. Ellipses represent 95% confidence intervals and are colored by treatment. Shapes denote treatment as well, cereal rye (circles), hairy vetch (triangles), sorghum (diamonds), octagons (rapeseed) and control (square) metabolomes and colors within shapes represent time. B-E) Line graphs show concentration of 4 abundant phytohormones colored by treatment. Circles represent the average concentration at that timepoint and error bars represent standard deviation. (B) Indole-3-acetic

acid (IAA) (C) 1-aminocyclopropane carboxylic acid (ACC) (D) gibberellic acid 4 (GA4) (E) Indole-3-butyric acid (IBA). F) PLS-DA biplot shows the relative contribution of each phytohormone to each treatment. Colored stars indicate a treatment and colored circles indicate a phytohormone. Colored circles with labels indicate phytohormones discussed in the text.

4.2.4. Specific Microbial Phytohormone Metabolisms are Captured in Laboratory Microcosms

To survey the metabolites and metabolite classes contributing to the changing treatment metabolomes, we combined unannotated and annotated metabolites and taxonomically classified compounds using ClassyFire²¹⁴ into compound superclasses. We detected metabolites across 14 superclass: Lipids and lipid-like molecules (179), organoheterocyclic compounds (106), organic acids and derivatives (97), phenylpropanoids and polyketides (64), organic oxygen compounds (52), benzenoids (38), nucleosides, nucleotides, and analogues (20), alkaloids and derivatives (18), organic nitrogen compounds (15), lignans, neolignans and related compounds (7), organosulfur compounds (3), organic 1,3-dipolar compounds (1), organic polymers (1) (**Appendix C**).

From our targeted methods, we quantified 8 phytohormones in total: indole-3-acetic acid (IAA), indole-3-butyric acid (IBA), gibberellic acid 4 (GA), 1-aminocyclopropane carboxylic acid (ACC), salicylic acid (SA), methyl salicylate (mSA), benzoic acid (BA), jasmonic acid (JA). In light of phytohormone production and consumption patterns observed in the microcosms, we focused on microbial responses to phytohormones as these compounds can have a significant impact in shaping plant-microbe interactions, soil health, and ecosystem functioning^{62,215}. IAA, IBA, BA, GA, and ACC had the strongest effect on treatment metabolomes out of the phytohormones surveyed (**Figure 15F**). Specifically, IAA and ACC, were detected in each cover crop treatment but not in the control (**Figure 15B-C**). Within all exudate treatment microcosms, ACC accumulated during the exudate addition phase and decreased in the post exudate addition phase. IAA, on the other hand, accumulated during the exudate addition phase in all treatments, but in sorghum exudate amended microcosms, IAA concentrations increased during the post-addition phase as well.

Cereal rye microcosms enriched production for two different phytohormones, GA4 and IBA (one-way ANOVA, $p < 0.05$); IBA was found exclusively in the post-exudate addition phase and GA4 was found at all timepoints following day 0 (**Figure 15D-E**). These data suggest root exudate treatment stimulated microbial phytohormone biosynthesis, which motivated us to investigate microbial genomes and coupled transcriptomes for members capable of these biosynthetic routes.

4.2.3. Microbial Production of ACC-Deaminase in Response to Addition of ACC

The non-proteinogenic 1-aminocyclopropane carboxylic acid (ACC) is an amino acid derivative and can act as an assembly cue for plant growth promoting (PGP) bacteria (PGPB) in the rhizosphere^{216,217}. Here, we detected the presence of ACC in each exudate treatment at a significantly higher amount than in the control (**Figure 15C**; one-way ANOVA, $p < 0.001$). The abundance of ACC increased with time during the exudate addition phase and at day 7, and then decreased or stayed the same, in the post-exudate addition phase indicating the degradation of the metabolite, with the largest conversion of ACC in the rapeseed treatment. Many PGPB metabolize ACC via deamination, thereby converting ACC into ammonium and α -ketobutyrate using the enzyme ACC deaminase (*acdS*), resulting in the production of microbial nitrogen and carbon sources, respectively. In fact, this gene has been cited as an important trait for PGPB rhizosphere competence and plants that harbor *acdS* are more stress resistant²¹⁸.

We mined the ARM MAGs for ACC deaminase genes and found 66 genes from 63 MAGs annotated as *acdS*. As ACC deaminase genes can be misannotated as D-cysteine sulfhydrase^{219,220}, we curated these putative hits with key active site residues to identify 13 high confidence *acdS* genes from distinct MAGs in the *Actinomycetota* ($n=9$) and the *Pseudomonadota* (formerly *Proteobacteriota*, $n=4$). We next wanted to examine the expression of these genes in the metatranscriptomes. ACC deaminase expression was detected in 2 *Actinomycetota* MAGs across

treatments and timepoints. One MAG in the genus *Pseudonocardia* expressed *acdS* at day 0 in the control soils, and at day 5 in the control soils and the sorghum and cereal rye treatments. The other MAG, a member of the genus *Ornithinibacter*, expressed *acdS* at day 0 in the sorghum treatment, at day 5 in the hairy vetch treatment, and at day 21 in the cereal rye treatment. Analysis of the broader gene expression of these two MAGs suggests they may be using the 2-oxobutanoate generated from ACC deaminase activity in isoleucine biosynthesis. To the best of our knowledge, this represents the first report of ACC deaminase potential for members of these two genera. In summary, we show the power of combining metabolite data with microbial genomic expression for highlighting how phytohormones may support soil microbial lifestyles.

4.2.4. Production of IAA & IAA-like Indole Derivatives is Stimulated by Rapeseed, Hairy Vetch and Sorghum Treatments, but not Cereal Rye.

Microbial biosynthesis of indole derivatives, like auxins, can have important impacts on plant health which makes characterizing microbial indole production an important avenue for evaluating the impact of PGP microbes^{62,138}. Mining our untargeted LC-MS data, complemented by targeted approaches (LC-MS/MS), we utilized *in silico* chemical structural predictions from MS-FINDER to determine taxonomic lineage of unknown compounds that shared structural similarity to the known phytohormones captured in our targeted methods (i.e., IAA and IBA). We first completed this analysis for indole compounds and found two indoles that shared structural similarity with IAA and thus were annotated as unknown indoles classified at a level 3 annotation¹⁰⁴. These indoles were from the subclasses (i) indole and (ii) indolyl carboxylic acids and derivatives and could be further categorized as a 3-alkylindole compound and an IAA derivative, respectively, at a parent level 1 according to ClassyFire²¹⁴.

First, we tracked changes in known phytohormones, IAA and IBA. IAA concentrations fluctuated in all treatment microcosms over time, but was detected at a higher relative abundance

in sorghum by day 21 compared to the control (**Figure 15B**) although this was not statistically significant due to the low abundance of this compound (2-way ANOVA, $p>0.05$). IBA, on the other hand, accumulated significantly by days 15 and 21 in cereal rye treated microcosms, indicating this auxin was likely produced during the post-exudate addition phase (2-way ANOVA, $p<0.002$; **Figure 15E**). The unknown IAA derivative was significantly more abundant in the rapeseed treatment compared to the control at each timepoint following day 1 (2-way ANOVA, $p<0.0001$, **Figure 16A**). Exhibiting a broader response, the unknown 3-alkylindole was significantly enriched at each timepoint following the first day of exudate addition in rapeseed, sorghum, and hairy vetch compared to the control (2-way ANOVA, $p<0.0001$, **Figure 16B**).

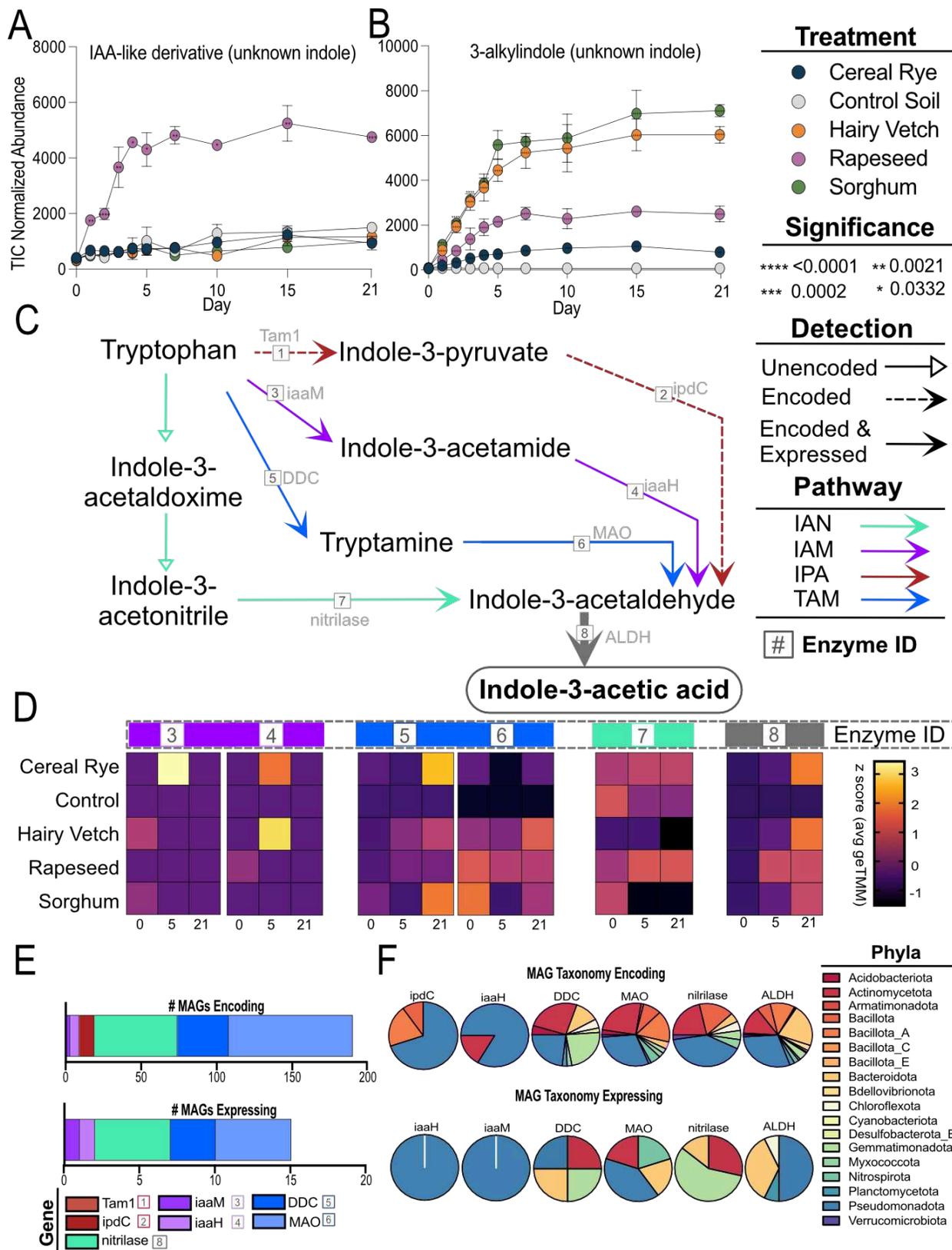


Figure 16. Production of Indole-3-Acetic Acid (IAA) & IAA-like compounds from diverse lineages in rapeseed, sorghum, and hairy vetch microcosms. A-B) Line plots represent TIC normalized abundances of two unknown IAA-like compounds, an unknown IAA-like derivative and an unknown indole, respectively, annotated at a level 3 annotation. Color corresponds to treatment, circles represent the mean, error bar represent standard deviation. Significance is indicated by an asterisk (*): 0.0332 (*), 0.0021 (**), 0.0002 (***), 0.0001 (****). C) IAA biosynthetic pathways. 4 routes for IAA were detected by metagenomics (metaG) or metatranscriptomics (metaT), detection is indicated by arrow type: hollow arrows indicate the gene was not detected via metaG, dotted arrows indicate the gene was only encoded, solid arrows indicate the gene was encoded and expressed (metaT). Enzymes names are in grey next to the corresponding reaction arrow, and given an enzyme #. Colors correspond to IAA pathway. The final product, IAA, is indicated as a boxed metabolite. D) Heatmaps show the zscored geTMM value for each summed gene abundance across all MAGs expressing the gene (corresponding to the enzyme number in (C)) within a treatment and timepoint. Enzyme number corresponds to biosynthetic routes in C. E) Summed abundance of all MAGs across all treatments either encoding an IAA-production gene (top) or expressing the gene (bottom). Colors correspond to IAA pathway and number corresponds to enzyme ID. F) pie charts show the distribution of each IAA-production gene summed at the phyla level for MAGs encoding (top) or expressing (bottom) the gene. Color corresponds to phyla assignment.

IAA, and indole intermediates, can be synthesized via five major pathways derived from tryptophan (Trp) metabolism resulting in the shared terminal oxidation step converting indole acetaldehyde (IAAld) to IAA by aldehyde dehydrogenase (*ALDH*) (**Figure 16C**)^{138,172}. We profiled genes for IAA pathways in the ARM database and found 4 pathways were represented in our MAGs. IAA can be synthesized from Trp via transamination (*Tam1*) to indole-3-pyruvate (IPA) and decarboxylated to IAAld via the IPA decarboxylase (*ipdC*) gene. Despite 10 MAGs carrying *ipdC*, we did not detect any MAGs encoding *Tam1*, suggesting this pathway is not fully represented in these soil genomes. Providing further evidence for the lack of IAA biosynthesis from the IPA pathway, there was no expression of *Tam1* or *ipdC* enzymes in any transcriptomes. The second pathway, indoleacetonitrile (IAN) pathway, is partially unknown for proximal oxidoreductases in bacteria (i.e., Trp → indoleacetaloxime), but the last step of the pathway hydrolyzing IAN to IAA catalyzed by nitrilase, was encoded in 55 MAGs. Nitrilases expressed only in treatment and not controls at days 5 and 21, were found in 5 MAGs from three genera in *Actinomycetota*, one in *Pseudomonadota*, and one in *Chloroflexota*. Expression of nitrilase,

averaged across these MAGs within a treatment revealed nitrilase expression was most abundant in cereal rye and rapeseed microcosms (**Figure 16D**).

Completing the two step IAM pathway, 3 MAGs encoded tryptophan 2-monoxygenase (*iaaM*) and 6 MAGs encoded indoleacetamide (IAM) hydrolase (*iaaH*). Of these MAGs, 1 MAG from the *Geminococcaceae* family expressed *iaaM* while a second MAG from the *Burkholderiales* order expressed *iaaH*. Both genes were expressed across the rapeseed, hairy vetch, and cereal rye exudate treatments at day 5, but not in the control (**Figure 16D**). There were no MAGs carrying both *iaaM* and *iaaH* to complete the two-step IAM pathway, suggesting the possibility MAGs may catalyze Trp into disparate intermediates, sharing only partial responsibility in IAA synthesis, as others have shown²²¹. The expression of the IAM pathway may elude to the utilization of this pathway by *Geminococcaceae* family or *Burkholderiales* order members as stimulated by cereal rye, hairy vetch, and rapeseed exudates, but not sorghum, contributing to IAA synthesis in these microcosms.

34 MAGs encoded the first step of the tryptamine (TAM) pathway, decarboxylating Trp to TAM by an aromatic-L-amino-acid/L-tryptophan decarboxylase (*DDC*) and 3 MAGs expressed *DDC* in the cereal rye, hairy vetch, and sorghum treatment microcosms but not the controls at day 21. Average expression per day of *DDC* revealed this enzyme was most active at day 21 in cereal rye and sorghum microcosms (**Figure 16D**). TAM is then oxidized to IAAld by monoamine oxidase (*MAO*), a gene encoded by 82 MAGs and expressed in 5 MAGs found in hairy vetch, rapeseed, and sorghum treatment microcosms at day 21 (**Figure 16D**). Finally, 215 MAGs from ARM encoded the NAD-dependent *ALDH* enzyme oxidizing IAAld to IAA. 15 of those MAGs expressed *ALDH* in each treatment microcosms, but not in the control. On average per day, MAG expression of *ALDH* in cereal rye and hairy vetch microcosms expressed higher relative amounts of *ALDH*, while rapeseed and sorghum produce less by day 21. The production of *ALDH* in these

15 MAGs at day 21, and not in the control, suggests these MAGs contributed the synthesis of IAA and IAA-like molecules over time.

Previous studies have demonstrated that bacterial genomes are primed to produce IAA via the IPA or IAM pathways, but our results highlight that the soil microbes analyzed here more dominantly encoded and expressed genes from the TAM and IAN pathways (**Figure 16 E-F**). Out of the 190 MAGs expressing at least one gene across the IAA biosynthesis pathways, 61% of MAGs encoded genes from the TAM pathway, while only 5% carried genes for the IAM pathway, and 29% encoded for nitrilase (**Figure 16E**). Out of the 15 MAGs expressing genes for IAA synthesis in treatment microcosms, 20% and 33% expressed *DDC* and *MAO*, respectively, as a part of the TAM pathway, while only 7% expressed *iaaM* and *iaaH*, and 33% expressed nitrilase (**Figure 16E**). These results suggest that these soil bacteria have very high potential to synthesize IAA via the TAM pathway, followed by IAN, IAM, and then the IPA pathway. Different from other studies that suggest the IPA is the primary bacterial IAA synthetic route, we found bacteria from varying phyla preferred different pathways (**Figure 16F**). Our results suggest that MAGs capable of metabolizing IAA from Trp via IPA or IAM were primarily from the *Pseudomonadota*, *Acidobacteriota*, or *Bacillota* phyla while MAGs synthesizing IAA from TAM or IAN were much more diverse (**Figure 16F**). This result indicates the IAA-producing trait is widely distributed in soil MAGs, and that multiple heterogenous root exudation strategies stimulate MAGs carrying these genes to upregulate indole biosynthetic routes.

4.2.5. Cereal Rye Exudates Stimulate Potentially New Bacterial Species with Gibberellic Acid 4 Biosynthetic Capacity.

We next examined the impact of cereal rye root exudate treatments on microbial gibberellic acid 4 (GA4) biosynthesis. A significant difference in the cereal rye metabolome as compared to the control was observed (**Figure 17A**; PERMANOVA; $p < 0.0001$), suggesting

unique capacity of cereal rye root exudates to modulate microbial community response. In particular, we observed GA4 was produced over time (**Figure 15D**) and two additional unknown gibberellins detected via nontargeted LC-MS/MS were detected at higher abundances in the cereal rye treatment microcosms over time compared to the control (**Figure 17A-B**). These two unknown gibberellins were classified as type C20 and C19 gibberellins according to ClassyFire and were significantly different between cereal rye metabolomes versus the control at all timepoints following the first day of exudate addition (two-way ANOVA, $p < 0.0001$; VIP > 1.9).

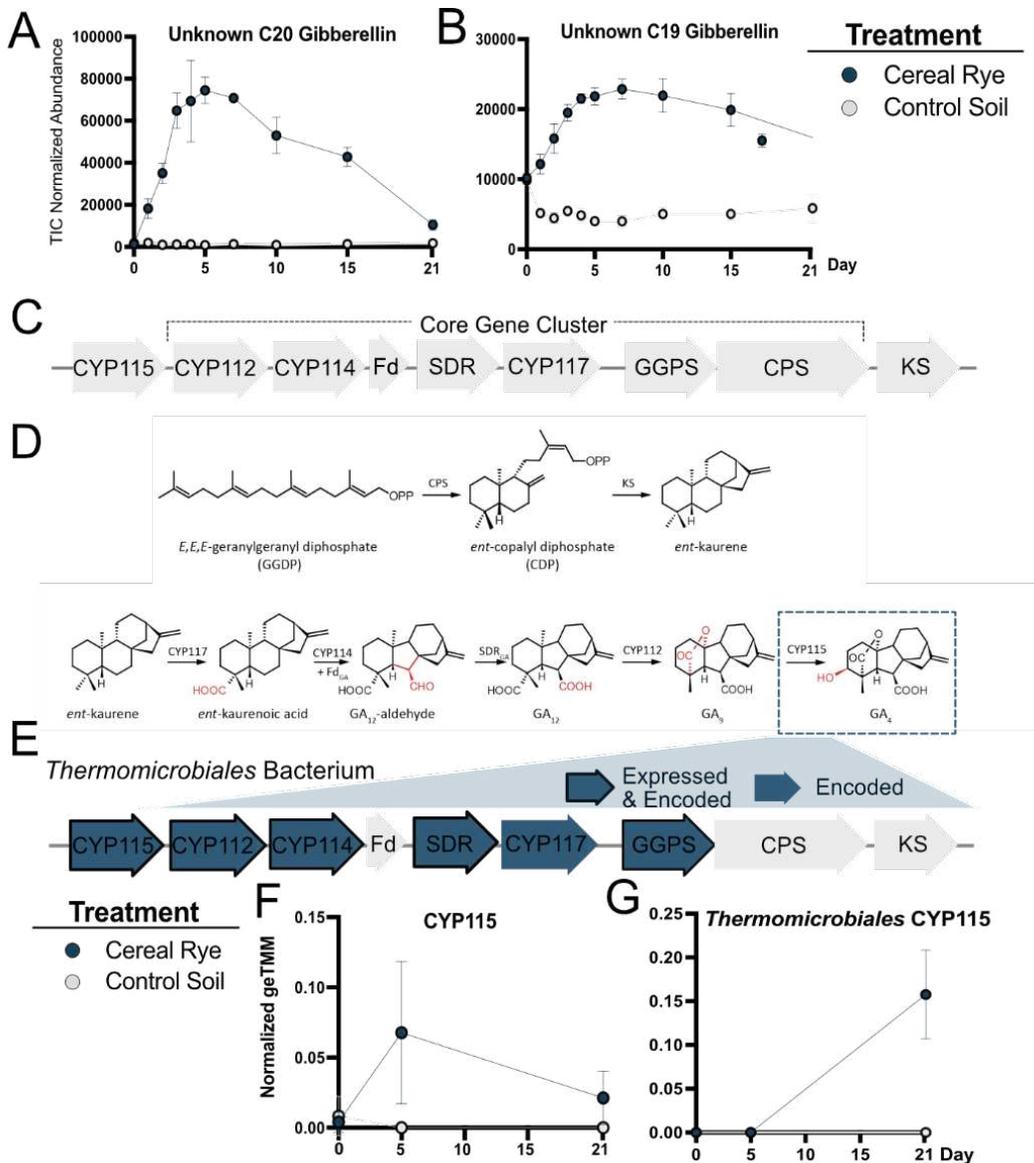


Figure 17. Gibberellic acid (GA) biosynthesis in cereal rye microcosms. A-B) Line plots represent TIC normalized abundances of two unknown GA-like compounds, an unknown C20 type GA and an unknown C19 type GA, respectively, annotated at a level 3 annotation. Color corresponds to treatment, circles represent the mean, error bar represent standard deviation. C) GA biosynthetic operon. The core operon CYP → KS is primarily conserved across bacteria, while *CYP115* and *KS* have only been found in certain bacterial strains. D) GA biosynthetic route beginning with the 20-carbon isoprenyl precursor (E,E,E)-geranylgeranyl diphosphate (GGDP) and ending with the bioactive GA₄, outlined with a blue dotted square. E) A potentially-new bacterium of the *Thermomicrobiales* order expressing 5 out of 9 GA biosynthetic proteins required to produce GA₄, including the final enzyme converting GA₉ → GA₄, the GA found most abundant in cereal rye microcosms at day 21. F) Averaged geTMM abundance of MAGs producing *CYP115* protein in cereal rye and control microcosms. G) Expression of *CYP115* of a novel *Thermomicrobiales* bacterium.

GAs are first derived from isoprenoid diphosphates which are converted to the 20-carbon isoprenyl precursor (E,E,E)-geranylgeranyl diphosphate (GGDP), by the geranylgeranyl diphosphate synthase (*GGPS* or *idsA*) gene. GGPP is transformed into *ent*-copalyl diphosphate (CDP) by *ent*-CPP synthase (*CPS*) followed by cyclization to form *ent*-kaurene by *ent*-kaurene synthase (*KS*), creating a key intermediate in the GA biosynthesis pathway. *ent*-kaurene is then oxidized to *ent*-kaurenoic acid by cytochrome P450 (CYP) monooxygenase 117 (*CYP117*) before a final oxidation to GA₁₂-aldehyde by *CYP114*. GA₁₂-aldehyde is modified to form GA₁₂ by a short chain dehydrogenase (*SDR*), converted to GA₉ by *CYP112*, and finally bioactivated to GA₄ by GA-3-oxidation (*CYP115*)²²². Finally, GAs can be inactivated gibberellin 2-oxidases (**Figure 17D**).

To investigate possible producers and inactivators of GA₄ to understand its role as a plant signaling molecule for growth and development, we first used a BLAST analysis with individual protein sequences of each gene as the query to survey MAGs containing genes located in the GA operon²²³. MAGs that contained sequence matches were considered if the BLAST E value was 10⁻¹⁰ or less and the bitscore was >150. We found 176 MAGs that encoded at least one enzyme in the GA operon. We mined cereal rye metatranscriptomes for *CYP115* to determine MAGs producing GA₄, as this gene encode the necessary step to produce GA₄ (**Figure 17D**). We found 5 MAGs expressing *CYP115* uniquely in the cereal rye treatment at days 5 and 21,

suggesting these MAGs are important for GA4 production in response to cereal rye exudate (**Figure 17F**). Interestingly, we found 1 MAG which encoded 6 of the 9 GA operon genes, and actively expressed 5 (*CYP112*, *CYP114*, *CYP115*, *GGPS*, *SRD*) of the proteins at day 21 in the cereal rye treatment and not in the control, corresponding to a high concentration of GA4 in the cereal rye microcosms (**Figure 17G**) indicating this MAG may be important in the production of GA4 under the conditions tested here. This MAG, from the *Thermomicrobiales* order represents a potentially-new bacterium with GA4 producing capabilities. As such, GA4, and related GA, biosynthesis by the soil microbiome and its relationship with cereal rye's utility as a cover crop, are areas of research that warrant more investigation as microbial production of GA4 is influential plant growth and may play an emerging role in shaping microbial communities and functions in the rhizosphere.

4.3 Conclusions

Integrating a simplified and manageable laboratory soil system with advanced high-resolution multiomics techniques afforded new perspectives into the intricate interplay between plant root exudates and the soil microbiome. We first provide ARM, a public agricultural genome database, creating a community resource for use in related microbiome-cover crop workflows. We then demonstrated that root exudates from 4 cover crop amendments applied at a physiologically-relevant concentration can impact microbial metabolisms and membership function, in particular the feedback loop between the microbial degradation of the phytohormone ACC by ACC deaminase. We also highlight the impact of cereal rye root exudation on the microbial production of GA and IBA as well as the impact all 4 treatments on microbial indole biosynthesis. Future studies can scale these hypotheses to the field to analyze the impact of plant genotype, root exudates, and associated microbial communities on overall plant performance. In summary, the results of this study represent an important step toward decoding the complex

chemical language between plants and their rhizosphere microbial communities, a translation required to optimize these interactions for enhanced agricultural management and production.

4.4 Materials & Methods

4.4.1 Cover Crop Root Exudate Collections

Mature seeds from the four chosen cover crops (cereal rye, hairy vetch, sorghum, and rapeseed) were first surface sterilized in a sterile tissue culture hood by placing seeds in 15 mL conical tubes with (A) 2 mL of Clorox bleach solution (3% sodium hypochlorite. 1:1 bleach to water) + 1 drop of Tween-20 to reduce surface tension and improve sterilization or (B) 2 mL of Clorox bleach solution (3% sodium hypochlorite. 1:1 bleach to water) (see **Appendix C** for sterilization optimization). Sterilization method A or B was chosen based on optimization of germination rate for each cover crop species (unpublished data). Seeds suspended in sterilization solution were shaken for 1 min. In a sterile tissue culture hood, bleach was removed using sterile techniques and the seeds were rinsed five times with sterile deionized (DI) water. Seeds were either (A) plated to agar plates with MS basal salt mixture (2.16 mg Murashige and Skoog (MS) media in 500 mL sterile DI water; MP Biomedicals, Santa Ana, CA) germinated in the dark for 3 days, and then transferred to sterile tubes with 3 mL liquid MS media after 3 days; or (B) placed directly in sterile growth vessel with 3 mL MS media. Growth vessels were either glass tubes which were used for monocotyledon species (cereal rye and sorghum) or Magenta boxes which were used for dicotyledon species (rapeseed and hairy vetch) based on preliminary experiments to optimize growth of each seedling (unpublished data). All seeds were incubated in a growth chamber with photoperiod 16 h light/8 h night at 25°C for 14 days. The 14-day-old seedlings did not introduce any microbes to the system as no bacterial growth was observed on the MS agar plates or growth containers. Cover crop seedlings were grown for 14-days before root exudate collections. Root exudates were collected by first rinsing the roots in sterile DI water 3 times and transferring to a new, sterile vessel filled with 5mL sterile DI water for a 24-h root exudate

collection period. The suspensions containing the root exudates were then filtered through a 0.2µm filter membrane to remove root detritus. Enough root exudates were collected and pooled together to achieve a treatment addition of 1.5 mg of pure exudates per day (36 mg total). 1.5mg of exudates was chose based on preliminary experiments and averaged the amount of exudate mass exuded per day across crop species thus adding a relevant amount of exudates per day according to what was collected over a 24-h collection period³⁶. Root exudates suspensions were frozen at -20°C in 200 mL sterile glass jars. Frozen root exudate samples were lyophilized completely (~72 h), weighed, before resuspension in sterile HPLC-grade water (Thermo Fisher Scientific, Waltham, MA). Resuspended exudates were vortexed thoroughly to remove all residue from the bottle (1min of vortexing along all edges of the bottle). A volume equivalent to 1.5mg was aliquoted into sterile glass jars for daily addition and frozen at -20C until use.

4.4.2 Cover Crop Root Exudate Microcosm Experimental Setup

Microcosms were established and sampled as previously described^{18,158}. The soil (microbial inoculum) was collected from agricultural fields at the Colorado State University Agricultural Research and Education Center (CSU-ARDEC) near Fort Collins, CO on June 21 2021. The climate at the site is semi-arid, with 408 mm mean annual precipitation and a mean annual temperature of 10.2°C (1981-2010 average, <https://usclimatedata.com/>). The soil is classified as an Aridic Haplustalf. Three 2-cm diameter soil core samples to approximately 12cm depth were collected at 5 random locations within the plot. The soil was stored at -20°C until microcosm construction. 20g of soil from each plot was pooled, homogenized, and sieved (2mm) to create a representative bulk soil repository used in the following microcosms experiment. 4 g of homogenized soil and 45 mL of sterilized water were added to quadruplet sterile 100-mL glass bottles to construct each microcosm (n=20). Microcosm slurries were vortexed and allowed to settle for 5 min. Then, day 0 samples were taken by removing 3 mL of soil slurry for parallel DNA/RNA extraction, and 1mL for metaproteomics analysis which was flash frozen in liquid N.

DNA/RNA samples were centrifuged at 10,000 rcf for 10 mins to separate DNA and RNA from the metabolites present in the supernatant. The supernatant was then transferred to a new tube as the metabolome sample. After baseline sampling, 4 mL of exudate treatment (1.5mg pure, dried, root exudates in 4mL sterilized LC-MS-grade water) was added to each microcosm replicate and vortexed. At this point, for each microcosm the bottle caps were removed and replaced with a sterile foam stopper for the rest of the experiment to maintain oxic conditions and prevent colonization by contaminating microbes. Microcosms were incubated in an orbital shaker set at 200 rpm at 24°C for 21 days. Each exudate treatment was added to microcosms on days 0, 1, 2, 3, 4, and 5 at approximately the same time every day. After day 5, no additional exudate treatments were applied, but microcosms were maintained until day 21, which afforded additional samples taken at days 7, 10, 15, and 21. Samples were collected at roughly the same time each day and collected with aseptic techniques to ensure no additional microbial influence was introduced. Metabolomic and DNA/RNA samples were immediately frozen at -80°C until processing. Once ready for analysis, metabolomic samples were thawed overnight at 4°C, lyophilized, and weighed to get the amount of metabolites in each sample. Because the metabolome was surveyed using LC-MS and LC-MS/MS, we aliquoted the sample into two 0.25mg subsamples for each analysis.

4.4.2 Metabolomics: Targeted UPLC-MS/MS for Phytohormone Analysis

0.25 mg subsamples were extracted in 75 µL of a spiked methanol solution containing 100% methanol with 65.2 ng/mL ABA-d6, 62.5 ng/mL salicylic acid-d6, and 90.0 ng/mL jasmonic acid-d5 (Sigma). After solvent addition, samples were placed on a shaker plate for 1 hour at the highest speed setting, centrifuged at 3500 x g at 4°C for 5 minutes, and transferred to glass inserts. A final centrifuge step at 3500 x g for 15 minutes at 4°C was completed to ensure any precipitate was in the bulb of the vial insert. Five microliters of plant extract was injected onto an LX50 UHPLC System, equipped with an LX50 Precision Sampling Module (20-µL sample loop,

partial loop injection mode) (PerkinElmer, Waltham, MA, USA). An ACQUITY UPLC T3 column (1 × 100 mm, 1.8 μM; Waters Corporation) was used for chromatographic separation. Mobile phase A consisted of LC-MS grade water with 0.1% formic acid and mobile phase B was 100% acetonitrile. Elution gradient was initially 0.1% B for 1 min, which was increased to 55.0% B at 12 min and further increased to 97.0% B at 15 min, then decreased to 0.1% B at 15.5 min. The column was reequilibrated for 4.5 min for a total run time of 20 min. The flow rate was set to 200 μL/min and the column temperature was maintained at 45 °C. Samples were held at 4 °C in the autosampler. Detection was performed on a QSight™ 420 triple quadrupole mass spectrometer (MS) operated in selected reaction monitoring (SRM) mode. SRM transitions for each compound were optimized through analysis of authentic standards (**Appendix C**). The MS was operated with the ESI voltage 5000 V in positive mode and -5000 V in negative mode. Nebulizer gas flow was set at 350 arbitrary units and drying gas was set to 120 arbitrary units. The source temperature was 315 °C and hot-surface induced desolvation (HSID) temperature 200 °C.

4.4.4 Metabolomics: Nontargeted LC-MS and fastDDA Analysis

0.25 mg subsamples were extracted in 80 μL of 20% MeOH, sonicated for 1 h at 20 °C, centrifuged at max speed for 15 min, and transferred to glass inserts for analysis. From each sample, 5 μL was aliquoted into a separate vial to be used as a pooled QC to monitor proper instrument function and to detect any analytical variation. From each sample within a treatment, 5 μL was aliquoted into a 2 mL vial to be used as a treatment pool for fastDDA.

One microliter of each sample was injected onto a Waters Acquity UPLC system. Separation was achieved using a Waters ACQUITY UPLC Premier BEH Amide 1.7 μm Column (2.1 × 100 mm), using a gradient from solvent A (10 mM ammonium formate in water with 0.125% formic acid) to solvent B (95:5 acetonitrile:water, 10 mM ammonium formate in water with 0.125% formic acid) and a flow rate of 0.5 mL/min. Column eluent was infused into a Waters Xevo G2-XS Q-TOF-MS with an electrospray source in positive ion, sensitivity mode, with data dependent

acquisition. For individual samples, the following parameters were used for MS1 survey scan: 50-1200 m/z at 0.2 seconds per scan, switching to MS/MS after individual ion intensity rises above 10000. MS/MS acquisition occurred at a scan rate of 0.2 seconds, or at accumulated TIC threshold of 100000, with 1 MS/MS event per MS1 scan. No inclusion list was used, but dynamic peak exclusion was used with a 30 second and 100 ppm mass difference exclusion window. A pooled QC standard was run after every 7 normal sample injections.

In addition, pooled cereal rye, control soil, control water, and hairy vetch samples were run separately with 10 replicate injections per pool in iterative exclusion mode using AutoCat_V1 processing. In this case the following parameters were used for MS1 survey scan: 50-1200 m/z mass range at 0.1 seconds per scan, switching to MS/MS after individual ion intensity rises above 5000. MS/MS acquisition occurred at a scan rate of 0.5 seconds, or after accumulated TIC threshold of 100000, with 5 MS/MS event per MS1 scan. For all experiments, collision energy for MS/MS was ramped from 15 to 30 V. Calibration was performed using sodium formate with 1 ppm mass accuracy. The capillary voltage was held at 700 V in positive mode. The source temperature was held at 150 °C and the nitrogen desolvation temperature at 450 °C with a desolvation flow rate of 1000 L/hr. Lockspray reference mass was used to correct for drift, with 40 seconds interval between scans, 0.1 seconds/scan and signal averaged over 3 scans. LeuEnk was used for mass correction, with reference mass of 556.2771 m/z. For pooled iterative exclusion samples, lockspray signal was collected but correction was not applied until post-processing. These latter data were processed using Waters Tynebridge to produce mass calibrated mzML files compatible with GNPS server. The column and samples were held at 30 °C and 6 °C, respectively.

4.4.5 Metabolomics: Data Analysis

For the analysis of low abundance phytohormones, Simplicity 3Q software (Version 3.0.2, PerkinElmer, Waltham, MA) was used to detect and integrate peak areas and to calculate linear regression of analytical standards used for quantification. Each peak was normalized to an

appropriate internal standard (IS). The corresponding linear regression equation was used for quantification (ng/mL) for each analyte, which was then adjusted for precise volume of slurry (ng/mL). The limit of detection (LOD) was calculated as 3 times the standard deviation of the blank divided by the slope of the calibration curve. Likewise, the limit of quantitation (LOQ) was calculated as 10 times the standard deviation of the blank divided by the slope of the calibration curve.

For LC-MS data, mzML files were processed through the following workflow: 1) XCMS software was used for preprocessing to identify molecular features¹²⁴; 2) features were further normalized to total ion current (TIC); 3) the package RAMClustR¹²⁵ was used for clustering features into spectra. MSFinder²²⁴ was used for spectral matching, formula inference, and tentative structure assignment. MSFinder InChiKey results were exported and uploaded to ClassyFire¹⁰⁶ for batch annotation of chemical taxonomy²¹⁴.

For LC-MS/MS fastDDA data, mzML files for each iterative exclusion across all treatments (n=50) were uploaded to GNPS for molecular networking and annotation. A molecular network was created using the online workflow (<https://ccms-ucsd.github.io/GNPSDocumentation/>) on the GNPS website (<http://gnps.ucsd.edu>). The data was filtered by removing all MS/MS fragment ions within +/- 17 Da of the precursor m/z. The precursor ion mass tolerance was set to 0.5 Da and a MS/MS fragment ion tolerance of 0.02 Da. A network was then created where edges were filtered to have a cosine score above 0.5 and more than 5 matched peaks. Further, edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 20 most similar nodes. Finally, the maximum size of a molecular family was set to 100, and the lowest scoring edges were removed from molecular families until the molecular family size was below this threshold. The spectra in the network were then searched against GNPS' spectral libraries. The library spectra were filtered in the same manner as the input data. All matches kept

between network spectra and library spectra were required to have a score above 0.5 and at least 4 matched peaks.

Annotation of compounds to the level 2 classification was completed through manual matching between the MS1 data from all samples and the MS2 data collected from the pooled QCs. Specifically, annotation results from GNPS molecular networking were exported and precursor masses were matched to MS1 spectra within an 11s retention time window (to account for analytical drift). Quantitative data (untargeted LC-MS and targeted LC-MS/MS for phytohormone analysis) were zscored and combined for the final quantitative dataset (**Appendix C**)

4.4.6 Metagenomics: Sample Preparation

For days 0, 3, 5, 7, 10, and 21, we obtained a metagenome for each treatment from a single sample (n = 26 metagenomes). For this, genomic DNA was prepared for metagenomic sequencing using the ZymoBiomics DNA/RNA Miniprep kit and the Tecan Ovation Ultralow System V2 library prep kit and was sequenced at the University of Colorado Anschutz Sequencing Core on the Illumina NovaSeq 6000 with 2x150bp chemistry at 20 million read pairs per sample.

4.4.7 Metagenomics: Sample Preparation

Fastq files were trimmed using Sickle (v1.33)¹⁹¹, and trimmed reads were assembled using MEGAHIT (v1.2.9)²²⁵. To maximize assembly, we performed (1) co-assemblies, combining reads from each treatment metagenome to increase assembly coverage (2) Iterative assembly, iteratively assembling reads that did not map to assembled scaffolds ≥ 3 kb at 97% identity on all metagenomes. Information for metagenome statistics, including assembly information, are found in **Appendix C**. For each assembly, scaffolds ≥ 2.5 kb were binned using MetaBAT2 (v2.12.1)¹⁹³, and MAG completion was assessed using checkM (v1.1.2)¹⁹⁴. MAGs were kept in the database if they were $>50\%$ complete and $<10\%$ contaminated. MAGs were dereplicated at 99% identity using dRep (v2.6.2)^{195, 160}. MAG taxonomy was assigned using GTDB-tk (v2.3.0). MAGs and

assemblies were annotated using DRAM (v1.4.4)¹⁹⁸. CAZymes were inferred from the DRAM hits. To quantify MAG relative abundance in each temporal sample and treatment, trimmed metagenomic reads were mapped to the dereplicated MAG set using bbmap (v38.90)¹⁹⁶ at `minidfilter=95`, and output as sam files which were converted to sorted bam files using samtools (v1.9)¹⁹⁷. We had two requirements for a MAG to be found in a sample: first we required reads to map to at least 75% of a MAG in a given sample, and second the MAG had to have at least 3X coverage in that sample. To determine MAGs that had reads mapped to at least 75% of the MAG, we used coverM (v0.6.0)²²⁶ in genome mode to output MAGs that passed this threshold (`--min-covered-fraction 75`). To obtain MAG coverage, we used coverM (v0.3.2) in genome mode to output `reads_per_base` (reads mapped/genome length), and from this calculated MAG coverage as `reads_per_base x 151 bp`. A bin was “present” in a treatment or in control if it was found with at least 3X average coverage across the MAG and had reads mapped to at least 75% of the MAG in any of the timepoints or was “present” in the treatments if these two criteria were met.

4.4.8 Metatranscriptomics: Sample Preparation

For days 0, 5, and 21, we obtained a metatranscriptomes for each treatment from a triplicate sample (n=45 metagenomes). RNA was prepared for sequencing using the ZymoBionics DNA/RNA Miniprep kit and cleaned using ZymoBionics RNA Clean & Concentrate Kit and sent to the Department of Energy Joint Genome Institute for sequencing on the Illumina NovaSeq 6000 with 2x151bp chemistry at a target depth of 100M reads per sample. rRNA was depleted from an input of 10 ng of total RNA using QIAseq FastSelect™ – 5S/16S/23S, rRNA Plant and rRNA Yeast Kits (Qiagen). Using TruSeq stranded mRNA kit (Illumina), the 300 bp - 400 bp heat fragmented RNA was reverse transcribed to create the first strand of cDNA with random hexamers and SuperScript™ II Reverse Transcriptase (Thermo Fisher Scientific) followed by second strand synthesis. The double stranded cDNA fragments were treated with A-

tailing, ligation with JGI's unique dual indexed adapters (IDT) and enriched using 13 cycles of PCR.

4.4.9 Metatranscriptomics: Data Analysis

Fastq files were trimmed, and adapters were removed using bbdut (v38.90) with the parameters ktrim=r, k=23, mink=11, hdist=1, and filtered using rqcfilter2 (v38.90). Trimmed, filtered reads were then mapped via Bowtie2²²⁷ to the MAG database (dereplicated to 99% ID). Sam files were transformed to bam files using samtools, filtered to 95% id using reformat.sh and name sorted using samtools. Transcripts were counted each gene with feature-counts²²⁸. Counts were transformed to geTMM (gene length corrected trimmed mean of M-values) in R using edgeR package²²⁹. Genes and bins were considered if they were expressed in 5% of samples.

4.4.10 Statistics

Peak area files for each metabolomic analysis were combined into one file (**Appendix C**) and normalized by z-scoring. Line graphs of targeted and nontargeted data were visualized in Prism using absolute concentrations and TIC normalized values, respectively. Two-way ANOVAs (multiple comparisons were corrected using Tukey's multiple comparison test) were calculated using log₁₀ transformed data to satisfy the assumption of data normality in order to determine if a compound was significantly different between a treatment and the control at each timepoint. Multivariate metabolomic statistics were conducted in SIMCA (v17.0.1) to generate PLS-DA models and gather VIP scores (here, we used a cutoff of VIP >1.9). PLS-DA was performed using annotated and unannotated metabolites using z-scored and UV-scaled data. The list of scores and loadings coordinates were then plotted in Prism. Significance between metabolic changes across time and treatment was quantified using the adonis2 commands from the vegan package R (v4.3.1) with a PERMANOVA model. Pairwise comparisons were completed using the pairwise_permanova commands.

Metatranscriptomic NMDS plots were generated in R (v4.3.1) to estimate beta diversity across treatments. We utilized Bray-Curtis dissimilarity matrix visualized by non-metric multidimensional scaling (NMDS) in R with the ggplot2 package with stress of the non-parametric fit for the goodness of fit for the model. NMDS scores were exported from R and imported for visualization into Prism. Significance of compositional differences across treatments and the interaction of treatment and time, was quantified using a multi-response permutation procedure, mrpp, and the betadisper commands from the vegan package (v2.6-4) with an ANOVA model in R. Metatranscriptomic hierarchical clustering was completed in R using the hclust function from the stats package (v4.3.1). Visualization of the clustering used the as.dendrogram function in the stats package.

REFERENCES

- 1 Basche, A. D. & DeLonge, M. S. Comparing infiltration rates in soils managed with conventional and alternative farming methods: A meta-analysis. *PLOS ONE* **14**, e0215702, doi:10.1371/journal.pone.0215702 (2019).
- 2 Blanco-Canqui, H. & Ruis, S. J. Cover crop impacts on soil physical properties: A review. *Soil Science Society of America Journal* **84**, 1527-1576, doi:<https://doi.org/10.1002/saj2.20129> (2020).
- 3 Austin, E. E., Wickings, K., McDaniel, M. D., Robertson, G. P. & Grandy, A. S. Cover crop root contributions to soil carbon in a no-till corn bioenergy cropping system. *GCB Bioenergy* **9**, 1252-1263, doi:<https://doi.org/10.1111/gcbb.12428> (2017).
- 4 Lawley, Y. E., Teasdale, J. R. & Weil, R. R. The Mechanism for Weed Suppression by a Forage Radish Cover Crop. *Agronomy Journal* **104**, 205-214, doi:<https://doi.org/10.2134/agronj2011.0128> (2012).
- 5 Garland, G. *et al.* Crop cover is more important than rotational diversity for soil multifunctionality and cereal yields in European cropping systems. *Nature Food* **2**, 28-37, doi:10.1038/s43016-020-00210-8 (2021).
- 6 Elhakeem, A. *et al.* Do cover crop mixtures give higher and more stable yields than pure stands? *Field Crops Research* **270**, 108217, doi:<https://doi.org/10.1016/j.fcr.2021.108217> (2021).
- 7 Otto, R. *et al.* Planting legume cover crop as a strategy to replace synthetic N fertilizer applied for sugarcane production. *Industrial Crops and Products* **156**, 112853, doi:<https://doi.org/10.1016/j.indcrop.2020.112853> (2020).

- 8 Fierer, N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology* **15**, 579-590, doi:10.1038/nrmicro.2017.87 (2017).
- 9 Cazzaniga, S. G. *et al.* On the legacy of cover crop-specific microbial footprints. *Soil Biology and Biochemistry* **184**, 109080, doi:<https://doi.org/10.1016/j.soilbio.2023.109080> (2023).
- 10 Barel, J. M. *et al.* Winter cover crop legacy effects on litter decomposition act through litter quality and microbial community changes. *Journal of Applied Ecology* **56**, 132-143, doi:<https://doi.org/10.1111/1365-2664.13261> (2019).
- 11 Larkin, R. P. & Honeycutt, C. W. Effects of Different 3-Year Cropping Systems on Soil Microbial Communities and Rhizoctonia Diseases of Potato. *Phytopathology*® **96**, 68-79, doi:10.1094/phyto-96-0068 (2006).
- 12 York, L. M., Carminati, A., Mooney, S. J., Ritz, K. & Bennett, M. J. The holistic rhizosphere: integrating zones, processes, and semantics in the soil influenced by roots. *Journal of Experimental Botany* **67**, 3629-3643, doi:10.1093/jxb/erw108 (2016).
- 13 Sasse, J., Martinoia, E. & Northen, T. Feed Your Friends: Do Plant Exudates Shape the Root Microbiome? *Trends Plant Sci* **23**, 25-41, doi:10.1016/j.tplants.2017.09.003 (2018).
- 14 Zhalnina, K. *et al.* Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nature Microbiology* **3**, 470-480, doi:10.1038/s41564-018-0129-3 (2018).
- 15 Canarini, A., Kaiser, C., Merchant, A., Richter, A. & Wanek, W. Root Exudation of Primary Metabolites: Mechanisms and Their Roles in Plant Responses to Environmental Stimuli. *Frontiers in Plant Science* **10**, doi:10.3389/fpls.2019.00157 (2019).
- 16 Chaparro, J. *et al.* Root Exudation of Phytochemicals in Arabidopsis Follows Specific Patterns That Are Developmentally Programmed and Correlate with Soil Microbial

- Functions. *PLoS ONE* **8**, e55731, doi: <https://doi.org/10.1371/journal.pone.0055731> (2013).
- 17 Chaparro, J. M., Badri, D. V. & Vivanco, J. M. Rhizosphere microbiome assemblage is affected by plant development. *The ISME Journal* **8**, 790-803, doi:10.1038/ismej.2013.196 (2014).
 - 18 Seitz, V. A. *et al.* Variation in Root Exudate Composition Influences Soil Microbiome Membership and Function. *Appl Environ Microbiol* **88**, e0022622, doi:10.1128/aem.00226-22 (2022).
 - 19 Micallef, S. A., Channer, S., Shiaris, M. P. & Colón-Carmona, A. Plant age and genotype impact the progression of bacterial community succession in the Arabidopsis rhizosphere. *Plant Signal Behav* **4**, 777-780, doi:10.1093/jxb/erp053 (2009).
 - 20 Micallef, S. A., Shiaris, M. P. & Colón-Carmona, A. Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. *Journal of Experimental Botany* **60**, 1729-1742, doi:10.1093/jxb/erp053 (2009).
 - 21 French, E., Kaplan, I., Iyer-Pascuzzi, A., Nakatsu, C. H. & Enders, L. Emerging strategies for precision microbiome management in diverse agroecosystems. *Nature Plants* **7**, 256-267, doi:10.1038/s41477-020-00830-9 (2021).
 - 22 Bardgett, R. D. & van der Putten, W. H. Belowground biodiversity and ecosystem functioning. *Nature* **515**, 505-511, doi:10.1038/nature13855 (2014).
 - 23 Ågren, G. I., Wetterstedt, J. Å. M. & Billberger, M. F. K. Nutrient limitation on terrestrial plant growth – modeling the interaction between nitrogen and phosphorus. *New Phytologist* **194**, 953-960, doi:<https://doi.org/10.1111/j.1469-8137.2012.04116.x> (2012).
 - 24 Soltangheisi, A., Teles, A. P. B., Sartor, L. R. & Pavinato, P. S. Cover Cropping May Alter Legacy Phosphorus Dynamics Under Long-Term Fertilizer Addition. *Frontiers in Environmental Science* **8**, doi:10.3389/fenvs.2020.00013 (2020).

- 25 dos Santos, R. M., Diaz, P. A. E., Lobo, L. L. B. & Rigobelo, E. C. Use of Plant Growth-Promoting Rhizobacteria in Maize and Sugarcane: Characteristics and Applications. *Frontiers in Sustainable Food Systems* **4**, doi:10.3389/fsufs.2020.00136 (2020).
- 26 Meena, V. S., Maurya, B. R. & Verma, J. P. Does a rhizospheric microorganism enhance K⁺ availability in agricultural soils? *Microbiological Research* **169**, 337-347, doi:<https://doi.org/10.1016/j.micres.2013.09.003> (2014).
- 27 Hu, X., Chen, J. & Guo, J. Two Phosphate- and Potassium-solubilizing Bacteria Isolated from Tianmu Mountain, Zhejiang, China. *World Journal of Microbiology and Biotechnology* **22**, 983-990, doi:10.1007/s11274-006-9144-2 (2006).
- 28 Wen, Z., White, P. J., Shen, J. & Lambers, H. Linking root exudation to belowground economic traits for resource acquisition. *New Phytologist* **233**, 1620-1635, doi:<https://doi.org/10.1111/nph.17854> (2022).
- 29 Rose, T. J., Damon, P. & Rengel, Z. Phosphorus-efficient faba bean (*Vicia faba* L.) genotypes enhance subsequent wheat crop growth in an acid and an alkaline soil. *Crop and Pasture Science* **61**, 1009-1016, doi:<https://doi.org/10.1071/CP10205> (2010).
- 30 Lambers, H. *et al.* in *Annual Plant Reviews online* 289-335.
- 31 Liu, H., White, P. J. & Li, C. Biomass partitioning and rhizosphere responses of maize and faba bean to phosphorus deficiency. *Crop and Pasture Science* **67**, 847-856, doi:<https://doi.org/10.1071/CP16015> (2016).
- 32 Zhang, D. *et al.* Increased soil phosphorus availability induced by faba bean root exudation stimulates root growth and phosphorus uptake in neighbouring maize. *New Phytologist* **209**, 823-831, doi:<https://doi.org/10.1111/nph.13613> (2016).
- 33 Strehmel, N., Böttcher, C., Schmidt, S. & Scheel, D. Profiling of secondary metabolites in root exudates of *Arabidopsis thaliana*. *Phytochemistry* **108**, 35-46, doi:<https://doi.org/10.1016/j.phytochem.2014.10.003> (2014).

- 34 Weston, L. A. & Mathesius, U. in *Root Engineering: Basic and Applied Concepts* (eds Asunción Morte & Ajit Varma) 221-247 (Springer Berlin Heidelberg, 2014).
- 35 Wink, M. Evolution of secondary metabolites in legumes (Fabaceae). *South African Journal of Botany* **89**, 164-175, doi:<https://doi.org/10.1016/j.sajb.2013.06.006> (2013).
- 36 Seitz, V. A., Chaparro, J. M., Schipanski, M. E., Wrighton, K. C. & Prenni, J. E. Cover Crop Cultivar, Species, and Functional Diversity is Reflected in Variable Root Exudation Composition. *Journal of Agricultural and Food Chemistry*, doi:10.1021/acs.jafc.3c02912 (2023).
- 37 Falcone Ferreyra, M. L., Rius, S. & Casati, P. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Frontiers in Plant Science* **3**, doi:10.3389/fpls.2012.00222 (2012).
- 38 Coskun, D., Britto, D. T., Shi, W. & Kronzucker, H. J. How Plant Root Exudates Shape the Nitrogen Cycle. *Trends in Plant Science* **22**, 661-673, doi:<https://doi.org/10.1016/j.tplants.2017.05.004> (2017).
- 39 Subbarao, G. V. *et al.* Biological nitrification inhibition (BNI) activity in sorghum and its characterization. *Plant and Soil* **366**, 243-259, doi:10.1007/s11104-012-1419-9 (2013).
- 40 Finzi, A. C. *et al.* Rhizosphere processes are quantitatively important components of terrestrial carbon and nutrient cycles. *Global Change Biology* **21**, 2082-2094, doi:<https://doi.org/10.1111/gcb.12816> (2015).
- 41 Bressan, M. *et al.* Exogenous glucosinolate produced by *Arabidopsis thaliana* has an impact on microbes in the rhizosphere and plant roots. *The ISME Journal* **3**, 1243-1257, doi:10.1038/ismej.2009.68 (2009).
- 42 Lazzeri, L., Curto, G., Leoni, O. & Dallavalle, E. Effects of Glucosinolates and Their Enzymatic Hydrolysis Products via Myrosinase on the Root-knot Nematode *Meloidogyne incognita* (Kofoid et White) Chitw. *Journal of Agricultural and Food Chemistry* **52**, 6703-6707, doi:10.1021/jf030776u (2004).

- 43 Bangarwa, S. K. & Norsworthy, J. K. in *Glucosinolates* (eds Jean-Michel Mérillon & Kishan Gopal Ramawat) 1-35 (Springer International Publishing, 2016).
- 44 Duff, J., Sprang, C. v., O'Halloran, J. & Hall, Z. *Guide to Brassica Biofumigant Cover Crops: Managing soilborne diseases in vegetable production systems*. (The Queensland Government through the Department of Environment and Science, 2020).
- 45 Tagele, S. B., Kim, R.-H. & Shin, J.-H. Interactions between Brassica Biofumigants and Soil Microbiota: Causes and Impacts. *Journal of Agricultural and Food Chemistry* **69**, 11538-11553, doi:10.1021/acs.jafc.1c03776 (2021).
- 46 Siebers, M. *et al.* Disruption of microbial community composition and identification of plant growth promoting microorganisms after exposure of soil to rapeseed-derived glucosinolates. *PLOS ONE* **13**, e0200160, doi:10.1371/journal.pone.0200160 (2018).
- 47 Hu, P., Hollister, E. B., Somenahally, A. C., Hons, F. M. & Gentry, T. J. Soil bacterial and fungal communities respond differently to various isothiocyanates added for biofumigation. *Frontiers in Microbiology* **5**, doi:10.3389/fmicb.2014.00729 (2015).
- 48 Neal, a. & Ton, J. Systemic defense priming by *Pseudomonas putida* KT2440 in maize depends on benzoxazinoid exudation from the roots. *Plant Signal Behav* **8**, e22655, doi:10.4161/psb.22655 (2013).
- 49 Neal, A. L., Ahmad, S., Gordon-Weeks, R. & Ton, J. Benzoxazinoids in Root Exudates of Maize Attract *Pseudomonas putida* to the Rhizosphere. *PLOS ONE* **7**, e35498, doi:10.1371/journal.pone.0035498 (2012).
- 50 Hu, L. *et al.* Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nature Communications* **9**, 2738, doi:10.1038/s41467-018-05122-7 (2018).
- 51 Cotton, T. E. A. *et al.* Metabolic regulation of the maize rhizobiome by benzoxazinoids. *The ISME Journal* **13**, 1647-1658, doi:10.1038/s41396-019-0375-2 (2019).

- 52 Etzerodt, T. *et al.* 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one (DIMBOA) inhibits trichothecene production by *Fusarium graminearum* through suppression of Tri6 expression. *Int J Food Microbiol* **214**, 123-128, doi:10.1016/j.ijfoodmicro.2015.07.014 (2015).
- 53 Schulz, M., Marocco, A., Tabaglio, V., Macias, F. A. & Molinillo, J. M. G. Benzoxazinoids in Rye Allelopathy - From Discovery to Application in Sustainable Weed Control and Organic Farming. *Journal of Chemical Ecology* **39**, 154-174, doi:10.1007/s10886-013-0235-x (2013).
- 54 Ye, X. *et al.* Effects of green manure continuous application on soil microbial biomass and enzyme activity. *Journal of Plant Nutrition* **37**, 498-508, doi:10.1080/01904167.2013.867978 (2014).
- 55 Tsiafouli, M. A. *et al.* Intensive agriculture reduces soil biodiversity across Europe. *Global Change Biology* **21**, 973-985, doi:<https://doi.org/10.1111/gcb.12752> (2015).
- 56 Poeplau, C. & Don, A. Carbon sequestration in agricultural soils via cultivation of cover crops – A meta-analysis. *Agriculture, Ecosystems & Environment* **200**, 33-41, doi:<https://doi.org/10.1016/j.agee.2014.10.024> (2015).
- 57 Taffner, J. *et al.* What Is the Role of Archaea in Plants? New Insights from the Vegetation of Alpine Bogs. *mSphere* **3**, doi:10.1128/mSphere.00122-18 (2018).
- 58 Taffner, J., Cernava, T., Erlacher, A. & Berg, G. Novel insights into plant-associated archaea and their functioning in arugula (*Eruca sativa* Mill.). *Journal of Advanced Research* **19**, 39-48, doi:<https://doi.org/10.1016/j.jare.2019.04.008> (2019).
- 59 Knief, C. *et al.* Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *The ISME Journal* **6**, 1378-1390, doi:10.1038/ismej.2011.192 (2012).
- 60 Conrad, R. in *Advances in Agronomy* Vol. 96 1-63 (Academic Press, 2007).

- 61 Remans, R. *et al.* Effect of Rhizobium–Azospirillum coinoculation on nitrogen fixation and yield of two contrasting Phaseolus vulgaris L. genotypes cultivated across different environments in Cuba. *Plant and Soil* **312**, 25-37, doi:10.1007/s11104-008-9606-4 (2008).
- 62 Spaepen, S. & Vanderleyden, J. Auxin and plant-microbe interactions. *Cold Spring Harb Perspect Biol* **3**, a001438, doi:10.1101/cshperspect.a001438 (2011).
- 63 Karlsson, A. E., Johansson, T. & Bengtson, P. Archaeal abundance in relation to root and fungal exudation rates. *FEMS Microbiology Ecology* **80**, 305-311, doi:10.1111/j.1574-6941.2012.01298.x (2012).
- 64 Jansa, J., Bukovská, P. & Gryndler, M. Mycorrhizal hyphae as ecological niche for highly specialized hypersymbionts – or just soil free-riders? *Frontiers in Plant Science* **4**, doi:10.3389/fpls.2013.00134 (2013).
- 65 Akiyama, K., Matsuzaki, K.-i. & Hayashi, H. Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature* **435**, 824-827, doi:10.1038/nature03608 (2005).
- 66 Tsai, S. M. & Phillips, D. A. Flavonoids Released Naturally from Alfalfa Promote Development of Symbiotic *Glomus* Spores In Vitro. *Applied and Environmental Microbiology* **57**, 1485-1488, doi:doi:10.1128/aem.57.5.1485-1488.1991 (1991).
- 67 Buee, M., Rossignol, M., Jauneau, A., Ranjeva, R. & Bécard, G. The Pre-Symbiotic Growth of Arbuscular Mycorrhizal Fungi Is Induced by a Branching Factor Partially Purified from Plant Root Exudates. *Molecular Plant-Microbe Interactions*® **13**, 693-698, doi:10.1094/mpmi.2000.13.6.693 (2000).
- 68 Steinkellner, S. *et al.* Flavonoids and Strigolactones in Root Exudates as Signals in Symbiotic and Pathogenic Plant-Fungus Interactions. *Molecules* **12**, 1290-1306 (2007).
- 69 Besserer, A. *et al.* Strigolactones Stimulate Arbuscular Mycorrhizal Fungi by Activating Mitochondria. *PLOS Biology* **4**, e226, doi:10.1371/journal.pbio.0040226 (2006).

- 70 Smith, S. E., Jakobsen, I., Grønlund, M. & Smith, F. A. Roles of Arbuscular Mycorrhizas in Plant Phosphorus Nutrition: Interactions between Pathways of Phosphorus Uptake in Arbuscular Mycorrhizal Roots Have Important Implications for Understanding and Manipulating Plant Phosphorus Acquisition. *Plant Physiology* **156**, 1050-1057, doi:10.1104/pp.111.174581 (2011).
- 71 Hou, E. *et al.* Global meta-analysis shows pervasive phosphorus limitation of aboveground plant production in natural terrestrial ecosystems. *Nature Communications* **11**, 637, doi:10.1038/s41467-020-14492-w (2020).
- 72 Zhou, X. G. & Everts, K. L. Effects of Host Resistance and Inoculum Density on the Suppression of Fusarium Wilt of Watermelon Induced by Hairy Vetch. *Plant Disease* **91**, 92-96, doi:10.1094/pd-91-0092 (2007).
- 73 Murrell, E. G., Ray, S., Lemmon, M. E., Luthe, D. S. & Kaye, J. P. Cover crop species affect mycorrhizae-mediated nutrient uptake and pest resistance in maize. *Renewable Agriculture and Food Systems* **35**, 467-474, doi:10.1017/S1742170519000061 (2020).
- 74 Waller, F. *et al.* The endophytic fungus *Piriformospora indica* reprograms barley to salt-stress tolerance, disease resistance, and higher yield. *Proceedings of the National Academy of Sciences* **102**, 13386-13391, doi:doi:10.1073/pnas.0504423102 (2005).
- 75 Benitez, M.-S., Taheri, W. I. & Lehman, R. M. Selection of fungi by candidate cover crops. *Applied Soil Ecology* **103**, 72-82, doi:<https://doi.org/10.1016/j.apsoil.2016.03.016> (2016).
- 76 Trinchera, A. *et al.* Effects induced by living mulch on rhizosphere interactions in organic artichoke: The cultivar's adaptive strategy. *Renewable Agriculture and Food Systems* **32**, 214-223, doi:10.1017/S1742170516000119 (2017).
- 77 Bowles, T. M., Jackson, L. E., Loeher, M. & Cavagnaro, T. R. Ecological intensification and arbuscular mycorrhizas: a meta-analysis of tillage and cover crop effects. *Journal of Applied Ecology* **54**, 1785-1793, doi:<https://doi.org/10.1111/1365-2664.12815> (2017).

- 78 Sikder, M. M. & Vestergård, M. Impacts of Root Metabolites on Soil Nematodes. *Frontiers in Plant Science* **10**, doi:10.3389/fpls.2019.01792 (2020).
- 79 Zasada, I. A., Meyer, S. L. F., Halbrecht, J. M. & Rice, C. Activity of Hydroxamic Acids from *Secale cereale* Against the Plant-Parasitic Nematodes *Meloidogyne incognita* and *Xiphinema americanum*. *Phytopathology*® **95**, 1116-1121, doi:10.1094/phyto-95-1116 (2005).
- 80 D'Addabbo, T. *et al.* Control of plant parasitic nematodes with active saponins and biomass from *Medicago sativa*. *Phytochemistry Reviews* **10**, 503-519, doi:10.1007/s11101-010-9180-2 (2011).
- 81 Aires, A., Carvalho, R., Da Conceição Barbosa, M. & Rosa, E. Suppressing Potato Cyst Nematode, *Globodera rostochiensis*, with Extracts of Brassicacea Plants. *American Journal of Potato Research* **86**, 327-333, doi:10.1007/s12230-009-9086-y (2009).
- 82 Ryan, P. R., Dessaux, Y., Thomashow, L. S. & Weller, D. M. Rhizosphere engineering and management for sustainable agriculture. *Plant and Soil* **321**, 363-383, doi:10.1007/s11104-009-0001-6 (2009).
- 83 Vessey, J. K. Plant growth promoting rhizobacteria as biofertilizers. *Plant and Soil* **255**, 571-586, doi:10.1023/A:1026037216893 (2003).
- 84 Chakraborti, S., Bera, K., Sadhukhan, S. & Dutta, P. Bio-priming of seeds: Plant stress management and its underlying cellular, biochemical and molecular mechanisms. *Plant Stress* **3**, 100052, doi:<https://doi.org/10.1016/j.stress.2021.100052> (2022).
- 85 Kawasaki, A. *et al.* Manipulating exudate composition from root apices shapes the microbiome throughout the root system. *Plant Physiology* **187**, 2279-2295, doi:10.1093/plphys/kiab337 (2021).
- 86 Krauss, M. *et al.* Enhanced soil quality with reduced tillage and solid manures in organic farming - a synthesis of 15 years. *Sci Rep* **10**, 4403, doi:10.1038/s41598-020-61320-8 (2020).

- 87 Wetzel, W. C., Kharouba, H. M., Robinson, M., Holyoak, M. & Karban, R. Variability in plant nutrients reduces insect herbivore performance. *Nature* **539**, 425-427, doi:10.1038/nature20140 (2016).
- 88 Zhao, Q. *et al.* Long-Term Coffee Monoculture Alters Soil Chemical Properties and Microbial Communities. *Scientific Reports* **8**, 6116, doi:10.1038/s41598-018-24537-2 (2018).
- 89 Shackelford, G. E. K., Rodd; Dicks, Lynn V. Effects of cover crops on multiple ecosystem services: Ten meta-analyses of data from arable farmland in California and the Mediterranean. *Land Use Policy* **88** (2019).
- 90 Kim, N., Zabaloy, M. C., Guan, K. & Villamil, M. B. Do cover crops benefit soil microbiome? A meta-analysis of current research. *Soil Biology and Biochemistry* **142**, 107701, doi:<https://doi.org/10.1016/j.soilbio.2019.107701> (2020).
- 91 Chapagain, T., Lee, E. A. & Raizada, M. N. The Potential of Multi-Species Mixtures to Diversify Cover Crop Benefits. *Sustainability* **12**, 2058 (2020).
- 92 Baraibar, B., Hunter, M. C., Schipanski, M. E., Hamilton, A. & Mortensen, D. A. Weed Suppression in Cover Crop Monocultures and Mixtures. *Weed Science* **66**, 121-133, doi:10.1017/wsc.2017.59 (2018).
- 93 Mitchell, J. P. *et al.* Cover cropping and no-tillage improve soil health in an arid irrigated cropping system in California's San Joaquin Valley, USA. *Soil and Tillage Research* **165**, 325-335, doi:<https://doi.org/10.1016/j.still.2016.09.001> (2017).
- 94 Büchi, L., Wendling, M., Amossé, C., Nepalova, M. & Charles, R. Importance of cover crops in alleviating negative effects of reduced soil tillage and promoting soil fertility in a winter wheat cropping system. *Agriculture, Ecosystems & Environment* **256**, 92-104, doi:<https://doi.org/10.1016/j.agee.2018.01.005> (2018).
- 95 Chaparro, J., Sheflin, A., Manter, D. & Vivanco, J. Manipulating the soil microbiome to increase soil health and plant fertility. *Biol Fertil Soils* **48**, 489-499 (2012).

- 96 SARE. Annual Report 2019-2020 National Cover Crop Survey (Sustainable Agriculture Research and Education (SARE), 2020).
- 97 SARE. Annual report 2016-2017: Cover Crop Survey. (Sustainable Agriculture Research and Education (SARE) and the Conservation Technology Information Center (CTIC), 2017).
- 98 Hudek, C., Putinica, C., Otten, W. & De Baets, S. Functional root trait-based classification of cover crops to improve soil physical properties. *European Journal of Soil Science* **73**, e13147, doi:<https://doi.org/10.1111/ejss.13147> (2022).
- 99 De Baets, S., Poesen, J., Meersmans, J. & Serlet, L. Cover crops and their erosion-reducing effects during concentrated flow erosion. *CATENA* **85**, 237-244, doi:<https://doi.org/10.1016/j.catena.2011.01.009> (2011).
- 100 Williams, A. *et al.* Root functional traits explain root exudation rate and composition across a range of grassland species. *Journal of Ecology* **110**, 21-33, doi:<https://doi.org/10.1111/1365-2745.13630> (2022).
- 101 Iannucci, A., Canfora, L., Nigro, F., De Vita, P. & Beleggia, R. Relationships between root morphology, root exudate compounds and rhizosphere microbial community in durum wheat. *Applied Soil Ecology* **158**, 103781, doi:<https://doi.org/10.1016/j.apsoil.2020.103781> (2021).
- 102 Wen, Z. *et al.* Tradeoffs among root morphology, exudation and mycorrhizal symbioses for phosphorus-acquisition strategies of 16 crop species. *New Phytologist* **223**, 882-895, doi:<https://doi.org/10.1111/nph.15833> (2019).
- 103 Oburger, E. & Jones, D. Sampling root exudates – Mission impossible? *Rhizosphere* **6**, doi:10.1016/j.rhisph.2018.06.004 (2018).
- 104 Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **3**, 211-221, doi:10.1007/s11306-007-0082-2 (2007).

- 105 Ernst, M. *et al.* MolNetEnhancer: Enhanced Molecular Networks by Integrating Metabolome Mining and Annotation Tools. *Metabolites* **9**, 144 (2019).
- 106 Djoumbou Feunang, Y. *et al.* ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *Journal of Cheminformatics* **8**, 61, doi:10.1186/s13321-016-0174-y (2016).
- 107 Dietz, S. *et al.* Root exudate composition of grass and forb species in natural grasslands. *Scientific Reports* **10**, 10691, doi:10.1038/s41598-019-54309-5 (2020).
- 108 Gifford, I. *et al.* Distinctive Patterns of Flavonoid Biosynthesis in Roots and Nodules of *Datisca glomerata* and *Medicago* spp. Revealed by Metabolomic and Gene Expression Profiles. *Frontiers in Plant Science* **9**, doi:10.3389/fpls.2018.01463 (2018).
- 109 Okutani, F. *et al.* Rhizosphere modelling reveals spatiotemporal distribution of daidzein shaping soybean rhizosphere bacterial community. *Plant Cell Environ* **43**, 1036-1046, doi:10.1111/pce.13708 (2020).
- 110 Hunter, P., Teakle, G. & Bending, G. Root traits and microbial community interactions in relation to phosphorus availability and acquisition, with particular reference to Brassica. *Frontiers in Plant Science* **5**, doi:10.3389/fpls.2014.00027 (2014).
- 111 Shimoyamada, M., Kudo, S., Okubo, K., Yamauchi, F. & Harada, K. Distributions of Saponin Constituents in Some Varieties of Soybean Plant. *Agricultural and Biological Chemistry* **54**, 77-81, doi:10.1080/00021369.1990.10869887 (1990).
- 112 Jun, H. S., Kim, S. E. & Sung, M. K. Protective effect of soybean saponins and major antioxidants against aflatoxin B1-induced mutagenicity and DNA-adduct formation. *J Med Food* **5**, 235-240, doi:10.1089/109662002763003393 (2002).
- 113 Kang, J.-H. *et al.* Soybean saponin inhibits tumor cell metastasis by modulating expressions of MMP-2, MMP-9 and TIMP- 2. *Cancer Letters* **261**, 84-92, doi:<https://doi.org/10.1016/j.canlet.2007.11.006> (2008).

- 114 Tsuno, Y., Fujimatsu, T., Endo, K., Sugiyama, A. & Yazaki, K. Soyasaponins: A New Class of Root Exudates in Soybean (*Glycine max*). *Plant and Cell Physiology* **59**, 366-375, doi:10.1093/pcp/pcx192 (2017).
- 115 Vincken, J.-P., Heng, L., de Groot, A. & Gruppen, H. Saponins, classification and occurrence in the plant kingdom. *Phytochemistry* **68**, 275-297, doi:<https://doi.org/10.1016/j.phytochem.2006.10.008> (2007).
- 116 Augustin, J. M., Kuzina, V., Andersen, S. B. & Bak, S. Molecular activities, biosynthesis and evolution of triterpenoid saponins. *Phytochemistry* **72**, 435-457, doi:<https://doi.org/10.1016/j.phytochem.2011.01.015> (2011).
- 117 Sugiyama, A. Flavonoids and saponins in plant rhizospheres: roles, dynamics, and the potential for agriculture. *Bioscience, Biotechnology, and Biochemistry* **85**, 1919-1931, doi:10.1093/bbb/zbab106 (2021).
- 118 Turner, T. R. *et al.* Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. *The ISME Journal* **7**, 2248-2258, doi:10.1038/ismej.2013.119 (2013).
- 119 Fujimatsu, T., Endo, K., Yazaki, K. & Sugiyama, A. Secretion dynamics of soyasaponins in soybean roots and effects to modify the bacterial composition. *Plant Direct* **4**, e00259, doi:<https://doi.org/10.1002/pld3.259> (2020).
- 120 Ha, T. J. *et al.* Rapid characterisation and comparison of saponin profiles in the seeds of Korean Leguminous species using ultra performance liquid chromatography with photodiode array detector and electrospray ionisation/mass spectrometry (UPLC–PDA–ESI/MS) analysis. *Food Chemistry* **146**, 270-277, doi:<https://doi.org/10.1016/j.foodchem.2013.09.051> (2014).
- 121 Sheflin, A. M. *et al.* High-throughput quantitative analysis of phytohormones in sorghum leaf and root tissue by ultra-performance liquid chromatography-mass spectrometry.

- Analytical and Bioanalytical Chemistry* **411**, 4839-4848, doi:10.1007/s00216-019-01658-9 (2019).
- 122 Sheflin, A. M. *et al.* Metabolomics of sorghum roots during nitrogen stress reveals compromised metabolic capacity for salicylic acid biosynthesis. *Plant direct* **3**, e00122-e00122, doi:10.1002/pld3.122 (2019).
- 123 R: A language and environment for statistical computing. v. Available online at <https://www.R-project.org/>. (R Foundation for Statistical Computing, Vienna, Austria, 2020).
- 124 Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Analytical Chemistry* **78**, 779-787, doi:10.1021/ac051437y (2006).
- 125 Broeckling, C. D., Afsar, F. A., Neumann, S., Ben-Hur, A. & Prenni, J. E. RAMClust: a novel feature clustering method enables spectral-matching-based annotation for metabolomics data. *Anal Chem* **86**, 6812-6817, doi:10.1021/ac501530d (2014).
- 126 Tsugawa, H. *et al.* MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods* **12**, 523-526, doi:10.1038/nmeth.3393 (2015).
- 127 Nothias, L. F. *et al.* Feature-based molecular networking in the GNPS analysis environment. *Nat Methods* **17**, 905-908, doi:10.1038/s41592-020-0933-6 (2020).
- 128 da Silva, R. R. *et al.* Propagating annotations of molecular networks using in silico fragmentation. *PLOS Computational Biology* **14**, e1006089, doi:10.1371/journal.pcbi.1006089 (2018).
- 129 Thompson, L. R. *et al.* A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457-463, doi:10.1038/nature24621 (2017).

- 130 Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S. & Vivanco, J. M. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annual Review of Plant Biology* **57**, 233-266, doi:10.1146/annurev.arplant.57.032905.105159 (2006).
- 131 Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nature Microbiology* **3**, 470-480, doi:10.1038/s41564-018-0129-3 (2018).
- 132 Korenblum, E. *et al.* Rhizosphere microbiome mediates systemic root metabolite exudation by root-to-root signaling. *Proceedings of the National Academy of Sciences* **117**, 3874-3883, doi:10.1073/pnas.1912130117 (2020).
- 133 Berg, G. & Smalla, K. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiology Ecology* **68**, 1-13, doi:10.1111/j.1574-6941.2009.00654.x (2009).
- 134 Jiao, X., Takishita, Y., Zhou, G. & Smith, D. L. Plant Associated Rhizobacteria for Biocontrol and Plant Growth Enhancement. *Frontiers in plant science* **12**, 634796-634796, doi:10.3389/fpls.2021.634796 (2021).
- 135 Singh, R. K. *et al.* Diversity of nitrogen-fixing rhizobacteria associated with sugarcane: a comprehensive study of plant-microbe interactions for growth enhancement in *Saccharum* spp. *BMC Plant Biology* **20**, 220, doi:10.1186/s12870-020-02400-9 (2020).
- 136 Manoj, S. R. *et al.* Understanding the molecular mechanisms for the enhanced phytoremediation of heavy metals through plant growth promoting rhizobacteria: A review. *J Environ Manage* **254**, 109779, doi:10.1016/j.jenvman.2019.109779 (2020).
- 137 Boiero, L. *et al.* Phytohormone production by three strains of *Bradyrhizobium japonicum* and possible physiological and technological implications. *Appl Microbiol Biotechnol* **74**, 874-880, doi:10.1007/s00253-006-0731-9 (2007).

- 138 Spaepen, S., Vanderleyden, J. & Remans, R. Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiology Reviews* **31**, 425-448, doi:10.1111/j.1574-6976.2007.00072.x (2007).
- 139 Kawasaki, A. *et al.* Microbiome and Exudates of the Root and Rhizosphere of *Brachypodium distachyon*, a Model for Wheat. *PLoS ONE* **11**, e0164533, doi:<https://doi.org/10.1371/journal.pone.0164533> (2016).
- 140 Hassan, S. & Mathesius, U. The role of flavonoids in root–rhizosphere signalling: opportunities and challenges for improving plant–microbe interactions. *Journal of Experimental Botany* **63**, 3429-3444, doi:10.1093/jxb/err430 (2012).
- 141 Cai, T. *et al.* Host legume-exuded antimetabolites optimize the symbiotic rhizosphere. *Molecular Microbiology* **73**, 507-517, doi:<https://doi.org/10.1111/j.1365-2958.2009.06790.x> (2009).
- 142 Khan, A. L. *et al.* Indole acetic acid and ACC deaminase from endophytic bacteria improves the growth of *Solanum lycopersicum*. *Electronic Journal of Biotechnology* **21**, 58-64, doi:<https://doi.org/10.1016/j.ejbt.2016.02.001> (2016).
- 143 Maki, Y. *et al.* 3-Phenyllactic acid, a root-promoting substance isolated from Bokashi fertilizer, exhibits synergistic effects with tryptophan. *Plant Biotechnol (Tokyo)* **38**, 9-16, doi:10.5511/plantbiotechnology.20.0727a (2021).
- 144 Finney, D. M., Buyer, J. S. & Kaye, J. P. Living cover crops have immediate impacts on soil microbial community structure and function. *Journal of Soil and Water Conservation* **72**, 361-373, doi:10.2489/jswc.72.4.361 (2017).
- 145 Root exudate metabolites drive plant-soil feedbacks on growth and defense by shaping the rhizosphere microbiota. *Nature Communications* **9**, 2738, doi:10.1038/s41467-018-05122-7 (2018).
- 146 SARE. *Managing Cover Crops Profitably*. (2012).

- 147 Cooper, E. A. *et al.* A new reference genome for *Sorghum bicolor* reveals high levels of sequence similarity between sweet and grain genotypes: implications for the genetics of sugar metabolism. *BMC Genomics* **20**, 420-420, doi:10.1186/s12864-019-5734-x (2019).
- 148 Ratnavathi, C. V. & Chavan, U. D. in *Sorghum Biochemistry* (eds C. V. Ratnavathi, J. V. Patil, & U. D. Chavan) 253-310 (Academic Press, 2016).
- 149 Manzoni, S., Taylor, P., Richter, A., Porporato, A. & Ågren, G. I. Environmental and stoichiometric controls on microbial carbon-use efficiency in soils. *New Phytologist* **196**, 79-91, doi:<https://doi.org/10.1111/j.1469-8137.2012.04225.x> (2012).
- 150 Steiger, M. G., Blumhoff, M. L., Mattanovich, D. & Sauer, M. Biochemistry of microbial itaconic acid production. *Front Microbiol* **4**, 23, doi:10.3389/fmicb.2013.00023 (2013).
- 151 Leonardi, R. & Jackowski, S. Biosynthesis of Pantothenic Acid and Coenzyme A. *EcoSal Plus* **2**, 10.1128/ecosalplus.1123.1126.1123.1124, doi:10.1128/ecosalplus.3.6.3.4 (2007).
- 152 He, M. Pipecolic acid in microbes: biosynthetic routes and enzymes. *Journal of Industrial Microbiology and Biotechnology* **33**, 401-407, doi:10.1007/s10295-006-0078-3 (2006).
- 153 Cavaliere, M., Feng, S., Soyer, O. S. & Jiménez, J. I. Cooperation in microbial communities and their biotechnological applications. *Environmental Microbiology* **19**, 2949-2963, doi:<https://doi.org/10.1111/1462-2920.13767> (2017).
- 154 Whipps, J. M. Microbial interactions and biocontrol in the rhizosphere. *Journal of Experimental Botany* **52**, 487-511, doi:10.1093/jexbot/52.suppl_1.487 (2001).
- 155 Parray, J. A. & Shameem, N. in *Sustainable Agriculture* (eds Javid Ahmad Parray & Nowsheen Shameem) 137-179 (Academic Press, 2020).
- 156 Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology* **35**, 725-731, doi:10.1038/nbt.3893 (2017).

- 157 Howe, A. C. *et al.* Tackling soil diversity with the assembly of large, complex metagenomes. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4904-4909, doi:10.1073/pnas.1402564111 (2014).
- 158 McGivern, B. B. *et al.* Decrypting bacterial polyphenol metabolism in an anoxic wetland soil. *Nature Communications* **12**, 2466, doi:10.1038/s41467-021-22765-1 (2021).
- 159 Narrowe, A. B. *et al.* Uncovering the Diversity and Activity of Methylophilic Methanogens in Freshwater Wetland Soils. *mSystems* **4**, e00320-00319, doi:doi:10.1128/mSystems.00320-19 (2019).
- 160 Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925-1927, doi:10.1093/bioinformatics/btz848 (2019).
- 161 Gaille, C., Reimann, C. & Haas, D. Isochorismate Synthase (PchA), the First and Rate-limiting Enzyme in Salicylate Biosynthesis of *Pseudomonas aeruginosa*. *Journal of Biological Chemistry* **278**, 16893-16898, doi:10.1074/jbc.M212324200 (2003).
- 162 Mishra, A. K. & Baek, K.-H. Salicylic Acid Biosynthesis and Metabolism: A Divergent Pathway for Plants and Bacteria. *Biomolecules* **11**, 705, doi:10.3390/biom11050705 (2021).
- 163 Bakker, P. A. H. M., Ran, L. & Mercado-Blanco, J. Rhizobacterial salicylate production provokes headaches! *Plant and Soil* **382**, 1-16, doi:10.1007/s11104-014-2102-0 (2014).
- 164 Islam, M. N., Ali, M. S., Choi, S.-J., Park, Y.-I. & Baek, K.-H. Salicylic Acid-Producing Endophytic Bacteria Increase Nicotine Accumulation and Resistance against Wildfire Disease in Tobacco Plants. *Microorganisms* **8**, 31 (2020).
- 165 Tufail, A., Arfan, M., Gurmani, A. R., Khan, A. & Bano, A. Salicylic acid induced salinity tolerance in maize (*Zea mays*). *Pak. J. Bot* **45**, 75-82 (2013).
- 166 Meyer, J. M., Azelvandre, P. & Georges, C. Iron metabolism in *Pseudomonas*: salicylic acid, a siderophore of *Pseudomonas fluorescens* CHAO. *Biofactors* **4**, 23-27 (1992).

- 167 Li, W. *et al.* Promysalin, a Salicylate-Containing *Pseudomonas putida* Antibiotic, Promotes Surface Colonization and Selectively Targets Other *Pseudomonas*. *Chemistry & Biology* **18**, 1320-1330, doi:<https://doi.org/10.1016/j.chembiol.2011.08.006> (2011).
- 168 Noda, S., Kitazono, E., Tanaka, T., Ogino, C. & Kondo, A. Benzoic acid fermentation from starch and cellulose via a plant-like β -oxidation pathway in *Streptomyces maritimus*. *Microbial cell factories* **11**, 49, doi:10.1186/1475-2859-11-49 (2012).
- 169 Widhalm, J. R. & Dudareva, N. A familiar ring to it: biosynthesis of plant benzoic acids. *Mol Plant* **8**, 83-97, doi:10.1016/j.molp.2014.12.001 (2015).
- 170 Allah, M. M. S., El-Bassiouny, H. M. S., Elewa, T. A. E. & El Sebai, T. Effect of salicylic acid and benzoic acid on growth, yield and some biochemical aspects of quinoa plant grown in sandy soil. **8**, 216-225 (2015).
- 171 Anjum, S. *et al.* Exogenous benzoic acid (BZA) treatment can induce drought tolerance in soybean plants by improving gas-exchange and chlorophyll contents. *Australian Journal of Crop Science* **7**, 555-560 (2013).
- 172 Patten, C. L. & Glick, B. R. Bacterial biosynthesis of indole-3-acetic acid. *Canadian Journal of Microbiology* **42**, 207-220, doi:10.1139/m96-032 %M 8868227 (1996).
- 173 Carreño-Lopez, R., Campos-Reales, N., Elmerich, C. & Baca, B. E. Physiological evidence for differently regulated tryptophan-dependent pathways for indole-3-acetic acid synthesis in *Azospirillum brasilense*. *Mol Gen Genet* **264**, 521-530, doi:10.1007/s004380000340 (2000).
- 174 Khalid, A., Arshad, M. & Zahir, Z. Screening plant growth-promoting rhizobacteria for improving growth and yield of wheat. *Journal of Applied Microbiology* **96**, 473-480, doi:<https://doi.org/10.1046/j.1365-2672.2003.02161.x> (2004).
- 175 Luziatelli, F. *et al.* A Genetic and Metabolomic Perspective on the Production of Indole-3-Acetic Acid by *Pantoea agglomerans* and Use of Their Metabolites as Biostimulants in Plant Nurseries. *Frontiers in Microbiology* **11**, doi:10.3389/fmicb.2020.01475 (2020).

- 176 Fu, S. F. *et al.* Indole-3-acetic acid: A widespread physiological code in interactions of fungi with other organisms. *Plant Signal Behav* **10**, e1048052, doi:10.1080/15592324.2015.1048052 (2015).
- 177 Lebrazi, S., Fadil, M., Chraibi, M. & Fikri-Benbrahim, K. Screening and optimization of indole-3-acetic acid production by *Rhizobium* sp. strain using response surface methodology. *Journal of Genetic Engineering and Biotechnology* **18**, 21, doi:10.1186/s43141-020-00035-9 (2020).
- 178 Islam, S., Akanda, A. M., Prova, A., Islam, M. T. & Hossain, M. M. Isolation and Identification of Plant Growth Promoting Rhizobacteria from Cucumber Rhizosphere and Their Effect on Plant Growth Promotion and Disease Suppression. *Frontiers in Microbiology* **6**, doi:10.3389/fmicb.2015.01360 (2016).
- 179 Pan, L. *et al.* Complete genome sequence of *Mycobacterium* Mya-zh01, an endophytic bacterium, promotes plant growth and seed germination isolated from flower stalk of *Doritaenopsis*. *Archives of Microbiology* **202**, 1965-1976, doi:10.1007/s00203-020-01924-w (2020).
- 180 Xie, H., Pasternak, J. J. & Glick, B. R. Isolation and Characterization of Mutants of the Plant Growth-Promoting Rhizobacterium *Pseudomonas putida* GR12-2 That Overproduce Indoleacetic Acid. *Current Microbiology* **32**, 67-71, doi:10.1007/s002849900012 (1996).
- 181 Bunsangiam, S., Thongpae, N., Limtong, S. & Srisuk, N. Large scale production of indole-3-acetic acid and evaluation of the inhibitory effect of indole-3-acetic acid on weed growth. *Scientific Reports* **11**, 13094, doi:10.1038/s41598-021-92305-w (2021).
- 182 Eliasson, L., Bertell, G. & Bolander, E. Inhibitory action of auxin on root elongation not mediated by ethylene. *Plant Physiol* **91**, 310-314, doi:10.1104/pp.91.1.310 (1989).

- 183 Fontaine, S., Mariotti, A. & Abbadie, L. The priming effect of organic matter: a question of microbial competition? *Soil Biology and Biochemistry* **35**, 837-843, doi:[https://doi.org/10.1016/S0038-0717\(03\)00123-8](https://doi.org/10.1016/S0038-0717(03)00123-8) (2003).
- 184 Pantigoso, H. A., He, Y., DiLegge, M. J. & Vivanco, J. M. in *The Plant Microbiome: Methods and Protocols* (eds Lilia C. Carvalhais & Paul G. Dennis) 291-303 (Springer US, 2021).
- 185 Borton, M. A. *et al.* Coupled laboratory and field investigations resolve microbial interactions that underpin persistence in hydraulically fractured shales. *Proceedings of the National Academy of Sciences* **115**, E6585-E6594, doi:10.1073/pnas.1800155115 (2018).
- 186 Chaparro, J. M., Holm, D. G., Broeckling, C. D., Prenni, J. E. & Heuberger, A. L. Metabolomics and Ionomics of Potato Tuber Reveals an Influence of Cultivar and Market Class on Human Nutrients and Bioactive Compounds. *Front Nutr* **5**, 36, doi:10.3389/fnut.2018.00036 (2018).
- 187 Broeckling, C. D. *et al.* Enabling Efficient and Confident Annotation of LC-MS Metabolomics Data through MS1 Spectrum and Time Prediction. *Anal Chem* **88**, 9226-9234, doi:10.1021/acs.analchem.6b02479 (2016).
- 188 Thompson, L. R. *et al.* A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457-463, doi:10.1038/nature24621 (2017).
- 189 Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* **37**, 852-857, doi:10.1038/s41587-019-0209-9 (2019).
- 190 Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**, 581-583, doi:10.1038/nmeth.3869 (2016).
- 191 Sickel: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available at <https://github.com/najoshi/sickle> (2011).

- 192 Peng, Y., Leung, H. C. M., Yiu, S. M. & Chin, F. Y. L. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**, 1420-1428, doi:10.1093/bioinformatics/bts174 (2012).
- 193 Kang, D. D. *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359, doi:10.7717/peerj.7359 (2019).
- 194 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**, 1043-1055, doi:10.1101/gr.186072.114 (2015).
- 195 Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *The ISME Journal* **11**, 2864-2868, doi:10.1038/ismej.2017.126 (2017).
- 196 BMap: A Fast, Accurate, Splice-Aware Aligner (Available at: sourceforge.net/projects/bbmap/, 2014).
- 197 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
- 198 Shaffer, M. *et al.* DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Research* **48**, 8883-8900, doi:10.1093/nar/gkaa621 (2020).
- 199 Blin, K. *et al.* antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res* **49**, W29-w35, doi:10.1093/nar/gkab335 (2021).
- 200 vegan: Community Ecology Package v. R package version 2.5-7. (2020).
- 201 ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York, 2016).

- 202 Vukicevich, E., Lowery, T., Bowen, P., Úrbez-Torres, J. R. & Hart, M. Cover crops to increase soil microbial diversity and mitigate decline in perennial agriculture. A review. *Agronomy for Sustainable Development* **36**, 48, doi:10.1007/s13593-016-0385-7 (2016).
- 203 Osipitan, O. A. *et al.* Impact of Cover Crop Management on Level of Weed Suppression: A Meta-Analysis. *Crop Science* **59**, 833-842, doi:<https://doi.org/10.2135/cropsci2018.09.0589> (2019).
- 204 Strickland, M. S. *et al.* Short-Term Effects of Cover Crops on Soil Microbial Characteristics and Biogeochemical Processes across Actively Managed Farms. *Agrosystems, Geosciences & Environment* **2**, 180064, doi:<https://doi.org/10.2134/age2018.12.0064> (2019).
- 205 Sasse, J., Martinoia, E. & Northen, T. Feed Your Friends: Do Plant Exudates Shape the Root Microbiome? *Trends in Plant Science* **23**, 25-41, doi:<https://doi.org/10.1016/j.tplants.2017.09.003> (2018).
- 206 Goldfarb, K. *et al.* Differential Growth Responses of Soil Bacterial Taxa to Carbon Substrates of Varying Chemical Recalcitrance. *Frontiers in Microbiology* **2**, doi:10.3389/fmicb.2011.00094 (2011).
- 207 Huang, A. C. *et al.* A specialized metabolic network selectively modulates *Arabidopsis* root microbiota. *Science* **364**, eaau6389, doi:doi:10.1126/science.aau6389 (2019).
- 208 Kudjordjie, E. N., Sapkota, R. & Nicolaisen, M. *Arabidopsis* assemble distinct root-associated microbiomes through the synthesis of an array of defense metabolites. *PLOS ONE* **16**, e0259171, doi:10.1371/journal.pone.0259171 (2021).
- 209 Gargallo-Garriga, A. *et al.* Root exudate metabolomes change under drought and show limited capacity for recovery. *Scientific Reports* **8**, 12696, doi:10.1038/s41598-018-30150-0 (2018).

- 210 Mönchgesang, S. *et al.* Natural variation of root exudates in *Arabidopsis thaliana*-linking metabolomic and genomic data. *Scientific Reports* **6**, 29033, doi:10.1038/srep29033 (2016).
- 211 Liu, M. *et al.* Plant growth-promoting rhizobacteria *Burkholderia vietnamiensis* B418 inhibits root-knot nematode on watermelon by modifying the rhizosphere microbial community. *Scientific Reports* **12**, 8381, doi:10.1038/s41598-022-12472-2 (2022).
- 212 Fernández-González, A. J. *et al.* The rhizosphere microbiome of burned holm-oak: potential role of the genus *Arthrobacter* in the recovery of burned soils. *Scientific Reports* **7**, 6008, doi:10.1038/s41598-017-06112-3 (2017).
- 213 Gkorezis, P. *et al.* Draft Genome Sequence of *Arthrobacter* sp. Strain SPG23, a Hydrocarbon-Degrading and Plant Growth-Promoting Soil Bacterium. *Genome Announcements* **3**, 10.1128/genomea.01517-01515, doi:doi:10.1128/genomea.01517-15 (2015).
- 214 Wishart, D. S. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *Journal of Cheminformatics* **8**, 61, doi:10.1186/s13321-016-0174-y (2016).
- 215 Egamberdieva, D., Wirth, S. J., Alqarawi, A. A., Abd_Allah, E. F. & Hashem, A. Phytohormones and Beneficial Microbes: Essential Components for Plants to Balance Stress and Fitness. *Frontiers in Microbiology* **8**, doi:10.3389/fmicb.2017.02104 (2017).
- 216 Gao, X. *et al.* Enhancing the 1-Aminocyclopropane-1-Carboxylate Metabolic Rate of *Pseudomonas* sp. UW4 Intensifies Chemotactic Rhizocompetence. *Microorganisms* **8**, 71 (2020).
- 217 Li, T., Zhang, J., Shen, C., Li, H. & Qiu, L. 1-Aminocyclopropane-1-Carboxylate: A Novel and Strong Chemoattractant for the Plant Beneficial Rhizobacterium *Pseudomonas putida* UW4. *Molecular Plant-Microbe Interactions*® **32**, 750-759, doi:10.1094/mpmi-11-18-0317-r (2019).

- 218 Orozco-Mosqueda, M. D. C., Glick, B. R. & Santoyo, G. ACC deaminase in plant growth-promoting bacteria (PGPB): An efficient mechanism to counter salt stress in crops. *Microbiol Res* **235**, 126439, doi:10.1016/j.micres.2020.126439 (2020).
- 219 Todorovic, B. & Glick, B. R. The interconversion of ACC deaminase and d-cysteine desulfhydrase by directed mutagenesis. *Planta* **229**, 193-205, doi:10.1007/s00425-008-0820-3 (2008).
- 220 Nascimento, F. X., Rossi, M. J., Soares, C. R. F. S., McConkey, B. J. & Glick, B. R. New Insights into 1-Aminocyclopropane-1-Carboxylate (ACC) Deaminase Phylogeny, Evolution and Ecological Significance. *PLOS ONE* **9**, e99168, doi:10.1371/journal.pone.0099168 (2014).
- 221 Zhang, P. *et al.* The Distribution of Tryptophan-Dependent Indole-3-Acetic Acid Synthesis Pathways in Bacteria Unraveled by Large-Scale Genomic Analysis. *Molecules* **24**, 1411, doi:10.3390/molecules24071411 (2019).
- 222 Nett, R. S. *et al.* Elucidation of gibberellin biosynthesis in bacteria reveals convergent evolution. *Nat Chem Biol* **13**, 69-74, doi:10.1038/nchembio.2232 (2017).
- 223 Nett, R. S. *et al.* Unraveling a Tangled Skein: Evolutionary Analysis of the Bacterial Gibberellin Biosynthetic Operon. *mSphere* **5**, 10.1128/msphere.00292-00220, doi:doi:10.1128/msphere.00292-20 (2020).
- 224 Tsugawa, H. *et al.* Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. *Anal Chem* **88**, 7946-7958, doi:10.1021/acs.analchem.6b00770 (2016).
- 225 Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674-1676, doi:10.1093/bioinformatics/btv033 (2015).
- 226 CoverM: Read coverage calculator for metagenomics (Available at: <https://github.com/wwood/CoverM>).

- 227 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
- 228 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).
- 229 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2009).

APPENDICES

Appendix A: Chapter 2 Supplementary Files

See supplemental files in AppendixA.zip

File S1: Plant, seed, germination information, root morphology and exudate measurements.

File S2: Combined z-scored metabolomics data including annotations of level of identification for each metabolite.

File S3: Additional experimental details, tables of ANOVA analyses, soyasaponin molecular networking, additional statistical significance tests and networking Figures, and SRM transitions.

Appendix B: Chapter 3 Supplementary Files

See supplemental files in AppendixB.zip

File S1: Exudate and Exometabolite data.

File S2: 16S rRNA gene amplicon sequencing summary, ANCOM statistics, and feature table; metagenome sequencing and assembly statistics, metagenome-assembled genome summaries and abundance data.

File S3: DRAM annotation outputs for MAGs.

File S4: Synthetic exudate calculations for sugars

File S5: Synthetic exudate calculations for organic acids + final exudate treatment calculations

Appendix C: Chapter 4 Supplementary Files

See supplemental files in AppendixC.zip

File S1: Metabolomics data. Measurement of phytohormones from previously published root exudate characterization experiment used to formulate root exudate treatments used in this study (sheet: root_exudate_phyto_report), untargeted LC-MS/MS analysis of root exudates from previously published root exudate characterization experiment used to formulate root exudate treatments used in this study (sheet: root_exudate_untargeted), untargeted LC-MS HILIC results from metabolomics profiling of microcosm soil (sheet: microcosms_untargeted_lc-ms), targeted measurement of phytohormones from microcosm soil (sheet: microcosms_targeted_phyto), zscore normalized combined metabolite data (sheet: microcosms_zscored_data).

File S2: Metagenome data. Metagenome stats, total sequencing (sheet: metagenome stats), assembly stats (sheet: assembly stats), MAG IDs of medium and high-quality MAGs along with their completeness and contamination scores, bin size, and contig information (sheet: MAG_stats). Relative abundance of each MAG in a metagenome (sheet: MAG_rel_Abund).

File S3: Supplementary figures and tables.

File S4: Metatranscriptomic Data. Metatranscriptome stats (sheet: metadata) and geTMM normalized gene abundances (sheet: geTMM)

File S5: Cover crop seed metadata. Seed name, germination rate, sterilization method, and purchasing link for supplier.

File S6: DRAM gene annotations. DRAM gene annotations of the Medium and High-quality MAGs in the ARM database.