DISSERTATION

INSIDE THE BLACK BOX:

ILLUMINATING THE SOIL MICROBIOME FROM AGRICULTURAL FIELDS TO URBAN ROOFTOPS

Submitted by

Lady Grant

Department of Soil & Crop Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2026

Doctoral Committee:

Advisor: Kelly C. Wrighton

Meagan Schipanski
Steven Fonte
Joshua Sbicca

ABSTRACT

INSIDE THE BLACK BOX:
ILLUMINATING THE SOIL MICROBIOME FROM AGRICULTURAL FIELDS TO URBAN ROOFTOPS

Soil processes are central to agricultural productivity and sustainability, however conventional assessments tend to emphasize bulk physicochemical measures and often fail to capture the full complexity of microbial processes driving nutrient cycling, phytohormone production, carbon stabilization, and ecosystem resilience. Multi-omics approaches, which integrate genome-resolved metagenomics, metatranscriptomics, and amplicon sequencing with geochemical and enzymatic measurements, offer a more holistic means of linking microbial community diversity, composition, functional potential, and activity to critical soil processes and outcomes across diverse agroecosystems. This dissertation applies high-resolution microbiome tools to both production-scale field soils and engineered rooftop agrivoltaic systems, providing the first glimpse into how microbial community structure, assembly, and function respond to contrasting environmental and management contexts, in ways that modulate soil processes.

In production scale maize cropping systems, metatranscriptomic and amplicon data in Chapter 2 revealed that the soil conservation management practice of strip tillage promoted microbial strategies favoring carbon cycling and reduced nitrogen loss, whereas conventional tillage enhanced nitrogen turnover and potential gaseous losses. These functional differences were not apparent in standard soil health metrics, highlighting the value of microbial gene expression as a more sensitive indicator of functional change. By coupling multi-omics with the

Soil Management Assessment Framework (SMAF), this research provided one of the first field-scale demonstrations that molecular-level data can expose management effects not yet captured in conventional soil health assessments.

In contrast to the established field system, incipient rooftop agrivoltaic systems offered a unique window into early microbial community diversity, assembly and ecological filtering in an engineered urban agroecosystem. Chapter 3 used amplicon sequencing to reveal that rooftop substrate microbial communities diverged rapidly from their initial greenhouse inocula but gradually converged toward the structure of mature rooftop soils over three years. Photovoltaic shading acted as an ecological filter on community composition rather than richness: alpha diversity increased uniformly across shaded and full-sun plots, whereas beta diversity remained distinct over time, indicating persistent compositional divergence between treatments.

Pairing genome-resolved metagenomics with geochemical and enzyme activities in Chapter 4 revealed that photovoltaic rooftop shading influenced both functional potential and realized activity over time. Carbohydrate-active enzyme genes were enriched under shaded conditions, and β-glucosidase activity, which was initially lower than in full-sun plots, increased over time in the most extreme shading and ultimately surpassed full-sun activity in the second year, indicating delayed but amplified carbon cycling under moderated moisture and temperature regimes. In contrast, full-sun substrates exhibited lower moisture, and heat stress–adapted and potentially pathogenic taxa were detected. Together, these findings indicate that photovoltaic rooftop microclimates exert strong selective pressures on microbial assembly and function, with direct implications for engineered substrate resilience.

By integrating multi-omics with soil health frameworks, this dissertation reveals how microbial communities function as both indicators and agents of ecological processes across

conventional and novel agroecosystems. Field-based tillage experiments highlight the sensitivity of microbial functional responses to management, while rooftop agrivoltaic systems demonstrate how environmental design and microclimate shape microbial succession and activity. Together, these findings provide a framework for incorporating microbial ecology into soil health assessments and guiding resilient agroecosystem design. These advances pave the way for scalable monitoring and management of soil function in a changing world.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, **Dr. Kelly C. Wrighton**, for her guidance, support, and encouragement throughout this project. I am also thankful to my committee members, **Dr. Joshua Sbicca**, **Dr. Meagan Schipanski**, and **Dr. Steven Fonte**, for their thoughtful feedback and mentorship.

I am especially grateful to my lab group: **Dr. Jared Ellenbogen**, **Dr. Bridget McGivern**, **Dr. Ikaia Leleiwi**, **Dr. Katherine Kokkinias**, **Dr. Mikayla Borton**, **Dr. Laura Mason**, **Christopher Chorpenning**, and **Djennyfer Ferreira**, for their friendship and collaboration. A special thank you goes to the **Wilkins Lab**, especially **Dr. Emily Bechtold** and **Dr. Kaela Amundson**, who answered every question without hesitation, and **Julie Fowler**, **Raegan Paul**, **Kya Sparks**, and **Will James**, for welcoming me into your cohort.

I would also like to thank the **Agricultural Research, Development, and Education Center (ARDEC)** team and the **Agricultural Water Quality Program** team, **Erik Wardle**, **Troy Bauder**, **Emmanuel Deleon**, and **AJ Brown**, for their dedication in helping my team sample year after year. My sincere appreciation also goes to **Dr. Tad Trimarco**, **Dr. Jim Ippolito**, and **Dr. Gene Kelly** for their invaluable scientific guidance and moral support.

Finally, I am grateful to **Dr. Jennifer Bousselot** and her team, **Dr. Maria Chavez**, **Eliza Gross**, and **Armando Villa-Ignacia** for the opportunity to collaborate on additional research projects that brought both joy and inspiration to my dissertation work.

This research was supported by the **USDA Food and Agricultural Sciences National Needs Graduate and Postgraduate Fellowship Grants Program (NNF)** and the **Colorado State Microbiome Network (CoSMic)**. I am also thankful to the **Department of Soil and Crop**

Sciences for providing an intellectual home and the resources necessary to complete this dissertation.

Lastly, I would like to thank my parents, **Elaine and Biran Mosley**, who have supported me as I took the path less traveled to get where I am today. Your unwavering encouragement throughout my life, even when I pursued things you did not always understand, showed me that I could remain authentically myself and still have people in my corner rooting for me every step of the way.

To my best friends, **Rebecca Holmes**, **Colby Whalen**, **Dr. Kenneth Marincin**, and **Mark Whittaker Allen**, it looks like we did well for a group of troublemakers.

DEDICATION

This dissertation is dedicated to:

Shadow Killer Mosley (2005 – February 23, 2022) who walked beside me at the start of this journey, and Orb Monster Kitty (2006 – June 27, 2025), who tried her hardest to stay until the end. You both gracefully ushered me from blunder years into my first phase of adulthood.

I will carry you with me, always.

TABLE OF CONTENTS

Chapter 1: Overview of the Current State of Soil Microbiome Science in Agriculture

1.1 Defining Soil Health in Agricultural Systems

Soil health has emerged as a unifying concept in both agricultural research and practice, offering a framework that integrates the biological, chemical, and physical dimensions of soil function. The U.S. Department of Agriculture (USDA) defines soil health as "the continued capacity of soil to function as a vital living ecosystem that sustains plants, animals, and humans"1. This definition reflects a shift away from viewing soils primarily as inert growth media toward recognizing them as dynamic, living systems whose integrity underpins agricultural sustainability. Central to this framing is the idea that soil health is not a fixed state but rather a continuum influenced by management practices, climate, and ecological interactions.

In agricultural systems, soil health is typically evaluated through multiple lenses. The chemical dimension often includes nutrient availability, pH balance, cation exchange capacity, and the presence of contaminants. The physical dimension encompasses soil texture, aggregation, porosity, and water-holding capacity, all of which directly affect plant growth and resilience to stress. Lastly, the biological dimension traditionally includes variables such as soil organic carbon, microbial biomass carbon, organic matter content, and an enzyme assay such as beta-glucosidase or autoclaved-citrate extractable (ACE) soil protein or potentially mineralizable nitrogen assay. Together, these indicators spanning three dimensions begin to profile the capacity of soil to deliver ecosystem services, such as nutrient cycling, carbon sequestration, water regulation, and disease suppression.

Despite this integrative framing, defining soil health in practice remains contested. Various indicators and indices have been developed to quantify soil health, ranging from field-

based assessments of structure and infiltration to measurements of soil fauna and plant performance, to laboratory assays of microbial activity and biodiversity. Each approach captures only part of the picture. Standardized frameworks like the Soil Management Assessment Framework (SMAF) or the Comprehensive Assessment of Soil Health (CASH) attempt to synthesize multiple indicators into a composite score, but such metrics are still limited by context dependency[2,3]. Soils that function well in one agroecosystem may be considered degraded in another due to differing climatic, cultural, or management expectations.

An important tension lies in balancing productivity goals with sustainable soil management. Conventional yield-focused measures often overlook long-term degradation processes such as organic matter loss or compaction, while overuse of synthetic fertilizers can decouple ecological processes from production, as well as cause considerable nitrogen loss through gaseous emissions and losses in water that cause contamination. Moreover, most soil health frameworks have been designed with conventional row-crop agriculture in mind, with unmeasured utility across diverse agroecosystems such as urban farms, indigenous agricultural systems, or perennial polycultures[4,5]. This problem, however, may be mitigated with greater emphasis on the biological or ecological dimension of soils, especially more holistic measurements that do not rely on single measurements or crude proxies of complex processes[6]. Yet the question of which biological indicators, and what level of aggregation, most accurately reflect soil resilience and function remains an active and evolving area of research, where microbial ecological frameworks coupled with multi-omic data can play a critical discovery role.

1.2 Microbial Contributions to Soil Function and Resilience

Microbes play a central role in soil health and fertility, yet current soil health indices often fall short in capturing the full biological dimension. For example, it is well established that microbial communities perform essential functions such as nutrient cycling (e.g., nitrogen fixation, nitrification, and denitrification)[7–9], organic matter decomposition through extracellular enzyme activity[10–12], and soil structure stabilization via fungal hyphae and microbial extracellular polymeric substances[13,14].

In terms of plant growth promotion, the organisms carrying out these processes are taxonomically diverse, and in some cases, functional roles have been characterized for particular lineages or representative taxa. For example, well-studied diazotrophs such as *Rhizobium* and free-living *Azotobacter* can drive biological nitrogen fixation in the rhizosphere. Likewise, bacterial genera such as *Bacillus* and *Pseudomonas* have been associated with nutrient mobilization and pathogen suppression[15,16]. However, these examples represent only a small subset of the functional breadth present in soil microbial communities, which often exhibit multiple, context-dependent strategies within and across taxonomic groups. While a handful of critical taxa have been historically associated with plant growth and soil health attributes, our understanding of the full diversity of microbial groups contributing to these key processes, and the mechanisms by which they sustain soil fertility and resilience, remains incomplete.

In human medicine, microbial communities can act as early indicators of health, disease, or recovery[17–19]. A similar principle may apply to agricultural systems, where microbiomes often respond more rapidly to management changes than bulk soil properties like organic matter content. Shifts in community membership, enrichment of specific bioindicator taxa, or changes in functional gene potential related to key soil processes may signal nutrient depletion,

contamination, or restoration well before such changes are detectable through conventional soil health tests. As a result, holistic microbial indicators offer a powerful complement to existing physical and chemical measures of soil health.

Yet despite their centrality to soil and crop health, microbial contributions have historically been underrepresented in soil health frameworks. Biological indicators often account for only a small fraction of composite soil health scores, and the metrics selected, such as microbial biomass carbon or potentially mineralizable nitrogen capture only a fraction of the metabolic complexity that the soil system as a whole is capable of. The recognition that microbes regulate both immediate productivity and long-term resilience has led to increasing calls for integrating more holistic measures of microbiology into soil health assessments, particularly through advanced molecular and multi-omic approaches. These tools allow researchers to move beyond bulk activity measures toward detailed characterizations of microbial composition, function, and interactions, laying the groundwork for the next phase of soil health.

1.3 Overview of Omic Technologies in Agricultural Research

For much of the past two decades, agricultural microbiome studies have relied heavily on amplicon sequencing of marker genes such as 16S rRNA for Bacteria and Archaea or ITS for Fungi. Amplicon sequencing remains popular because it is relatively inexpensive, scalable, and computationally accessible. It has provided valuable insights into microbial diversity, community structure, and how these patterns shift across soil types, management practices, and cropping systems. However, amplicon sequencing is limited to taxonomic identification and relative abundance. It cannot directly reveal the functional capacities of microbial communities,

making it insufficient for connecting soil microbial ecology to soil health processes, especially those of carbon and nitrogen mineralization.

To address these limitations, more comprehensive omic approaches have been adopted. Metagenomics (MetaG) profiles the total DNA present in soil, providing information about both community composition and the genetic potential for nutrient cycling, organic matter degradation, or pathogen suppression. Metatranscriptomics (MetaT) builds on this by sequencing RNA to capture actively expressed genes, offering a snapshot of functional processes under specific environmental or management conditions. Metaproteomics and metabolomics extend further by characterizing expressed proteins and small-molecule metabolites, respectively, which reflect microbial activity at the molecular and biochemical level.

Although technically demanding and more resource-intensive, these approaches provide insights into microbial function that amplicon sequencing alone cannot deliver. Increasingly, researchers are combining multiple omic layers in multi-omic frameworks such as, allowing integration of taxonomic, genetic, and functional information. While applications in agriculture remain fewer compared to human microbiome research, the work presented in this dissertation leverages these emerging tools to directly link soil management practices with microbial processes and soil health outcomes.

1.4 Multi-omics application to the agricultural soil microbiome: from concept to capacity

Multi-omic approaches have expanded the ways in which soil microbial communities can be studied in agricultural systems. Among these, amplicon sequencing of the Bacterial (16S rRNA) marker gene remains the most widely applied method. This approach provides insight into microbial community composition, diversity, and increasingly, network architecture across

soils and cropping systems. Large-scale biogeographic surveys have been particularly impactful, revealing continental-scale patterns in microbial assemblages and their relationships to soil health and productivity. For example, Wilhelm et al. (2022) identified archaeal nitrifiers as indicators of degraded soils, while members of *Chloroflexi* (class KD4-96) were associated with higher soil health scores[20]. More recently, Mo et al. (2024) revealed that fertility source and management practices strongly influence microbial assembly and co-occurrence patterns across soil depth in agricultural systems[21]. Together, these studies underscore the continuing importance of amplicon surveys for benchmarking microbiome structure, identifying potential indicator taxa, and informing the integration of more functionally resolved approaches.

Despite these advances in amplicon surveys, there is growing interest in moving beyond community structure to directly assess microbial functional potential through metagenomics. Historically, metagenomic approaches were limited by cost and low genome recovery in complex soils. However, advances in assembly algorithms and the construction of curated gene and genome catalogs from ecosystems ranging from wetlands to agricultural soils (e.g., Borton et al., 2025; Bechtold et al., 2025; Seitz et al., 2022)[22–24] have made these approaches increasingly tractable. This has opened new opportunities to link microbial metabolism with key soil processes such as nitrogen transformation and carbon stabilization. These functional datasets enable more mechanistic interpretations of soil processes, complementing amplicon-based indicators.

While metagenomics reveals functional potential, it does not indicate which genes are actively expressed in situ. Metatranscriptomics, in contrast, provides direct evidence of gene expression and active pathways. Although rarely applied in agricultural contexts, largely due to challenges in RNA extraction, sample stability, sequencing costs, and computational demands,

metatranscriptomic studies in other soil systems have shown remarkable promise. For example, Bechtold et al., 2025 demonstrated that metatranscriptomic data offered greater predictive power than amplicon or metagenomic data for methane fluxes, revealing active metabolic networks underlying greenhouse gas emissions[23]. These results highlight the potential of metatranscriptomics to uncover dynamic microbial responses to management and environmental change in agricultural soils.

A key barrier to broader adoption of these higher-resolution approaches is the lack of optimized bioinformatic pipelines for diverse soil microbiomes. Many existing tools were designed for medical or model systems, resulting in low recovery rates and limited functional resolution for environmental samples. Functional annotation is particularly challenging because many soil processes are underrepresented in reference databases. To address these gaps, tools such as the Distilled and Refined Annotation of Metabolism (DRAM and DRAM2) have been developed to improve functional annotation and pathway reconstruction from microbial genomes, particularly those related to biogeochemical cycling[25]. Although this dissertation employs DRAM1 for functional annotation, it also contributed to the development of DRAM2, which represents the next generation of tools for improving ecological interpretation of soil omics data.

These advances open a pathway to developing microbiome-based metrics for soil health, moving beyond proxy enzyme assays to mechanistic, multi-trophic indicators grounded in microbial metabolism. Yet a key open question, and the focus of this dissertation, is which microbial metrics are most informative and scalable across the inherent spatial, temporal, and ecological complexity of agricultural soils. Amplicon, metagenomic, and metatranscriptomic approaches offer complementary strengths: amplicon sequencing is cost-effective and scalable,

while metagenomics and metatranscriptomics provide deeper insight into microbial functional capacity and activity. Used in tandem, these tools bridge the gap between community structure and function, enabling more mechanistic linkages between soil management practices, microbial ecology, and soil health outcomes.

1.5 The Value of Comparing Large-Scale and Urban Small-Scale Systems

Agricultural systems operate across diverse scales and contexts, from industrial monocultures that dominate global food production to urban agriculture initiatives that address local food security and climate resilience. Each context shapes soil microbial communities in distinct ways and comparing them provides opportunities to evaluate how management practices, scale, and environmental factors influence soil health.

The large-scale system examined in this dissertation (Chapter 2) represents a semi-arid, monocropped maize-based field managed with synthetic fertilization and both conventional and soil conservation tillage practices for over a decade. This type of system exemplifies many of the challenges facing modern production agriculture: declining soil organic matter, heavy nutrient inputs, and physical disturbance. At the same time, such systems are critical for sustaining global food supply, making it essential to understand how soil microbial communities respond to management decisions that affect long-term soil health.

In contrast, the small-scale system examined in this study is an urban rooftop agrivoltaic system, where crops are cultivated beneath photovoltaic panels in an engineered substrate. Unlike the conventional, field-based soils discussed in Chapter 2, these systems are designed to maximize limited urban space while advancing sustainability goals such as local food production, stormwater mitigation, and renewable energy generation[26]. Because weight is a

critical design constraint, the material supporting plant growth is not technically soil in the pedological sense[27]. Instead, the plant growth medium is an engineered substrate, typically composed of blends of compost, sand, perlite, or other lightweight materials. These engineered substrates lack the mineral weathering profiles, structured horizons, and legacy organic matter characteristic of mature agricultural soils. As constructed, shallow systems, they are subject to unique environmental stressors, including limited rooting depth, rapid wet-dry cycles, and greater exposure to temperature extremes. These material and environmental differences are likely to shape microbial and biogeochemical dynamics in the rhizosphere. Despite these challenges, urban agroecosystems such as rooftop agrivoltaic systems are increasingly recognized as valuable testbeds for innovative approaches to soil management, microbial monitoring, and climate adaptation, offering a forward-looking model for resilient food production in built environments.

Comparing these two systems highlights the contrasting pressures and opportunities that shape microbial communities in agricultural contexts. In large-scale row-crop systems, the emphasis is often on reconciling productivity with conservation, while in urban rooftop systems the priority is building soil function and resilience from a relatively undeveloped starting point. Multi-omic analyses provide a universal lens, enabling researchers to move beyond context-specific metrics toward a mechanistic understanding of microbial contributions to soil health that is relevant across scales.

While these two systems represent distinct agricultural contexts, we recognize that the results presented here are based on two study sites, both located in Colorado, and thus broader generalizations should be made with caution. Nonetheless, these contrasting systems played a critical role in demonstrating the tractability of diverse soil environments to current microbiome-

based analytical approaches. Together, they served as testbeds for evaluating and refining methods of microbial community characterization and functional annotation, including the frameworks developed here to assess microbial contributions to soil health. Insights gained from these studies provided foundational proof-of-concept for applying multi-omics and genome-resolved analyses to soil systems, approaches that are now being scaled to broader, continental datasets to inform future soil health monitoring and management strategies (e.g., US. Department of Energy IN-RICHES project).

By applying similar methodologies across both systems, this dissertation demonstrates how amplicon, metagenomic, and metatranscriptomic approaches can uncover the black box of the soil microbiome, providing a deeper, functional understanding of how agricultural management practices and environmental gradients shape soil microbial function. In doing so, this thesis paves a path toward soil assessments that are both mechanistically informed and broadly applicable across agroecosystems.

1.6 Goals of this Dissertation

Using various estimates of the microbiome, this dissertation explores holistic features of soil microbial systems including diversity, community composition, and functional potential to begin opening the black box of how microbiomes mediate agricultural soil processes. These molecular measurements are coupled with traditional process-based metrics, offering new ways to connect microbial traits with soil function and management outcomes. Together, this work develops microbiome workflows for complex agricultural soils, spanning both conventional and engineered systems. The overarching goals are to: **Chapter 2**) evaluate how tillage practices influence microbial features in a large-scale maize-based system; **Chapter 3**) characterize

mechanisms for microbial community assembly and succession in a newly established rooftop agrivoltaic system; and **Chapter 4**) identify functional traits responsive to different rooftop agrivoltaic environments. We conclude with **Chapter 5**, which synthesizes findings across systems and discusses their broader implications for biological soil health assessment. This final chapter emphasizes the value of multi-omics as a complement to traditional soil health measurements and makes a compelling case for their applicability across diverse agroecosystems, from large-scale monocultures to small-scale urban rooftops.

Chapter 1: References

1. Soil Health | Natural Resources Conservation Service. https://www.nrcs.usda.gov/conservation-basics/natural-resource-concerns/soil/soil-health (2025).
2. Andrews, S. S., Karlen, D. L. & Cambardella, C. A. The Soil Management Assessment Framework. *Soil Sci. Soc. Am. J.* **68**, 1945–1962 (2004).
3. Idowu, O. J. *et al.* Use of an integrative soil health test for evaluation of soil management impacts. *Renew. Agric. Food Syst.* **24**, 214–224 (2009).
4. Salomon, M. J., Watts-Williams, S. J., McLaughlin, M. J. & Cavagnaro, T. R. Urban soil health: A city-wide survey of chemical and biological properties of urban agriculture soils. *J. Clean. Prod.* **275**, 122900 (2020).
5. Hu, X. *et al.* Soil health quantification via SMAF and CASH across diverse land uses. *Geoderma* **461**, 117492 (2025).
6. Bhaduri, D. *et al.* A review on effective soil health bio-indicators for ecosystem restoration and sustainability. *Front. Microbiol.* **13**, (2022).
7. Kuypers, M. M. M., Marchant, H. K. & Kartal, B. The microbial nitrogen-cycling network. *Nat. Rev. Microbiol.* **16**, 263–276 (2018).
8. Ichihashi, Y. *et al.* Multi-omics analysis on an agroecosystem reveals the significant role of organic nitrogen to increase agricultural crop yield. *Proc. Natl. Acad. Sci.* **117**, 14552–14560 (2020).
9. Wright, C. L. & Lehtovirta-Morley, L. E. Nitrification and beyond: metabolic versatility of ammonia oxidising archaea. *ISME J.* **17**, 1358–1368 (2023).

10. Wallenstein, M. D. & Burns, R. G. Ecology of Extracellular Enzyme Activities and Organic Matter Degradation in Soil: A Complex Community-Driven Process. in *Methods of Soil Enzymology* 35–55 (John Wiley & Sons, Ltd, 2011). doi:10.2136/sssabookser9.c2.
11. Bai, X. *et al.* Extracellular enzyme activity and stoichiometry: The effect of soil microbial element limitation during leaf litter decomposition. *Ecol. Indic.* **121**, 107200 (2021).
12. Luo, L., Meng, H. & Gu, J.-D. Microbial extracellular enzymes in biogeochemical cycling of ecosystems. *J. Environ. Manage.* **197**, 539–549 (2017).
13. Costa, O. Y. A., Raaijmakers, J. M. & Kuramae, E. E. Microbial Extracellular Polymeric Substances: Ecological Function and Impact on Soil Aggregation. *Front. Microbiol.* **9**, (2018).
14. Redmile-Gordon, M., Gregory, A. S., White, R. P. & Watts, C. W. Soil organic carbon, extracellular polymeric substances (EPS), and soil structural stability as affected by previous and current land-use. *Geoderma* **363**, 114143 (2020).
15. Mehmood, N. *et al.* Multifaceted Impacts of Plant-Beneficial *Pseudomonas* spp. in Managing Various Plant Diseases and Crop Yield Improvement. *ACS Omega* **8**, 22296–22315 (2023).
16. Etesami, H., Jeong, B. R. & Glick, B. R. Potential use of Bacillus spp. as an effective biostimulant against abiotic stresses in crops—A review. *Curr. Res. Biotechnol.* **5**, 100128 (2023).
17. Wu, G. *et al.* A core microbiome signature as an indicator of health. *Cell* **187**, 6550-6565.e11 (2024).
18. Tegegne, H. A. & Savidge, T. C. Leveraging human microbiomes for disease prediction and treatment. *Trends Pharmacol. Sci.* **46**, 32–44 (2025).

19. Aggarwal, N. *et al.* Microbiome and Human Health: Current Understanding, Engineering, and Enabling Technologies. *Chem. Rev.* **123**, 31–72 (2023).
20. Wilhelm, R. C., van Es, H. M. & Buckley, D. H. Predicting measures of soil health using the microbiome and supervised machine learning. *Soil Biol. Biochem.* **164**, 108472 (2022).
21. Mo, Y. *et al.* Agricultural practices influence soil microbiome assembly and interactions at different depths identified by machine learning. *Commun. Biol.* **7**, 1349 (2024).
22. Borton, M. A. *et al.* A functional microbiome catalogue crowdsourced from North American rivers. *Nature* **637**, 103–112 (2025).
23. Bechtold, E. K. *et al.* Metabolic interactions underpinning high methane fluxes across terrestrial freshwater wetlands. *Nat. Commun.* **16**, 944 (2025).
24. Seitz, V. A. *et al.* Variation in Root Exudate Composition Influences Soil Microbiome Membership and Function. *Appl. Environ. Microbiol.* **88**, e00226-22 (2022).
25. Shaffer, M. *et al.* DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* **48**, 8883–8900 (2020).
26. Huang, M. *et al.* Turning rooftops green with agrivoltaics: A path to carbon reduction and mitigating climate change. *Energy Build.* **343**, 115882 (2025).
27. Kader, S., Chadalavada, S., Jaufer, L., Spalevic, V. & Dudic, B. Green roof substrates—A literature review. *Front. Built Environ.* **8**, (2022).
28. Ning, D. *et al.* A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat. Commun.* **11**, 4717 (2020).

Chapter 2: Microbial Functions Under Different Tillage Practices

2.1 Summary

Tillage practices shape soil microbial communities and their functional roles in biogeochemical cycling, with implications for soil health and sustainability. While conventional tillage (CT) is commonly used worldwide, its impacts on microbial processes underlying nitrogen cycling and carbon stabilization remain poorly resolved. Here, we integrate amplicon sequencing (16S rRNA, ITS), multi-omics (metagenomics, metranscriptomics), and traditional soil health data to examine the effect of a decade of CT versus the soil conservation practice of strip tillage (ST) in a semi-arid, continuously cropped maize-based system. Amplicon data revealed persistent treatment differences in bacterial and fungal communities, especially in rhizosphere soils. Genome resolved metatranscriptomics showed elevated transcription of ammonia oxidation and nitrite reduction genes under CT, primarily by Thermoproteota (syn. Nitrososphaerota) and Nitrospirota, suggesting enhanced nitrogen loss. In contrast, ST soils exhibited enhanced carbon depolymerization activity and increased expression of microaerophilic respiration genes (i.e., genes that enable respiration under limited oxygen). Despite these shifts in microbial function, biological soil health parameters such as soil organic carbon, microbial biomass carbon, potentially mineralizable nitrogen and beta-glucosidase showed no significant treatment effects. Our results highlight microbial gene expression as a sensitive indicator of soil functional shifts that may precede detectable changes in soil health metrics. This work underscores the value of integrating multi-omics with soil health assessments to guide sustainable agricultural practices and improve early detection for soil ecosystem responses.

2.2 Introduction: Tillage, soil health, and microbial communities

Global agricultural systems face mounting pressure to remain productive amid climate change, land degradation, and growing resource constraints[1]. Enhancing the resilience and sustainability of food production is critical to maintaining yields and supporting long-term ecosystem function. Improving soil health is viewed as a key strategy for sustainable food production under changing climates[2].

Soil health is defined as "the ability of soil to sustain the productivity, diversity, and environmental services of terrestrial ecosystems"[3,4] and can be improved by management decisions[5]. The main principles of managing soil health are minimizing soil disturbance, maximizing plant diversity, maintaining living roots, and maximizing soil coverage[6]. An increasingly popular management strategy is adoption of conservation tillage approaches, like strip tillage (ST, Figure 1). This approach minimizes soil disturbance by limiting plowing to narrow seed rows and maintaining at least 30% crop residue cover on the soil surface[7]. This contrasts with conventional tillage (CT, Figure 1), which involves soil inversion, often through moldboard plowing, and reduces soil cover by mixing crop residues into the soil. Together, these contrasting practiced create distinct soil environments that can shape nutrient cycling, microbial activity and soil functioning.

Assessing soil health remains challenging due to the complexity of biological, chemical, and physical interactions in soils. There are various indices that collapse multiple chemical, physical, and biological indicators into an overall score[8]. These include, but are not limited to, pH, bulk density, soil organic carbon (SOC), aggregate stability, nutrient availability, microbial biomass, and enzyme activities[9]. While these metrics can offer valuable insights, they often

overlook biological drivers of soil function[10,11], and operate at a much lower resolution than the processes they aim to capture.

The soil microbiome plays a critical role in soil health by regulating organic matter decomposition, contributing to soil aggregation and structure, and influencing nitrogen cycles[12]. Given their rapid responses to environmental changes[13,14], microbial communities may enable earlier detection of soil degradation or recovery when combined with traditional soil health metrics. Microbial communities can be measured at varying levels of resolution, looking at diversity, composition, functional potential, and gene expression. Recent work suggests that microbial functional potential and transcriptional activity may serve as early and sensitive indicators of soil health change, offering a promising complement to conventional assessments[15–17]. However, most existing studies, rely on gene abundance alone, which lacks insight into transcriptional activity, ecological context, or the specific taxa responsible for functional responses. This limits mechanistic understanding of how microbial processes drive soil health outcomes. Therefore, integrating microbiome multi-omics approaches with traditional soil health assessments may provide a more comprehensive understanding of microbial contributions to soil function and sustainability and improve measurements of soil health.

Here, we leverage a long-term tillage agricultural experiment to assess the soil microbiome's potential as an indicator of soil health. By integrating amplicon sequencing and genome-resolved metatranscriptomes from soils under CT and ST management strategies[18], we evaluated dynamics in microbial community structure and function. Collectively, we provide a deeper understanding of how soil management shapes the biological mechanisms underpinning soil health, knowledge which can be leveraged to make food production more sustainable.

2.3 Results & Discussion

2.3.1 Tillage alters microbial community composition

We sampled soils from a decade-long experiment in Northern Colorado, US, which evaluated crop production under CT and ST. Before the experiment began in 2005, the field site had been managed uniformly under a sunflower (*Helianthus annus*), corn (*Zea mays* L.), dry beans (*Phaseolus vulgaris*) rotation using full moldboard-plow tillage comparable to the current CT treatment. Both treatments were managed using consistent practices for conventional and conservation tillage, including utilizing distinct fertilization strategies, with a dry fertilizer broadcast in CT and liquid banded application in ST (Figure 1). To investigate how tillage management shapes soil microbial communities, we assessed bacterial and fungal diversity across 207 samples collected from surface (0-10 cm) and deep soils (10-30 cm), as well as rhizosphere soils spanning three maize growth stages: emergence (VE, seedling stage), vegetative tasseling (VT, tassel development), and reproductive (R5, kernel filling stage) (Supplementary Figure 1a). A total of 5.7 million bacterial and archaeal 16S rRNA gene amplicon reads and 2.6 million fungal ITS amplicon reads were obtained, resulting in 39,855 bacterial/archaeal and 24,154 fungal amplicon sequence variants (ASVs).

Bacterial, archaeal and fungal alpha diversity patterns varied across treatments, soil compartments, and timepoints (Supplementary Figure 2 a-d). When looking at bacterial and archaeal communities, ST had significantly higher ASV richness and Shannon diversity across bulk soils (surface and deep soils combined), however within the rhizosphere, CT had significantly higher richness while Shannon diversity showed no significant difference between the two treatments. Fungal communities followed the same pattern, to a lesser extent

(Supplementary Figure 2 e-h). Fungal functional guild (Supplementary Figure 3) showed no consistent treatment effects. Although bacterial and archaeal alpha diversity revealed significant differences in some instances, high variability across individual blocks and directions suggests that richness alone does not consistently reflect treatment effects.

There were however clear and consistent treatment effects in beta diversity, particularly in rhizosphere bacterial and archaeal communities (Figure 2a) (Supplementary Figure 4-5). In July rhizosphere samples, tillage treatment explained 42.72% of the variance in bacterial and archaeal community composition (PERMANOVA, $p < 0.05$), compared to 22.46% in surface soils (Figure 1a). In bulk soils, tillage treatment and depth showed significant variance (PERMANOVA, $p < 0.05$), though depth accounted for the largest share of variance (11.58%) (Figure 1a). Fungal communities showed weaker structure overall (Supplementary Figure 5 e-h), with treatment explaining only 8.45% of the variance in bulk soils, but 17.17% in surface samples and 17.66% in the rhizosphere. These significant shifts in beta diversity indicate that while the total number of species (alpha diversity) within each treatment may not differ drastically, the identities and relative abundances of taxa do. These differences are likely driven by well-documented contrast between CT and ST systems, including greater surface reside retention, improved soi structure, higher infiltration rated and enhanced biological activity[19], all of which create distinct microbial niches and resource gradients. Such contrasts are especially pronounced in the rhizosphere, where root inputs further amplify treatment-specific microbial community assembly.

Given the strong treatment signal in the rhizosphere at July (VT) (Figure 1a), we next identified microbial lineages that were differentially abundant between CT and ST in this subset of samples (Figure 1 b-c). Nine bacterial and one archaeal phyla were significantly more

abundant in CT, including the Actinomycetota, Chloroflexota, Desulfobacteria_B, Myxococcota, Nitrospirota, Tectomicrobia, Thermoplasmatota, Thermoproteota, and Verrucomicrobiota phyla (Figure 1c). In contrast, Dependentiae, Patescibacteria, and Pseudomonadota phyla were significantly more abundant in ST (ANOVA, $p < 0.05$; Figure 1c). At a finer taxonomic resolution, discriminant ASVs in (MaAsLin2, FDR < 0.05) in CT included *Nitrospira_C japonica_A*, an ASV within Nitrososphaera, and a member of the genus Nonomuraea. Discriminant ASVs enriched in ST included representatives of the class Gammaproteobacteria, the family Sphingomonadaceae, and the order Pseudomonadales. Overall, the enrichment of nitrifiers and stress-adapted lineages in CT, compared to rhizosphere-associated and resource-responsive taxa in ST, reinforces that the two tillage systems foster fundamentally different microbial strategies, especially within the rhizosphere.

While many ASVs were treatment-specific, the overlap between communities was minimal: only 9.7% of ASVs were shared across management regimes. We defined these shared taxa as the core rhizosphere microbiome, present in ≥85% of both CT and ST rhizosphere samples. These core taxa were members of Pseudomonadota, Gemmatimonadota, Chloroflexota, Actinomycetota, Planctomycetota, and Acidobacteriota.

These results indicate that long-term tillage management not only shifts the relative abundance of dominant bacterial groups but also leads to divergent community structures with a minimal shared core. Treatment effects were strongest in the July rhizosphere (VT), a period when root activity and nutrient demands peak. This heightened sensitivity likely reflects the combined effects of tillage-driven differences in soil structure and residue distribution together with management-specific nutrient placement, as fertilizers are banded directly into the root zone in ST but broadcast more broadly under CT. These interacting factors create sharper contrasts in

rhizosphere resource environments, offering a clear window into how conserved microbial lineages shift their functional priorities in resources in response to tillage-induced changes in soil physical conditions and nutrient availability.

2.3.2 Genome-resolved analysis reveals management-specific functional activity

Given the strong treatment effect observed in the bacterial communities of July rhizosphere samples, we used metagenomic sequencing to investigate the underlying functional gene and genome potential of these soils. We obtained deep metagenomic sequencing (range 6.3–160.6 Gbp/sample) on 12 representative samples spanning tillage treatments, soil depths, and time points generating a total of 367.74 Gbp of sequencing.

From these data, we reconstructed 108 medium- and high-quality (MQ-HQ)[20] metagenome-assembled genomes (MAGs). After dereplication at 99% identity, these MAGs were integrated with 342 MQ-HQ genomes from the Agricultural Exudate-Responsive Metagenomic (ARM) database, which includes enriched genomes derived from agricultural soils in the same research complex[21]. The final dataset comprised 378 dereplicated MQ-HQ MAGs, including 49 newly reconstructed in this study. This integrated genome catalog provides a high-resolution view of the microbial functional landscape under long-term tillage regimes and enables direct comparison of both community membership and encoded metabolic traits. Catalog MAGs from 23 different phyla recruited metagenome reads, with members of the Pseudomonadota being the most abundant (28.3% abundance). This catalog represents one of the first genome-resolved resources for agroecosystems under long-term tillage management, providing a foundation for linking microbial community structure to ecosystem function at high resolution.

Many of the MAGs in our catalog represent uncharacterized or poorly resolved microbial lineages, raising the question of which of these taxa are metabolically active under different tillage regimes. To address this, we obtained metatranscriptome sequencing on quadruplet July rhizosphere samples from CT and ST treatments (mean = 13.6 Gbp per sample). Mapping the resulting transcripts to our MAG database revealed that 43 MAGs recruited metatranscriptomic reads, indicating active gene expression *in situ* (Figure 3a). These transcriptionally active genomes spanned 10 bacterial and one archaeal phyla, with the highest representation from Actinomycetota, Thermoproteota, and Pseudomonadota (Figure 3c). Strikingly, a large proportion of the most active MAGs belonged to poorly classified lineages, with taxonomic assignments limited to alphanumeric identifiers at higher taxonomic ranks: one at the class level, five at order, eight at family, and 27 at genus (Figure 3b). These unnamed (i.e., not classified at that level by the GTDB) but transcriptionally active lineages may play unexplored roles in rhizosphere biogeochemistry and represent promising candidates for future cultivation and functional investigation.

Underscoring the distinct functional trajectories of microbial communities under different tillage intensities, 13 MAGs showed significant treatment-specific expression patterns (Figure 3a) (MaAsLin2, $q < 0.05$). Among the thirteen treatment-associated MAGs, five MAGs were enriched in ST, while eight were enriched in CT (Figure 3a), revealing a distinct functional expression pattern. Across both treatments, a single Actinomycetota MAG (*JARDZU01 sp029240695*) dominated expression profiles, accounting for 44–72% of total transcripts (Figure 3c). This organism likely plays a central role in rhizosphere carbon turnover, as it actively transcribed genes for polyphenol oxidation (CAZy AA1), cellulose degradation (GH16, GH5), and uptake of plant-derived sugars, including transporters for alpha-glucoside, rhamnose, xylose,

and raffinose-family oligosaccharides (Supplementary Data 1). Its high gene expression across tillage regimes suggests it is a core contributor to the microbial decomposition of plant residues, a critical process for maintaining soil organic matter and nutrient availability in agroecosystems.

The second most transcriptionally active MAG, a member of the Desulfobacterota_B (*WHTF01 sp021323355*), contributed 10–34% of transcripts and expressed genes involved in carbon fixation (reductive TCA cycle) as well as key nitrogen cycling pathways, including nitrite oxidoreductase (*nxr*) and nitric oxide reductase (*nor*) (Figure 3c; Supplementary Data 2). These functions suggest a role in both inorganic nitrogen transformation and redox balance, with potential implications for nitrogen retention and greenhouse gas fluxes. Together, these highly active but poorly described taxa appear to be central players in rhizosphere carbon and nitrogen cycling in tillage-managed soils. Their strong transcriptional signals, coupled with functional roles in residue decomposition and nitrogen transformation, underscore their potential importance in regulating nutrient turnover and maintaining soil function in production agriculture.

2.3.3 Thermoproteota transcriptional activity is increased in conventional tillage

Among the 43 transcriptionally active MAGs, 13 belonged to the archaeal family Nitrososphaeraceae. All 13 encoded the capacity for ammonia oxidation, with two genomes exhibiting significantly higher transcription under CT management, suggesting a shift in metabolic activity linked to tillage regime. Within Nitrososphaeraceae, a single genus, TH1177, dominated both amplicon and metatranscriptomic datasets. To further explore this lineage, we constructed a family-level phylogeny with transcription overlaid, revealing that the majority of Thermoproteota activity was attributed to genus TH1177 (Figure 4a).

To assess functional differences associated with management, we analyzed the expression of nitrogen-cycling genes within Nitrososphaeraceae, focusing on *amoCAB* (ammonia monooxygenase), Urease, *amt* (ammonium transporter), and *nirK* (nitrite reductase) (Figure 4 b-e). Expression of all selected genes was elevated in CT relative to ST. Notably, urease expression was significantly higher in CT, consistent with the higher proportion of urea-based fertilizer applied in this system (Figure 1a). As urease catalyzes the hydrolysis of urea into ammonium and carbon dioxide elevated transcription suggests increased microbial processing of fertilizer-derived nitrogen in conventionally tilled soils.

TH1177 exhibited higher expression of ammonia monooxygenase genes under CT, consistent with increased ammonia oxidation activity. This lineage also showed elevated transcription of carbon fixation and electron transport chain genes, suggesting enhanced chemoautotrophic metabolism under conventional tillage. Together, these findings indicate that Thermoproteota, particularly the TH1177 lineage, are key mediators of nitrogen transformation in maize rhizospheres, with metabolic activity tightly coupled to tillage-driven shifts in nitrogen inputs and availability. These responses have direct implications for fertilizer efficiency, nitrogen retention, and nutrient cycling in agroecosystems.

2.3.4 Transcription of carbon and nitrogen cycling genes across tillage is distinct

To ensure that key functional processes were not overlooked due to incomplete genome reconstruction, we also mapped metatranscriptomic reads to our full non-redundant gene catalog derived from all metagenome assemblies (440,445 genes). This analysis identified 12,387 transcribed genes, which we categorized into functional groups involved in carbon, nitrogen, and oxygen cycling (Figure 5b). These transcripts clustered by tillage treatment, forming functionally

distinct groups in a hierarchical clustering analysis (Figure 5a), indicating that microbial metabolic activity diverges meaningfully between CT and ST soils.

Given the differences in plant residue persistence between treatments (Figure 1), we aimed to understand carbon depolymerization processes, particularly those associated with the breakdown of cellulose, hemicellulose, starch, polyphenols, and lignin, which dominate corn litter[22]. While general carbon cycling transcripts were shared across treatments, plant cell wall degradation genes, notably xylanases, cellulases (GH5, GH9), and lignin-degrading CAZymes (AA1), were significantly more expressed in ST (MaAsLin2, FDR < 0.05; Figure 5b). These genes were primarily transcribed by members of the Actinomycetota and Pseudomonadota (Supplementary Figure 6a), suggesting that residue-degrading specialists are more active in low-disturbance soils.

To assess treatment-specific responses in organic nitrogen cycling, we examined expression of extracellular peptidases, key enzymes involved in nitrogen mineralization[23]. Aspartic and serine peptidases were more highly expressed in CT than ST, although some lacked clear annotations (Figure 5b). These functions were primarily attributed to Actinomycetota and Thermoproteota MAGs, suggesting that conventionally managed soils favor microbial strategies for rapid nitrogen mineralization, potentially as a response to urea input or greater disturbance-induced substrate turnover. These findings align with previous studies reporting priming effects under tillage, where disturbance stimulates microbial degradation of existing organic nitrogen pools[24–26]. Transcript-level evidence for nitrification matched our genome-resolved findings, with significantly higher expression of ammonia oxidation (amoCAB) and nitrite oxidation (nxrAB) genes in CT (Supplementary Figure 6 b-c, Figure 5b). These transcripts were

predominantly assigned to *TH1177* and *Nitrospira*, respectively. Although denitrification genes (*nirK, norB*) were also more expressed in CT, these trends were not statistically significant.

Consistent with evidence that ST promotes more stable aggregates, greater macrofaunal activity, and improved infiltration[19], we observed higher expression of high-affinity cytochrome oxidases in ST than CT (Figure 5b). These genes, which support respiration under microaerophilic microsites[27] that can form within protected soil aggregates or residue-rich zones, even when soil aeration is high. These transcripts were primarily transcribed by the same Pseudomonadota and Actinomycetota lineages involved in carbon degradation, suggesting that decomposition in ST is carried out by organisms adapted to oxygen limited microhabitats created by residue retention and aggregate protection.

2.3.5 Enzyme assays mirror transcriptional genes and provide whole pathway insights

To directly compare metatranscriptome activity with enzymatic soil health metrics, we extracted transcripts with EC numbers corresponding to common enzyme assays and evaluated their alignment with Soil Management Assessment Framework (SMAF) measurements collected concurrently from the same plots and at the same time point[18] (Figure 6, Supplementary Fig 8). Overall, metatranscriptome patterns generally mirrored SMAF measurements. SMAF beta-glucosidase (βG) enzyme activity trended higher in ST (137 ± 3 PNP/kg/hr) than in CT (122 ± 52 PNP/kg/hr) (Figure 6a), consistent with elevated βG-related gene expression in ST (Figure 6b), though these were not significant. This agreement between enzymatic assays and transcriptional activity highlights how combining SMAF indicators with multi-omic measurements can reveal both the magnitude of soil health responses and the underlying

microbial processes driving them, providing researchers and managers with a more mechanistic and actionable understanding of treatment effects.

Interestingly, when transcriptional genes involved in all three steps in cellulose degradation are combined, gene expression was significantly higher in ST (Figure 6d, p-value = 0.029). This shows how metatranscripts can provide more insight in the degradation of soil organic matter. In particular, cellulose degradation requires multiple enzymatic steps, and pooled transcriptional data revealed significant differences not captured by single assays. This highlights how metatranscriptomics can uncover mechanistic patterns in organic matter turnover that extend beyond the scope of traditional biochemical assays.

Across other SMAF biological indicators, soil organic carbon (SOC), and potentially mineralizable nitrogen (PMN), values did not differ significantly between CT and ST, yet all trended higher in ST (Supplementary Figure 8 a-d). This limited statistical separation is expected because SOC and PMN reflect broad, aggregate soil properties influenced by mineral associations, residue inputs, and long-term management history. In contrast, transcriptomic profiles capture the specific microbial pathways that contribute to these bulk properties, providing mechanistic insight into carbon and nitrogen transformations that SMAF metrics integrate but do not resolve. In contrast, transcriptomic profiles resolve the specific microbial pathways that contribute to these bulk properties, offering mechanistic insight into carbon and nitrogen transformations that SMAF metrics integrate but cannot disaggregate.

This distinction illustrates a key point: traditional soil health metrics and metatranscriptomics quantify different but complementary dimensions of soil function. SMAF indicators capture ecosystem-level outcomes, whereas transcriptomics identifies the microbial processes that help generate those outcomes. As a result, functional divergence between tillage

systems can be evident at the transcript level even when bulk soil-health indicators appear similar.

Together, these findings demonstrate that integrating multi-omic approaches with soil health frameworks yields a more complete understanding of how tillage influences soil function. SMAF benchmarks offer valuable agronomic context, while metatranscriptomics uncovers the mechanistic underpinnings of carbon and nitrogen cycling. When combined, these tools provide a more actionable and biologically grounded picture of soil health, improving our ability to evaluate and manage agroecosystem sustainability.

While genome and transcriptome resolved comparisons of carbon cycling across tillage systems are scarce, our observations are broadly consistent with previous work showing that reduced disturbance enhances residue retention and aggregate protection, slowing carbon turnover, whereas intensive tillage increases exposure of particulate and mineral-associated organic matter and stimulates enzyme-mediated decomposition[28–30].

2.4 Conclusions

Our study demonstrates that tillage significantly influences microbial community composition and function, with pronounced effects on nitrogen and carbon cycling pathways. By integrating multi-omics approaches with traditional soil health metrics, we identified distinct microbial and functional responses to conventional vs. strip tillage, revealing key differences in ammonia oxidation, carbon fixation, and nitrogen retention mechanisms. Importantly, our findings suggest that strip tillage fosters microbial strategies that promote carbon stabilization and reduce nitrogen loss, supporting its role as a more sustainable soil management practice. This study highlights the need for a function-first approach in soil health assessment, moving

beyond alpha diversity and taxonomic composition to incorporate transcriptional activity and biogeochemical function. By bridging high-resolution microbial insights with soil health frameworks, we can refine soil management practices that enhance long-term sustainability and nutrient efficiency in agroecosystems.

2.5 Materials & Methods

2.5.1 Site description and soil health data

The experimental site was Colorado State University's Agricultural Research Development and Education Center (ARDEC, 40°40'40" N 104°59'51" W, elevation = 1567 m). Soils at the site are dominated by Garrett sandy-loams (fine-loamy, mixed mesic Pachic Argiustoll)[31]. Monthly mean maximum temperatures range from approximately 30°C in July to around 5°C in December. The region receives ~270 mm of annual precipitation, most of which falls between April and October. The long-term tillage experiment was established as a randomized complete block design with two replicate blocks, each containing production-scale CT and ST plots. Each plot measured approximately 320 m in length by 28 m in width and consisted of 36 furrows spaced ~77 cm from bed to crest to crest (Supplementary Fig1b). All plots received equivalent nitrogen inputs in 2021 (~180 kg N ha^{-1}), with slightly different phosphorus rates (CT: ~34 kg P ha^{-1}; ST: 25 kg P ha^{-1}). CT plots received all N and P in early April as a single pre-plant application. ST plots received 38% of their N and all P in early April, followed by the remaining 62% of N in mid-June. The long-term tillage experiment and associated soil health results are described in Trimarco et al[18].

2.5.2 Field campaign and sample collection

Samples for microbial analyses and basic soil chemistry were collected across the 2021 corn growing season. Bulk soil cores were collected in May at the time of emergence (VE), July at the time of tasseling (VT), and September at the time of reproduction (R5) using a 3.12 cm diameter step probe to a depth of 30 cm. Within each tillage treatment (CT and ST), samples were collected from two replicate plots (Block 1 and Block 2), and from two fixed positions within each plot (North and South). At each plot-position, four soil cores were collected from every other crop row, with cores taken from the side of the bed to avoid rhizosphere influence. This yielded a total of 96 bulk cores across the three sampling time points. Cores were immediately sub-sectioned in the field into 96 surface (0-10 cm) and 96 deep (20-30 cm) samples (total of 192 bulk subsamples). Further, four rhizosphere samples per treatment were taken during July (VT) and September (R5) from Block 1 North (Supplementary Figure 2b) to generate 16 samples. For rhizosphere collection, roots were excavated by inserting a shovel to the depth of the blade (~20-25 cm) and loosening the surrounding soil until the root system could be lifted intact. Loose soil was gently shaken off, and the soil adhering firmly to the roots was defined as rhizosphere fraction. In total, 208 individual soil samples were collected. All soil samples were immediately placed on ice in the field following collection, except those intended for metatranscriptomic analyses (n = 8, rhizosphere, Block 1 North) which were flash-frozen in the field using an ethanol and dry ice bath. Flash frozen samples for metatranscriptomics were stored at -80°C in the laboratory, while others were stored at -20°C until analysis.

2.5.3 DNA extraction, 16S rRNA gene and ITS amplicon sequencing

Total DNA was extracted from 0.5 g of soil (n=208) using Zymo Research Quick-DNA fecal/soil Microbe kits (cat #D6010) following the manufacturer's 'soil' protocol. Bacterial/archaeal 16S rRNA (V4 region, primers 515F/806R[32]) and fungal ITS (primers ITS1f/ITS2[33]) genes were PCR amplified following the Earth Microbiome Project (EMP) PCR protocol[34] and then sequenced on the Illumina MiSeq platform using 251 bp paired-end reads at the University of Colorado Boulder Center for Microbial Exploration. Raw amplicon sequence data were processed using the QIIME2 (release qiime2-2021.2) pipeline[35]. Data were trimmed to a length of (248 bp), then reads were demultiplexed then denoised with DADA2[36]. ASV taxonomy was assigned using the naïve Bayes sklearn classifier trained with the GTDB database[37] (GTDB 220 v2.4.0) and the UNITE database[38] (UNITE v9.0) for bacteria/archaea and fungi, respectively. ASVs classified as mitochondria or chloroplast and unclassified ASVs were removed before further analysis. Ecological guilds were assigned to fungal ASVs using FUNGuild[39] (v1.2). Following FUNGuild recommendations, only guild classifications identified as 'highly probable' or 'probable' were accepted to avoid possible overinterpretation.

2.5.4 Metagenomic Sequencing, assembly and binning and annotation

Samples from Block 1-North were chosen across treatments within this field system including six surface, three deep, and three rhizosphere samples. Metagenomic DNA sequencing (n=12) was conducted at the Department of Energy Joint Genome Institute, using the Illumina NovaSeq 6000 platform (Illumina, v.1.5 chemistry, S4 flow cell, 2 × 150 bp). Raw reads were trimmed using sickle[40] (v1.2.9) in paired-end mode, using default parameters (minimum quality score = 20, minimum length = 20) and assuming Sanger-encoded quality scores (-t sanger).

Trimmed reads were then assembled using MEGAHIT[41] (v1.2.9), with all samples assembled individually except for three deeply sequenced samples that were co-assembled to improve contig recovery, and contigs > 2500bp were binned into MAGs using MetaBAT2[42] (v2.12.1) with default settings. JGI-generated MAGs via their metagenome pipeline[43] were downloaded and included in analyses. MAG completeness and contamination were assessed via CheckM[44] (v1.1.2), resulting in 108 medium and high-quality MAGs (MQ-HQ) per MIMAG[20] standards. To construct MAG database, the 108 MAGs generated in this study were dereplicated at 99% identity using dRep[45] (v.2.6.2) alongside 342 MQ-HQ MAGs from the Agricultural Exudate-Responsive Metagenomic (ARM) database[21] resulting in a combined set of 378 MAGs. Beyond MAGs, we also constructed a complementary gene database by calling genes from a concatenated set of contigs greater than 2500 base pairs from individual metagenome assemblies using Prodigal[46] (v.2.6.2) with the -p meta flag to account for fragmented metagenomic assemblies. These genes were then clustered at 95% nucleotide identity and 80% coverage using CD-HIT[47] (v4.6). The cluster representative genes (n = 440,445) were annotated with DRAM (v1.4.4).

To confirm functional annotations of nitrogen cycle genes, we constructed phylogenetic trees using reference sequences and gene database sequences. Gene alignments were performed using MUSCLE[48] v3.8.31, manually curated, and phylogenetic trees were built with ProtPipeliner (https://github.com/WrightonLabCSU/columbia_river/blob/67dff3bd8f6af3323ba75d4b7145dc799e03782d/8_protpipeliner_phylogenetic_trees.py#L4). Phylogenetic trees were generated for key nitrogen cycling enzymes, including ammonia monooxygenase (*amoCAB*) and nitrite

reductases (*narG* and *narH*). Additionally, an AOA MAG tree was generated using GTDB-Tk denovo workflow.

2.5.5 Metatranscriptome sequencing and analysis

Rhizosphere samples (n=8) from Block 1-North, CT and ST, were chosen for metatranscriptomics at the July time point. Total RNA was extracted from the rhizosphere samples, using 0.5g of flash frozen soil with the Zymo Research ZymoBIOMICS RNA Miniprep Kit (cat #R2001). Metatranscriptomic libraries were prepared using the Zymo-Seq RiboFree® Total RNA Library Kit (cat #R3000, #R3003) according to the manufacturer's protocol, which includes a step to deplete ribosomal RNA. Following reverse transcription, cDNA libraries were sequenced at the Genomics Shared Resource Facility at the University of Colorado Anschutz Medical Campus on an Illumina NextSeq2000. Sequencing was performed using a dual-indexed run configuration that employed an 8×8 index scheme with the i5 index read in reverse complement. Raw metatranscriptome reads were trimmed, and adapters were removed using bbduk[49] (v38.90) with the parameters ktrim=r k=23 mink=11 hdist=1 qtrim=rl trimq=20 minlength=75, then sub-sampled to 5.5 Gbp to standardize sequence depth across samples.

The reads were mapped separately using Bowtie2[50] (v2.4.5) (-D 10 -R 2 -N 0 -L 22 -i S,0,2.50) to two reference databases: (1) a dereplicated gene catalog derived from the MAG database, and (2) a gene catalog constructed from the metagenome assemblies. Resulting SAM files were transformed to BAM files using samtools[51] then filtered to 93% ID using reformat.sh[49] and name sorted using samtools. Transcripts were counted for each gene using feature-counts[52]. Counts were transformed to geTMM[53] values (gene length corrected trimmed mean of M-values) in R using edgeR[54] package. For genome level analyses of the MAG database, gene expression

was averaged within a genome. Taxonomy was assigned to MAGs via GTDB-Tk[31] (v2.4.0 r220), and MAGs and genes were annotated using DRAM[55] (v1.4.4) with default settings.

Transcribed genes related to soil health were grouped by function across the gene database identifiers and annotation informed metabolism calls (Supplementary Data 9). First, transcribed genes were classified by enzyme, then by functional process, and finally by biogeochemical group. For example, in carbon, a gene was classified as glycoside hydrolase 3 (GH3), then cellulose, then carbon. In nitrogen, a gene was classified as *amoA*, then nitrification, then nitrogen. This gene hierarchy was employed across the dataset, enabling collation of bulk microbiome function at various levels. For linkage to SMAF-based indicators, E.C. numbers were also profiled across the transcribed gene set. For example, the beta-glucosidase assay pertains to E.C. 3.2.1.21, thus all genes with this annotation could be grouped. These aggregated transcript-level functional groupings are visualized in Supplementary Figure 7**.**

2.5.6 Statistical Analysis

MaAsLin2[56] was used for differential expression analysis for (1) log-transformed geTMM values for MAG expression, (2) log-transformed geTMM values for gene expression and subsequent functional groupings and (3) relative abundance of 16S rRNA and ITS amplicon data. Significance was determined if the FDR was < 0.05. To assess differences in bacterial richness and Shannon diversity index, ANOVA followed by Tukey's HSD was applied for pairwise comparisons. Non-metric multidimensional scaling (NMDS) was conducted on Bray-Curtis dissimilarities using the vegan R package (v 2.7-0) to examine broad differences in microbial communities. Permutational multivariate analysis of variance (PERMANOVA) was

used to test for significant differences in community composition across treatments, depths and time. All reported statistical analyses were performed in R.[57]

Chapter 2 Figures

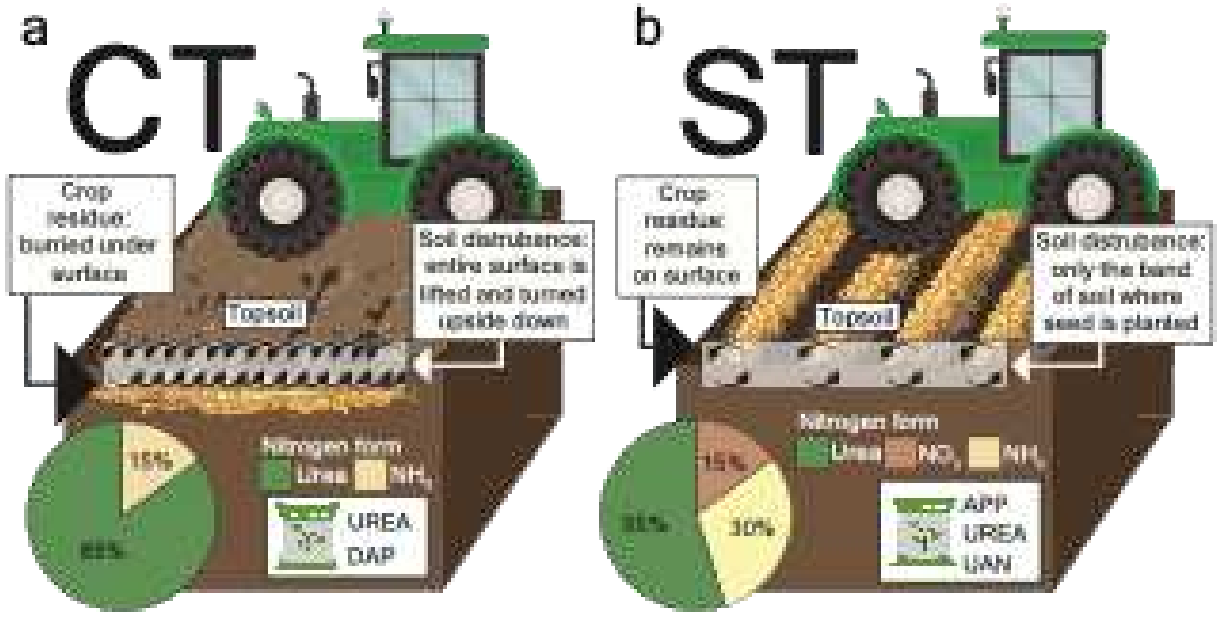


Figure 1. Differences in conventional tillage verses strip tillage management . Overview of the primary differences between conventional tillage (CT) and strip tillage (ST). Treatments received different fertilizer applications. CT was added via broadcast, a mix of solid 46-0-0 Urea and 18-46-0 Diammonium phosphate (DAP). ST received a banded application of consisted of 46-0-0 Urea, 32-0-0 Urea ammonium nitrate (UAN) and 10-34-0 ammonium polyphosphate (APP).

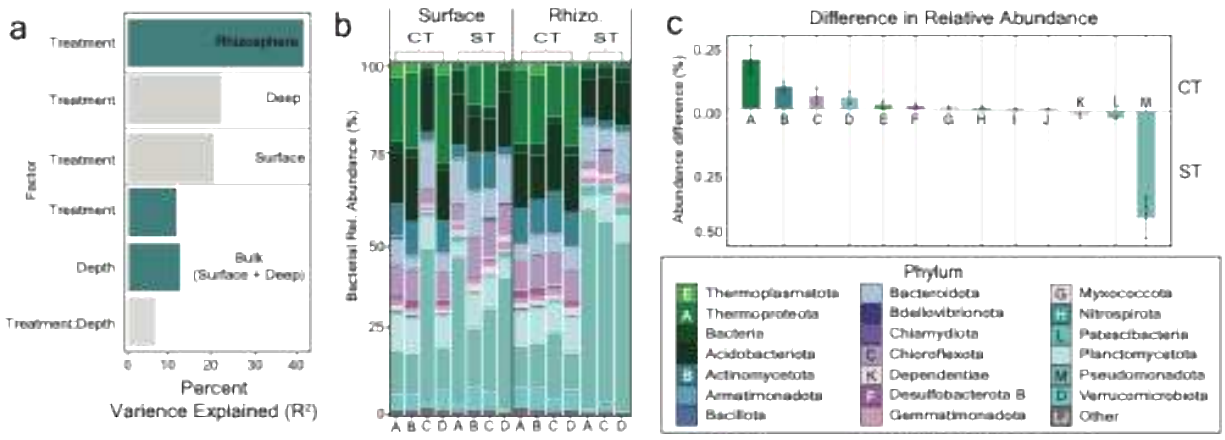


Figure 2. Tillage practice shows the largest microbial shift in the rhizosphere. a Percent variance in bacterial/archaeal ASV community composition explained (R^2) via PERMANOVA per soil compartment (Rhizosphere, Deep, Surface, Bulk) for July timepoint. Significant factors (p-value < 0.05) are shown by blue bars, while grey bars were not significant. **b** Bacterial and archaeal relative abundance from 16S rRNA gene amplicon analysis for July surface and rhizosphere soil samples. **c** *Absolute differences in relative abundance of significantly different taxa between treatments in the July rhizosphere.* Bars represent the mean percentage difference in relative abundance (%), calculated from pairwise Tukey's HSD tests following ANOVA. Only taxa with significant differences ($p < 0.05$) are shown. Error bars indicate 95% confidence intervals. Values are presented as absolute differences to highlight the magnitude of change, regardless of which treatment had higher abundance. Letters show taxonomic identification.

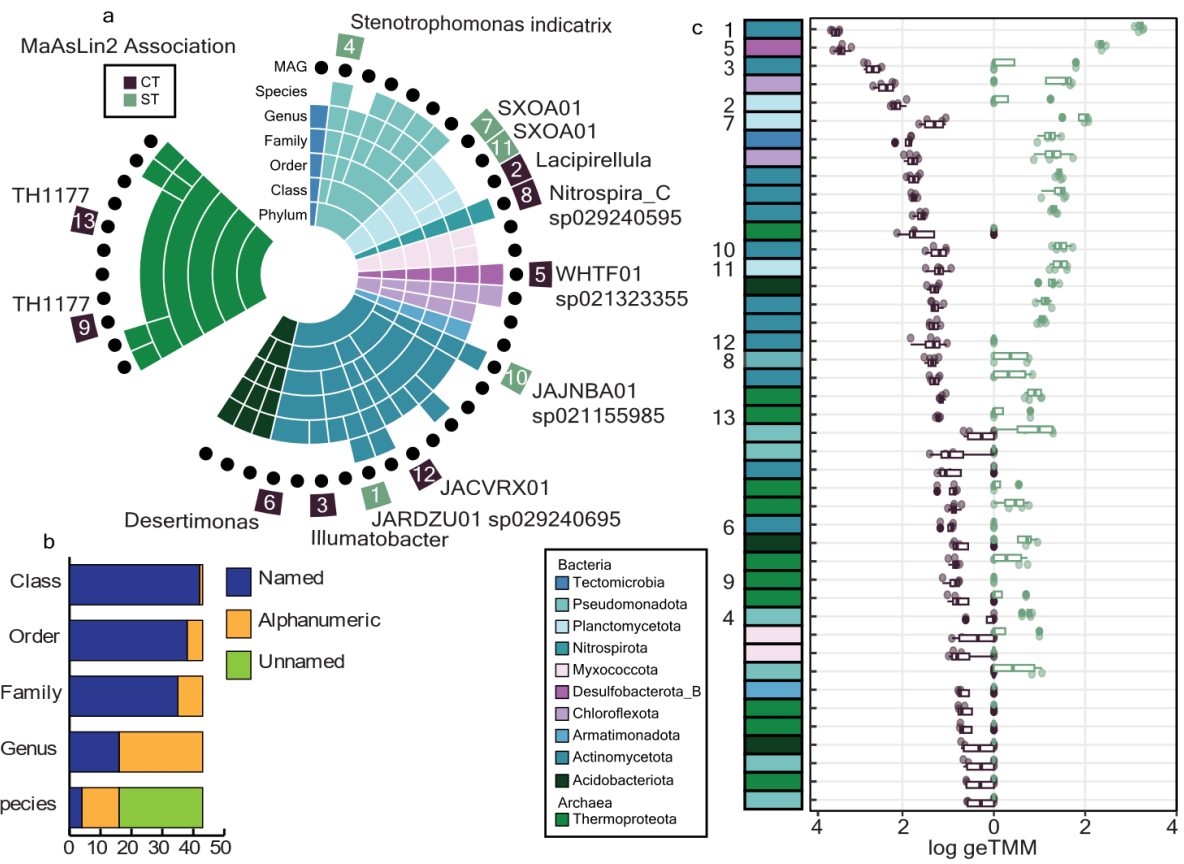


Figure 3. Tillage treatment influences taxonomic composition and transcriptional activity of active MAGs. a Sunburst chart displaying the taxonomy of the 43 metagenome-assembled genomes (MAGs) that recruited metatranscriptome reads. Each ring represents a subsequent level of taxonomy, starting from phylum at the inner ring. Blank levels correspond to unnamed taxonomic levels. Each MAG is marked by a black dot. MAGs that exhibited significant associations with CT or ST management (MaAsLin2, $q<0.05$) are noted in the outer ring, with the lowest named level of taxonomy labelled. Numbers correspond to MAGs in C. **b** Bar chart of the novelty represented in the 43 active MAGs from taxonomic level class and lower. "Named" corresponds to MAGs that were classified with a valid Latin name. "Alphanumeric" corresponds to MAGs that received a classification at that level from GTDB, but the names were alphanumeric. "Unnamed" corresponds to MAGs that were not classified at that level by the GTDB. **c** Butterfly plot of MAG metatranscriptome expression across treatments. MAGs are given along the x-axis, and the y-axis is the log10 transformed MAG transcription (geTMM). The lower and upper boxplot edges represent the 25th and 75th percentiles, respectively, and the middle line is the median. The whiskers extend from the median to 1.5× the interquartile range.

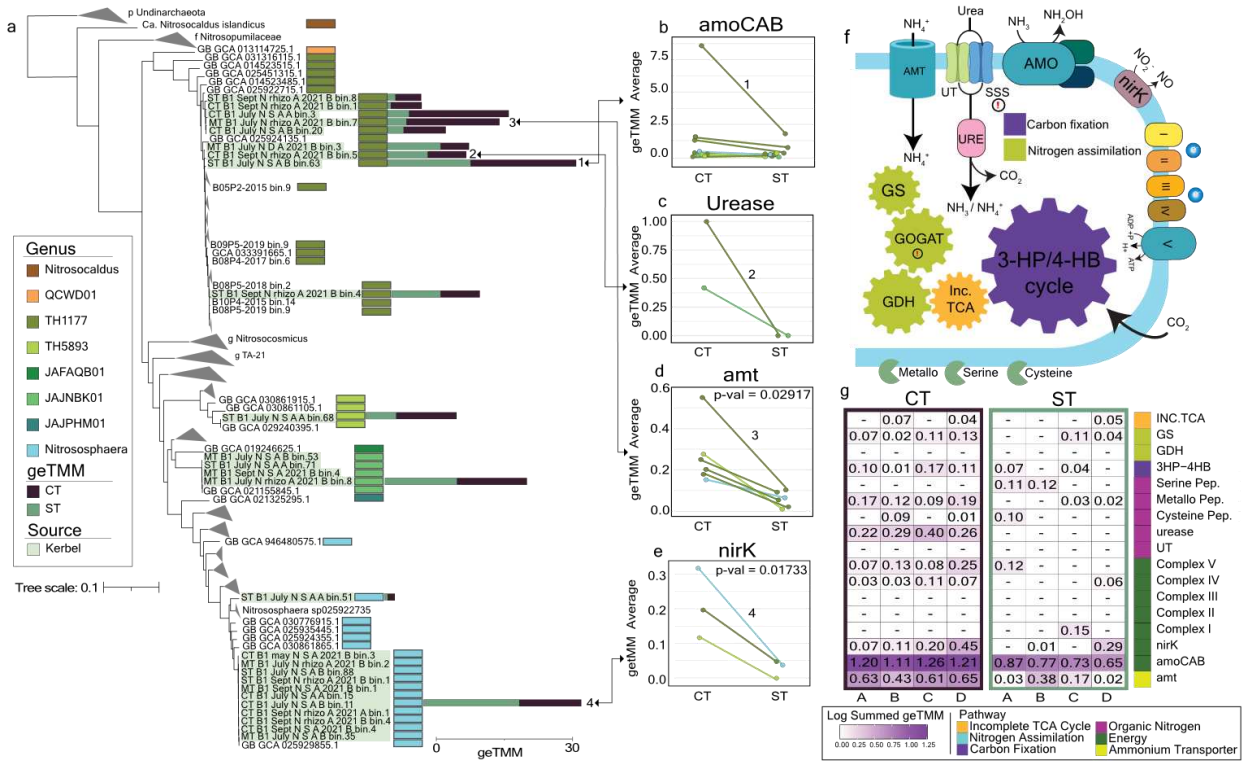


Figure 4. Genome phylogeny and transcriptional activity of ammonia-oxidizing Archaea (AOA) MAGs in agricultural soils. a Maximum likelihood tree of recovered ammonia oxidizing Archaea (AOA) MAGs (n = 26, green) with GTDB and NCBI reference genomes, built from 53 phylogenetic markers using GTDB-Tk and rooted with Undinarchaeota. Bars at tips indicate total geTMM values per treatment recruiting metatranscriptomic reads, with numbers identifying MAGs expressing the highest enzyme levels in b-e. **b-e** Averaged activity of nitrogen-cycling genes by treatment. Each point represents a MAG. The numbers correlate to the MAG with the highest geTMM values. **b** Ammonia monooxygenase (amoCAB), **c** Urease (all subunits), **d** Ammonium transporter (amt, ANOVA p = 0.02917), **e** Nitrite reductase (nirK, ANOVA p = 0.01733). **f** Predicted genome schematic of AOA genus TH1177; absent genes marked with a red exclamation point in a circle. SSS = solute/sodium symporter, GOGAT = glutamate synthase, INC = incomplete TCA cycle, GS = glutamine synthetase, GDH = glutamate dehydrogenase, 3HP-4HB = hydroxypropionate/hydroxybutyrate cycle, UT = urea transporter, I-V = electron transport chain units. **g** Heatmap of expression of selected genes in TH1177 MAGs with geTMM values reported for each gene, Dashes indicate annotated genes with no geTMM value (value = 0).

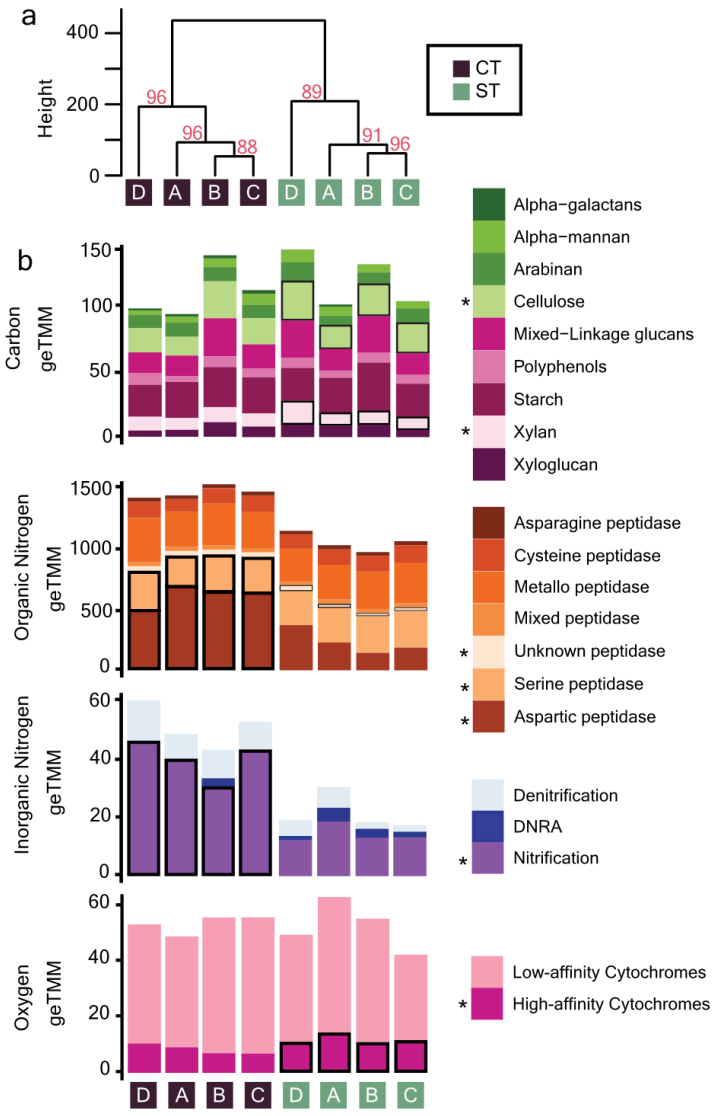


Figure 5. Metatranscriptomes reveal disparities in microbial soil health functions. **a** Hierarchical clustering dendrogram of samples from two treatments, ST (green boxes: D, A, B, C) and CT (dark purple boxes: D, A, B, C). The dendrogram shows the hierarchical relationships between samples based on their dissimilarity. Numbers in red indicate Approximately Unbiased (AU) p-values derived from multiscale bootstrap resampling. Clusters with AU p-values ≥ 95 are statistically significant. **b** Stacked bar plots show the summed transcription values (geTMM) for genes relevant to soil health. Significance (FDR ≤ 0.05, MaAsLin2) of each gene is indicated by a black box on the bar plot and an asterisk next to the gene name.

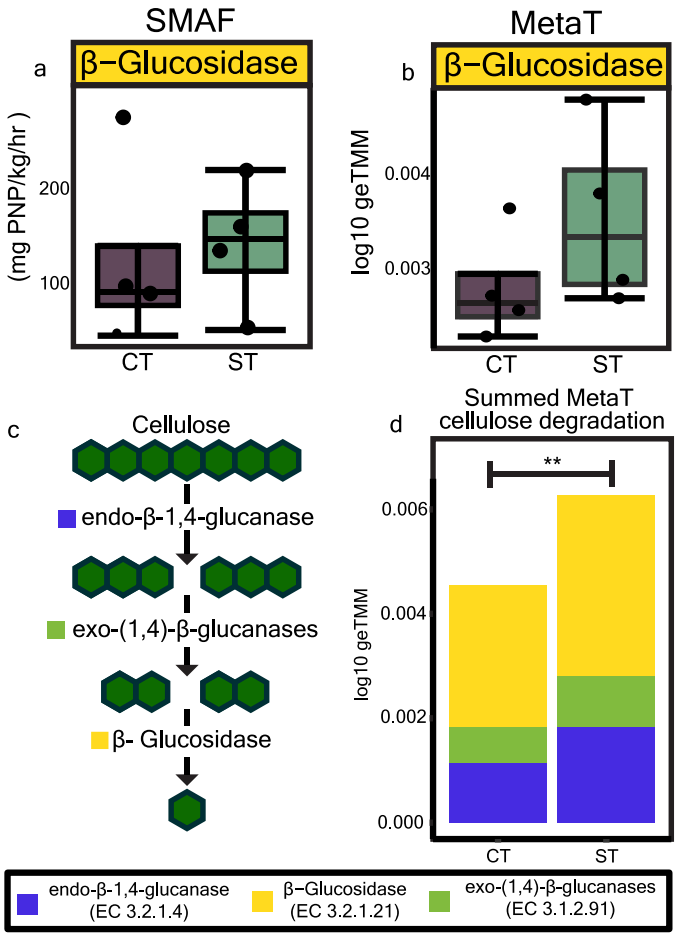


Figure 6. Transcripts provide deeper insight to traditional soil health metrics a Beta-glucosidase enzyme activity (mg PNP/kg/hr) used as a biological indicator in the Soil Management Assessment Framework (SMAF). Boxplots display the median (horizontal line), interquartile range (IQR; the box), and whiskers extending to 1.5×IQR from the quartiles; outliers are plotted individually. **b** Metatranscript total (log10 geTMM value) of genes annotated as beta-glucosidase under the EC number 3.2.1.21. No significance was found for a or c. **c** Diagram showing steps in cellulose degradation. **d** Stacked bar plots of metatranscript recruitment of genes involved in cellulose degradation. (ANOVA, p-value <0.05).

Chapter 2 References

1. Hossain, A. *et al.* Agricultural Land Degradation: Processes and Problems Undermining Future Food Security. in *Environment, Climate, Plant and Vegetation Growth* (eds Fahad, S. et al.) 17–61 (Springer International Publishing, Cham, 2020). doi:10.1007/978-3-030-49732-3_2.
2. Doran, J. W. Soil health and global sustainability: translating science into practice. *Agric. Ecosyst. Environ.* **88**, 119–127 (2002).
3. Lehmann, J., Bossio, D. A., Kögel-Knabner, I. & Rillig, M. C. The concept and future prospects of soil health. *Nat. Rev. Earth Environ.* **1**, 544–553 (2020).
4. Towards a definition of soil health. https://openknowledge.fao.org/items/c9075777-81b6-45cf-80bb-e1a8582fdf44.
5. Teng, J. *et al.* Conservation agriculture improves soil health and sustains crop yields after long-term warming. *Nat. Commun.* **15**, 8785 (2024).
6. Soil Health | Natural Resources Conservation Service. https://www.nrcs.usda.gov/conservation-basics/natural-resource-concerns/soils/soil-health.
7. Resource Display :: Conservation Technology Information Center. https://www.ctic.org/resource_display/?id=322&title=Tillage+Type+Definitions.
8. Nunes, M. R. *et al.* The soil health assessment protocol and evaluation applied to soil organic carbon. *Soil Sci. Soc. Am. J.* **85**, 1196–1213 (2021).
9. Allen, D. E., Singh, B. P. & Dalal, R. C. Soil Health Indicators Under Climate Change: A Review of Current Knowledge. in *Soil Health and Climate Change* (eds Singh, B. P., Cowie, A. L. & Chan, K. Y.) 25–45 (Springer, Berlin, Heidelberg, 2011). doi:10.1007/978-3-642-20256-8_2.

10. Fierer, N., Wood, S. A. & Bueno de Mesquita, C. P. How microbes can, and cannot, be used to assess soil health. *Soil Biol. Biochem.* **153**, 108111 (2021).
11. Bhaduri, D. *et al.* A review on effective soil health bio-indicators for ecosystem restoration and sustainability. *Front. Microbiol.* **13**, (2022).
12. Dubey, A. *et al.* Soil microbiome: a key player for conservation of soil health under changing climate. *Biodivers. Conserv.* **28**, 2405–2429 (2019).
13. Wilhelm, R. C., van Es, H. M. & Buckley, D. H. Predicting measures of soil health using the microbiome and supervised machine learning. *Soil Biol. Biochem.* **164**, 108472 (2022).
14. Trivedi, P., Delgado-Baquerizo, M., Anderson, I. C. & Singh, B. K. Response of Soil Properties and Microbial Communities to Agriculture: Implications for Primary Productivity and Soil Health Indicators. *Front. Plant Sci.* **7**, (2016).
15. Jia, J. *et al.* Unlocking soil health: Are microbial functional genes effective indicators? *Soil Biol. Biochem.* **204**, 109768 (2025).
16. Zhang, L. & Lv, J. Metagenomic analysis of microbial community and function reveals the response of soil respiration to the conversion of cropland to plantations in the Loess Plateau of China. *Glob. Ecol. Conserv.* **23**, e01067 (2020).
17. Sharma, P. K. *et al.* Comparative metatranscriptome analysis revealed broad response of microbial communities in two soil types, agriculture versus organic soil. *J. Genet. Eng. Biotechnol.* **17**, 6 (2019).
18. Tad Trimarco *et al.* Conservation Tillage Under Furrow Irrigation Improves Health. (2025).
19. Deleon, E., Bauder, T. A., Wardle, E. & Fonte, S. J. Conservation tillage supports soil macrofauna communities, infiltration, and farm profits in an irrigated maize-based cropping system of Colorado. *Soil Sci. Soc. Am. J.* **84**, 1943–1956 (2020).

20. Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**, 725–731 (2017).
21. Seitz, V. A. *et al.* Cover crop root exudates impact soil microbiome functional trajectories in agricultural soils. *Microbiome* **12**, 183 (2024).
22. Gáspár, M., Kálmán, G. & Réczey, K. Corn fiber as a raw material for hemicellulose and ethanol production. *Process Biochem.* **42**, 1135–1139 (2007).
23. Nguyen, T. T. H., Myrold, D. D. & Mueller, R. S. Distributions of Extracellular Peptidases Across Prokaryotic Genomes Reflect Phylogeny and Habitat. *Front. Microbiol.* **10**, (2019).
24. Bei, S., Li, X., Kuyper, T. W., Chadwick, D. R. & Zhang, J. Nitrogen availability mediates the priming effect of soil organic matter by preferentially altering the straw carbon-assimilating microbial community. *Sci. Total Environ.* **815**, 152882 (2022).
25. Wyngaard, N., Franklin, D. H., Habteselassie, M. Y., Mundepi, A. & Cabrera, M. L. Legacy Effect of Fertilization and Tillage Systems on Nitrogen Mineralization and Microbial Communities. *Soil Sci. Soc. Am. J.* **80**, 1262–1271 (2016).
26. Mganga, K. Z., Rolando, J. L., Kalu, S., Biasi, C. & Karhu, K. Priming effect depending on land use and soil types in a typical semi-arid landscape in Kenya. *Biogeochemistry* **163**, 49–63 (2023).
27. Pitcher, R. S. & Watmough, N. J. The bacterial cytochrome *cbb*3 oxidases. *Biochim. Biophys. Acta BBA - Bioenerg.* **1655**, 388–399 (2004).
28. Just, C. *et al.* Soil organic carbon sequestration in agricultural long-term field experiments as derived from particulate and mineral-associated organic matter. *Geoderma* **434**, 116472 (2023).

29. Blanco-Canqui, H. & Lal, R. Soil structure and organic carbon relationships following 10 years of wheat straw management in no-till. *Soil Tillage Res.* **95**, 240–254 (2007).
30. Six, J., Elliott, E. T. & Paustian, K. Soil macroaggregate turnover and microaggregate formation: a mechanism for C sequestration under no-tillage agriculture. *Soil Biol. Biochem.* **32**, 2099–2103 (2000).
31. U.S. Department of Agriculture, Natural Resources Conservation Service. Soil Survey. *USDA Natural Resources Conservation Services* http://websoilsurvey.sc.egov.usda.gov/ (2025).
32. Apprill, A., McNally, S., Parsons, R. & Weber, L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* **75**, 129–137 (2015).
33. Bokulich, N. A. & Mills, D. A. Improved Selection of Internal Transcribed Spacer-Specific Primers Enables Quantitative, Ultra-High-Throughput Profiling of Fungal Communities. *Appl. Environ. Microbiol.* **79**, 2519–2526 (2013).
34. 16S Illumina Amplicon Protocol : earthmicrobiome. https://earthmicrobiome.org/protocols-and-standards/16s/.
35. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
36. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
37. Parks, D. H. *et al.* GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **50**, D785–D794 (2022).

38. Abarenkov, K. *et al.* The UNITE database for molecular identification and taxonomic communication of fungi and other eukaryotes: sequences, taxa and classifications reconsidered. https://dx.doi.org/10.1093/nar/gkad1039.

39. Nguyen, N. H. *et al.* FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol.* **20**, 241–248 (2016).

40. NA najoshi & Frass, JN. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. (2025).

41. 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 (2015).

42. Kang, D. D. *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).

43. Huntemann, M. *et al.* The standard operating procedure of the DOE-JGI Metagenome Annotation Pipeline (MAP v.4). *Stand. Genomic Sci.* **11**, 17 (2016).

44. 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 (2015).

45. 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. *ISME J.* **11**, 2864–2868 (2017).

46. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).

47. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
48. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
49. BBMap. *SourceForge* https://sourceforge.net/projects/bbmap/ (2025).
50. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
51. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
52. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma. Oxf. Engl.* **30**, 923–930 (2014).
53. Smid, M. *et al.* Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. *BMC Bioinformatics* **19**, 236 (2018).
54. 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 (2010).
55. Shaffer, M. *et al.* DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* **48**, 8883–8900 (2020).
56. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLOS Comput. Biol.* **17**, e1009442 (2021).
57. R Core Team. R: A Language and Environment for Statistical Computing. https://www.r-project.org/ (2024).

Chapter 3 Successional dynamics and community variation in rooftop agrivoltaics

3.1 Summary

Rooftop agrivoltaic systems integrate food and energy production in urban environments, yet their belowground ecology remains poorly understood. Because newly installed engineered substrates lack soil legacies, understanding how microbial communities assemble is critical for predicting early ecosystem function and resilience. This study tested three ecological mechanisms hypothesized to shape microbial assembly in newly constructed rooftop agrivoltaic engineered substrates: abiotic filtering imposed by rooftop microclimates and engineered substrates, functional compensation by stress-tolerant taxa, and temporary decoupling between plant productivity and microbial diversity. Across two growing seasons, microbial communities transitioned from low-diversity, greenhouse-influenced assemblages to more complex and differentiated communities that increasingly resembled those of mature rooftop engineered substrates. Diversity increased most strongly in rhizospheres, while phylogenetic turnover (βNTI) revealed a shift from deterministic, substrate-driven control toward mixed stochastic assembly. Enrichment of spore-forming *Firmicutes* and metabolically versatile *Actinobacteriota* indicated early functional compensation under environmental constraints. Together, these results demonstrate that rooftop engineered substrates rapidly develop biologically active communities through environmental filtering, microbial adaptation, and early ecosystem function in engineered urban agroecosystems.

3.2 Introduction: Urban rooftops as engineered microbial habitats

Urban rooftop agrivoltaic systems (RAV), which integrate food production with solar energy capture, are emerging as multifunctional green infrastructure for addressing land scarcity, climate resilience, and supporting sustainability initiatives, including expanded access to fresh food in cities.[1,2,3] These systems offer clear aboveground benefits, including enhancing food production capacity, creating critical habitats for pollinators, moderating building temperatures through insulation and evapotranspiration, and generating renewable energy via photovoltaic panels.[4] The growing medium in these systems typically consists of engineered substrates that are lightweight artificial soil-like blends designed for rooftop use, rather than native field soils. Yet, the engineered substrates and the microbial communities that establish within them remain poorly characterized, despite their critical role in supporting these functions.

Green roofs are engineered ecosystems designed to support plant growth on building rooftops, primarily in urban environments. They are typically categorized as extensive (defined by shallow substrate depth (<15cm), low-maintenance vegetation such as sedums or grasses) or intensive (defined by deeper substrates that support a wider variety of plant types including shrubs, perennials, and even small trees).[5] These systems, both extensive and intensive, offer numerous environmental benefits, though to varying degrees, including stormwater retention[6], insulation[7], urban heat mitigation[8], and provision of urban pollinator habitat and other beneficial organisms.[9] Despite these well-documented advantages, the underlying engineered substrates remain poorly characterized biologically, with limited understanding of how microbial communities influence substrate function and plant performance. In natural soils, microbes drive nutrient mineralization, stabilize organic matter, form aggregates, mediate plant nutrient acquisition, and anchor trophic interactions through well-established networks.[10] In engineered

substrates, many of these processes are newly initiated and may operate differently due to low organic matter, minimal horizon development, and a lack of legacy carbon pools, potentially shifting microbial function toward stress tolerance, opportunistic nutrient capture, and early successional pathways rather than mature soil biogeochemistry.

From a microbial perspective, green roof engineered substrates impose unique constraints on establishment and persistence. These substrates are composed of lightweight, low-organic-matter mixtures of expanded shale, perlite, and sand, sometimes blended with compost[11], optimized for drainage and load-bearing capacity rather than biological function. Thus, they lack the mineral weathering profiles that contribute to silt and clay components, horizon development, and legacy organic matter that underpin long-term biological activity in pedogenic soils. Additionally, newly installed systems are often sterilized to remove weed seeds, resulting in a biologically simplified substrate that lacks an established microbial community and must rely on slow natural colonization from the surrounding environment. This limited inoculation pool constrains both microbial diversity and early functional capacity, delaying the development of a stable food web and nutrient cycling network, which often requires additional microbial inoculants or organic amendments.[12,13] Moreover, rooftop environments expose these substrates to extreme conditions like high winds, temperature extremes, rapid wet–dry cycles, and ultraviolet radiation that further challenge microbial establishment. As a result, microbial communities in green roof systems are typically less diverse and functionally constrained in early stages compared to agricultural or natural soils, and strongly shaped by habitat design.[14]

Agrivoltaics refers to the co-location of crop production and photovoltaic (PV) panels on the same piece of land. Originally conceptualized as a solution to land-use competition between food and energy production[15], this system offers an elegant and sustainable approach to

maximizing productivity in a changing climate. Multiple benefits have been reported in traditional, at-grade agrivoltaic systems, including increased farm value compared to conventional agriculture[3], higher yields in selected shade-tolerant crops[16,17], and improved water use efficiency[18] which is a particularly important outcome in drought-prone regions such as the western United States. The benefits do not stop above ground. The microclimates created under PV panels have been associated with increased soil enzyme activity[19], stable fungal networks[20], and enrichment of microbial taxa associated with roles in carbon and nutrient cycling.[21] Together, this shows potential synergies between PV infrastructure and belowground microbiome ecological function.

As rooftop agrivoltaic (RAV) systems expand to meet sustainability and climate resilience goals, it becomes increasingly important to understand how dual-use designs influence microbial community assembly, nutrient cycling, and long-term substrate function. Evidence from green roofs shows that engineered substrates experience greater physical stress and biological limitation than natural soils, yet in many cases sustain productive vegetation, while studies of at-grade agrivoltaic systems demonstrate that photovoltaic shading creates microhabitats with distinct temperature and moisture regimes, altering the microbiome composition and potentially biogeochemical cycling. As a result, we draw on findings from two related systems, green roofs and ground-based (at grade) agrivoltaics, to provide context for our study. These comparisons help infer potential microbial succession, functional roles, and highlight knowledge gaps specific to rooftop agrivoltaic substrates. Understanding these dynamics can help design dual-use systems that optimize both ecological and agricultural performance in urban environments.

In conventional soils, strong plant-microbe feedbacks tightly couple aboveground productivity to belowground biological complexity. Plants release root exudates that stimulate microbial guild formation, accelerate nutrient cycling, and build trophic interactions within the rhizosphere; these processes, in turn, influence plant performance. In newly constructed rooftop substrates, however, microbial communities assemble from a limited inoculation pool and initial organic matter reserves are low, reducing carbon exchange between plants and microbes. As a result, early plant productivity is often supported directly through nutrient inputs rather than microbially mediated cycling, while microbial networks develop gradually through dispersal-limited colonization and succession. Under these conditions, plant growth can increase more rapidly than microbial community complexity, creating a temporary decoupling between above- and belowground dynamics until stable trophic interactions and nutrient cycling networks emerge.[12,13]

Together, insights from these separate green roof and agrivoltaic systems suggest three likely, interacting mechanisms that may shape microbial structure and function in dual-use rooftop agrivoltaics: (1) weak plant-microbe feedbacks, leading to decoupling between plant productivity and belowground biological complexity in engineered RAV substrates, (2) abiotic filtering imposed by the heterogeneous microhabitats created within the RAV environment that promote convergence toward stress-tolerant microbial assemblages, and (3) functional compensation by specialized or stress-adapted microbes that sustain key ecosystem processes despite reduced community diversity.

This study represents the first detailed investigation of rooftop agrivoltaic microbiomes, bridging soil ecology, green roof design, and agrivoltaic management. We characterize microbial diversity, community assembly processes, and successional trajectories in engineered rooftop

substrates, and interpret these patterns through three ecological lenses motivated by theory and prior observations: (1) the strength of early plant-microbe feedbacks, (2) abiotic filtering by rooftop microhabitats, and (3) functional compensation by stress-adapted taxa. These mechanisms are central to determining how rapidly engineered substrates develop functional soil properties and how resilient nutrient cycling networks emerge under rooftop stress, which are core considerations for designing reliable, multifunctional urban agrivoltaic systems. By linking microbial ecology to urban sustainability frameworks, our work provides a foundation for managing rooftop agroecosystems that unite food production, renewable energy, and ecological performance within a single engineered environment.

3.3 Results & Discussion

3.3.1 Physicochemical gradients that shape microbial assembly in rooftop substrates

Microclimatic differences among rooftop agrivoltaic (RAV) treatments define the physical environment in which microbial communities assemble. To characterize this abiotic context, we quantified variation in substrate temperature, incoming light under photovoltaic structures, and soil moisture across bifacial, opaque, and full-sun beds over two growing seasons. Full-sun treatments received unmodified solar radiation, bifacial panels created intermediate shading, and opaque panels produced the highest shade intensity. These structural contrasts generated consistent differences in substrate heating, light availability, and moisture retention within each growing season, creating distinct microhabitats that are expected to influence microbial establishment, dispersal, and successional trajectories in newly installed engineered substrates.

We first quantified substrate temperature patterns across treatments, as temperature is a key microclimatic factor shaping microbial establishment in newly installed engineered

substrates. Substrate temperatures aligned with shading intensity. In 2023, mean temperatures were broadly similar among treatments (19.7–21.0 °C) (Table 1), but thermal contrasts emerged during daily extremes: the full sun plot reached 38.2 °C, while shaded plots (bifacial and opaque) peaked near 33–34 °C. Full sun also showed greater diurnal variability (SD = 7.49 °C) compared to opaque and bifacial treatments (~6.6 °C) (Table 1).

In 2024, seasonal conditions were warmer overall, and treatment differences became more pronounced. The full-sun plot averaged 25.0 °C, approximately 3–4 °C higher than opaque (21.9 °C) and bifacial (20.9 °C) plots (Table 1). Variability also diverged by treatment, with shaded plots showing reduced variation (SD = 3.12–3.30 °C) relative to full sun (SD = 5.21 °C). Maximum temperatures again separated by treatment, with full sun reaching 39.4 °C versus 32–36 °C under shade (Table 1). These differences reflect both panel-driven microclimate effects and interannual climatic variation, rather than a continuous temporal trend tied to substrate development.

Together, these data show that panel shading consistently reduces substrate heating and daily thermal fluctuations, creating cooler and more thermally buffered microhabitats relative to full sun, roof top conditions. Short-term thermal environments are ecologically meaningful because microbes experience temperature as a distribution of daily extremes, not as seasonal means.[22] Unshaded substrates regularly exceeded 35 °C, favoring microbial traits that support persistence under rapid heating and drying, including spore formation, dormancy, protective pigments, and membrane modifications. In contrast, reduced variability under opaque and bifacial panels provides buffered conditions that support continuous metabolic activity rather than boom-bust cycles driven by thermal stress. Over time, these patterns act as abiotic filters: extreme temperatures constrain the establishment of heat-sensitive taxa, while shaded conditions

promote the assembly of communities with stress-tolerant life-history strategies typical of early successional environments. In subsequent sections, we examine how these temperature differences interact with moisture dynamics and rhizosphere geochemistry to structure the pace and direction of microbial community development in rooftop agrivoltaic substrates.

Solar radiation patterns reflected the intended shading gradient across treatments, with full-sun plots receiving an order of magnitude higher incident light than beds located beneath bifacial or opaque photovoltaic panels. Across both growing seasons, the full-sun plots experienced mean radiation levels of ~240 W m^{-2}, compared to ~22–32 W m^{-2} in shaded plots (Table 2). These differences were stable across years, even as interannual weather variation produced minor shifts in mean values. Bifacial plots received slightly higher mean values than opaque treatments, consistent with partial light transmission and edge scattering from bifacial panel design, but these differences were small relative to full-sun exposure and were not statistically significant. As a result, both shaded treatments effectively represented a common low-radiation microclimate. Minimum radiation values were uniformly near zero across treatments due to diurnal cycles, while maximum values ranged from ~500–650 W m^{-2} under shaded panels to >1,200 W m^{-2} under full sun, illustrating the magnitude of daily peaks that distinguish these microclimates.

These light environments constitute a primary abiotic filter for microbial assembly, both indirectly through plant responses and directly through substrate temperature and moisture dynamics.[23] Lower radiation under shaded treatments reduces surface heating and slows evaporation, creating cooler and more moisture-stable microhabitats that favor continuous microbial metabolism rather than heat-driven boom-bust activity.[24] In contrast, full sun conditions expose the substrate to intense midday radiation, producing rapid warming and drying

cycles that can select for microbes with protective traits such as dormancy, pigment production, and stress-responsive metabolism.[25–27] Because engineered substrates initially lack legacy carbon pools and aggregate structure, early microbial food webs depend heavily on root-derived carbon inputs.[28,29] Therefore, shading-induced changes in root exudation, photosynthetic rate, and rhizosphere moisture control the pace at which microbial communities can increase in diversity and initiate nutrient cycling.[30] In subsequent sections, we examine how the interaction between incident light, substrate temperature, and nutrient availability shapes rhizosphere geochemistry and microbial succession trajectories in rooftop agrivoltaic substrates.

Substrate moisture, measured as volumetric water content ($m^3\ m^{-3}$), varied systematically with shading intensity across treatments (Table 3). In both growing seasons, shaded plots beneath opaque and bifacial panels maintained higher mean moisture than full sun plots. In 2023, mean moisture ranged from 0.062 $m^3\ m^{-3}$ in full sun to 0.117 $m^3\ m^{-3}$ in opaque treatments, with bifacial beds at an intermediate level (0.080 $m^3\ m^{-3}$). This pattern persisted into 2024, with opaque substrates again showing the highest average moisture (0.131 $m^3\ m^{-3}$), followed by bifacial (0.087 $m^3\ m^{-3}$) and full sun (0.094 $m^3\ m^{-3}$). Variability also differed by treatment: shaded substrates exhibited lower standard deviations across both years, indicating more stable moisture conditions through time, whereas full-sun substrates showed greater temporal fluctuations around the mean. Negative minimum values occasionally recorded by sensors likely reflect calibration drift or reduced soil-sensor contact under very dry conditions, which is common in coarse, low-organic-matter media.

These moisture contrasts represent a key abiotic filter that shapes early microbial assembly in engineered substrates. More frequent low-moisture conditions in full sun environments favor microbes with traits supporting survival during desiccation, including

dormancy, osmolyte production, and extracellular protective compounds, which are characteristic of early successional strategies in newly formed soils. In contrast, the more persistent moisture availability observed under opaque and bifacial panels provides longer periods for active metabolism, allowing microbial taxa to utilize root-derived carbon and establish trophic interactions rather than cycling rapidly between growth and dormancy. Because rooftop substrates can lack legacy carbon pools and stable aggregate structure, the duration and stability of moderate moisture conditions are particularly influential for determining which microbial guilds can establish and initiate nutrient cycling processes during initial ecosystem development.

Together, these physicochemical gradients define a predictable hierarchy of abiotic filters shaping microbial assembly in rooftop agrivoltaic substrates. Full sun plots combine intense radiation, higher substrate temperatures, and more frequent desiccation events, generating a selective environment that favors stress tolerant, opportunistic taxa capable of persisting through rapid boom-bust metabolic cycles. In contrast, shaded plots beneath opaque and bifacial panels experience reduced thermal stress, lower evaporative demand, and more stable moisture availability, creating buffered microsites that support continuous growth and earlier formation of trophic interactions. Because engineered substrates lack legacy organic matter and aggregate structure, early community development depends on the interaction between these microclimate filters and plant derived carbon inputs. In the following section, we evaluate how these contrasting microhabitats influence rhizosphere geochemistry and the successional dynamics of microbial communities during the first two years of rooftop agrivoltaic establishment.

3.3.2 Rhizosphere geochemistry reflects microclimate driven differences in substrate development

Rhizosphere geochemical properties showed early signs of substrate development between 2023 and 2024, but overall carbon and nutrient pools remained low and frequently declined over time. Total carbon, total nitrogen, and organic matter remained low in absolute terms, reflecting the absence of legacy pools, and in most cases declined from 2023 to 2024, with the largest declines evident in full sun plots. (Figure 1). This pattern is consistent with rapid consumption and leaching of amendment derived carbon and nitrogen in a highly drained, low buffer substrate, where plant and microbial demand exceed the rate at which new organic inputs are stabilized.[31,32] Shaded treatments beneath opaque and bifacial panels exhibited smaller year-to-year shifts, suggesting that cooler, moister microclimates modestly slow the loss of labile carbon and nutrients during early rooftop substrate development.

Plant available nutrient pools (P, K, and NO_3^-) were low across all treatments and years (Figure 1). Both available phosphorus and potassium declined from 2023 to 2024 in all treatments, indicating a uniform downward trend independent of shading. These decreases were modest in magnitude and likely reflect rapid utilization or leaching in the coarse, low cation exchange substrate rather than microclimate driven differences. Nitrate exhibited a less consistent pattern: concentrations increased slightly in the opaque plot, remained the same in the bifacial plot and slightly decreased in the full sun plot in 2024, although overall values remained low. Overall, the patterns indicate that shading did not strongly influence plant available nutrient pools, instead, nutrient dynamics were primarily governed by the inherent properties of the engineered substrate, which has limited capacity to retain soluble ions and is prone to leaching and rapid biological uptake.[32]

Soil chemistry also shifted between years, with the most pronounced change occurring in pH (Figure 1). In both opaque and full sun plots, rhizosphere pH increased from ~6.6–7.0 in 2023 to ~7.6–7.8 in 2024 while in the bifacial plot the pH showed no change. In contrast, electrical conductivity (EC), which reflects soluble ion concentrations, decreased from 2023 to 2024 in all treatments. This divergence of a higher pH but lower EC, indicates that the pH increase was likely not driven by an accumulation of soluble salts. Such a pattern is consistent with progressive depletion of acidic functional groups (e.g., organic acids) as initial compost-derived organic matter is mineralized and leached from the system, coupled with the inherently low buffering capacity of lightweight engineered substrates. As organic acids decline and cation exchange sites become depleted, substrate pH tends to drift upward toward the alkalinity of irrigation water or carbonate-containing minerals in the substrate mix, both well documented behavior in green roof media.[32,33] Declining EC during the same period suggests continued leaching of soluble ions rather than nutrient enrichment.

Despite clear geochemical shifts between 2023 and 2024, these geochemical patterns did not translate into detectable microbiome linkages at the community level. Environmental fitting of substrate geochemical properties onto rhizosphere PCoA ordinations did not yield significant associations, suggesting that early microbial community structure was not strongly driven by these variables. Nonetheless, the geochemical trends provide important context for interpreting microbial dynamics: losses of organic matter, total C, and total N indicate diminishing resource availability for heterotrophs, and the upward drift in pH coupled with declining EC suggests a shift toward more oligotrophic, weakly buffered conditions. Collectively, these changes point to an early-stage rooftop rhizosphere in which microbial communities are likely constrained by carbon and nutrient scarcity, weakly differentiated by shading treatment, and still in the process

of responding to the emerging microclimates created by the photovoltaic structures. As substrate development continues and organic inputs accumulate over time, stronger microclimate–microbial linkages may emerge.

3.3.3 Rhizosphere communities show increases in alpha diversity across treatments and time

Microbial diversity provides a complementary measure of system development to plant productivity, reflecting how biological complexity builds in parallel with early rooftop agrivoltaic crop function. Across the two growing seasons, α-diversity, measured by ASV richness and Shannon diversity, increased significantly from 2023 to 2024 in all rooftop treatments (Figure 2). The most pronounced gains occurred in rhizosphere samples, where richness nearly doubled in some treatments (e.g., Opaque) and Shannon diversity rose significantly (Dunn's test with Benjamini–Hochberg correction, $p < 0.05$). Bulk substrates at planting and harvest exhibited more modest changes, suggesting that rhizosphere-root interactions were the principal driver of microbial diversification. This interpretation is supported by the observation of the greenhouse soil, from which transplants originated, had the lowest richness compared to all other sample types (Figure 2).

These temporal trends mirror patterns observed in other constructed ecosystems such as green roofs, where microbial richness tends to rise with substrate age and plant establishment.[34,35] The increase in diversity likely reflects both substrate development and expanding recruitment niches driven by root exudation, organic residue accumulation, and moderated microclimates under vegetation cover. Early in succession (2023), rhizosphere communities exhibited relatively low richness and evenness across shading treatments,

indicating strong environmental filtering and limited colonization potential. By 2024, these communities became more even and taxonomically diverse, consistent with a transition from abiotic to biotic control as substrates matured. [34,35]

The annual planting regime may also have contributed to these diversity gains. Pepper plants were re-transplanted from the greenhouse into each rooftop plot at the start of each season, and at harvest plants were destructively sampled by cutting stems at the substrate surface while leaving roots in place. Residual root-associated taxa and microbial residues from the prior season may have enriched the 2024 planting bulk substrates, providing a more diverse starting point for subsequent community assembly.

3.3.4 Community composition diverges by year and converges toward mature substrates

Building on the alpha diversity trends, we also examined how microbial community composition changed over time to evaluate whether increased richness corresponded with directional assembly and environmental stabilization. β-diversity analyses revealed distinct shifts in community structure between 2023 and 2024 across rooftop agrivoltaic plots (Figure 3a). Principal Coordinates Analysis (PCoA) based on Bray–Curtis dissimilarity showed clear temporal structuring, with samples from each year forming distinct clusters. Communities in 2024 appeared more cohesive and directionally displaced from their 2023 counterparts, which were more dispersed, suggesting that rooftop microbiomes are undergoing non-random, directional changes consistent with ecological succession. These temporal patterns were further supported by ANOSIM results ($R = 0.488$, $p = 0.001$), indicating moderate to strong separation in microbial community structure between years. Importantly, this separation was observed

across all sample types and locations, indicating microbial assembly driven by consistent environmental filtering.

This temporal trajectory suggests that roof top engineered substates follow a successional pattern in which communities diverge rapidly from their greenhouse inoculum and gradually converge toward the structure of established substrates. The tighter clustering of 2024 samples (regardless of the age of the material) relative to 2023 (Figure 3a) suggests reduced compositional variability which are hallmarks of ecological stabilization. This pattern supports the early stages of the physicochemical and geochemical constraints filtering, where rooftop microbiomes transition from environmentally constrained assemblages toward more similarly structured communities. Similar convergence patterns have been observed in other engineered substrates, including early-stage green roofs, where initially variable microbial communities become more compositionally similar over time as shared environmental constraints select for stress-tolerant taxa.[32]

The ordination also highlighted contrasts among control soils. Samples from the greenhouse (GH) clustered distinctly from all rooftop and at-grade soils, reaffirming its role as the shared microbial inoculum source rather than a stable endpoint community. In contrast, soils from the mature backyard (BY) and the engineered substrate from the Terra (T) plots clustered closely with the 2024 rooftop samples, suggesting that environmental filtering and microbial stabilization occur progressively over time under rooftop agrivoltaic conditions. Terra's consistent alignment with 2024 samples, even when compared to the first-year rooftop soils, further indicates that its microbial community had already reached a more mature, stabilized state by the start of the study.

To quantify the relative influence of experimental factors on microbial community composition, we conducted a PERMANOVA using Bray–Curtis dissimilarities (Figure 3b). Although temporal succession (Year) structured the primary axis of separation in the ordination ($R^2 = 0.10161$, $p = 0.001$) (consistent with ANOSIM results), shade regime (Shade Type) explained the largest proportion of overall variance in community composition ($R^2 = 0.21564$, $p = 0.001$). Sample type (rhizosphere (RZ) vs. bulk substrate (PB, HB) explained a smaller but still significant portion of variance ($R^2 = 0.06804$, $p = 0.001$), reflecting distinct assembly pathways in plant-associated versus free-living microbial communities. All interaction terms were significant ($p \leq 0.001$), indicating that the influence of shading regime and microhabitat type was modulated by time. Together, these results highlight a dual pattern: year effects dominate the temporal structuring of communities, while shade-driven microclimatic conditions account for the largest share of total variation across the experiment. It should be noted that the difference between the ANOSIM and PERMANOVA results reflects the complementary nature of these tests: ANOSIM emphasizes relative group separation in multivariate space, highlighting clear temporal clustering in the ordination, whereas PERMANOVA partitions the total variance, revealing that shade treatments explain a larger share of the overall community variation even though their separation is less visually pronounced.

3.3.5 Ecological taxon groupings: persistent, broad and specialist

Microbial ASVs were classified into ecological groupings to evaluate how different taxa contributed to early rooftop community assembly. These groupings: persistent taxa**,** broad rooftop colonizers, and specialized taxa, reflect ecological behavior rather than taxonomic exclusivity. Taxa were designated as persistent if they were present in the greenhouse inoculum

and detected across all rooftop 2023 and 2024 samples; broad colonizers if absent from the greenhouse but widespread across rooftop plots in either year; and specialists if their occurrence was restricted to specific microhabitats or years. Although several families contained ASVs assigned to multiple ecological roles, reflecting intra-family functional diversity, these categories captured distinct assembly pathways shaping rooftop microbiomes.

Persistent taxa (Figure 4a) were relatively few in number (n = 586 ASVs) compared to the broad and specialist groups, but still taxonomically diverse. Families unique to this group included Caldilineaceae and the uncultured JG30-KF-CM45 lineage from the Chloroflexi phylum; Micropepsaceae and Rhodanobacteraceae from the Proteobacteria phylum and Rubinisphaeraceae from the Planctomycetota phylum. Many Caldilineaceae family members are chemoheterotrophs known for being thermophilic; however they can also be found in mesophilic, oligotrophic habitats, suggesting broad tolerance of low-nutrient conditions.[36] Micropepsaceae include mildly acidophilic taxa capable of obligate or facultative anaerobic metabolism, consistent with ecological flexibility under fluctuating moisture and oxygen availability.[37] Rhodanobacteraceae are frequently associated with nitrate-rich, acidic, or otherwise chemically stressed soils.[38,39] Taken together, the persistent-group families represent organisms with broad environmental tolerances rather than specialized functional roles, which is consistent with what would be expected during the earliest stages of colonization in a low-organic-matter engineered substrate. Their presence across both growing seasons indicates that despite the harsh physical transition from greenhouse to rooftop, a small but stable consortium of taxa successfully established foundational microbial functions and contributed to early substrate colonization.

Broad rooftop colonizers (Figure 4b) represented the largest ecological group (n = 14,349 ASVs), consisting of ASVs not detected in the greenhouse inoculum but widespread across rooftop plots in Year 1 and/or Year 2. The only unique family to this group is the family Pedosphaeraceae, which is a family within the phylum Verrucomicrobiota. Pedosphaeraceae tend to be found in conditions with higher moisture, stable hydrologic conditions, and organic-rich micro-zones like the rhizosphere.[40] The other lineages found within this group, include members from the phylum Bdellovibrionota, Chloroflexi, Proteobacteria, and, Planctomycetota. The taxa within this group likely arrived via atmospheric deposition, dust, rainfall, and building-associated dispersal, reflecting the high connectivity of rooftop environments.[41] Ecologically, these taxa represent the "generalist wave" of colonization typical of early successional habitats, filling available niche space before more specialized interactions emerge.

Specialist taxa (Figure 4c) comprised ASVs whose presence was restricted to particular shading treatments or to a single year, indicating strong environmental filtering by rooftop microhabitats. Although restricted in distribution, this group contained the largest number of ASVs (n = 25,351), reflecting fine-scale niche partitioning across heterogeneous rooftop environments rather than broad habitat generalism. Specialist ASVs could be further subdivided by treatment and year at the ASV level; however, taxonomic resolution below the family rank was dominated by uncultured or alphanumeric lineages, limiting meaningful ecological interpretation at finer scales. Consequently, specialist dynamics are interpreted here at the family and functional level, which more robustly reflects habitat filtering in early-stage engineered substrates. The only family unique to this group is Anaerolineaceae (phylum Chloroflexi). Members of this family are typically slow growing, chemoheterotrophic, mesophilic or thermophilic and strictly anaerobic.[42] Their restricted occurrence suggests the development of

localized anaerobic microenvironments within rooftop substrates.[43,44] Their enrichment in either shaded or full-sun plots suggests that microhabitat differences in temperature, moisture availability, and radiative exposure select for specific microbial functions, even at these early stages of substrate development. The specialist group highlights how fine-scale physicochemical heterogeneity introduced by photovoltaic structures can support distinct, spatially restricted microbial populations within an otherwise shared engineered environment.

Together, these ecological groupings illustrate how early rooftop microbiome assembly proceeds through multiple, overlapping pathways rather than a single dominant trajectory. Persistent taxa provide continuity from the greenhouse inoculum, broad colonizers rapidly occupy newly available niche space through high dispersal and ecological flexibility, and specialist taxa emerge in response to fine-scale physicochemical heterogeneity imposed by rooftop microclimates. While these groups differ in their spatial and temporal distributions, they collectively underscore the importance of microhabitat diversity in shaping early community structure within engineered substrates.

3.3.6 Treatment associated shifts in dominant rhizosphere genera

Building on the ecological taxon groupings described above, we next examined how dominant microbial taxa varied across shading treatments and time, with a focus on relative abundance shifts in the rhizosphere (Figure 5). This analysis aimed to identify microbial lineages associated with specific rooftop microhabitats and to assess whether consistent treatment-associated patterns could be observed across years. We acknowledge that inferring stress tolerance or functional traits directly from taxonomic identity is limited for many ASVs, particularly given the prevalence of uncultured lineages in amplicon datasets. Accordingly, only

a subset of taxa can be interpreted in an ecological or physiological context. Nevertheless, several notable lineages exhibited consistent abundance shifts across shading treatments, suggesting potential associations with rooftop or microhabitat-specific conditions.

Notable lineages included members of thermophilic anaerobic Chloroflexi genera, which showed treatment specific temporal dynamics. For example, *AKYG1722* significantly declined (MaAsLin2, p-val < 0.05) across all shading treatments, *OLB13* significantly declined (MaAsLin2, p-val < 0.05) in both the opaque and full-sun treatments, and *SBR1031* significantly increased (MaAsLin2, p-val < 0.05) only under opaque panels. Collectively, these Chloroflexi taxa had the highest relative abundance in full-sun plots but were present across all samples, with overall abundance declining from 2023 to 2024. This pattern suggests that members of these Chloroflexi groups act as early colonizers adapted to rooftop conditions characterized by heat and desiccation but are gradually displaced as plant-associated communities become more structured and competitive. In contrast, thermophilic and halotolerant anaerobes within the Firmicutes, including *Geobacillus*, *Hydrogenispora*, and *Planifilum*, increased over time across all treatments, with significant (MaAsLin2, p-val < 0.05) enrichment in the opaque and full-sun plots. Importantly, electrical conductivity declined across all treatments and substrates were well drained (see Section 3.3.2), indicating that the occurrence of anaerobic or halotolerant taxa does not reflect bulk salinity or sustained saturation. Instead, their presence is consistent with the formation of transient, microscale low-oxygen niches within the rhizosphere. Adaptation to warm, intermittently anoxic environments in these taxa is often attributed to traits such as endospore formation and stress tolerance, which confer resilience to fluctuating rooftop conditions. This divergence between Chloroflexi and Firmicutes highlights differences in ecological strategy: slow-growing, non-spore-forming Chloroflexi suited to stable environments

versus spore-forming Firmicutes capable of surviving disturbance and may be responsive to rhizosphere stimulation.

Nitrogen-cycling taxa also exhibited treatment-specific dynamics within rhizosphere communities. The ammonia-oxidizing archaeon *Candidatus Nitrocosmicus* increased significantly over time in the opaque treatment, indicating that shaded, more thermally buffered microhabitats may favor ammonia-oxidizing archaea during early substrate development consistent with their preference for oxic environments and reduced competition with heterotrophic root-associated microbes.[45,46] In contrast, *SH-PL14*, a lineage associated with anaerobic ammonium oxidation (anammox), declined over time in the opaque treatment, suggesting that rhizosphere conditions under sustained shading became progressively less favorable for anaerobic nitrogen-transforming taxa. Although bifacial and opaque treatments received comparable mean solar radiation (see Section 3.3.1), opaque panels produced more thermally buffered and temporally stable microhabitats, which may more strongly favor ammonia-oxidizing archaea. These opposing trends indicate that shading-mediated microclimate differences can shift the balance of nitrogen-cycling strategies within rhizosphere communities over time, even in the absence of strong differences in bulk nutrient pools.

Overall, discriminant taxon analyses revealed that rooftop agrivoltaic substrates exhibit clear, directional patterns of microbial succession shaped by rooftop microclimates and substrate conditions. Specific stress-resilient lineages were repeatedly associated with these environments: thermophilic Chloroflexi appeared to function as early colonizers adapted to heat and desiccation, while spore-forming Firmicutes increased through successive growing seasons, consistent with recruitment of taxa capable of persisting through temperature and moisture fluctuations. Stable populations of Actinobacteriota were maintained across treatments and time,

suggesting a persistent functional backbone potentially contributing to organic matter turnover and degradation of recalcitrant compounds. In parallel, nitrogen-cycling taxa showed treatment-specific dynamics within rhizosphere communities, indicating that microclimate-mediated shifts can alter biogeochemical strategies even in the absence of large changes in bulk nutrient pools. Collectively, these patterns demonstrate that rooftop substrates undergo predictable microbial succession driven by physicochemical filtering and plant microbe interactions, highlighting the importance of early-stage substrate design and microclimate management for guiding community development and sustaining rooftop soil function.

3.3.7 Microbial Community Succession from Greenhouse to Rooftop

To assess whether temporal and treatment-associated shifts in rooftop microbial communities were driven by deterministic environmental filtering or stochastic processes, we applied the β-nearest taxon index (βNTI), which quantifies phylogenetic turnover relative to a null model. Negative βNTI values (<-2) indicate homogeneous or deterministic selection, which is where strong, consistent environmental filtering leads to similar community composition across sites. βNTI values near 0 (-2 to +2) indicate stochastic processes, where communities are shaped by random dispersal, drift, or weak selection. Positive βNTI values (>+2) indicate variable selection where divergent environmental conditions select for different taxa, through a combination of strong through heterogenous selection and deterministic processes.

Across all early-stage samples, median βNTI values were strongly negative in the greenhouse and first-year rooftop rhizosphere communities, indicating dominance of homogeneous selection during early community assembly (Figure 6). This pattern was consistent across shading treatments in 2023, with all rooftop plots exhibiting βNTI values well below −2.

In 2024, rooftop rhizosphere communities remained under strong deterministic control, with median βNTI values still firmly negative across opaque, bifacial, and full-sun treatments. Although slight shifts in median decreases were observed between years and among treatments, these differences were not statistically significant, indicating persistence of homogeneous selection through the second growing season. In contrast, mature rooftop (Terra) rhizosphere substrates exhibited markedly higher and more variable βNTI values, spanning from near-zero to positive values, consistent with a transition toward variable selection and increased stochasticity among replicates. The at-grade backyard rhizosphere soil, however, maintained strongly negative βNTI values similar to early stage rooftop substrates, indicating continued deterministic control in this established system.

These patterns indicate that rooftop microbiome assembly remains dominated by strong environmental filtering throughout the first two growing seasons, with relaxation of deterministic control emerging only in longer established rooftop soils. From a management perspective, these results indicate that the early stages of rooftop soil development are dominated by strong environmental filtering driven by substrate design, irrigation practices, and microclimatic conditions. This sustained period of deterministic control likely represents the most influential window for intervention, during which substrate composition, moisture regimes, and planting strategies may have lasting effects on microbial community structure and functional potential as rooftop soils continue to develop.

3.4 Conclusions

This study provides an integrated view of microbial succession in rooftop agrivoltaic engineered substrates by linking microclimate variation, substrate geochemistry, community

diversity, taxonomic composition, and phylogenetic assembly processes. The results show that microbial communities assemble rapidly under strong physicochemical constraints imposed by engineered substrates and photovoltaic-driven microclimates, while gradually increasing in complexity and differentiation over time.

Rooftop agrivoltaic treatments created consistent and biologically meaningful microclimatic gradients. Shading altered substrate temperature, light exposure, and moisture stability, with full-sun plots experiencing greater thermal extremes and desiccation, while opaque and bifacial panels produced cooler, more buffered environments. These gradients established the abiotic template for early microbial assembly in substrates lacking legacy organic matter and soil structure.

Despite these microclimatic contrasts, rhizosphere geochemical pools remained uniformly low and often declined over time, reflecting rapid consumption and leaching of amendment derived inputs in coarse, weakly buffered media. While shading modestly influenced the rate of resource loss, geochemical variables did not strongly structure microbial communities at the community level, indicating that early microbial assembly occurs under broadly constrained and oligotrophic conditions.

Within this physicochemical framework, microbial diversity increased significantly between years, particularly in rhizosphere samples, highlighting the growing influence of plant–microbe interactions during early substrate development. Community composition shifted directionally from 2023 to 2024, with second year communities becoming more cohesive and converging toward the structure of mature rooftop substrates. Shade regime explained the largest proportion of overall community variation, underscoring the importance of microclimate in shaping successional trajectories.

Ecological taxon grouping analyses revealed that early assembly proceeds through multiple, overlapping pathways. A small group of persistent taxa bridged the transition from greenhouse to rooftop, broad colonizers dominated overall richness through high dispersal and ecological flexibility, and numerous specialist taxa reflected fine-scale niche partitioning driven by microhabitat heterogeneity. Analyses of dominant rhizosphere genera further showed predictable lineage-level responses to rooftop conditions, with thermophilic Chloroflexi declining over time and spore-forming Firmicutes increasing across treatments, consistent with shifts in stress tolerance and successional strategy. Nitrogen-cycling taxa exhibited treatment-specific dynamics, indicating that subtle microclimate differences can alter biogeochemical strategies within the rhizosphere.

Phylogenetic turnover analyses confirmed that microbial assembly during the first two growing seasons was dominated by strong environmental filtering. Negative βNTI values across greenhouse and rooftop rhizospheres indicate persistent homogeneous selection driven by engineered substrate properties and rooftop microclimates, with relaxation of deterministic control emerging only in longer-established rooftop soils.

Overall, these findings demonstrate that rooftop agrivoltaic substrates undergo predictable microbial succession governed primarily by physicochemical filtering and microclimate structure during early establishment. Rather than rapid stochastic assembly, these systems remain under strong environmental control for multiple growing seasons, with gradual increases in diversity and functional differentiation occurring within this deterministic framework. This sustained window of strong filtering represents a critical opportunity for management interventions, as substrate composition, irrigation, and photovoltaic design are

likely to exert lasting influence on microbial community development and long-term soil function in rooftop agrivoltaic systems.

3.5 Methods

3.5.1 Site Description

The experiment was conducted at the Colorado State University Spur Campus (CSU Spur) located at 4777 National Western Dr., Denver, CO 80216, USA (39°47.0092′ N, 104°58.4378′ W; 1609 m above sea level). The experimental design included four rooftop treatments (three newly installed engineered substrates and one mature rooftop plot), one at-grade field plot, and a greenhouse treatment that served as the starting microbial inoculum prior to rooftop exposure.

Two of the rooftop plots were planted beneath photovoltaic (PV) arrays: one under opaque silicon-based PV panels (LR4-72HPH, Longi, Sydney, Australia) and one under bifacial silicon-based PV panels, which contain solar cells on both sides and allow partial light transmission. Both PV arrays were mounted using standard ground-mounted racking on the rooftop at a 44° east-facing tilt. A third rooftop plot was maintained in full sun with no PV array. An additional mature rooftop plot ("Terra"), also in full sun and without PV coverage, was established one year prior to this study and served as a comparison representing a more developed rooftop microbial community. In 2024, an at-grade, full sun field plot was included as an additional comparison to rooftop systems. All rooftop plots were constructed as part of an Intensive Garden Roof® system with a substrate depth of 45.72 cm and were filled with LITETOP® Intensive Agriculture 17 growing media (American Hydrotech Inc., Chicago, IL, USA). According to manufacturer specifications, LITETOP® growing media consist of varying

ratios of lightweight aggregate and aggregate fines, low-clay sandy loam, medium-coarse sand, compost, peat moss, and perlite, formulated to balance load-bearing constraints, drainage, and plant growth requirements. The Intensive blend is designed to exhibit water-holding characteristics comparable to loamy garden soils while remaining well drained, supporting a wide range of plant types including shrubs, perennials, and shallow-rooted woody species. Exact component ratios are proprietary.

The first year of this study (2023) corresponded to the first growing season following installation and planting of the rooftop agrivoltaic system. The mature rooftop treatment was established one year earlier, in 2022, providing a reference point for more developed rooftop substrate conditions.

3.5.2 Planting and management practices

Four cultivars of peppers were planted in each treatment plot however all microbial sampling in this study was conducted beneath a single pepper cultivar, 'Mosco' (*Capsicum annuum L.*), to ensure consistency in plant-microbe interactions across treatments and years. The Mosco pepper is well adapted to semi-arid conditions in Colorado and is commonly used in containerized and rooftop production systems, making it a suitable model crop for rooftop agrivoltaic studies.

Pepper seedlings were started in the greenhouse approximately six weeks prior to transplanting using a commercial germination mix (Lambert LM-GPS; Lambert, Rivière-Ouelle, Québec, Canada) in plug trays. At the beginning of each growing season (late May), seedlings were transplanted into rooftop and at-grade plots and planted in discrete rows (plants were not mixed across rows). Within each plot, plants were spaced 30.5 cm apart within rows, and rows

were separated by 30.5 cm. This uniform planting design was maintained across all treatments and years to standardize plant density and root zone overlap. At the end of each season, aboveground biomass was removed by cutting plants at the substrate surface, while root systems were left intact to decompose in situ, allowing residual root-associated microbial communities to persist between growing seasons.

All plots received identical fertilization and irrigation regimes. At transplanting, a single application of slow-release granular fertilizer (Dynamite® 18–6–8 N–P–K; Sun Bulb Co. Inc., Arcadia, FL, USA) was applied uniformly to each planting hole at transplant. Approximately 30g of fertilizer was applied by hand per plant. No additional fertilizer applications were made during the growing season. During the transplanting of seedlings into each research plot (May 31, 2023), 250mL of EcoGro compost (A1 Organics: Commerce City, CO, USA) was mixed into the substrate in each planting hole to provide nutrients and organic matter. Irrigation was supplied once daily uniformly across treatments using an automated irrigation system. Differences in substrate moisture among treatments therefore reflect shading-induced microclimatic variation rather than differential water application.

3.5.3 Substrate sample collection

Substrate samples for microbial and soil geochemical analyses were collected during the 2023 and 2024 growing seasons at two time points: at transplanting (May) and at final harvest (October). Three substrate fractions were targeted: planting bulk (PB), harvest bulk (HB), and rhizosphere (RZ), which were operationally defined as described below. All samples were collected using a stainless steel garden trowel (blade length 15.7 cm), which was sterilized with

70% ethanol between each sample to prevent cross-contamination. Samples were collected to a depth of approximately 0-15 cm.

Planting bulk (PB) samples were collected at transplanting from the planting holes prior to the addition of compost, fertilizer, or seedlings. For each treatment, planting holes were excavated to accommodate transplant plugs, and substrate was collected directly from the hole before any amendments were applied. In 2023, three PB samples were collected from randomly selected planting holes in each rooftop treatment (opaque, bifacial, full sun) and from the mature rooftop plot (Terra). In 2024, PB samples were collected in the same manner from all rooftop treatments (opaque, bifacial, full sun, Terra), with three replicates per treatment. Harvest bulk (HB) samples were collected at final harvest from substrate located at least 30.5 cm away from any plant stem to minimize rhizosphere influence and represent non–plant-associated substrate. In 2023, HB samples were collected only from the mature rooftop (Terra) plot (n = 3). In 2024, HB samples were collected from all rooftop treatments (opaque, bifacial, full sun, Terra), with three replicates per treatment. Rhizosphere (RZ) samples were collected at final harvest by carefully excavating entire plants. Plants were removed by loosening the surrounding substrate and lifting the root system intact. Loosely adhering soil was gently shaken off, and soil remaining tightly bound to the roots was operationally defined as the rhizosphere fraction. In 2023, rhizosphere samples were collected from three plants per rooftop treatment (opaque, bifacial, full sun, Terra) and from three greenhouse-grown plants, yielding three RZ replicates per treatment. In 2024, rhizosphere samples were collected from three plants per treatment for all rooftop plots (opaque, bifacial, full sun, Terra), and from an at-grade field plot, which was included that year as an additional comparison and sampled only for rhizosphere soil. Across both growing seasons, a total of 69 substrate samples were collected: 30 samples in 2023 and 39

samples in 2024. All samples were placed on ice immediately after collection, transported to the laboratory, and stored at −20 °C until further processing.

3.5.4 16S rRNA gene amplicon sequencing

Total DNA was extracted from 0.5 g of substrate (n = 69) using the Zymo Research Quick-DNA Fecal/Soil Microbe Kit (cat. no. D6010) following the manufacturer's recommended soil protocol. Bacterial and archaeal 16S rRNA genes (V4 region; primers 515F/806R32) were PCR-amplified following the Earth Microbiome Project (EMP) protocol[47] and sequenced on an Illumina MiSeq platform using 2 × 251 bp paired-end chemistry at the Colorado State University Sequencing Core. Raw amplicon sequence data were processed using QIIME2 (version 2023.9)[48]. Reads were demultiplexed, quality filtered, and truncated to 250 bp prior to denoising with DADA2[49] to infer amplicon sequence variants (ASVs). Taxonomy was assigned using a naïve Bayes classifier trained on the SILVA 138 database[50] for bacteria and archaea. ASVs classified as chloroplasts, mitochondria, or unassigned at the domain level were removed prior to downstream analyses.

3.5.5 Soil geochemical analysis

Plant-available phosphorus (P) and potassium (K) were determined using Mehlich-3 extraction. Nitrate (NO_3^-) concentrations were measured colorimetrically. Organic matter (OM) was quantified by loss on ignition, and total carbon (C) and nitrogen (N) were measured by dry combustion. Soil pH and electrical conductivity (EC) were determined in a 1:1 substrate-to-deionized water slurry. All geochemical analyses were conducted at the Soil, Water, and Plant Testing Laboratory located at the CSU Spur campus in Denver, Colorado.

3.5.6 Microclimate and sensor data

Substrate temperature, solar radiation, and substrate moisture data were obtained from continuous sensor measurements collected as part of a parallel study conducted at the same site (Gross, 2024).[51] Air and substrate temperatures were measured using HOBO 12-Bit Temperature Smart Sensors connected to HOBO H21-USB Micro Station data loggers (Onset Computer Corporation, Bourne, MA, USA). Sensors measuring below-panel temperatures were mounted approximately 45 cm above the substrate surface, while above-panel sensors were installed ~15 cm above photovoltaic panels in the bifacial and opaque treatments. Substrate moisture was measured using EC-5 Soil Moisture Smart Sensors (Onset Computer Corporation, Bourne, MA, USA) installed within each plot. Incoming solar radiation was monitored using HOBO Silicon Pyranometer Smart Sensors mounted approximately 45 cm above the substrate surface near the center of each plot.

All microclimate variables were recorded at 15-minute intervals throughout each growing season. These data were used to quantify treatment-level differences in temperature, moisture, and solar exposure and to provide environmental context for interpreting microbial community assembly and succession. Full details of sensor placement, calibration, and data processing are provided in Gross (2024).[51]

3.5.7 Statistical Analysis

To assess phylogenetic turnover and underlying assembly processes, we calculated the beta Nearest Taxon Index (βNTI) using the bNTIn.p function in the iCAMP package. Analyses were performed on a rarefied ASV table (rarified 10,000 reads per sample, 3 samples per treatment) and a phylogenetic tree reconstructed from representative sequences. The phylogenetic tree was constructed from aligned representative sequences using default QIIME2[48]

methods. βNTI was computed with 1,000 randomizations, weighted by relative abundance, and corrected for special cases using the "MNTD" method. For interpretation, thresholds of $|\beta NTI| < 2$ were taken to indicate stochastic assembly processes (drift, dispersal, neutral dynamics), $\beta NTI < -2$ to indicate homogeneous selection, and $\beta NTI > +2$ to indicate variable selection. We compared within-treatment βNTI distributions using nonparametric tests (Kruskal–Wallis followed by Dunn's post hoc test with Benjamini–Hochberg correction for multiple comparisons). To assess microbial community dynamics, several complementary statistical approaches were used. Differential abundance of bacterial genera was tested using MaAsLin2[52] on 16S rRNA amplicon data from rhizosphere and bulk substrate samples in the Hydro rooftop treatments, with significance determined at FDR < 0.05. Alpha diversity (ASV richness and Shannon diversity) was evaluated using ANOVA with Tukey's HSD for pairwise comparisons; when assumptions of normality were violated, nonparametric Dunn's tests with Benjamini–Hochberg correction were applied. Beta diversity was visualized using Principal Coordinates Analysis (PCoA) based on Bray–Curtis dissimilarities, implemented in the vegan R package (v. 2.7-0). Differences in community composition across treatments, years, and sample types were assessed using PERMANOVA. PERMANOVA was conducted with 999 permutations, and homogeneity of group dispersions was assessed using betadisper prior to interpretation. ANOSIM was also applied to test for temporal clustering between years. All statistical analyses were conducted in R[53].

Chapter 3 Figures and tables

Table 1: Substrate temperatures (°C) across 2023 and 2024 growing seasons

Treatment	Mean Temp (°C)	SD	Min Temp	Max Temp	Year
bifacial	19.7	6.62	-1.1	34.3	2023
opaque	19.8	6.58	-0.93	33.8	2023
full sun	21	7.49	1.34	38.2	2023
bifacial	20.9	3.12	11.8	36.4	2024
opaque	21.9	3.3	12.1	32.5	2024
full sun	25	5.21	11.2	39.4	2024

Table 1. Substrate temperatures (°C) across rooftop agrivoltaic shading treatments during the 2023 and 2024 growing seasons. Mean daily substrate temperature ± standard deviation (SD), with observed seasonal minimum and maximum temperatures, are shown for bifacial, opaque, and full-sun treatments installed on the rooftop agrivoltaic system. Data reproduced with permission from Gross (2025).[51]

Table 2. Summary of solar radiation (W m^{-2}) across 2023 and 2024 growing seasons

Treatment	Mean solar rad.	SD	Min solar rad.	Max solar rad.	Year
bifacial	25.5	64.8	0.6	629	2023
opaque	30.1	70.3	0.6	577	2023
full sun	240	333	0.6	1277	2023
bifacial	22.2	60.1	0.6	504	2024
opaque	31.7	74.2	0.6	653	2024
Full sun	240	315	0.6	1277	2024

Table 2. Summary of solar radiation (W m^{-2}) measured in full sun, bifacial, and opaque treatments during the 2023 and 2024 growing seasons. Values represent the mean, standard deviation (SD), minimum, and maximum irradiance recorded for each treatment and year. Continuous logging was available for all treatments in 2023. In 2024, a logging interruption affected the bifacial sensor, so 2024 values are derived from months with complete records and should be interpreted as seasonal summaries rather than full time-series statistics. Data reproduced with permission from Gross (2025).[51]

Table 3. Summary of substrate moisture (m^3/m^3) across 2023 and 2024 growing seasons

Treatment	Mean Moisture	SD	Min Moisture	Max Moisture	Year
bifacial	0.0801	0.0603	-0.128	0.254	2023
opaque	0.117	0.0212	0.0867	0.216	2023
full sun	0.0621	0.032	-0.0055	0.199	2023
bifacial	0.0865	0.0311	-0.126	0.246	2024
opaque	0.131	0.0147	-0.0941	0.251	2024
full sun	0.0941	0.00616	0.0816	0.521	2024

Table 3. Summary of substrate volumetric moisture content (m^3/m^3) across bifacial, opaque, and full sun rooftop treatments during the 2023 and 2024 growing seasons. Values represent the mean, standard deviation (SD), minimum, and maximum moisture recorded for each treatment within each growing year. Moisture was measured continuously using in-situ soil sensors; the summary reflects all available recordings within the June through October monitoring period. Data reproduced with permission from Gross (2025).[51]

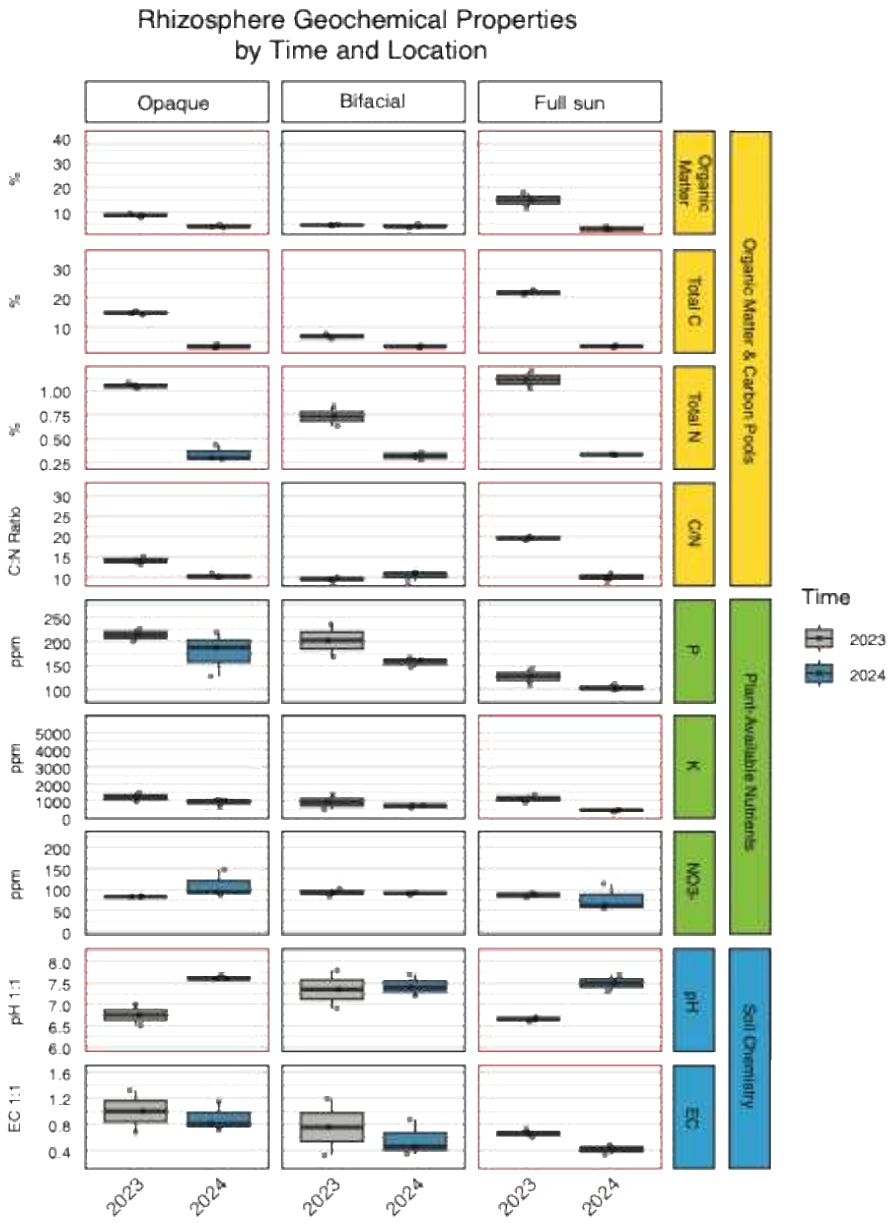


Figure 1. Rhizosphere geochemical properties by plot for 2023 and 2024. Boxplot displaying the distribution of values for a geochemical variable within a treatment and year for green house (GH), Hydro Opaque (H1) and Hydro full sun (H3). Boxes represent the interquartile range (IQR: 25th–75th percentile), the horizontal line indicates the median, and whiskers extend to the minimum and maximum values within 1.5× IQR. Individual points represent biological replicates (n = 3 per group). Colored boxes on the right denote geochemical categories: enzyme activity (orange), plant-available nutrients (green), organic matter and carbon pools (yellow), soil chemistry (blue), and physical properties (purple). Significance was tested with ANOVA after confirming normality with the Shapiro-Wilk test. Significance between years Is denoted by a red box outline. Abbreviations: EC = electrical conductivity, OM = organic matter, P = plant-available phosphorus, K = potassium, NO_3^- = nitrate, Total C = total carbon, Total N = total nitrogen, C:N = carbon-to-nitrogen ratio, pH = soil pH.

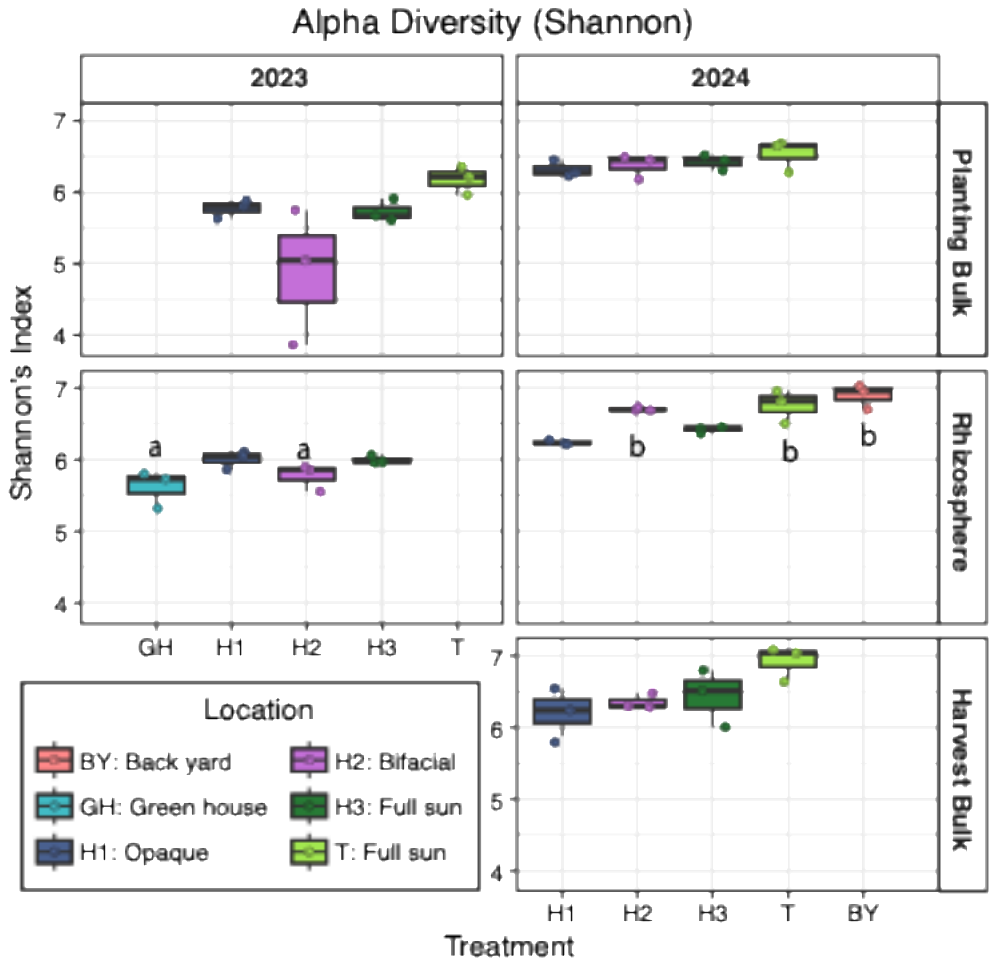


Figure 2. Alpha diversity (Richness) across rooftop agrivoltaic plots over two growing seasons. Boxplots display observed ASV richness for each treatment location in 2023 and 2024, faceted by sample type. For each boxplot, the horizontal line indicates the median, boxes represent the interquartile range (IQR), and whiskers extend to 1.5 × IQR; colored points represent individual samples. Letters above boxes indicate significant differences ($p < 0.05$, Dunn's test with Benjamini-Hochberg correction) within the rhizosphere samples across years and locations, where samples sharing the same letter are not significantly different.

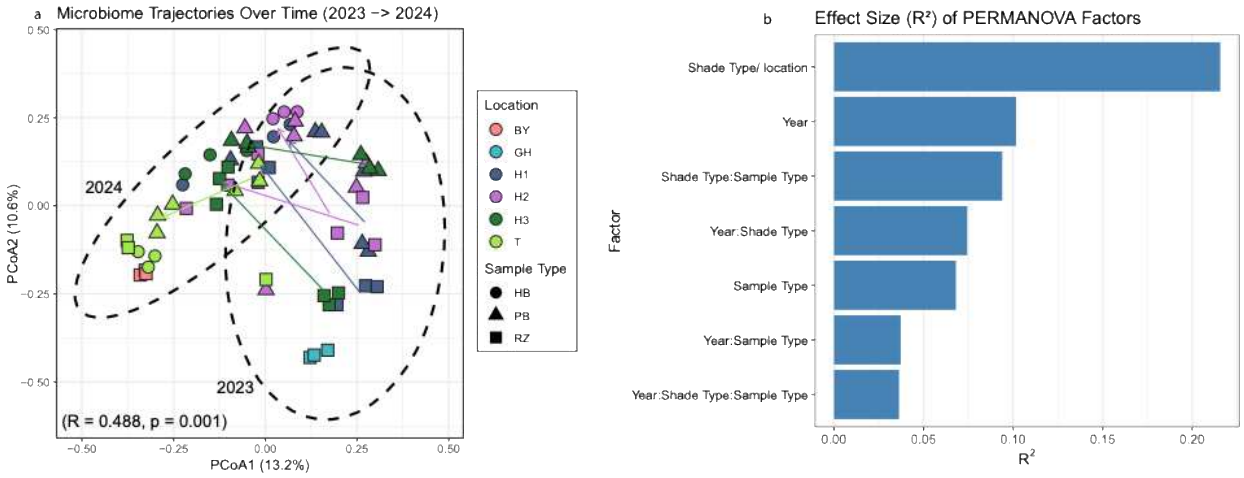


Figure 3. Beta diversity (PCoA) of rooftop agrivoltaic microbiomes across two years. a) Principal Coordinates Analysis (PCoA) based on Bray–Curtis dissimilarity showing distinct clustering of samples by year. Samples from 2023 and 2024 separated along the primary ordination axes, with 2024 samples forming a more cohesive cluster. Greenhouse (GH) samples clustered separately from all rooftop and at-grade soils. Backyard (BY) and Terra (T) samples grouped closely with 2024 rooftop samples. Lines connect group centroids for each location and sample type across years. ANOSIM indicated significant differences between years ($R = 0.488$, $p = 0.001$). b) Proportion of variation in community composition explained by experimental factors and interactions from a PERMANOVA (Bray–Curtis). Shade type (Opaque, Bifacial, full sun) explained the greatest variation ($R^2 = 0.216$), followed by Year ($R^2 = 0.102$) and Sample type (harvest bulk (HB), planting bulk (PB), rhizosphere (RZ)) ($R^2 = 0.068$). All factors and interaction terms were statistically significant ($p = 0.001$, 999 permutations).

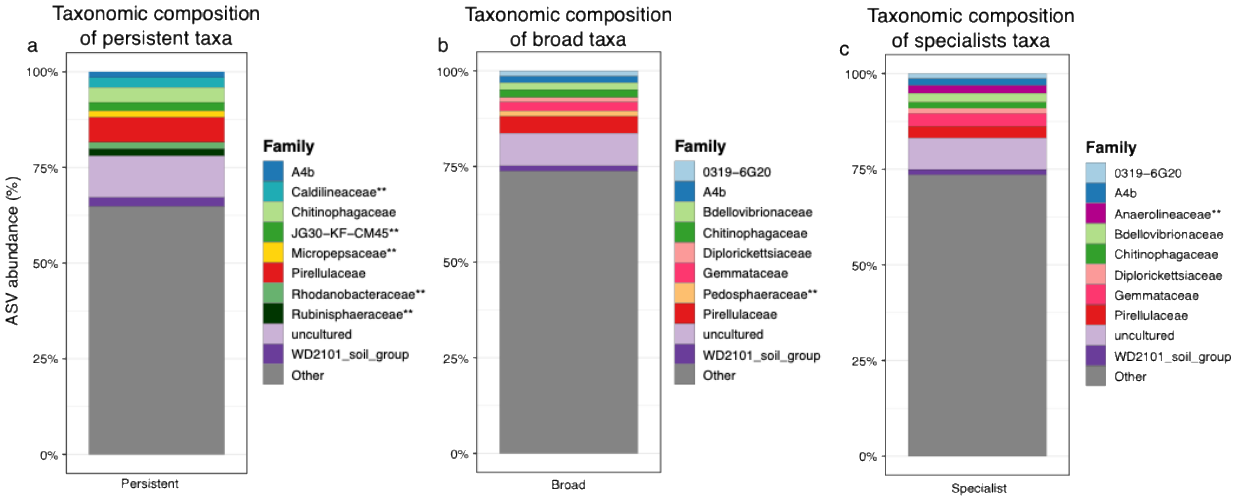


Figure 4. Taxonomic composition of ecological ASV groupings in rooftop agrivoltaic substrates. Stacked bar plots show the family-level taxonomic composition of ASVs classified as a) persistent taxa, b) broad rooftop colonizers, and c) specialist taxa. Bars represent the fraction of ASVs within each ecological grouping assigned to each family; families contributing <2% of ASVs within a group are aggregated as "Other." Families marked with a double asterisk (**) indicate lineages that were unique to that ecological grouping. Taxonomic assignments are shown at the family level to summarize dominant lineages while minimizing redundancy arising from limited genus- and species-level resolution in 16S rRNA gene data.

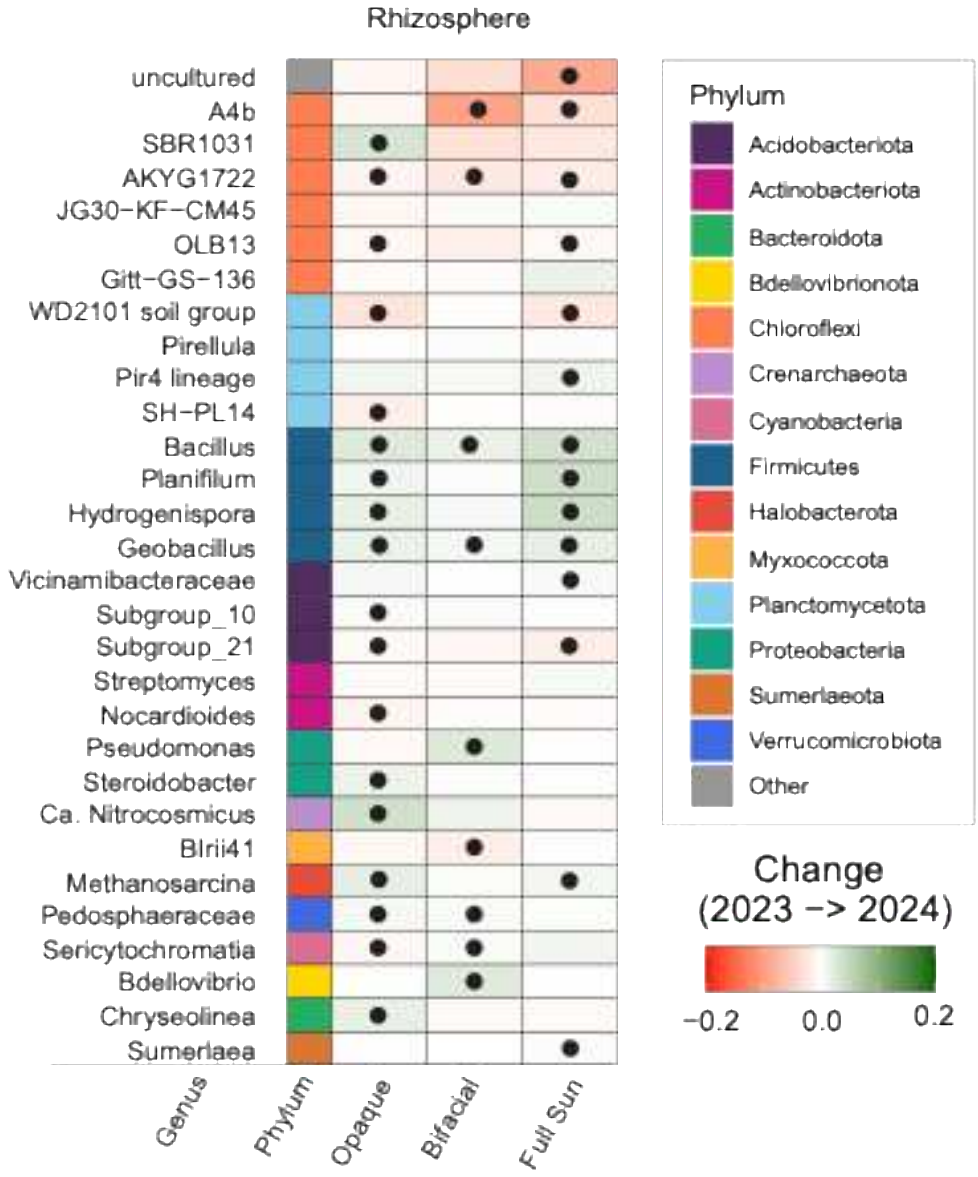


Figure 5. Changes in rhizosphere and bulk substrate genera between 2023 and 2024 in three rooftop treatments. The heatmaps display the difference in mean relative abundance for the top 30 most abundant genera in rhizosphere substrate samples across the three shading treatments: opaque, bifacial and full sun. Colors indicate change in relative abundance from 2023 to 2024, with green representing increased abundance and red representing decreased abundance. Black dots indicate genera identified as significantly different across years by MaAsLin2 ($q \leq 0.25$) using the full genus-level dataset (not limited to the top 30 genera shown). Genera are grouped by phylum, with colors along the "Phylum" column corresponding to phylum identity. Relative abundance differences are shown as raw differences (2024 mean – 2023 mean) in proportion. Color scaling is consistent across panels (±0.2).

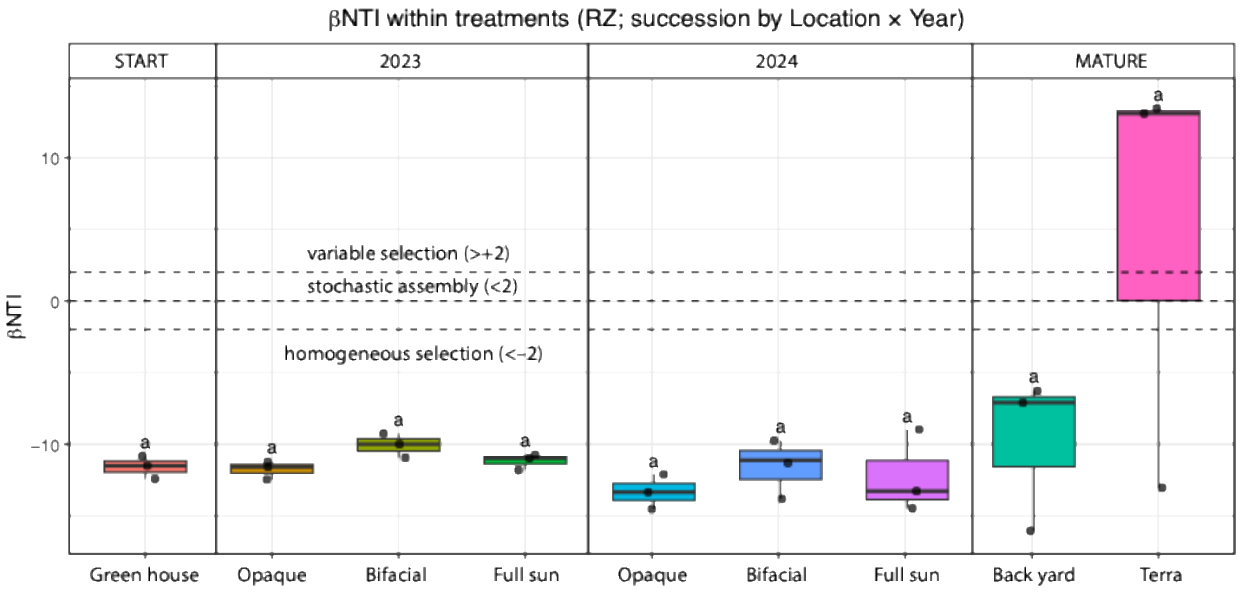


Figure 6. Rooftop rhizospheres substrates show homogeneous selection. Within-treatment βNTI values for rhizosphere samples across greenhouse (GH), rooftop plots (H1–H3), and mature soils (T = mature rooftop, BY = ground-level control) from 2023–2024. Dashed lines at βNTI = 0 and ±2 indicate thresholds for stochastic assembly (<2), variable selection (>+2), and homogeneous selection (<–2). Median βNTI values in the greenhouse and year-1 rooftops (2023) were strongly negative, consistent with homogeneous selection. Year-2 rooftop rhizospheres (2024) shifted toward less extreme values, while the mature rooftop (T) showed greater spread toward stochastic or variable assembly. The ground-level bulk yield soil (BY) remained strongly deterministic. Differences among treatments were not statistically significant (Kruskal–Wallis, $p > 0.05$).

Chapter 3 References

1. Sheikh, M. Z. *et al.* A Review of Roof Top Agro-Voltaic System Development: Integrating Renewable Energy and Urban Agriculture. *Adv. Res.* **26**, 95–106 (2025).
2. Maity, R., Hariram, N. P., Quazi, M. M. & Kumarasamy, S. Agrivoltaic systems for sustainability: An overview of emerging trends and practices. *Sol. Compass* **16**, 100148 (2025).
3. Dinesh, H. & Pearce, J. M. The potential of agrivoltaic systems. *Renew. Sustain. Energy Rev.* **54**, 299–308 (2016).
4. Walters, S. A. & Stoelzle Midden, K. Sustainability of Urban Agriculture: Vegetable Production on Green Roofs. *Agriculture* **8**, 168 (2018).
5. Vijayaraghavan, K. Green roofs: A critical review on the role of components, benefits, limitations and trends. *Renew. Sustain. Energy Rev.* **57**, 740–752 (2016).
6. VanWoert, N. D. *et al.* Green Roof Stormwater Retention. https://doi.org/10.2134/jeq2004.0364 doi:10.2134/jeq2004.0364.
7. Castleton, H. F., Stovin, V., Beck, S. B. M. & Davison, J. B. Green roofs; building energy savings and the potential for retrofit. *Energy Build.* **42**, 1582–1591 (2010).
8. Takebayashi, H. & Moriyama, M. Surface heat budget on green roof and high reflection roof for mitigation of urban heat island. *Build. Environ.* **42**, 2971–2979 (2007).
9. Passaseo, A., Pétremand, G., Rochefort, S. & Castella, E. Pollinator emerging from extensive green roofs: wild bees (Hymenoptera, Antophila) and hoverflies (Diptera, Syrphidae) in Geneva (Switzerland). *Urban Ecosyst.* **23**, 1079–1086 (2020).
10. Bardgett, R. D. & van der Putten, W. H. Belowground biodiversity and ecosystem functioning. *Nature* **515**, 505–511 (2014).

11. Ampim, P. A. Y., Sloan, J. J., Cabrera, R. I., Harp, D. A. & Jaber, F. H. Green Roof Growing Substrates: Types, Ingredients, Composition and Properties. *J. Environ. Hortic.* **28**, 244–252 (2010).
12. OBERNDORFER, E. *et al.* Green Roofs As Urban Ecosystems: Ecological Structures, Functions And Services. *BioScience* **57**, 823–833 (2007).
13. Nemergut, D. R. *et al.* Patterns and Processes of Microbial Community Assembly. *Microbiol. Mol. Biol. Rev.* **77**, 342–356 (2013).
14. Hénault, A. *et al.* Stressful, isolated, yet diverse: Green roofs have rich microbiomes that are not dominated by oligotrophic taxa. *Environ. Microbiol. Rep.* **14**, 766–774 (2022).
15. Nonhebel, S. Renewable energy and food supply: will there be enough land? *Renew. Sustain. Energy Rev.* **9**, 191–201 (2005).
16. Barron-Gafford, G. A. *et al.* Agrivoltaics provide mutual benefits across the food–energy–water nexus in drylands. *Nat. Sustain.* **2**, 848–855 (2019).
17. Hermelink, M. I., Maestrini, B. & de Ruijter, F. J. Berry shade tolerance for agrivoltaics systems: A meta-analysis. *Sci. Hortic.* **330**, 113062 (2024).
18. Mamun, M. A. A., Dargusch, P., Wadley, D., Zulkarnain, N. A. & Aziz, A. A. A review of research on agrivoltaic systems. *Renew. Sustain. Energy Rev.* **161**, 112351 (2022).
19. Liu, Z. *et al.* Potential benefits and risks of solar photovoltaic power plants on arid and semi-arid ecosystems: an assessment of soil microbial and plant communities. *Front. Microbiol.* **14**, (2023).
20. Liu, Y., Ding, C., Su, D., Wang, T. & Wang, T. Solar park promoted microbial nitrogen and phosphorus cycle potentials but reduced soil prokaryotic diversity and network stability in alpine desert ecosystem. *Front. Microbiol.* **13**, (2022).

21. Rumble, H., Finch, P. & Gange, A. C. Can microbial inoculants boost soil food webs and vegetation development on newly constructed extensive green roofs? *Urban For. Urban Green.* **75**, 127684 (2022).
22. Allison, S., Wallenstein, M. & Bradford, M. Soil-carbon Response to Warming Dependent on Microbial Physiology. *Nat. Geosci.* **3**, (2010).
23. De Vries, F. T. & Shade, A. Controls on soil microbial community stability under climate change. *Front. Microbiol.* **4**, (2013).
24. Schimel, J., Balser, T. C. & Wallenstein, M. Microbial Stress-Response Physiology and Its Implications for Ecosystem Function. *Ecology* **88**, 1386–1394 (2007).
25. Lennon, J. T. & Jones, S. E. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* **9**, 119–130 (2011).
26. Ruan, L. *et al.* Melanin pigment formation and increased UV resistance in Bacillus thuringiensis following high temperature induction. *Syst. Appl. Microbiol.* **27**, 286–289 (2004).
27. Allison, S. D. & Martiny, J. B. H. Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 11512–11519 (2008).
28. Chaparro, J. M., Badri, D. V. & Vivanco, J. M. Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* **8**, 790–803 (2014).
29. Jones, D. L., Hodge, A. & Kuzyakov, Y. Plant and mycorrhizal regulation of rhizodeposition. *New Phytol.* **163**, 459–480 (2004).
30. Siggers, J. A. *et al.* Environmental Heterogeneity Imposed by Photovoltaic Array Alters Grassland Soil Microbial Communities. *Glob. Change Biol.* **31**, e70376 (2025).

31. Cotrufo, M. F., Wallenstein, M. D., Boot, C. M., Denef, K. & Paul, E. The M icrobial E fficiency- M atrix S tabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Glob. Change Biol.* **19**, 988–995 (2013).
32. Oberndorfer, E. *et al.* Green Roofs as Urban Ecosystems: Ecological Structures, Functions, and Services. *BioScience* **57**, 823–833 (2007).
33. Ampim, P. A. Y., Sloan, J. J., Cabrera, R. I., Harp, D. A. & Jaber, F. H. Green Roof Growing Substrates: Types, Ingredients, Composition and Properties. *J. Environ. Hortic.* **28**, 244–252 (2010).
34. Hoch, J. M. K. *et al.* Soil Microbial Assemblages Are Linked to Plant Community Composition and Contribute to Ecosystem Services on Urban Green Roofs. *Front. Ecol. Evol.* **7**, 198 (2019).
35. Droz, A. G., Coffman, R. R., Eagar, A. C. & Blackwood, C. B. Drivers of fungal diversity and community biogeography differ between green roofs and adjacent ground-level green space. *Environ. Microbiol.* **24**, 5809–5824 (2022).
36. Sekiguchi, Y. & Hanada, S. Caldilineaceae. in *Bergey's Manual of Systematics of Archaea and Bacteria* 1–3 (John Wiley & Sons, Ltd, 2020). doi:10.1002/9781118960608.fbm00320.
37. Bräuer, S., Harbison, A. & Ueki, A. Micropepsaceae. in *Bergey's Manual of Systematics of Archaea and Bacteria* 1–5 (John Wiley & Sons, Ltd, 2018). doi:10.1002/9781118960608.fbm00311.
38. Green, S. J. *et al.* Denitrifying Bacteria from the Genus Rhodanobacter Dominate Bacterial Communities in the Highly Contaminated Subsurface of a Nuclear Legacy Waste Site. *Appl. Environ. Microbiol.* **78**, 1039–1047 (2012).

39. Campbell, B. J., Polson, S. W., Hanson, T. E., Mack, M. C. & Schuur, E. A. G. The effect of nutrient deposition on bacterial communities in Arctic tundra soil. *Environ. Microbiol.* **12**, 1842–1854 (2010).

40. Cordero, I., Leizeaga, A., Hicks, L. C., Rousk, J. & Bardgett, R. D. High intensity perturbations induce an abrupt shift in soil microbial state. *ISME J.* **17**, 2190–2199 (2023).

41. Fulthorpe, R., MacIvor, J. S., Jia, P. & Yasui, S.-L. E. The Green Roof Microbiome: Improving Plant Survival for Ecosystem Service Delivery. *Front. Ecol. Evol.* **6**, (2018).

42. Yamada, T. & Sekiguchi, Y. Anaerolineaceae. in *Bergey's Manual of Systematics of Archaea and Bacteria* 1–5 (John Wiley & Sons, Ltd, 2018). doi:10.1002/9781118960608.fbm00301.

43. Kuzyakov, Y. & Blagodatskaya, E. Microbial hotspots and hot moments in soil: Concept & review. *Soil Biol. Biochem.* **83**, 184–199 (2015).

44. Chen, X., Liu, R., Liu, D. & Xin, X. Analysis of Preferential Flow in Artificial Substrates with Sedum Roots for Green Roofs: Experiments and Modeling. *Water* **15**, 914 (2023).

45. Prosser, J. I. & Nicol, G. W. Archaeal and bacterial ammonia-oxidisers in soil: the quest for niche specialisation and differentiation. *Trends Microbiol.* **20**, 523–531 (2012).

46. Kits, K. D. *et al.* Low yield and abiotic origin of N2O formed by the complete nitrifier Nitrospira inopinata. *Nat. Commun.* **10**, 1836 (2019).

47. 16S Illumina Amplicon Protocol : earthmicrobiome. https://earthmicrobiome.org/protocols-and-standards/16s/.

48. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

49. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

50. Chuvochina, M. *et al.* SILVA in 2026: a global core biodata resource for rRNA within the DSMZ digital diversity. *Nucleic Acids Res.* gkaf1247 (2025) doi:10.1093/nar/gkaf1247.
51. Gross, E. (Capsicum annuum L.) PRODUCTION, GROWTH, AND CONSUMER ACCEPTANCE.
52. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLOS Comput. Biol.* **17**, e1009442 (2021).
53. R Core Team. R: A Language and Environment for Statistical Computing. https://www.r-project.org/ (2024).

Chapter 4: Functional processes and microbial traits in rooftop agrivoltaics

4.1 Summary

This chapter establishes a baseline understanding of soil microbial processes in a newly constructed rooftop agrivoltaic (RAV) system by integrating enzyme activity measurements with metagenomic analyses from the first year of establishment (2023). Rhizosphere data revealed significant temporal and treatment-specific shifts between shaded (opaque) and full-sun plots, providing a functional genomic perspective that complements the multi-year compositional and environmental analyses presented in Chapter 3. Together, these data capture both the early genomic potential and associated biochemical responses that underpin rooftop soil development under contrasting microclimatic conditions.

Metagenomic binning yielded 200 medium- and high-quality MAGs, with greenhouse soils recovering the most genomes, the opaque plot showing enrichment of predatory Myxococcota, and the full sun substrate containing a potential plant pathogen (*Pseudomonas_E mediterranea*). Functional profiling indicated that opaque substrates were statistically more enriched in carbohydrate-active enzymes, particularly glycoside hydrolases, consistent with patterns observed in enzyme activity assays. Full sun soils exhibited intermediate functional capacity but comparatively lower representation of carbon-degrading potential.

Collectively, these findings indicate that treatment-associated differences in rooftop microclimate coincide with distinct patterns of microbial functional potential during the first year of rooftop agrivoltaic establishment. Opaque substrates were characterized by greater representation of genes associated with carbon-degrading capacity, whereas full sun substrates showed genomic signatures consistent with stress tolerance and opportunistic strategies. These

patterns underscore the role of microclimatic filtering as a key contextual factor influencing early functional trajectories of microbial communities in engineered rooftop soils.

4.2 Introduction: Functional resilience in rooftop agrivoltaic systems

Studies examining the functional genomic potential of green roof microbial communities remain limited. Existing work has relied primarily on taxonomic surveys or targeted marker genes to infer potential microbial functions within soil-like rooftop substrates. For example, bacterial taxa such as *Bacillus*, *Azospirillum*, and *Bradyrhizobium* have been identified through taxonomic analyses, with inferred roles in nitrogen fixation and drought resilience.[1,2] Another study found that urban green roofs contained greater relative abundance of taxa associated with carbon decomposition and nitrogen cycling compared to other urban soils from parks and tree pits.[3] Functional gene quantitative measurements targeting nitrogen cycling pathways have also detected genes for bacterial ammonia oxidation (e.g., *Nitrospira*) and nitrogen fixation (*nifH*) in a subset of green roof substrates.[2]

Studies of agrivoltaic systems similarly indicate that photovoltaic (PV) induced microclimate variation is closely associated with shifts in soil microbial communities, although functional outcomes are often inferred rather than directly measured. In arid and semi-arid solar installations, soils beneath PV panels frequently exhibit altered microbial community composition, including enrichment of taxa such as Actinobacteriota and Proteobacteria, which have been interpreted as reflecting enhanced carbon cycling or adaptation to increased organic inputs and moderated temperature extremes[4,5] However, in certain dryland systems, shading coincided with reduced precipitation inputs, and in these zones, microbial metabolic activity (e.g., dehydrogenase enzyme activity) and the abundance of functionally important groups, such as Acidobacteriota, Myxococcota, and Glomeromycota, declined, which the authors indicated

signaled reduced nutrient cycling and potential disruption of plant–microbe symbioses.[6] Soil enzyme assays and taxonomic predictions further support changes in carbon, nitrogen, and phosphorus cycling genes.[7] Given these insights from green roofs and agrivoltaics systems primarily derived from taxonomic inference tools, or a handful of marker genes or enzymes, the functional potential of rooftop or agrivoltaics microbiomes has yet to be characterized with metagenomic tools, warranting the need for direct process-based and genome measurements.

As demonstrated in Chapter 3, rooftop agrivoltaic engineered substrates host distinct microbial communities that assemble under contrasting light and moisture regimes, reflecting strong abiotic filtering during early establishment. A critical next step is to determine whether these compositional differences correspond to meaningful differences in functional potential or whether key soil functions are maintained through compensatory mechanisms despite taxonomic divergence.

Accordingly, this chapter addresses two guiding questions. First, do shaded and full-sun rooftop agrivoltaic environments exhibit differences in genomic and biochemical indicators of carbon and nutrient cycling during early substrate development? Second, do the dominant microbial lineages associated with each microclimatic condition encode functional traits consistent with the observed enzyme activity patterns?

Addressing these questions contributes to the concept of functional compensation, in which stress-adapted or specialized microbial taxa sustain essential biogeochemical functions despite strong environmental filtering and reduced community complexity. In engineered rooftop substrates, where organic matter pools are limited and abiotic constraints are pronounced, functional outcomes may be maintained through shifts in microbial strategies rather than increases in diversity alone.

In this chapter, genome-resolved metagenomics is used to evaluate whether dominant microbial lineages associated with contrasting rooftop microclimates encode metabolic traits consistent with observed patterns in enzyme activity. By integrating a β-glucosidase enzyme assay with metagenomic reconstruction, this analysis links measurable activity to underlying genomic potential, providing a functional context for the successional and compositional patterns described in Chapter 3. Together, this approach establishes a baseline for understanding how early-stage rooftop agrivoltaic substrates maintain functional capacity under divergent microclimatic constraints.

4.3 Results & Discussion

4.3.1 Genome-resolved diversity hints at microclimatic filtering in rooftop substrates

Metagenomic binning from the first year of rooftop agrivoltaic soils yielded 200 medium- and high-quality MAGs across treatments (Figure 1). The greenhouse starter microbiome produced the largest number of MAGs (n = 104), while fewer bins were recovered from the full-sun (n = 47) and opaque (n = 49) rooftop substrates. Differences in MAG recovery among treatments should not be interpreted as direct proxies for microbial diversity, as amplicon-based analyses demonstrated that greenhouse soils had lower ASV richness relative to rooftop rhizospheres (Chapter 3, Section 3.3.3). Instead, reduced MAG recovery from engineered rooftop substrates likely reflects methodological and biological constraints, including lower DNA yield, greater genome fragmentation, and the enrichment of stress-adapted or spore-forming taxa observed in Chapter 3 (Sections 3.3.5–3.3.6), which are often more difficult to reconstruct into high-quality genomes from metagenomic data. Similar challenges in genome recovery have been reported in other low-biomass or highly stressed soil environments.

Accordingly, MAG recovery here is interpreted conservatively as a reflection of genome reconstruction rather than ecological abundance or diversity.

Assembly size also differed among treatments, with the full-sun plot producing the largest total assembly (30.67 gigabase pairs, Gbp), followed by the greenhouse (24.39 Gbp) and the opaque plot (20.32 Gbp) (Figure 1). Despite yielding fewer medium- and high-quality MAGs, the larger total assembly size in full-sun substrates indicates that a substantial proportion of assembled sequence could not be resolved into discrete genomes. This pattern is consistent with increased genome fragmentation, strain heterogeneity, or reduced binning efficiency under environmentally stressful conditions, rather than a greater number of recoverable genomes. In contrast, the opaque substrate yielded both fewer assembled bases and fewer MAGs, suggesting lower overall genomic recovery but potentially more homogeneous communities. Together, these patterns indicate that rooftop microclimates influence not only microbial composition but possibly the recovery and structure of genomic information during early substrate establishment.

Similar to the amplicon data in Chapter 3, taxonomic composition differed markedly among shaded and full-sun treatments. Greenhouse soils were enriched in Acidobacteriota, particularly members of the order Bryobacterales, lineages commonly associated with oligotrophic, resource-efficient strategies in more stable soil environments.[8] Their absence from rooftop full sun and opaque treatments suggests that early rooftop conditions rapidly disfavor these slower-growing taxa, consistent with the strong abiotic filtering observed during early substrate establishment (Chapter 3, Sections 3.3.4–3.3.7). This pattern has been reported in other engineered or disturbance-prone soils, where Acidobacteriota are often replaced by faster-growing or stress-tolerant groups during early succession.[9]

Full-sun soils were uniquely characterized by *Pseudomonas_E mediterranea*, a plant-associated taxon previously described in agricultural systems and reported to include pathogenic strains.[10] Its detection exclusively in full-sun rooftop substrates suggests that high light, temperature, and desiccation stress may favor opportunistic or disturbance-tolerant taxa during early establishment, though pathogenicity was not directly assessed in this study.

In contrast, the opaque treatment was distinguished by enrichment of Myxococcota (Polyangiaceae), a group of predatory bacteria known for complex social behavior and extracellular enzyme production (Figure 1).[11] Their recovery as MAGs is consistent with the more buffered, hydrated microclimates beneath opaque panels (Chapter 3, Section 3.3.1), which may facilitate surface-associated motility, enzyme secretion, and prey access within the soil matrix. These treatment-specific lineages illustrate how rooftop microclimates can foster distinct microbial assemblages, even within the first year of substrate establishment.

Comparison of the taxonomic composition of the metagenome-assembled genomes (MAGs) to the amplicon profiles presented in Chapter 3 revealed strong agreement across dominant lineages. Many of the same genera identified as environmentally responsive in the amplicon dataset, including *Hydrogenispora* and *Tepidimicrobium*, both thermophilic, fermentative Firmicutes enriched in shaded rooftop substrates were also recovered as MAGs in the metagenomic assemblies. Likewise, taxa such as *Pseudomonas_E mediterranea*, detected only in full-sun substrate, and Myxococcota, enriched in the opaque treatment, appeared consistently across both datasets. This concordance across methods strengthens the interpretation that shading-mediated microclimates impose consistent selective pressures detectable at both taxonomic and genomic resolutions.

Ultimately, this overlap demonstrates strong continuity between community patterns inferred from amplicon sequencing and the genome-resolved profiles obtained from metagenomics. Together, these complementary approaches reveal convergent responses to rooftop microclimatic conditions and support the use of genome-resolved data to contextualize functional potential within the successional framework established in Chapter 3. This agreement underscores that microclimatic filtering structures not only taxonomic composition but also the genomic landscape through which microbial communities adapt to engineered rooftop environments.

4.3.2 Functional gene profiles and signals of emerging functional compensation

MAGs from rooftop soils revealed clear treatment-level differences in functional gene inventories during the first year of establishment, despite all plots originating from the same greenhouse inoculum (Figure 2). When assessed by total gene counts, opaque (shaded) substrates contained the largest overall functional gene inventory, consistent with the enrichment patterns observed in the z-score heatmap (Figure 2b–e). Relative to the full sun treatment, the shaded substrate exhibited greater representation of genes associated with nitrogen, phosphorus, sugar utilization, central carbon metabolism, and plant growth regulator pathways, suggesting broader metabolic potential under buffered microclimatic conditions. Full sun substrates retained lower overall gene counts than opaque substrates but consistently exceeded those recovered from the greenhouse, indicating that early rooftop environments rapidly expand functional potential beyond the starting inoculum. Because gene inventories reflect genomic potential rather than realized activity, these patterns are interpreted as differences in encoded metabolic capacity rather than direct measures of process rates.

During the first year of establishment, genes encoding carbohydrate-active enzymes (CAZymes) and pathways associated with central carbon metabolism were enriched in the opaque substrate, whereas β-glucosidase enzyme activity was highest in the full sun treatment (Figure 2 a). In contrast to enzyme activity patterns, the relative abundance of β-glucosidase genes (EC 3.2.1.21) was higher in the opaque treatment compared to the full-sun, indicating a decoupling between genomic potential and realized enzymatic activity. Decoupling between genomic potential and realized enzymatic activity is widely reported in soil systems, where gene presence reflects long-term metabolic capacity, while enzyme activity responds rapidly to short-term environmental conditions such as substrate availability, temperature, moisture, and physiological regulation of expression.[12] The opaque substrate exhibited enrichment across a broader suite of carbon-degrading genes (Figure 2 b-e), including glycoside hydrolases, glycosyltransferases, carbohydrate esterases, and carbohydrate-binding modules, reflecting a more distributed genomic capacity for complex organic matter transformation.[13]

This divergence suggests that β-glucosidase activity in full-sun soils may be driven by regulatory, physiological, or environmental controls (e.g., substrate availability, temperature, or stress-induced enzyme expression) rather than gene abundance alone, whereas the opaque substrate encodes a wider array of carbon-degrading pathways that may not yet be fully expressed at the enzyme level. These results highlight that single-enzyme assays capture realized process rates under prevailing conditions, while metagenomic profiles reflect latent functional potential, emphasizing the importance of integrating both approaches to resolve how different rooftop microclimates shape carbon cycling capacity.[14]

Taken together, the opaque substrate encodes a more complete suite of genomic pathways associated with polysaccharide depolymerization and downstream carbon processing,

suggesting the potential to generate substrates necessary for terminal cellulose degradation across multiple metabolic steps. Whole pathway metagenomic profiles therefore provide a complementary perspective to single enzyme assays, capturing interconnected transformations that may precede or support detectable enzyme activity**.** Such pathway level resolution has been proposed as a more stable indicator of functional potential in early stage or environmentally constrained soils**.**[15]

These results address the central questions of whether distinct rooftop microclimates sustain comparable functional potential and whether dominant taxa encode genes necessary to support key biogeochemical processes. Although shaded and full sun substrates differed in the magnitude and composition of their functional gene inventories, both retained genomic capacity for core carbon and nutrient cycling pathways, and these capacities were encoded by distinct microbial assemblages shaped by microclimatic filtering. This pattern is consistent with emerging functional compensation, whereby contrasting communities maintain overlapping functional potential despite differences in taxonomic composition**.**[16]

Rather than demonstrating equivalent functional outcomes, these data indicate that multiple genomic configurations may support similar ecosystem functions during early rooftop soil development. From a management perspective, genome-resolved analyses offer a diagnostic framework for identifying functional gaps or redundancies during early establishment, potentially informing substrate design or targeted amendments before functional constraints emerge at the process level.[17]

4.4 Conclusions

Together, the compositional and genome-resolved analyses presented in Chapters 3 and 4 show that rooftop agrivoltaic substrates rapidly diverge from their greenhouse inoculum and

assemble distinct microbial communities shaped by time and shading-mediated microclimatic conditions. Although shaded and full-sun environments supported different microbial assemblages and functional gene inventories, both retained genomic capacity for core carbon and nutrient cycling processes during early establishment. This pattern indicates that strong abiotic filtering drives early compositional divergence without eliminating essential biogeochemical potential, consistent with emerging functional compensation in engineered rooftop soils. Moreover, the recovery of environmentally responsive taxa identified in amplicon analyses as metagenome-assembled genomes confirms that dominant community members encode the genes necessary to support key metabolic functions. Collectively, these findings suggest that while rooftop microclimates strongly structure early microbial assembly, functional resilience can emerge rapidly, allowing essential soil processes to persist across contrasting environmental conditions during rooftop agrivoltaic development.

4.5 Materials & Methods

4.5.1 Site Description

This chapter builds directly on the rooftop agrivoltaic (RAV) experiment described in Chapter 3, shifting the focus from microbial community composition to genome-resolved functional potential and enzyme activity. The experimental site, substrate composition, planting design, irrigation, fertilization, and microclimate treatments are described in detail in Chapter 3 (Section 3.5) and are not repeated here.

4.5.2 Soil Sampling and Processing

For the genome-resolved and enzyme activity analyses presented in Chapter 4, we focused on a subset of rhizosphere and bulk substrate samples collected during the first year of rooftop establishment (2023). Specifically, samples were drawn from the greenhouse inoculum

(GH), the opaque rooftop treatment (H1), and the full-sun rooftop treatment (H3), representing the inoculum source and two contrasting rooftop microclimates. Rhizosphere samples were prioritized for metagenomic sequencing and enzyme assays, as these plant-associated substrates exhibited the strongest treatment responses in Chapter 3. All samples were handled, transported, and stored following the protocols outlined in Chapter 3.

4.5.3 Enzyme assay analysis

β-glucosidase (BG) activity was quantified using a colorimetric assay based on the hydrolysis of p-nitrophenyl-β-D-glucopyranoside (pNPG), following standard soil enzyme assay protocols.[18] Briefly, substrate samples were incubated with pNPG under controlled conditions, and enzymatic activity was determined by measuring the release of p-nitrophenol (PNP) following alkaline termination of the reaction. Absorbance was measured spectrophotometrically, and enzyme activity was calculated using a PNP standard curve. Enzyme activity is reported as mg PNP kg^{-1} substrate h^{-1}. All assays were conducted at the Soil, Water, and Plant Testing Laboratory located at the Colorado State University SPUR campus in Denver, Colorado. For the analyses presented here, three biological replicates per treatment (greenhouse, opaque rooftop, and full-sun rooftop) were analyzed from the 2023 growing season.

4.5.4 DNA Extraction, Metagenomic Sequencing, and Assembly

Total DNA was extracted from 0.5 g of soil (n=9) using Zymo Research Quick-DNA fecal/soil Microbe kits (cat #D6010) following the manufacturer's 'soil' protocol. Metagenomic DNA sequencing (n=9) was conducted at the Next Generation Sequencing Core at Colorado State University, using the Illumina NovaSeq 6000 platform (Illumina, v.1.5 chemistry, S4 flow cell, 2 × 150 bp). Raw reads were trimmed using sickle[19] (v1.2.9) in paired-end mode, using default parameters (minimum quality score = 20, minimum length = 20) and assuming Sanger-

encoded quality scores (-t sanger). Trimmed reads were then assembled using MEGAHIT[20] (v1.2.9), with all samples assembled individually except for three deeply sequenced samples that were co-assembled to improve contig recovery, and contigs > 2500bp were binned into MAGs using MetaBAT2[21] (v2.12.1) with default settings. MAG completeness and contamination were assessed via CheckM[22] (v1.1.2), resulting in 200 medium and high-quality MAGs (MQ-HQ) per MIMAG[23] standards. To construct MAG database, the 200 MAGs generated in this study were dereplicated at 99% identity using dRep[24] (v.2.6.2)

4.5.6 Functional Annotation

Functional annotation was performed with DRAM1[25], which identifies metabolic pathways, carbohydrate-active enzymes (CAZy), and other gene categories. Gene presence/absence matrices were generated across treatments and visualized as row z-score heatmaps and stacked bar charts. Categories of interest (e.g., nitrogen cycling, phosphorus metabolism, CAZy, central carbon metabolism, hydrocarbon degradation, sugar utilization, plant growth regulators) were emphasized in comparative analyses.

4.5.7 Statistical Analyses

Metagenomic and functional gene profiles presented in this chapter were analyzed descriptively. Visual comparisons, including heatmaps, bar charts, and relative enrichment plots, were used to summarize patterns in functional gene inventories and enzyme activity across treatments. Inferential statistical testing was not applied to metagenomic data because genome-resolved analyses were conducted for a single establishment year (2023) with limited replication at the treatment level, precluding robust statistical inference. All data processing, visualization, and summary analyses were conducted in R (v4.3.2) using the packages vegan (v2.7-0), ggplot2

(v3.4.4), and ggpubr (v0.6.0). Where relevant, statistical analyses of microbial community composition, diversity, and physicochemicalvariables are described in detail in Chapter 3.

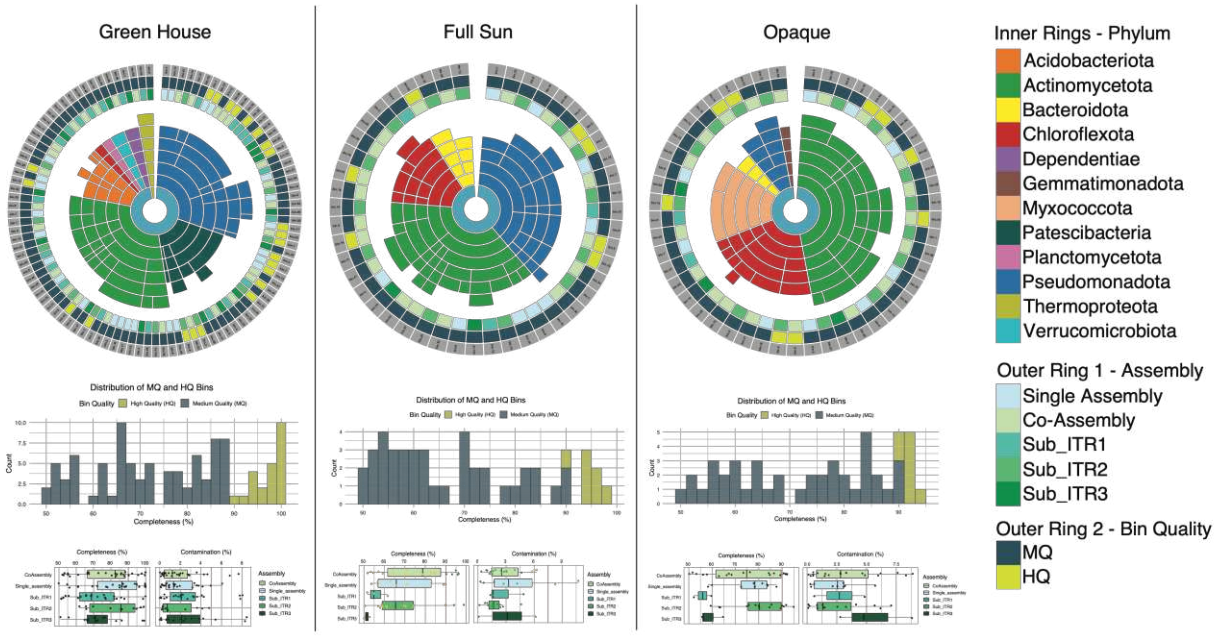


Figure 1. Metagenome-assembled genomes (MAGs) recovered from greenhouse, full sun, and shaded rooftop soils. Taxonomic composition and quality metrics of MAGs recovered from greenhouse (left), full sun (center), and opaque (right) substrates. Circular plots display taxonomic composition, with inner rings showing phylum-level classifications and outer rings indicating assembly source and bin quality (HQ = high quality, MQ = medium quality). Histograms illustrate the distribution of MAG completeness, with HQ (≥90% complete, ≤5% contamination) and MQ (≥50% complete, ≤10% contamination) bins distinguished by color. Boxplots show distributions of completeness and contamination across co-assemblies and single assemblies.

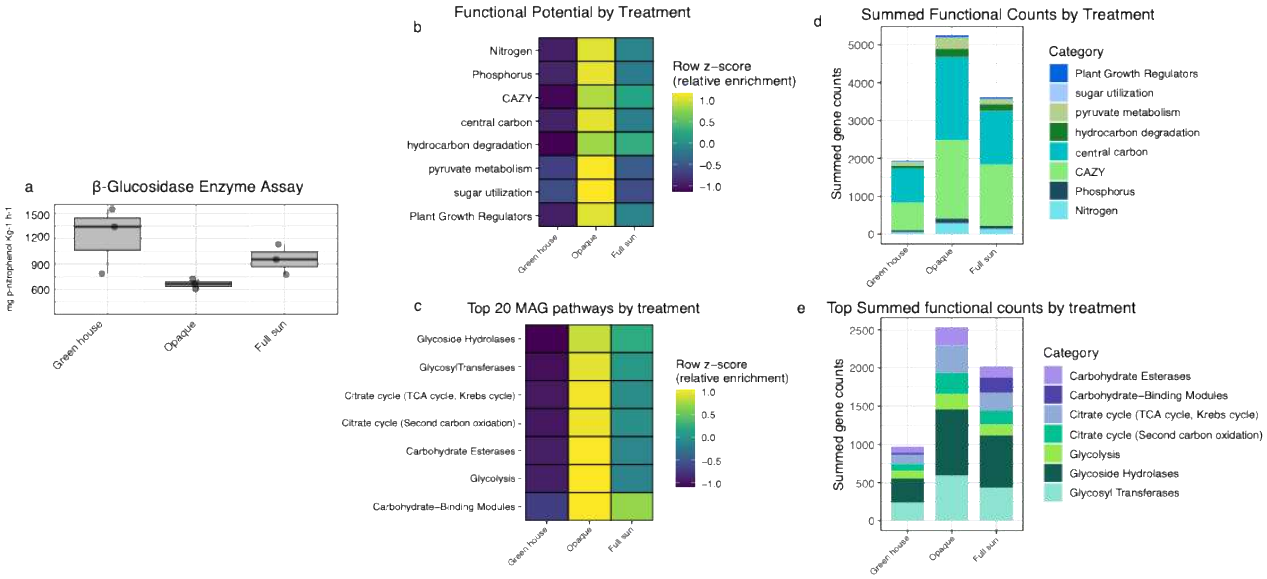


Figure 2. Enzyme activity and genome-resolved functional potential across rooftop agrivoltaic treatments. a) β-glucosidase enzyme activity measured in rhizosphere substrates from greenhouse, opaque (shaded), and full-sun treatments during the first year of rooftop establishment (2023). Boxplots represent the median (horizontal line), interquartile range (box; 25th–75th percentiles), and range excluding outliers (whiskers extending to 1.5× the interquartile range). Individual points represent biological replicates. β-glucosidase activity is expressed as mg *p*-nitrophenol kg^{-1} h^{-1} and reflects realized microbial potential for cellulose- and hemicellulose-derived carbon turnover at the time of sampling. b) Heatmap of relative functional enrichment across treatments based on genome-resolved annotations, summarized by major biogeochemical categories (row Z-score). Values represent relative enrichment within each functional category rather than absolute gene abundance. c) Heatmap showing the top 20 most abundant metabolic pathways recovered from metagenome-assembled genomes (MAGs), highlighting pathways associated with carbohydrate processing and central carbon metabolism. d) Summed gene counts by treatment across major functional categories, illustrating differences in overall functional gene representation among greenhouse, opaque, and full-sun substrates. e) Summed gene counts for carbohydrate-active enzyme (CAZyme) classes, including glycoside hydrolases, glycosyltransferases, carbohydrate esterases, carbohydrate-binding modules, and central carbon pathways.

Chapter 4 References

1. Hoch, J. M. K. *et al.* Soil Microbial Assemblages Are Linked to Plant Community Composition and Contribute to Ecosystem Services on Urban Green Roofs. *Front. Ecol. Evol.* **7**, 198 (2019).
2. Mitchell, M. E., Hamilton, T. L., Uebel-Niemeier, C., Hopfensperger, K. N. & Buffam, I. Nitrogen cycling players and processes in green roof ecosystems. *Appl. Soil Ecol.* **132**, 114–125 (2018).
3. Gill, A. S., Purnell, K., Palmer, M. I., Stein, J. & McGuire, K. L. Microbial Composition and Functional Diversity Differ Across Urban Green Infrastructure Types. *Front. Microbiol.* **11**, 912 (2020).
4. Bai, Z. *et al.* Photovoltaic panels have altered grassland plant biodiversity and soil microbial diversity. *Front. Microbiol.* **13**, (2022).
5. Luo, Z., Luo, J., Wu, S., Luo, X. & Sui, X. Soil bacterial community in a photovoltaic system adopted different survival strategies to cope with small-scale light stress under different vegetation restoration modes. *Front. Microbiol.* **15**, (2024).
6. Liu, Z. *et al.* Potential benefits and risks of solar photovoltaic power plants on arid and semi-arid ecosystems: an assessment of soil microbial and plant communities. *Front. Microbiol.* **14**, (2023).
7. Liu, Y., Ding, C., Su, D., Wang, T. & Wang, T. Solar park promoted microbial nitrogen and phosphorus cycle potentials but reduced soil prokaryotic diversity and network stability in alpine desert ecosystem. *Front. Microbiol.* **13**, (2022).
8. Dedysh, S. N. Bryobacter. in *Bergey's Manual of Systematics of Archaea and Bacteria* 1–5 (John Wiley & Sons, Ltd, 2019). doi:10.1002/9781118960608.gbm01667.

9. McReynolds, E., Elshahed, M. S. & Youssef, N. H. An ecological-evolutionary perspective on the genomic diversity and habitat preferences of the Acidobacteriota. *Microb. Genomics* **11**, 001344 (2025).
10. Licciardello, G. *et al.* Transcriptome analysis of *Pseudomonas mediterranea* and *P. corrugata* plant pathogens during accumulation of medium-chain-length PHAs by glycerol bioconversion. *New Biotechnol.* **37**, 39–47 (2017).
11. Saggu, S. K., Nath, A. & Kumar, S. Myxobacteria: biology and bioactive secondary metabolites. *Res. Microbiol.* **174**, 104079 (2023).
12. Zhao, Z., Baltar, F. & Herndl, G. J. Decoupling between the genetic potential and the metabolic regulation and expression in microbial organic matter cleavage across microbiomes. *Microbiol. Spectr.* **12**, e03036-23 (2024).
13. Trivedi, P. *et al.* Microbial regulation of the soil carbon cycle: evidence from gene–enzyme relationships. *ISME J.* **10**, 2593–2604 (2016).
14. Nannipieri, P. *et al.* Soil enzymology: classical and molecular approaches. *Biol. Fertil. Soils* **48**, 743–762 (2012).
15. Louca, S. *et al.* Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* **2**, 936–943 (2018).
16. Delgado-Baquerizo, M. *et al.* Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat. Commun.* **7**, 10541 (2016).
17. Enabling sustainable agriculture through understanding and enhancement of microbiomes - Trivedi - 2021 - New Phytologist - Wiley Online Library. https://nph.onlinelibrary.wiley.com/doi/full/10.1111/nph.17319.

18. Tabatabai, M. A. Soil Enzymes. in *SSSA Book Series* (eds Weaver, R. W. et al.) vol. 5 775–833 (Wiley, 1994).
19. NA najoshi & Frass, JN. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. (2025).
20. 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 (2015).
21. Kang, D. D. *et al.* MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).
22. 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 (2015).
23. Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**, 725–731 (2017).
24. 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. *ISME J.* **11**, 2864–2868 (2017).
25. Shaffer, M. *et al.* DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* **48**, 8883–8900 (2020).

Chapter 5: Synthesis and Conclusions

This dissertation demonstrates that soil microbial communities provide sensitive, mechanistic insight into agroecosystem function across both conventional agricultural fields and engineered urban rooftop systems. By integrating amplicon sequencing, genome-resolved metagenomics, metatranscriptomics, enzyme assays, and geochemical measurements, this work shows that microbial responses to management and environment often precede, and are not fully captured by, traditional soil health metrics.

In long-term maize production systems (Chapter 2), microbial community composition and transcriptional activity responded strongly to tillage intensity, revealing distinct nitrogen and carbon cycling strategies that were not detectable using standard biological soil health indicators alone. Conservation tillage promoted microbial functions associated with carbon depolymerization and microaerophilic metabolism, whereas conventional tillage favored rapid nitrogen turnover and ammonia oxidation. These findings demonstrate that microbial gene expression provides an early and sensitive indicator of functional change, reinforcing the value of multi-omic approaches for diagnosing soil processes before they manifest in bulk soil properties.[1,2]

In contrast, the rooftop agrivoltaic system (Chapter 3 and 4) offered a unique lens into early microbial assembly and functional constraint within an engineered substrate. Amplicon-based analyses showed rapid divergence from greenhouse inocula and persistent compositional differences driven by photovoltaic shading, while genome-resolved analyses revealed that shaded and full-sun microclimates encoded distinct functional gene inventories. Importantly, genomic potential did not consistently align with realized enzyme activity, highlighting a

decoupling between functional capacity and expression that is well documented in soil systems.[3,4]

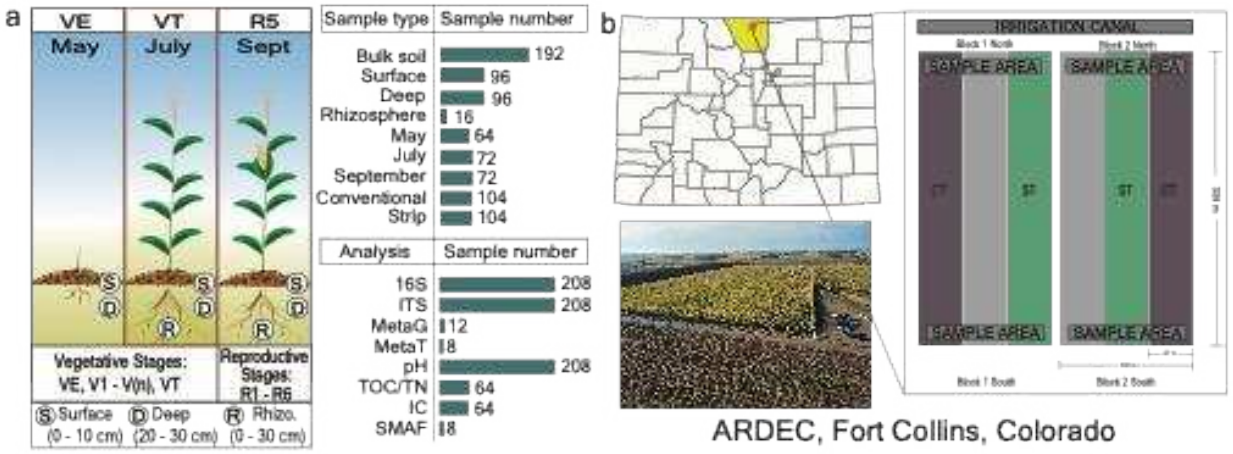
Together, these results indicate that functional compensation can emerge early in both natural and engineered soils, whereby distinct microbial communities retain overlapping genomic capacity for core biogeochemical processes. However, this compensation reflects potential rather than guaranteed functional equivalence. Environmental constraints, physiological regulation, and microclimatic filtering shape whether and when encoded pathways are expressed. As such, functional resilience should be understood as a dynamic and context-dependent property, not an inherent feature of microbial diversity alone.[5]

Across systems, this dissertation underscores that microbial communities act both as indicators of environmental change and as agents shaping soil processes. By moving beyond taxonomic inventories toward genome-resolved and transcriptional frameworks, this work contributes to a growing body of evidence that soil health assessment benefits from incorporating microbial functional traits alongside traditional physical and chemical indicators.[6–8]

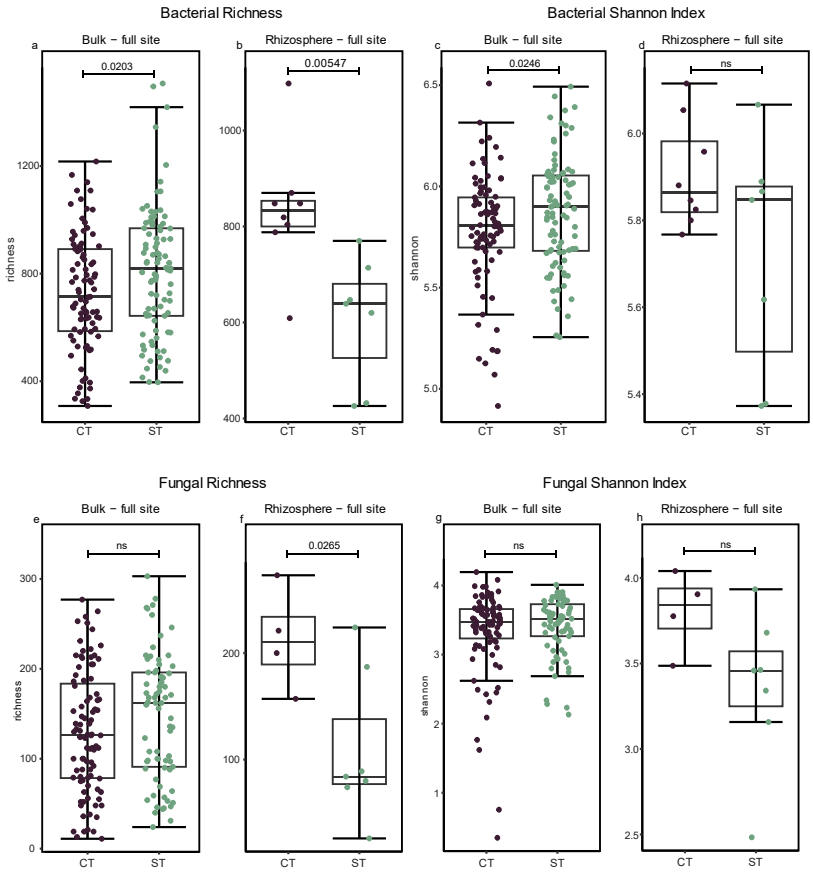
Ultimately, these findings support a shift toward function-first soil microbiome assessments, particularly in systems undergoing rapid environmental change or early development. As agricultural landscapes expand into novel contexts such as urban rooftops and agrivoltaic infrastructure multi-omic approaches provide a powerful, scalable toolkit for diagnosing constraints, guiding management interventions, and designing soils that sustain ecosystem function under increasing environmental pressure.

Chapter 5 References

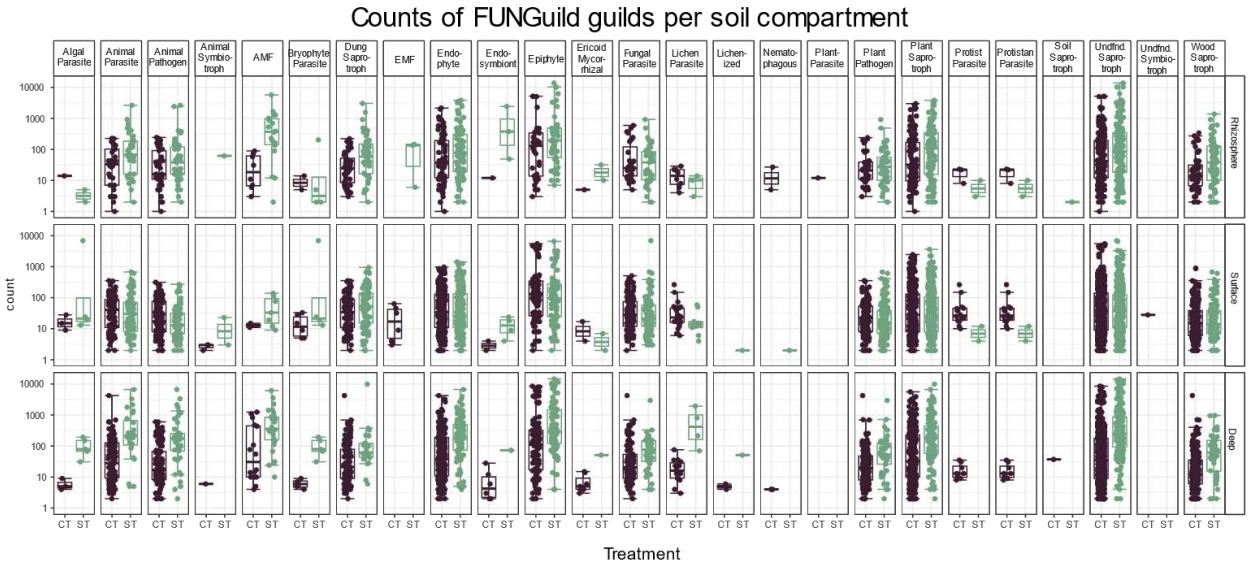
1. Prosser, J. I. Dispersing misconceptions and identifying opportunities for the use of 'omics' in soil microbial ecology. *Nat. Rev. Microbiol.* **13**, 439–446 (2015).
2. Trivedi, P., Delgado-Baquerizo, M., Anderson, I. C. & Singh, B. K. Response of Soil Properties and Microbial Communities to Agriculture: Implications for Primary Productivity and Soil Health Indicators. *Front. Plant Sci.* **7**, (2016).
3. Burns, R. G. *et al.* Soil enzymes in a changing environment: Current knowledge and future directions. *Soil Biol. Biochem.* **58**, 216–234 (2013).
4. Rocca, J. D. *et al.* Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. *ISME J.* **9**, 1693–1699 (2015).
5. Allison, S. D. & Martiny, J. B. H. Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci.* **105**, 11512–11519 (2008).
6. Fierer, N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* **15**, 579–590 (2017).
7. Doran, J. W. & Zeiss, M. R. Soil health and sustainability: managing the biotic component of soil quality. *Appl. Soil Ecol.* **15**, 3–11 (2000).
8. Lehmann, J., Bossio, D. A., Kögel-Knabner, I. & Rillig, M. C. The concept and future prospects of soil health. *Nat. Rev. Earth Environ.* **1**, 544–553 (2020).



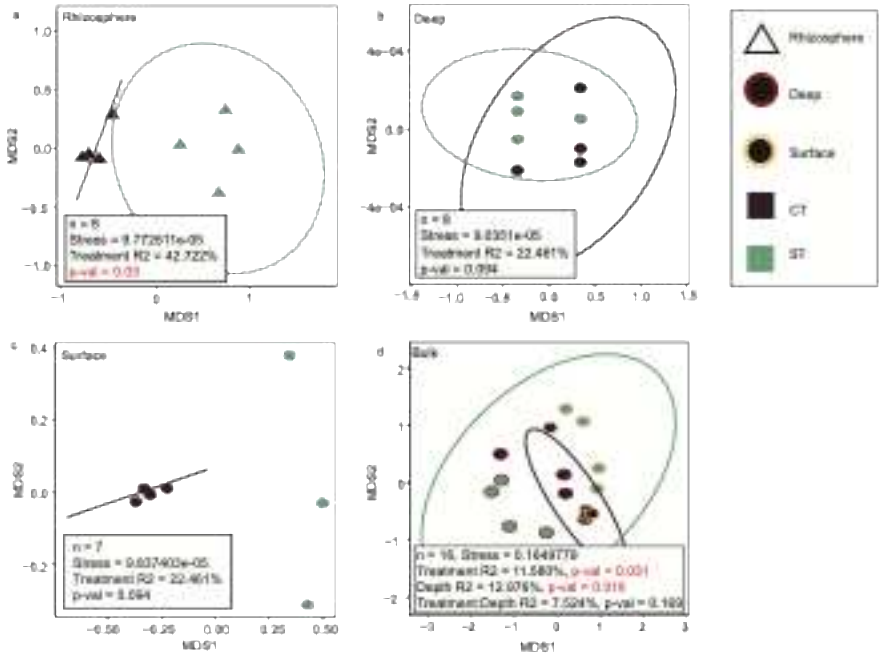
Supplemental Figure 1. Experimental design. a Sampling times and types throughout the corn growing season and breakdown of data types collected are shown by bar charts. Specifically, bulk soil and rhizosphere samples were taken from three stages of corn growth: emergence (VE), tassel (VT), and dent (R5). Bulk soil samples were taken at surface (S, 0-10 cm), and deep (D, 20-30 cm). Rhizosphere samples were taken root zone (R, 0-30 cm). Analyses of each sample type included 16S rRNA amplicon analysis (16S), metagenomics (MetaG), metatranscriptomics (MetaT), pH, total carbon and nitrogen (TOC/TN), ion chromatography (IC) and Soil Management Assessment Framework (SMAF). **b** Map of Colorado with Larimer County highlighted, the red star shows the ARDEC location in northern Fort Collins, Colorado. Photos of the field were provided by Emmanuel Deleon. Blow out plot shows the layout of the full field experiment with sample area highlighted.



Supplemental Figure 2. Opposing trends in microbial alpha diversity between bulk and rhizosphere soils under different tillage regimes. Full site (all blocks and directions) richness (left) and Shannon diversity index (right) for bacteria and archaea (top, 16S rRNA gene sequencing) and fungi (bottom, ITS sequencing). Significance is indicated above each boxplot with horizontal bars and corresponding p-values from analysis of variance (ANOVA). Average full site bacterial/archaeal richness (bulk and rhizosphere combined) was 754.94, and fungal richness averaged 138.61. Average Shannon diversity index for bacteria/archaea and fungi was 5.81 and 3.35, respectively. Boxplots display the median (horizontal line), interquartile range (IQR; the box), and whiskers extending to 1.5×IQR from the quartiles; outliers are plotted individually. **a** Bulk bacterial richness was significantly higher in ST (mean = 808.84) than CT (mean = 728.96) ($p < 0.05$). **b** Rhizosphere bacterial richness was significantly higher in CT (mean = 835.5) than ST (mean = 606.86) ($p < 0.05$). **c** Bulk bacterial Shannon diversity was significantly higher in ST (mean = 5.87) than CT (mean = 5.76) ($p < 0.05$). **d** Rhizosphere bacterial Shannon diversity did not differ significantly across treatments. **e** Bulk fungal richness showed no significant difference between CT (mean = 129.43) and ST (mean = 147.84). **f** Rhizosphere fungal richness was significantly higher in CT (mean = 212.75) than ST (mean = 109.14) ($p < 0.05$). **g–h** Fungal Shannon diversity did not differ significantly between treatments.



Supplemental Figure 3. Fungal guild composition does not differ significantly across treatments. Fungal guilds were assigned using FUNGuild. Significance is noted above each boxplot by horizontal bars and associated p-values, calculated via Permutational Multivariate Analysis of Variance (PERMANOVA, $p < 0.05$). Boxplots display the median (horizontal line), interquartile range (IQR; the box), and whiskers extending to 1.5×IQR from the quartiles; outliers are plotted individually. **a** Raw amplicon sequence variant (ASV) counts from all timepoints across fungal guilds for each soil compartment. No significant differences were detected.



Supplemental Figure 4. Strongest treatment effects were observed in the rhizosphere, guiding the focus of subsequent analyses. a-d Non-metric multidimensional scaling (NMDS) plots showing bacterial and archaeal (16S rRNA gene) community composition (Bray-Curtis dissimilarity) by soil compartment (rhizosphere, surface, deep) for Block 1 North soils collected in July. Points represent individual samples, colored by tillage treatment (CT: dark purple, ST: green). Ellipses represent 95% confidence intervals for each treatment group. PERMANOVA results (R^2 and p-values) are shown for each compartment. Across compartments, the rhizosphere exhibited the strongest treatment effect, with tillage treatment explaining a larger proportion of microbial community variation compared to surface, deep and bulk soils.

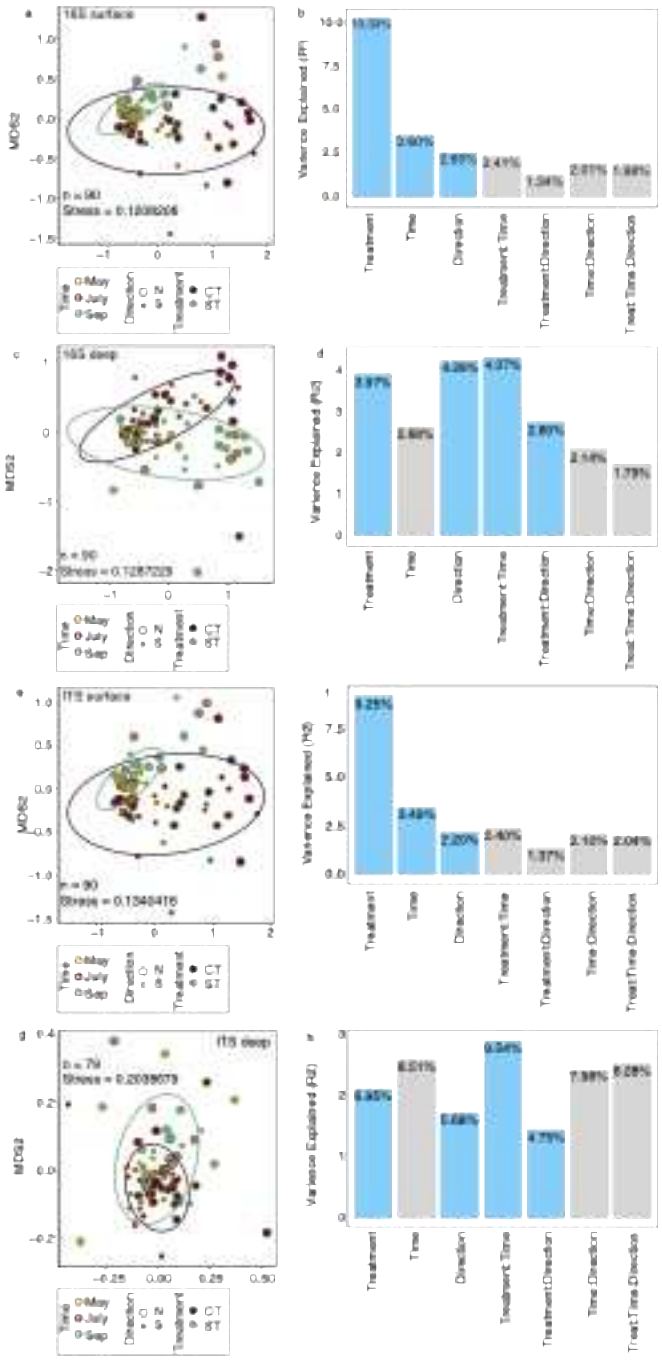


Figure 5. Treatment-driven microbial community differences persist across spatial and temporal variation. Non-metric multidimensional scaling (NMDS) ordinations and variance explained (R^2) from PERMANOVA analyses for microbial communities across surface and deep soils. **a, c** Panels show NMDS ordinations for 16S and **e, g** ITS datasets in **a, e** surface and **c, e** deep soil layers, including all timepoints (May, July, September) and both directional transects (North, South). Points are outlined by time, filled by treatment (CT = conventional till, ST = strip till), and sized by direction. Ellipses represent 95% confidence intervals around treatment centroids. Corresponding PERMANOVA R^2 values are shown in bar plots adjacent to each NMDS plot, indicating the proportion of variance explained by treatment, time, direction, and their interactions. Significant variances ($p < 0.05$) are shown in blue. Separate NMDS and PERMANOVA analyses were performed for each dataset and depth to visualize how treatment effects vary within soil compartments. Stress values for each ordination are provided.

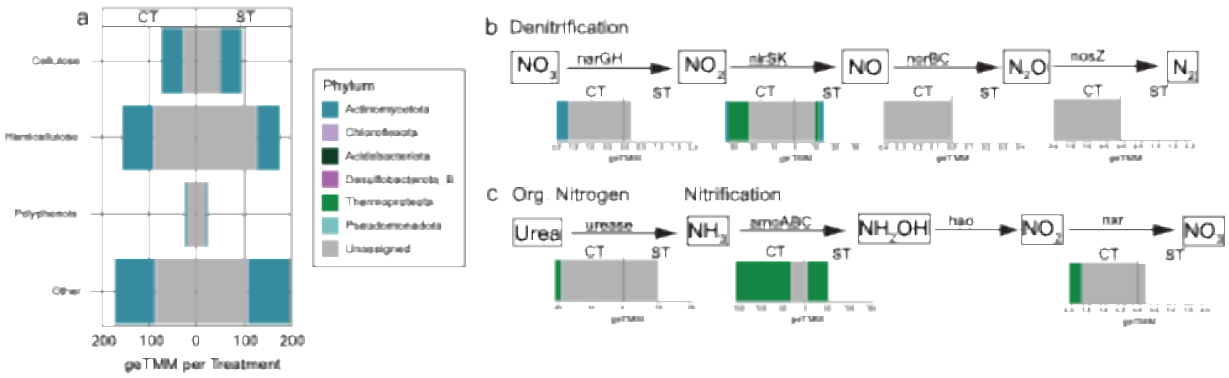


Figure 6. Expression of genes involved in carbohydrate and nitrogen cycling across tillage treatments. a Summed geTMM values for genes involved in the degradation of carbohydrates, grouped into cellulose, hemicellulose, polyphenols, and other substrates. Bar height indicates total expression (geTMM), with colors representing the phylum assignment. Grey bars indicate reads unassigned to metagenome-assembled genomes (MAGs), while colored bars are assigned. Bars are mirrored across the y-axis by treatment, with CT (conventional tillage) on the left and ST (strip tillage) on the right. Summed geTMM expression of key genes in **b** denitrification and **c** organic nitrogen and nitrification pathways. Boxes represent steps in the pathway, with bars underneath showing total gene expression (geTMM) by treatment. Colored bars represent gene expression assigned to phulym; grey bars represent unassigned reads. Genes include *narGH*, *nirSK*, *norBC*, and *nosZ* (denitrification), *urease* (organic nitrogen mineralization), and *amoABC*, *hao*, and *nxr* (nitrification).

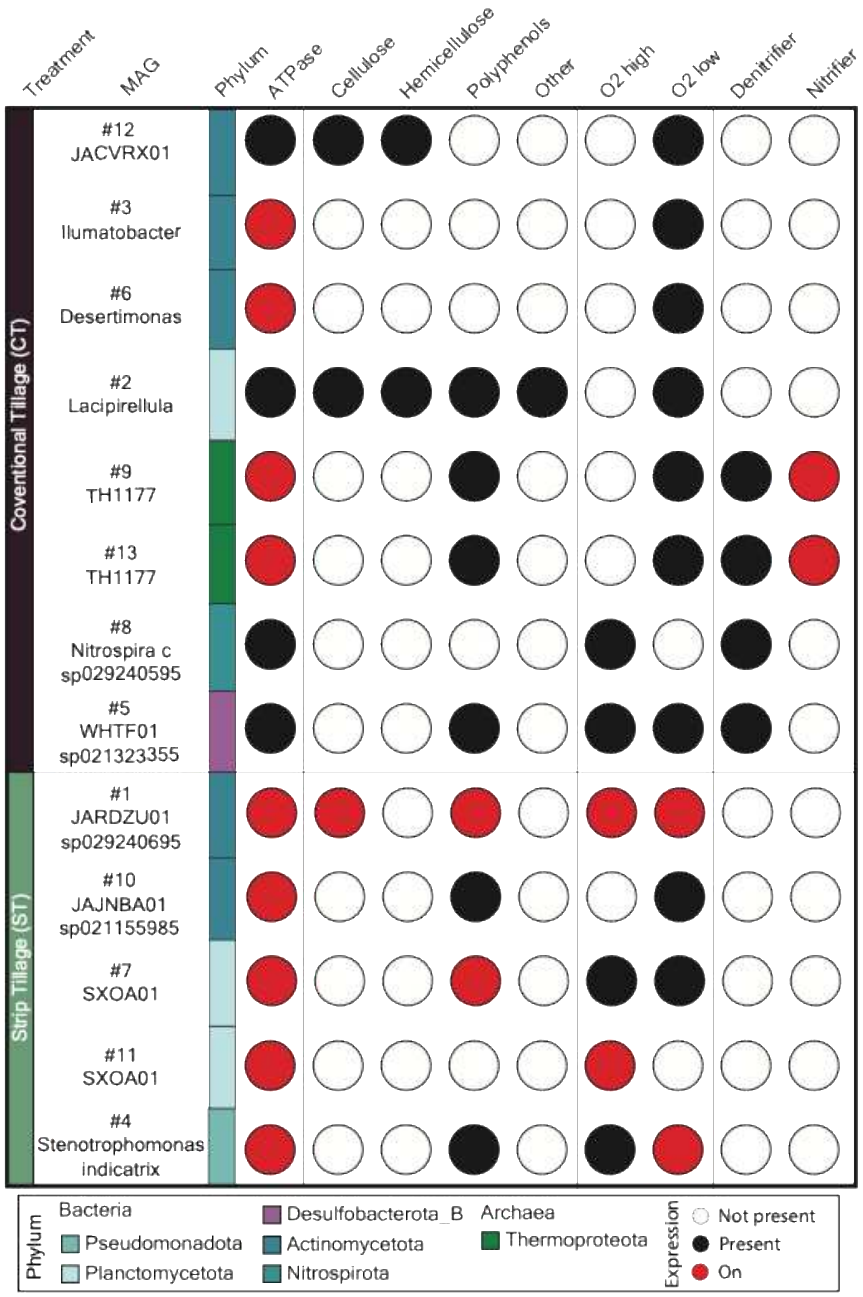
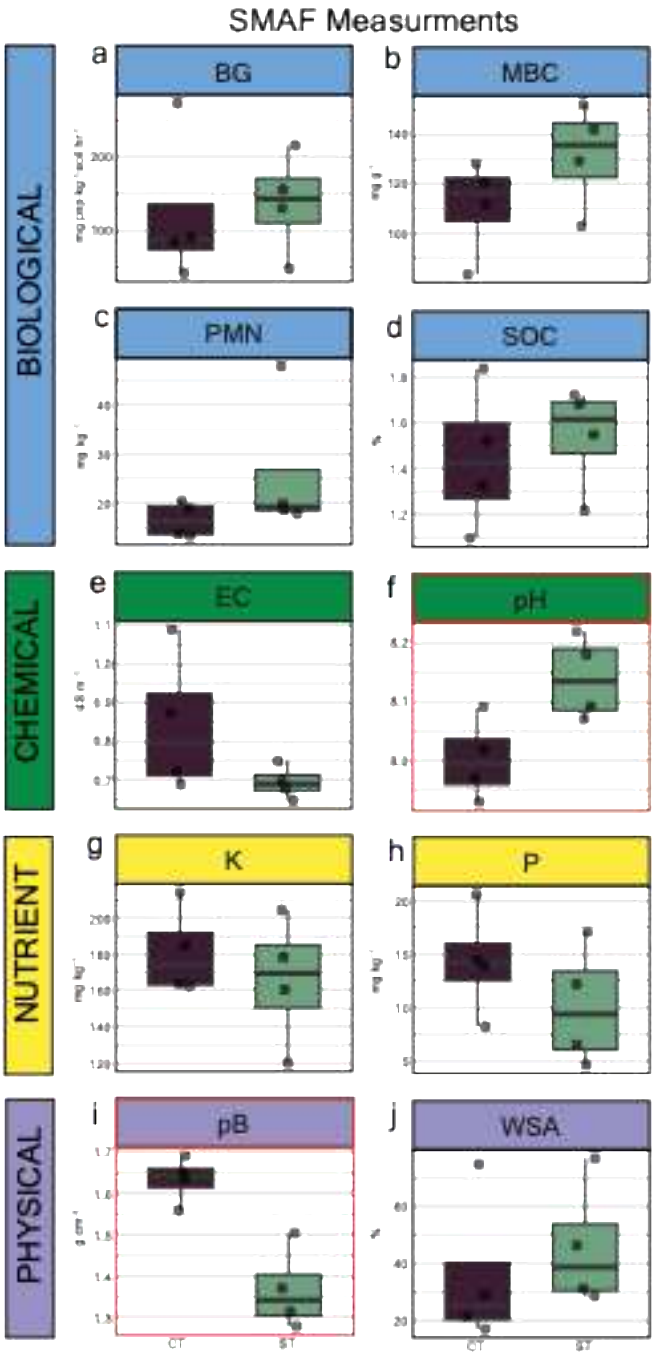


Figure 7. Expression of selected nitrogen cycling genes in 13 metagenome-assembled genomes (MAGs) that were significantly differentially abundant between conventional tillage (CT) and strip tillage (ST) treatments. MAGs are grouped by phylum, with bacterial and archaeal classifications indicated by color along the vertical bar. Gene presence is indicated by red circles, and absence by white circles. Expression was detected using metatranscriptomic read mapping, and selected genes span the organic nitrogen transformation, nitrification, and denitrification pathways.



Supplemental Figure 8. Tillage effects on SMAF soil health indices are limited to pH and bulk density. Boxplots show the distribution of ten soil health indicators across conventional till (CT) and strip till (ST) treatments. Each boxplot displays the median (central line), interquartile range (box), and whiskers extending to 1.5 times the interquartile range. Points represent individual soil sample replicates (n = 4 per treatment). **a** β-glucosidase (BG), **b** microbial biomass carbon (MBC), **c** potentially mineralizable nitrogen (PMN), **d** soil organic carbon (SOC) **e** electrical conductivity (EC), f pH, **g** Olsen potassium (K), **h** Olsen phosphorus (P), **i** bulk density (pB), **j** water-stable aggregates (WSA). Statistical significance was assessed using one-way ANOVA; significant differences ($p < 0.05$) between treatments are indicated in red outline.