DISS. ETH NO. 30203

# Mechanisms that link climate, soil properties and microbes to macroscale soil organic carbon dynamics: Insights from a geoclimatic gradient

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES

(Dr. sc. ETH Zurich)

presented by

*DANIEL WASNER*

*MSc. Ecology & Ecosystems, University of Vienna, Vienna, Austria*

Born on 24.02.1991

Accepted on the recommendation of

*Sebastian Doetterl*

*Pascal Boeckx*

*Joerg Schnecker*

*Samantha Weintraub-Leff*

2024

# Acknowledgements

I want to thank everyone who has supported me on this journey, both professionally and personally. Thanks to my supervisors, Sebastian, Jörg and Pascal. I thank you for all the trust and the scientific freedom that you gave me, and for all the valuable advice that you shared with me. It was great to learn from you, and I hope that these bonds will stay. It was a wild ridge up that mountain, at times it felt pretty darn exposed and tiring, lots of ups and downs, circles, detours, brittle rocks and rappels. On the other hand there were also plenty of amazing views, lots of sweet solid rock and fun climbs. To stick with the metaphor, you taught me how to use map and compass to always find the way. And, even more importantly, you taught me that there is always a way (forward, and up that science mountain, eventually). I certainly saw and learned more than if I had taken the cable-car, that's for sure! All that experience has become part of my scientific parent material, and it will never weather away. Thank you for that! I also want to sincerely thank Samantha for taking the time and effort to evaluate my thesis as the external examiner. Someone needs to check that I was actually on the summit, after all!

I want to thank the entire soil resources team, past and present, every single one of you! The friendship, support, interest and team spirit over the years was very beautiful. Thinking out loud together and throwing around ideas with you was one of my favorite aspects of work. Laura, Marco and Annina, you were around from the beginning until the end, I find that very cool!

I want to thank the co-authors, collaborators and colleagues who shared ideas with me, and who let me be part of their work as well. Being able to take part in other projects was very stimulating, and I appreciate it very much. I learned so many things that I would not have learned otherwise. Thank you, it really means a lot to me! Thank you also to Estelle and her team for hosting me so kindly in State College. I wish I could have stayed longer, maybe I can come and visit again!

No text can be written alone in a satisfying way (which always became evident when Sebastian started editing my drafts, and they turned into red strawberry fields). I therefore want to thank Lu, Inna, Moritz, Sophie and Marcus for their helpful feedback on sections of this thesis over the last weeks!

I want to thank all my friends in Switzerland, Austria and all over the place, the entire Wasner tribe and the syncopation climbing club (which is of course implicit in the friends section, but we all know that synco deserves special mention in the context of this thesis)! If any of you ever happens to stumble across these words, you know that I mean you, and I hope that you know how much you mean to me (I also hope that there are no more sentences in this thesis that are as complicated as this one). Without

keeping the balance between work and life, work would not always have been fun. And you, my friends, are the reason why the life part has been so overwhelmingly beautiful and rewarding and great and will always be. Thank you!

At that point, I also want to thank the Alps. I could not be without you mountains, so I had to write this here. Although you will probably never read it. Which is ok. Because you are just mountains, after all. But still. If there were no mountains, there would be no thesis. Very simple. Thank you!

Mama and Papa, you always encouraged me on my way and supported my choices and decisions. You opened all these doors wide, and I just had to choose one and start walking. I am very grateful for this. I also think that curiosity is one of my dearest treasures, and I think it was a gift from you. I would not have done any of this if there hadn`t been this curiosity. I thank you for that. In addition, I have to say that it is quite remarkable that you never questioned the fact that my formal education took 23 (?!) years. …Or did you?

My Lu. I could not have done any of this without you. Without your boundless support, help, patience and love, I would have either given up or I would have become hopelessly obsessed. Work and science and soil and all of that are nice, but the nicest part of the day is always when we come home, and when we are in our world. I cannot thank you enough. I love you as much as there are microbes in all of the worlds soil together!

PS: I just realized that in my metaphor from the beginning I am still stuck on that mountain top. Today I might actually take the cable-car down... Everyone look away!

# Table of Contents

# English summary

Adequate representation of the soil organic carbon (SOC) cycle at the macroscale (km to global) is important in order to understand and predict the implications of SOC dynamics for the climate and our planet. Such adequate representation is challenging, because mechanistic knowledge about soils is derived from the much smaller mesoscale (mm to m). However, the dominant mechanisms and controls of the SOC cycle vary across scales. Therefore, specific information is needed about which mesoscale processes affect SOC dynamics at the macroscale. This thesis uses a geoclimatic gradient of temperate grassland soils to investigate whether three well established SOC cycle mechanisms at the mesoscale translate to the macroscale, and therefore require scaling. The mechanisms of interest are (1) stable microaggregates as an SOC reservoir, (2) competitive exclusion and substrate specialization as determinants of bacterial community composition, and (3) microbial community composition as a driver of biomass-specific carbon (C) metabolism.

In Chapter 1, I first describe the role of SOC in the terrestrial C cycle and provide an overview of the three fundamental spatial scales at which the SOC cycle can be studied. I explain why the mechanisms that are relevant for the SOC cycle can differ across these scales and introduce the concept of "scaling" to bridge from meso- to macroscale. I then summarize the current consensus view of SOC dynamics at the macroscale and identify the knowledge gaps that this thesis will address. I provide a brief overview of methods which can be used to inform scaling and explain why gradient studies are particularly well suited for this purpose. Lastly, I introduce the geoclimatic gradient with which I worked and present the detailed research questions of this thesis.

In Chapter 2, together with coauthors I investigated whether stable microaggregates – a mechanism of spatial organization within soil – constitute a quantitatively relevant SOC fraction at the macroscale. For this, we applied a fractionation scheme which separates SOC into particulate organic matter, silt- and clay-sized particles and stable microaggregates. We found that stable microaggregates contained a large fraction of SOC, with environmental drivers and chemical characteristics that were distinct from particulate organic matter and silt- and clay-sized fractions. We concluded that stable microaggregates merit scaling to the macroscale.

In Chapter 3, we investigated whether competitive exclusion and bacterial substrate specialization – mechanisms of microbial community assembly – provide scalable links between soil bacterial community composition and the quantity as well as qualitative characteristics of SOC. For this, we measured qualitative characteristics of SOC and characterized bacterial community composition with

sequencing of the 16S rRNA gene. We found patterns in line with the mechanism of competitive exclusion but could not conclusively rule out other potentially underlying mechanisms. In addition, we did not find evidence that bacterial substrate specialization translated directly to the macroscale as a mechanism of bacterial community assembly.

In Chapter 4, we investigated how soil microbial traits and functions – which are products of microbial C metabolism – relate to the environment and to soil microbial community composition at the macroscale. For this, we measured microbial traits and functions important for the SOC cycle and characterized climatic and soil physicochemical conditions as well as bacterial and fungal community composition. We found that respiration and growth normalized for microbial biomass (*i.e.*, biomass-specific) were related to different features of microbial community composition, which resulted in strong effects on microbial C use efficiency. We concluded that biomass-specific C metabolism is a mechanism that merits scaling to the macroscale.

In Chapter 5, I compare the SOC dynamics of two contrasting systems along the geoclimatic gradient of temperate grassland soils in order to summarize the findings of Chapters 2 to 4, and to highlight interdisciplinary links among them. I then discuss the general conclusions from each of these Chapters in the context of the current state of the art. I further elaborate the implications of the findings for the respective research fields and propose future research directions that build on the insights of this thesis. Following these perspectives, I briefly provide guidance on how to place the findings of this thesis in a global context. Lastly, I conclude with a general outlook about future research with this dataset.

# German summary

Um die Auswirkungen des organischen Kohlenstoffs im Boden (SOC) auf das Klima und unseren Planeten verstehen und vorhersagen zu können, ist es wichtig, den SOC-Kreislauf auf der Makroskala (km bis global) angemessen darzustellen. Eine solche Darstellung ist herausfordernd, da mechanistisches Wissen über Böden auf der viel kleineren Mesoskala (mm bis m) gewonnen wird. Die vorherrschenden Mechanismen und beeinflussenden Faktoren des SOC-Kreislaufs variieren jedoch je nach Maßstab. Daher werden Informationen darüber benötigt, welche Prozesse der Mesoskala die SOC Dynamik auf der Makroskala beeinflussen. In dieser Arbeit wird anhand eines geoklimatischen Gradienten temperater Graslandböden untersucht, ob drei auf der Mesoskala gut etablierte Mechanismen des SOC-Kreislaufs auf die Makroskala übertragbar sind. Die Mechanismen von Interesse sind (1) stabile Mikroaggregate als SOC-Reservoir, (2) Konkurrenz-Ausschluss und Nahrungsspezialisierung als Treiber der Zusammensetzung bakterieller Gemeinschaft und (3) die Zusammensetzung mikrobieller Gemeinschaft als Treiber für den biomassespezifischen Kohlenstoff(C)-Stoffwechsel.

In Kapitel 1 beschreibe ich zunächst die Rolle des SOC im terrestrischen C-Kreislauf und gebe einen Überblick über die drei grundlegenden räumlichen Skalen, auf denen der SOC-Kreislauf untersucht werden kann. Ich erkläre, warum sich die relevanten Mechanismen für den SOC-Kreislauf auf diesen Skalen unterscheiden können, und führe das Konzept der "Skalierung" ein, um eine Brücke von der Meso- zur Makroskala zu schlagen. Anschließend fasse ich den derzeitigen Konsens über die SOC-Dynamik auf der Makroskala zusammen und zeige diejenigen Wissenslücken auf, die diese Arbeit schließen wird. Ich gebe einen kurzen Überblick über Methoden, die verwendet werden können, um die Notwendigkeit der Skalierung von Mechanismen zu beurteilen. Ausserdem erkläre ich, warum Gradientenstudien für diesen Zweck besonders gut geeignet sind. Abschließend stelle ich den dieser Arbeit zugrundeliegenden geoklimatischen Gradienten vor, und erläutere die detaillierten Forschungsfragen dieser Arbeit.

In Kapitel 2 untersuchte ich gemeinsam mit Koautoren, ob stabile Mikroaggregate - ein Mechanismus der räumlichen Organisation im Boden - eine quantitativ relevante SOC Fraktion auf der Makroskala darstellen. Dazu verwendeten wir ein Fraktionierungsschema, das den SOC in grobes organisches Material, schluff- und tonartige Partikel sowie stabile Mikroaggregate aufteilt. Wir fanden heraus, dass stabile Mikroaggregate einen großen Anteil des C eines Bodens enthalten können. Dabei unterscheiden sie sich vom groben organischen Material und von schluff- und tonhaltigen Partikeln durch ihre

chemischen Eigenschaften und durch die beeinflussenden Umweltfaktoren. Wir kommen zu dem Schluss, dass stabile Mikroaggregate auf der Makroskala repräsentiert werden sollten.

In Kapitel 3 untersuchten wir, ob bakterieller Konkurrenz-Ausschluss und Nahrungsspezialisierung skalierbare Verbindungen zwischen der Zusammensetzung der bakteriellen Bodengemeinschaft und der Menge und Qualität des SOC darstellen. Beide Mechanismen sind auf der Mesoskala wichtig für den Aufbau von Bakteriengemeinschaften. Zu diesem Zweck charakterisierten wir die qualitativen Merkmale des SOC sowie die Zusammensetzung der bakteriellen Gemeinschaft mittels Sequenzierung des 16S rRNA-Gens. Wir fanden Muster, die mit dem Mechanismus des Konkurrenz-Ausschusses übereinstimmen. Allerdings können wir nicht gänzlich ausschließen, dass andere Mechanismen für diese Muster verantwortlich sind. Wir fanden keine Hinweise darauf, dass sich die bakterielle Nahrungsspezialisierung als Mechanismus für den Aufbau von Bakteriengemeinschaften direkt auf die Makroskala übertragen lässt.

In Kapitel 4 untersuchten wir, wie Eigenschaften und Funktionen der Bodenmikroben - Produkte des mikrobiellen C-Stoffwechsels - mit der Umwelt und der Zusammensetzung der mikrobiellen Bodengemeinschaft auf der Makroskala zusammenhängen. Dazu wurden mikrobielle Eigenschaften und Funktionen gemessen, die für den SOC-Kreislauf wichtig sind. Genauer wurden die klimatischen und bodenphysikochemischen Bedingungen sowie die Zusammensetzung der Bakterien- und Pilzgemeinschaft charakterisiert. Wir fanden heraus, dass Atmung und Wachstum, normiert auf die mikrobielle Biomasse (d. h. biomassespezifisch), mit verschiedenen Merkmalen der mikrobiellen Gemeinschaftszusammensetzung zusammenhängen. Das wirkt sich stark auf die mikrobielle C-Nutzungseffizienz aus. Der biomassespezifische C-Stoffwechsel ist daher ein Mechanismus, der auf die Makroskala übertragen werden sollte.

In Kapitel 5 vergleiche ich die SOC-Dynamik von zwei kontrastierenden Systemen entlang des geoklimatischen Gradienten, um die Ergebnisse der Kapitel 2 bis 4 zusammenzufassen und Verbindungen zwischen ihnen aufzuzeigen. Darüber hinaus diskutiere ich die allgemeinen Schlussfolgerungen aus jedem dieser Kapitel im Kontext des aktuellen Stands der Wissenschaft. Des Weiteren erläutere ich die Implikationen der Ergebnisse für die jeweiligen Forschungsbereiche und schlage zukünftige Fragestellungen vor, die auf den Erkenntnissen dieser Arbeit aufbauen können. Im Anschluss daran gebe ich einen kurzen Leitfaden, wie die Ergebnisse dieser Arbeit in einen globalen Kontext gestellt werden können. Abschließend gebe ich einen allgemeinen Ausblick auf zukünftige Forschungsmöglichkeiten mit diesem Datensatz.

# Abbreviations

| | |
|---|---|
| ASV | Amplicon sequencing variant |
| BG | ß-Glucosidase |
| BP | Breusch-Pagan test |
| C | Carbon |
| CB | Cellobiosidase |
| CFE | Chloroform fumigation extraction |
| CUE | Carbon use efficiency |
| D.f. | Degrees of freedom |
| Decomp. | Decomposition |
| DI | Decomposition index |
| DOPA | L-3,4-dihydroxyphenylalanine |
| DRIFTS | Diffuse reflectance infrared fourier transformed spectroscopy |
| EE | Extracellular enzymes |
| ENV | Environmental setting |
| EPS | Extracellular polymeric substances |
| FDR | False discovery rate |
| FUN | Microbial function |
| MAOM | Mineral-associated organic matter |
| MAP | Mean annual precipitation |
| MAT | Mean annual temperature |
| MBC | Microbial biomass carbon |
| MCN | Microbial C:N ratio |
| MIC | Microbial community |
| MUF | 4-Methylumbelliferyl |
| N | Nitrogen |
| NAG | N-acetyl-β-glucosaminidase |
| NPP | Net primary productivity |
| OM | Organic matter |
| OXA | Ammonium oxalate |
| PCA | Principal component analysis |
| PCA | Principal component |
| Ped. oxides | Pedogenic oxides |
| PEEA | Potential extracellular enzyme activity |
| PET | Potential evapotranspiration |
| PEX | Peroxidase |
| POM | Particulate organic matter |
| POX | Phenoloxidase |
| PP | Sodium-pyrophosphate |
| qPCR | Quantitative polymerase chain reaction |
| RC | Rotated component |
| RMSE | Root mean squared error |
| RPA | Relative peak area |
| rPCA | Rotated principal component analysis |
| S+C | Silt- and clay-sized particles |
| SA | Stable microaggregates |
| SOC | Soil organic carbon |

| | |
|---|---|
| SOM | Soil organic matter |
| Spec. | (Biomass-)specific |
| SW | Shapiro-Wilk test |
| UV-Vis | Ultraviolet-visible spectroscopy |
| WHC | Water holding capacity |
| WRB | World reference base for Soil Resources |

# 1. Introduction

The overall aim of this thesis is to contribute to an improved understanding of soil organic carbon (SOC) dynamics at the macroscale. In Chapter 1, I will introduce the reader to the general topic of this thesis, to the questions that motivated this work, and to the approaches that were used to address these questions. I will first describe the role of SOC in the terrestrial carbon (C) cycle and introduce the fundamental factors that regulate the SOC cycle (Chapter 1.1). Next, I will provide an overview of the three fundamental spatial scales at which the SOC cycle can be studied, namely the micro-, meso- and macroscale (Chapter 1.2). The mechanisms that are relevant for the SOC cycle differ between these scales, and I will highlight the causes of these discrepancies and introduce the concept of "scaling" to bridge scales (Chapter 1.3). After providing this context, I will summarize the current consensus view of SOC dynamics at the macroscale (Chapter 1.4) and identify the knowledge gaps that this thesis will address (Chapter 1.5). This thesis will focus on the question whether three mechanisms that are well established at the mesoscale translate to the macroscale, and thereby merit scaling. The mechanisms of interest are (1) stable microaggregates as an SOC reservoir, (2) competitive exclusion and substrate specialization as determinants of bacterial community composition, and (3) microbial community composition as a driver of biomass-specific C metabolism. In Chapter 1.6 I will provide a brief overview of methods which can be used to inform the scaling of mechanisms, and I will give a rationale why gradient studies are particularly well suited for this purpose. In Chapter 1.7 I will introduce the geoclimatic gradient with which I worked in this thesis. Lastly, I will present the detailed research questions and hypotheses that I addressed, together with an outline of the synthesis (Chapter 1.8).

## 1.1 The role of soil organic carbon in the terrestrial carbon cycle

Globally, soils contain 1700 Gt SOC in the first meter, which is more than the stocks of terrestrial vegetation and the atmosphere combined (450 and 885 Gt C, respectively; Canadell et al., 2021). To most people, soil is just dirt, and gigatons are abstract numbers. An analogy might help to visualize the amount and tremendous value of global topsoil SOC. Only the upper 30 cm of the world's soil contain as much SOC (680 Gt C, FAO-UNESCO, 2018) as a pure diamond cube with edges of 6 km length! But why should SOC be considered that valuable? Through photosynthesis, terrestrial ecosystems are estimated to remove 2 – 4 Gt C per year from the atmosphere by conversion into C (Friedlingstein et al., 2023). On the timescales of years to centuries, much of this plant organic C passes through the soil. In soil, this organic C can either be converted back into $CO_2$ and return to the atmosphere, or it can become a more permanent part of the large SOC pool. Depending on a balance of intricate soil processes, soils can therefore either be a sink or a source of atmospheric $CO_2$. As such, SOC has the

potential to affect climate change through slowing down or accelerating the increase of atmospheric $CO_2$ concentrations. SOC is therefore a linchpin of the global C cycle, and the potential of soils to mitigate climate change has received much attention over the last years (e.g. Paustian et al., 2016, 2019; Amelung et al., 2020).

The specific processes and underlying mechanisms that regulate the fate of SOC depend on four fundamental factors: (1) the environmental setting, (2) the chemical characteristics of SOC, (3) the soil microbial community and (4) microbial functions. The environmental setting comprises the physical, chemical and biological context in which SOC is embedded, from the physicochemistry of a soil pore to the temperature regime of a climate zone. In addition to this environmental diversity, SOC can be present in soils at contents spanning two orders of magnitude, and with highly diverse physicochemical characteristics (Lehmann and Kleber, 2015). SOC is often investigated together with nitrogen (N) because biological cycling of C and N can be intertwined. When referring to properties that describe both C and N (for instance the C:N ratio), the term soil organic matter (SOM) is therefore used. To refer to the diverse physicochemical characteristics of SOC or SOM, the term “quality” is used. In part as a consequence of the tremendous breadth of environmental settings and SOM quality, soils are among the most biologically diverse ecosystems of the planet (Fierer, 2017). Soil microbial communities can therefore vary strongly in their functional composition and role (Don et al., 2017; Lehmann et al., 2020; Philippot et al., 2024). This is particularly central for SOC dynamics, because microbes carry out a large number of functions that are involved in SOC cycling. Subject to environmental and microbial controls, microbial functions are at the nexus between the environmental setting, SOM quality and the microbial community composition. The true key to understanding SOC dynamics – and thereby to harness the value of SOC in the global C cycle – is to understand how these four fundamental factors interact. These interactions are context-specific and depend on the scale. To better understand this scale-dependency, we first need to familiarize ourselves with the spatial scales of the SOC cycle.

## 1.2 The spatial scales of the soil organic carbon cycle

The SOC cycle can be described at three fundamentally different spatial scales (Figure 1-1) (Hinckley et al., 2014; O’Rourke et al., 2015; Wieder et al., 2015; Pachepsky and Hill, 2017; Blankinship et al., 2018). (1) The microscale, at which most biogeochemical processes take place. (2) The mesoscale, at which most processes in soil biogeochemistry are commonly measured. (3) The macroscale, at which large scale phenomena such as land use change or climate change are studied, and at which policy making needs to be informed. This spatial hierarchy typically covaries with temporal variability (Hinckley et al., 2014; Pachepsky and Hill, 2017). Mechanisms at the microscale can occur within seconds to minutes,

whereas mesoscale studies typically investigate processes over days to months. SOC dynamics at the macroscale are ultimately integrated over timescales from seasons to years or are even the product of processes that play out over millennia. While I acknowledge this temporal variability, I will focus the framing of this thesis on the spatial scales.

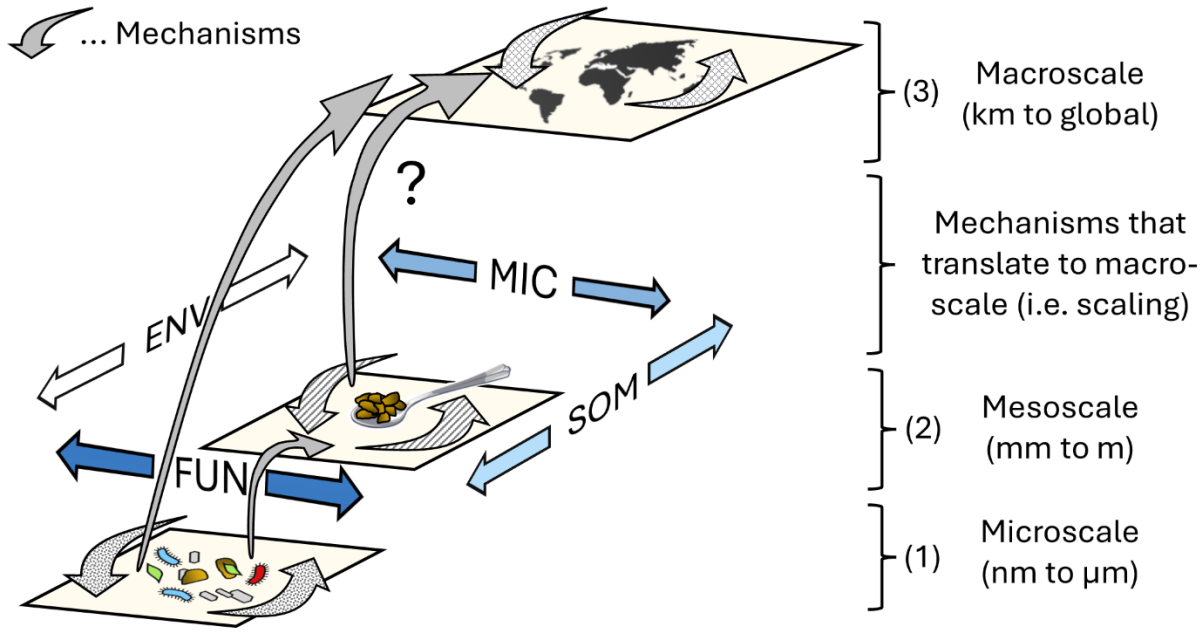



**Figure 1-1.** Schematic representation of the three fundamental spatial scales at which SOC cycling can be described: (1) The microscale, at which most biogeochemical processes take place. (2) The mesoscale, at which most SOC cycling processes are commonly measured. (3) The macroscale, at which the SOC cycle intersects with interdisciplinary aspects of global change and with policy making. SOC cycling takes place within diverse arrangements of the environmental setting (ENV), soil organic matter quantity and quality (SOM), microbial communities (MIC) and microbial functions (FUN). Grey arrows represent mechanisms that either operate within or across scales.

The microscale ranges from nanometers to micrometers. At this scale, chemical interactions such as the formation of OM-mineral associations take place, and microbes grow, respire and interact. Soil physicochemical characteristics such as soil structure directly affect individual biogeochemical reactions. At this point, I briefly want to define the term “mechanism”: Mechanisms are systems of causally interacting processes that produce regular and therefore predictable changes. Processes in turn are the cumulative result of underlying biological, physical or chemical mechanisms. Having defined what mechanisms are, it becomes clear that most fundamental mechanisms take place at the microscale. For example, physical mechanisms govern the distribution and movement of water, which affects microbial activity at pore scale via dynamic changes of aeration, connectivity and the spatial

arrangement of reactants. Targeted intervention on soil functioning always requires fundamental mechanistic understanding (Blankinship et al., 2018; Baveye, 2023). The microscale is therefore the scale at which we strive to understand the fundamental mechanisms that underly the SOC cycle (Smercina et al., 2021). As a consequence, this scale is subject to intense mechanistic investigation, mostly with spatially resolved imaging methods (Vos et al., 2013; Baveye et al., 2018; Smercina et al., 2021; Schweizer, 2022; Li et al., 2023) or microsensors and planar sensors (Pedersen et al., 2015). A large part of research at the microscale has been conducted within the confines of disciplines. However, interdisciplinary efforts increasingly manage to link soil physics, chemistry and biology, contributing to a holistic understanding of the microscale (Vos et al., 2013; Nunan, 2017; Tecon and Or, 2017; König et al., 2020).

The mesoscale ranges from millimeters to meters. A large part of soil biogeochemical research takes place at this scale. For example, many common analytical procedures use milligrams to grams of soil material to characterize pools and fluxes of C and nutrients, or to characterize microbial communities. Also at the mesoscale, soil profiles provide information on local soil formation, and experimental manipulations to understand mechanisms in plant-soil systems are conducted at most in plots of several meters. As a result, most information that feeds into macroscale models is measured at the mesoscale (Jungkunst et al., 2022). Mesocale measurements have several advantages: On the one hand, the mesoscale is practical in terms of feasibility, and on the other hand it is relevant because it is the smallest scale at which soils are typically managed (O'Rourke et al., 2015). However, observations at the mesoscale neither capture individual biogeochemical reactions, nor the exact microscale conditions in which they take place. Rather, observations at the mesoscale integrate units of soil (e.g. defined by soil layer or horizon, or by proximity to plant roots) and describe processes and properties that are weighted averages of these units or that emerge from the sum of multiple individual reactions within these units.

The macroscale ranges from kilometers to the global system. One major aim of studying the SOC cycle at the macroscale is to generate insights that can inform decision making. How should we manage soil systems? How will soil systems respond to global change? How do soil systems affect global change? It is evident why proper understanding of the SOC cycle at the macroscale is pressing and important. At the macroscale, information is usually extrapolated based on observations on the mesoscale. This can either be done through increasingly sophisticated machine-learning approaches (e.g. Hengl et al., 2017; Sanderman et al., 2021; Wang et al., 2024) or through numerical modelling (Jones, 2021). Machine-learning approaches are limited to the data range on which they were trained, because they only represent observed correlative connections instead of mechanisms. Such models are therefore not

reliable tools to predict future scenarios (Wadoux et al., 2020; Grunwald, 2022; Meyer and Pebesma, 2022; Baveye, 2023). For this, numerical models are necessary. Numerical models predict the behavior of systems based on the mathematical representation of processes and underlying mechanisms.

In the soil system, we usually observe processes (or correlations as the results of processes) and make inferences about the underlying mechanisms, which we formulate as concepts or mathematical equations. This can either be in the form of general simple assumptions such as first-order reactions, in combination with biogeochemical principles (e.g. the Arrhenius equation for the temperature dependency of chemical reactions), or through explicit representation of mechanisms of varying complexity (Campbell and Paustian, 2015; Wieder et al., 2015; Chandel et al., 2023; Schimel, 2023). However, while such numerical models are typically applied at the macroscale, information about the mechanisms on which they are based are typically derived at the micro- and mesoscales. These mechanisms must therefore be transferred across scales, and often into other biogeochemical conditions than those from which they were derived empirically (Wieder et al., 2015). Next, we will explore why this is not always as simple and straight forward as it may seem.

## 1.3 Scaling the key mechanisms of the soil organic carbon cycle

The transfer of concepts and mechanisms across scales (typically from smaller to larger spatial scales, (Figure 1-1) is generally referred to as scaling (Pachepsky and Hill, 2017). Scaling represents a major challenge in soil science and related fields (Davidson et al., 2014; Hinckley et al., 2014; O'Rourke et al., 2015; Wieder et al., 2015; Luo et al., 2016; Baveye et al., 2018; Blankinship et al., 2018; Getz et al., 2018; Martín et al., 2021; Wan and Crowther, 2022). In the context of macroscale models, parameterization, calibration and validation can be difficult aspects of scaling, particularly because of gaps in empirical data (Luo et al., 2016; Cameron et al., 2018; Kögel-Knabner and Amelung, 2021; Jungkunst et al., 2022; Le Noë et al., 2023). However, an even more fundamental challenge is to decide which micro- and mesoscale mechanisms should be represented at the macroscale (i.e., model structure formulation, Luo et al., 2016; Abramoff et al., 2018). This is because the relevant mechanisms - and as a consequence the relevant controls - of the SOC cycle depend on the spatial scale (Ali et al., 2018; González-Domínguez et al., 2019; Wiesmeier et al., 2019; Li et al., 2020; Nave et al., 2021; Tian et al., 2022). On the one hand, certain mechanisms and controls only emerge at larger scales. Mechanisms can emerge at larger scales when they result from the (complex) interaction of units at smaller scales. For example, leaching as a vertical SOC flux between soil horizons (Kindler et al., 2011; Nakhavali et al., 2021) only becomes evident at depth-explicit mesoscales, and ecological interactions of microbes may affect process rates at the mesoscale in ways that are currently difficult to predict

from the microscale (Kaiser et al., 2014; Falconer et al., 2015; Cordero and Datta, 2016; Buchkowski et al., 2017; Georgiou et al., 2017; E. K. Hall et al., 2018). On the other hand, not all mechanisms and controls translate from the micro- and mesoscale to macroscale. The reasons for this can be manifold. Spatial heterogeneity of conditions such as substrate availability or moisture may neutralize opposed processes or average variable process rates at larger scales (Nunan, 2017). As a consequence, the underlying mechanisms can mask and neutralize each other. Controls may also affect SOC cycling hierarchically across spatial scales, with variable effect sizes. For example, at a given level of water saturation, microscale heterogeneity within one soil can cause microsite-differences in respiration rates of several orders of magnitude due to local resource limitation or physiological stress (Yan et al., 2023). However, at the macroscale, climatic and pedogenic differences cause large differences in bulk soil microbial biomass, which – again at a given level of water saturation - can lead to differences in bulk soil respiration that are almost at the order of two magnitudes (Chapter 4). As a consequence, the variation of respiration rates may be determined by pore architecture at the microscale, and by the geoclimatic setting at the macroscale.

While the scale-dependence of natural systems has been acknowledged for centuries (Pachepsky and Hill, 2017), scaling poses an ongoing challenge in the context of the SOC cycle. This is evident by the large structural diversity of SOC cycle models (Campbell and Paustian, 2015; Luo et al., 2016; Chandel et al., 2023; Garsia et al., 2023; Schimel, 2023), and even more by the weak consensus among model projections (Tian et al., 2015; Shi et al., 2018; Sulman et al., 2018; Wieder et al., 2019; Georgiou et al., 2021, 2024; Hashimoto et al., 2023). One source of this problem is that relevant micro- and mesoscale mechanisms are missing in our conceptual understanding at the macroscale. As boiled down by Luo et al. (2016): *"[…] the art […] is to determine what should be explicitly represented […] and what can be ignored."* In order to be able to identify such knowledge gaps, we next need to get an overview of how the SOC cycle is currently represented at the macroscale.

## 1.4 The soil organic carbon cycle at the macroscale

Given how complex soil systems are, a comprehensive and exhaustive review of macroscale SOC dynamics would probably require writing a book. This chapter therefore aims to provide a general overview of the current consensus representation of the SOC cycle at the macroscale. This overview was guided by the structures of numerical SOC models that were designed to represent SOC dynamics at the scale of ecosystems or in the context of earth system models (e.g. CENTURY, Parton et al., 1987; Roth-C, Jenkinson et al., 1990 ; RESOM, Tang and Riley, 2015; CORPSE, Sulman et al., 2014; MIMICS, Wieder et al., 2014; COMISSION, Ahrens et al., 2015; MEMS 2.0, Zhang et al., 2021; MEND, G. Wang et

al., 2022; Millennial v2, Abramoff et al., 2022). In several reservoirs of the SOC cycle it is common to consider C and N jointly, which is why the term SOM is occasionally used instead of SOC. This overview will focus on C dynamics alone, but the term SOM will be used where appropriate.

C can enter and exit soil via various processes. The most important source for C input into soils are plants. Plants supply C to soils via aboveground litterfall, root turnover which results in belowground litter, rhizodeposition and through symbiotic relationships with mycorrhizal fungi (Jackson et al., 2017; Basile-Doelsch et al., 2020; Huang et al., 2021). A large part of C input comes from roots, and therefore roots can cause hotspots of microbial activity (Sokol et al., 2019; Smercina et al., 2021). For this reason, several conceptual frameworks consider the rhizosphere (*i.e.* soil close to roots) and "bulk soil" (*i.e.* soil distant from roots) to be functionally distinct zones (Sulman et al., 2014; Zhang et al., 2021). The main losses of C from soil systems are through respiration (Raich et al., 2002; Trumbore, 2006) (or methanogenesis in anaerobic conditions), erosion (Doetterl et al., 2016; Naipal et al., 2018) and leaching (Kindler et al., 2011; Nakhavali et al., 2021). Vertical SOC fluxes may also be an important source for C in subsoil systems (Sanderman and Amundson, 2008; Kaiser and Kalbitz, 2012), although this is debated (Sierra et al., 2024).

SOC can have diverse physical and chemical characteristics. Molecules that contain C can for example have low or high molecular weight, they can have a low or a high degree of reduction, and they can be dissolved in soil solution, readily exchangeable from mineral surfaces or bound strongly to mineral surfaces over long timescales. As a consequence, not all SOC is equally likely to contribute to C storage and loss. To categorize this diversity, SOC is commonly classified into functionally distinct groups (Figure 1-2, ①). Such groups of SOC can be called "pools" when they are assigned with a turnover rate, or "fractions" when they are isolated with physicochemical separation methods. However, some groups of SOC do not fit properly into either terminology (e.g. microbial biomass C, MBC), which is why I will use the more general term "reservoir" in this section. SOC reservoirs with distinct dynamics are the foundation of most macroscale SOC concepts ever since seminal first-order decay models such as CENTURY (Parton et al., 1987) and Roth-C (Jenkinson et al., 1990) demonstrated their utility. With the emergence of process-explicit models, there have been attempts to align these conceptual reservoirs with measurable SOC reservoirs. The reservoirs that are most commonly considered are microbially available organic matter (defined as dissolved OM or extractable OM, Kalbitz and Kaiser, 2008), MBC (Crowther et al., 2019), mineral-associated organic matter (MAOM) and particulate organic matter (POM) (Lavallee et al., 2020). The contributions of the individual reservoirs to total SOC vary widely, depending on the soil system and on the quantification methods.

C enters the mineral soil either directly in the form of low molecular weight (LMW, < 600 Da) dissolved OM compounds, or as coarse litter material (Kleber et al., 2015; Sokol et al., 2019) (Figure 1-2, ②). Decomposition through microbial activity (via extracellular enzymes) depolymerizes litter and produces available OM and POM (Grandy and Neff, 2008; Cotrufo et al., 2015) (Figure 1-2, ③). Available OM can either be taken up by microbes or form MAOM through sorption to mineral components (Liang et al., 2017; Sokol et al., 2019, 2022b) (Figure 1-2, ④). POM and MAOM are conceptually considered and practically isolated as "counterparts" after dispersion of physical soil structures (Lavallee et al., 2020) (Figure 1-2, ⑤). Common definitions for separation are either size (e.g. POM > 63 µm > MAOM), by density (POM < 1.6 – 1.8 g $cm^3$ < MAOM) or a combination thereof. POM consists of coarse plant-derived organic material with a high (chemically inherent) activation energy and a high C:N ratio. MAOM primarily consists of LMW compounds of plant and microbial origin, with lower (chemically inherent) activation energies and low C:N ratio. While POM has no strong physical or chemical protection from microbial decomposition, MAOM is considered physically or chemically protected through association with clay minerals, Al and Fe oxy-hydroxides or cations (von Lützow et al., 2006; Rowley et al., 2018; Kleber et al., 2021). In general, C in POM is considered to cycle faster than C in MAOM (Trumbore, 2000; Heckman et al., 2022). There are further approaches to separate fast- and slow-cycling SOC pools which are independent of soil isolation into POM and MAOM. For example, data from RockEval thermal analysis has been used to predict SOC reservoirs with different turnover times (Cécillon et al., 2021). In practice, all reservoirs still contain a diversity of SOC compounds of variable microbial accessibility, activation energy and turnover. For example, despite the overall stable nature of MAOM (Heckman et al., 2022), parts of it can be labile (S. W. Stoner et al., 2023). As a consequence, numerical models usually represent sorption and desorption rates (e.g. Ahrens et al., 2015; Tang and Riley, 2015; Abramoff et al., 2022) or subdivide functional reservoirs further into variable pools (*e.g.* Sulman et al., 2014; Zhang et al., 2021).

The soil microbial community is a very complex part of the SOC cycle that requires drastic simplification at the macroscale (Figure 1-2, ⑥). Trait-based concepts are a popular approach to reduce the complexity of microbial community composition. Such concepts aim to classify soil microbes into functional groups that affect SOC cycling differently (Fierer et al., 2007; Krause et al., 2014; Ho et al., 2017; Malik et al., 2020; Wan and Crowther, 2022). In recent years, metagenomic methods and stable isotope probing have greatly accelerated this field (Li et al., 2019; Guo et al., 2020; Y. Chen et al., 2021; Stone et al., 2023). However, this avenue of research remains challenging because large parts of the microbial community are inactive at any given time (Blagodatskaya and Kuzyakov, 2013; Couradeau et al., 2019) due to adverse conditions at the microscale, and microbial functions are affected by diverse ecological interactions (Cordero and Datta, 2016; Buchkowski et al., 2017). The soil microbial

community receives much attention because it conducts several functions that cause important SOC transformations. Microbes exude extracellular enzymes, which depolymerize litter and POM in order to produce available OM, which serves as a substrate that is taken up by microbes (Sinsabaugh et al., 2014; Zuccarini et al., 2023) (Figure 1-2, ③④). In microbial metabolism, this substrate can either undergo catabolism to be converted into energy, or it can undergo anabolism to build up microbial compounds (Russell and Cook, 1995; van Bodegom, 2007) (Figure 1-2, ⑦). Catabolism (in aerobic conditions) results in the production of $CO_2$ through heterotrophic respiration. This process is a major route of C loss from soils and has therefore been a center of interest for decades (Bond-Lamberty et al., 2024). C that enters anabolism is invested into extracellular enzymes, extracellular polymeric substances (EPS) that protect and connect microbes and into growth (Manzoni et al., 2012a; Flemming, 2016; Costa et al., 2018; Hagerty et al., 2018). Microbial growth and death (also conceptualized as microbial turnover) generate microbial necromass, which is a key source of MAOM (Kallenbach et al., 2016; Wang et al., 2021; Sokol et al., 2022b) (Figure 1-2, ⑧).

The environmental setting controls many processes of the SOC cycle. Climate acts mainly through precipitation and temperature. Precipitation and temperature affect different components of the SOC cycle over different timescales, all of which are important at the macroscale. At short timescales (days to months), soil moisture and temperature directly affect soil microbial process rates. Moisture is necessary to provide connectivity at the microscale, and absence of water additionally causes physiological drought stress for microbial cells (Tecon and Or, 2017; Schimel, 2018; Malik and Bouskill, 2022). On the other hand, water saturation leads to anaerobic conditions, which require the microbial community to switch to less efficient types of metabolism (Keiluweit et al., 2016, 2017). As a consequence, moisture cycles reverberate from the micro- to the macroscale. Temperature directly affects microbial process rates because all biochemical and many physical processes in soil are positively temperature-dependent in the range of soil conditions (Lloyd and Taylor, 1994; Schipper et al., 2014). At intermediate timescales (season to years), climatic conditions affect the balance between plant inputs and microbial decomposition activity, which can lead to the accumulation of SOC, particularly in the POM reservoir (Wiesmeier et al., 2019; García-Palacios et al., 2021; Rocci et al., 2021). Climate is also a key determinant for the type of vegetation that grows in an ecosystem, and the type of vegetation in turn determines the quantity and quality of C inputs (Jobbágy and Jackson, 2000; Guo and Gifford, 2002; Meier and Bowman, 2008). C inputs in turn can also increase or decrease the mineralization of native SOC, an effect called priming. There are plenty of different mechanisms that can underly priming (Bernard et al., 2022). As a result of this complexity, observed priming patterns are variable (Bastida et al., 2019), the exact mechanisms at work often remain elusive (Liu et al., 2020) and the relevance for SOC cycling at longer timescales is not always clear (Schiedung et al., 2023b). At

long timescales (decades to millennia), climate co-determines which type of vegetation dominates a given ecosystem, and how rapidly soil parent material is transformed into secondary minerals through the process of weathering. In wet and warm conditions, chemical weathering proceeds faster than in dry and cold conditions (Jenny, 1941; Rasmussen et al., 2018; Slessarev et al., 2022). Climate is thereby a major driver of soil formation. As a consequence, soil mineralogy at the macroscale is mainly understood as the product of climate, parent material and soil age (Doetterl et al., 2018; Delgado-Baquerizo et al., 2020; Slessarev et al., 2022). Soil formation alters soil physicochemical properties, and thereby the context of SOC dynamics. Soil physicochemical properties (*e.g.* texture, soil pH) and mineralogical properties (e.g. mineral reactivity, surface area) strongly affect the amount of stabilized SOC (Doetterl et al., 2015; Rasmussen et al., 2018; Georgiou et al., 2022; Heckman et al., 2022), the qualitative characteristics of stabilized SOC (Zhao et al., 2020; Mainka et al., 2022; Yu et al., 2022) and SOC fluxes (Tang and Riley, 2015; Xu et al., 2016; Finke et al., 2019). In addition, soil pH is well established as the most important macroscale driver of soil microbial community composition (Fierer, 2017; Delgado-Baquerizo et al., 2018).

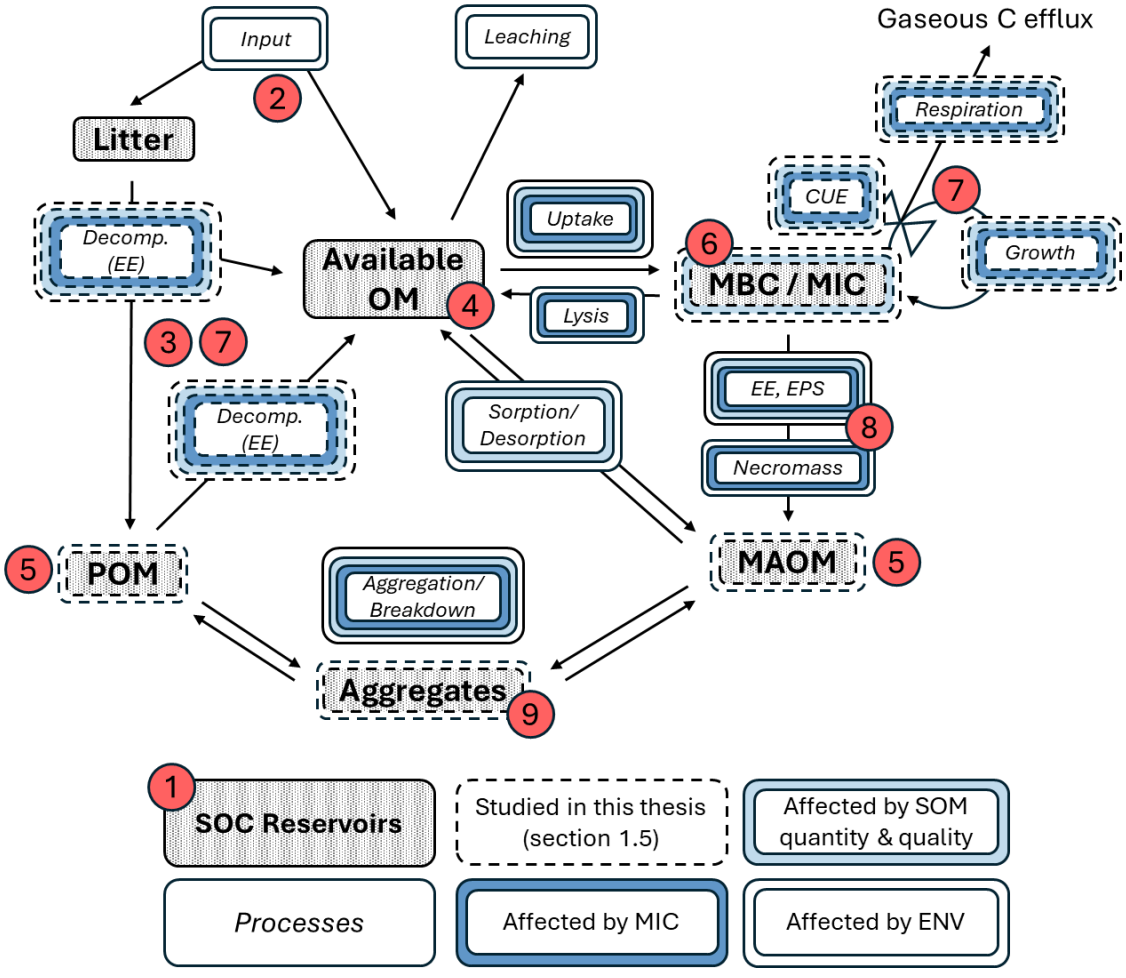



**Figure 1-2.** General overview of the current consensus representation of the SOC cycle of mineral soils at the macroscale. Microbial functions are represented explicitly as mechanisms. For simplification,

effects of ENV, SOM and MIC on mechanisms are not described explicitly. ENV = environmental setting, SOM = soil organic matter, MIC = microbial community composition. Decomp. = decomposition, EE = extracellular enzymes, CUE = carbon use efficiency, MBC = microbial biomass carbon, EPS = extracellular polymeric substances. The numbers in red circles (points ① to ⑨) are a visual aid to guide readers of Chapters 1.4 and 1.5 through the figure.

## 1.5 Knowledge gaps at the macroscale

Due to different foci and for the sake of simplification, none of the reviewed numerical models features all of the SOC reservoirs, mechanisms and controls which were highlighted in Chapter 1.4. Nevertheless, the relevance of these reservoirs, mechanisms and controls at the macroscale is widely accepted, and the described outline sketches the current consensus on how to conceptualize the SOC cycle at the macroscale (Figure 1-2). However, the relevance of several mechanisms which are recognized to be important at the mesoscale is yet unresolved or understudied at the macroscale.

Did we overlook scaling a relevant SOC reservoir? Stable microaggregates are soil structures which consist of diverse mineral and organic materials (including MAOM and POM), and which resist dispersion under natural conditions and are therefore bound together over long timescales (Tisdall and Oades, 1982; Totsche et al., 2018) (Figure 1-2, ⑨). There is considerable uncertainty regarding the potential role of stable microaggregates as an SOC reservoir at the macroscale. A conceptual framework that rapidly gained popularity in recent years promotes strong dispersal of physical structures before separation of SOC into MAOM and POM (Lavallee et al., 2020). The main rationales to disperse stable microaggregates between 63 and 250 µm size are to save labor, and to bypass the challenge of diverse fractionation schemes and difficult inter-study comparability. However, at the mesoscale it is well established that water-stable microaggregates constitute quantitatively important SOC fractions (Six et al., 2004; Totsche et al., 2018), and that SOC associated with aggregates follows distinct dynamics (Dungait et al., 2012; Schrumpf et al., 2013; Segoli et al., 2013; Laub et al., 2024). Moreover, distinct dynamics of stable microaggregates have also been shown at the macroscale (Poeplau and Don, 2013; Heckman et al., 2022; Edlinger et al., 2023). Nevertheless, few models and modelling approaches explicitly feature the mechanism of aggregate turnover or stable microaggregates as an SOC reservoir (Zimmermann et al., 2007; Abramoff et al., 2022). A systematic and methodically consistent investigation of the quantitative importance, qualitative characteristics and environmental controls at the macroscale is yet missing.

Is there a scalable link between quantity and qualitative characteristics of SOM and soil bacterial community composition? Soil bacterial alpha diversity (*i.e.* local diversity) and community composition are known to be strongly linked to soil pH and to the amount of SOC and N (jointly referred to as SOM) at the macroscale (Fierer, 2017). One proposed mechanism for this link is competitive exclusion, where taxa that are adapted to high substrate conditions (copiotrophic taxa) outcompete taxa that are adapted to low substrate conditions (oligotrophic taxa). This pattern commonly emerges at the mesoscale in response to substrate addition in experiments (e.g. Fierer et al., 2007; Geyer and Barrett, 2019; Stone et al., 2023). In simplified micro- and mesoscale study systems, bacterial taxa have also been shown to vary in their use of substrate with different qualitative characteristics (e.g. oxidation state, C:N ratio) (Goldfarb et al., 2011; Baran et al., 2015; Dolan et al., 2017). Soil bacteria may be evolutionary adapted to the preferential use of different substrates, which provides the basis for the mechanism of bacterial substrate specialization (Johnson et al., 2012; Trivedi et al., 2013; Y. Wang et al., 2022). However, it remains largely unexplored whether this mechanism translates to the macroscale. A tundra ecosystem model (SCAMPS, Sistla et al., 2014) successfully implemented a microbial community with shifting community composition in response to resource stoichiometry, but whether SOM quality affects soil bacterial diversity and community composition consistently across the macroscale is unresolved (Figure 1-2, ⑥).

Does biomass-specific microbial metabolism translate to the macroscale? The central role of soil microbes in the macroscale SOC cycle is well established. Microbial processes and properties (together referred to as microbial traits and functions) such as growth, respiration and extracellular enzymes mediate – or at least affect – almost all important SOC transformations (Figure 1-2, ⑦). Microbial traits and functions thereby actively drive the SOC cycle. However, there is considerable uncertainty how strongly the environmental setting, SOM quantity and quality and microbial community composition control these microbial traits and functions at the macroscale. Based on theoretical considerations of metabolic mechanisms, the rates of growth, respiration and extracellular enzyme activities normalized for MBC (*i.e.* biomass-specific) should at least partially depend on soil microbial community composition (van Bodegom, 2007; Roller and Schmidt, 2015; Hagerty et al., 2018; Malik et al., 2020). This is supported by experimental studies at the micro- and mesoscale (Domeignoz-Horta et al., 2020; Simon et al., 2020; T. P. Smith et al., 2021; Caro et al., 2023). However, to date it is unresolved to what degree microbial community composition affects microbial functions at the macroscale through metabolic mechanisms.

## 1.6 Gradient studies are a suitable tool to inform scaling

In the previous chapters we have identified the challenge of scaling mechanisms from small scales to the macroscale. We have further obtained an overview of the SOC cycle at the macroscale, and we have identified several knowledge gaps in our current understanding of the macroscale SOC cycle. Next, we need to briefly discuss the approaches that are available to address these knowledge gaps and to inform scaling. In general, mechanisms that merit scaling can be identified with numerical modelling, experimental research and various types of gradient studies. I will briefly elaborate the advantages and disadvantages of these approaches and provide a rationale why I think that geoclimatic gradients are particularly well suited to inform scaling.

Numerical models can be used as a tool to identify or test mechanisms that underlie complex systems (Campbell and Paustian, 2015; Marschmann et al., 2019; Le Noë et al., 2023). For this, hypothesized mechanisms are mathematically formalized, and comparison of model simulations with empirical observations can then be used to test them (Le Noë et al., 2023). However, the performance of numerical models is often hampered by the problem of equifinality (Marschmann et al., 2019; Schimel, 2023). Equifinality describes the phenomenon when multiple model structures produce patterns that fit the data equally well, because parameter calibration has multiple possible solutions (Sierra et al., 2015).

Scientific experiments are the gold-standard to reject or confirm hypotheses. Carefully designed experiments can be used to infer causality of processes, and therefore mechanisms. An experiment ideally manipulates the variable(s) of interest, and controls for all other variables. Because the complexity of soil systems increases with scale, it becomes increasingly harder to control experiments adequately as they are scaled up (Beier et al., 2012; De Boeck et al., 2015). Another challenge to use experiments for scaling is that sample sizes are often limited due to practical reasons. In order to increase statistical power, effect size can be increased through strong manipulation. However, action of a mechanism upon overly strong manipulation must not automatically imply that this mechanism is relevant at the macroscale under natural conditions. Similarly, observations from isolated experiments in the absence of environmental heterogeneity do not necessarily imply that the tested mechanisms are relevant in the context of environmental heterogeneity. Several of these challenges can be overcome with networks of experimental sites (Torn et al., 2015; Weintraub et al., 2019). Another tool to evaluate the macroscale relevance of experimental insights are meta-analyses. Sophisticated statistical tools allow for the comparison of results across experimental studies, but it remains challenging to cope with inconsistent methods and experimental approaches (Eysenck, 1994; Greco et

al., 2013; Mengist et al., 2020). In addition, meta-analyses are limited to the range of existing data, which does not always cover the full and unbiased range of relevant conditions.

Nested study approaches can help to link mesoscale proxies and mesoscale processes explicitly via the responsible mechanism at the microscale (e.g. Smercina et al., 2021; Mbé et al., 2022; Baveye, 2023; Ortega-Ramírez et al., 2023). One way to achieve this has been outlined in Baveye (2023): the soil microscale (*e.g.* the spatial organization) needs to be characterized on the exact same samples on which mesoscale processes of interest are measured. Next, microscale features which predict the mesoscale processes need to be identified. As a last step, other mesoscale features which can serve as proxies for the relevant microscale feature need to be found. This approach requires a high level of interdisciplinarity to combine sophisticated work at the micro- and mesoscale. In addition, it faces the same challenge as experiments: the effort required for such complex case studies practically obstructs the application across a large numbers of different soil systems.

Lastly, gradient studies can be used to investigate whether the effects of hypothesized mechanisms are reflected at the macroscale. For this, environmental gradients that represent relevant macroscale heterogeneity are selected. Along these macroscale gradients, the processes or properties of interest are measured on the micro- or mesoscale. In combination with underlying mechanistic knowledge, this approach allows to evaluate whether micro- and mesoscale mechanisms cause relevant patterns at the macroscale, i.e. whether they therefore “translate” to the macroscale. In addition, correlation analyses, regression approaches and data assimilation methods can be used to explore links between the spatial variability of target variables and relevant aspects of the environment.

In soil science, various types of gradients are common. The choice of gradient depends on the research question. Several regional types of gradients can control for environmental aspects that are not of interest, while maximizing those aspects that are in the focus of a research question. Chronosequences are used to study the effect of system age and are often interpreted with a space-for-time rationale (*e.g.* Bernasconi et al., 2011; Doetterl et al., 2018; Delgado-Baquerizo et al., 2020). Topographic sequences (catenas) are used to study lateral soil transfer and to efficiently capture local to regional soil diversity (Sommer and Schlichting, 1997; Schaetzl, 2013). Altitudinal gradients are used to maximize climatic differences while controlling for parent material (*e.g.* Whitaker et al., 2014; Rasmussen et al., 2018; De Jonge et al., 2024). Spatially nested sampling allows for direct comparison of variation and co-variation across the meso- and macroscale, or across several levels of macroscales (Corstanje et al., 2008; Fromin et al., 2013; Bradford et al., 2017; Doetterl et al., 2021; Nave et al., 2021; von Fromm et al., 2021). However, nested sampling requires large sample sizes. In a wider sense,

databases such as ISRaD for $^{14}C$ data (Lawrence et al., 2020), SRDB for soil respiration data (Jian et al., 2021) or SoDaH to harmonize soil data across research networks (Wieder et al., 2021) also constitute gradients. Databases collect measured data from different studies, but unlike in meta-analyses, data must not necessarily stem from experiments. The use of databases can come with the same challenges as meta-analyses, namely inconsistent methodology and limited (potentially biased) data range. (Geo-)climatic gradients combine several aspects of the approaches described above (Craine et al., 2010; Doetterl et al., 2015; Bradford et al., 2019; Hall et al., 2020; L. Chen et al., 2021; Yu et al., 2021). They are less spatially confined than regional gradients such as chronosequences, catenas or altitudinal gradients, but at the cost of higher environmental covariation. Networks of research sites (e.g. NEON United States, Weintraub et al., 2019; ICP Forests Europe, Schwaerzel et al., 2022) or soil monitoring networks (e.g. NABO Switzerland, Gubler et al., 2022; BZE Germany, Poeplau et al., 2020; LUCAS European Union, De Rosa et al., 2024) can also provide (geo-)climatic gradients, with the further benefits of available metainformation and research infrastructure. The main strength of (geo-)climatic gradients as a tool to inform scaling is that the variability of the variables of interest can be maximized. Depending on the research question, unbiased datasets that continuously range across natural variability can be sampled. Moreover, all types of gradient studies have the advantage that data collection can be planned a priori (in contrast to databases or meta-analyses). This allows to optimize the methods that are applied to address specific research questions, and it allows to apply consistent methodology. As an additional benefit, gradient studies thereby generate datasets that can be valuable for the parametrization and calibration of numerical models.

## 1.7 A Chilean geoclimatic gradient

In order to study SOC dynamics at the macroscale, we therefore developed a large geoclimatic soil gradient. Together with collaborators around Prof. Erick Zagal Venegas (Universidad de Concepción, Chile) we identified and selected a geoclimatic gradient of 35 sites along a 2300 km soil gradient. Site selection was aided by information from previous regional soil surveys by the Chilenean CIREN (Centro de Información de Recursos Naturales) (Figure 1-3). To constrain the gradient to a coherent type of ecosystem, sites had to be vegetated by extensive rangeland or natural grassland. With this choice, we controlled for consistent vegetation dynamics and a consistent type of OM input. In addition, sites were selected to be carbonate free (i.e. null HCl reaction) in order to ensure comparability across climatic conditions. Within this controlled type of system (grassland soils with neutral to acidic soil pH), sites were chosen to maximize climatic and soil physicochemical contrasts (see Tables S2-1 & S2-2). The wide span of geoclimatic settings that were considered resulted in a continuous range of SOC contents between 0.6 to 18.7 % SOC. This allows to evaluate the macroscale relevance of mechanisms involved

in SOC dynamics across almost the entire range of geoclimatic settings in which temperate grasslands occur (except for alkaline soil pH). The gradient features 30 sites located in the temperate Köppen-Geiger climate zone, with cold and warm summers, and with or without dry season (Cfb, Cfc, Csb, Csc). In the north, the gradient is concluded by arid steppe (Bsk), and in the south it is concluded by polar tundra (ET). The large range of conditions that this gradient encompasses is also reflected in the diversity of soil types: The gradient includes 10 (out of 32) Major Soil Groups based on the classification of the World Reference Base for Soil Resources (WRB). Soil groups span soils with less pronounced pedogenic features (Leptosol, Arenosol, Cambisol), soils characterized by the influence of water (Gleysol, Planosol), humus-rich soils (Kastanozem, Chernozem) as well as soils characterized by low (Acrisol) and high (Andosol, Luvisol) mineral reactivity. Notably, this gradient features 18 sites from a previous study which demonstrated that the interaction of climate and soil geochemistry can drive bulk SOC dynamics (Doetterl et al., 2015). To extend the covered physicochemical and climatic conditions, 17 new sites were added.

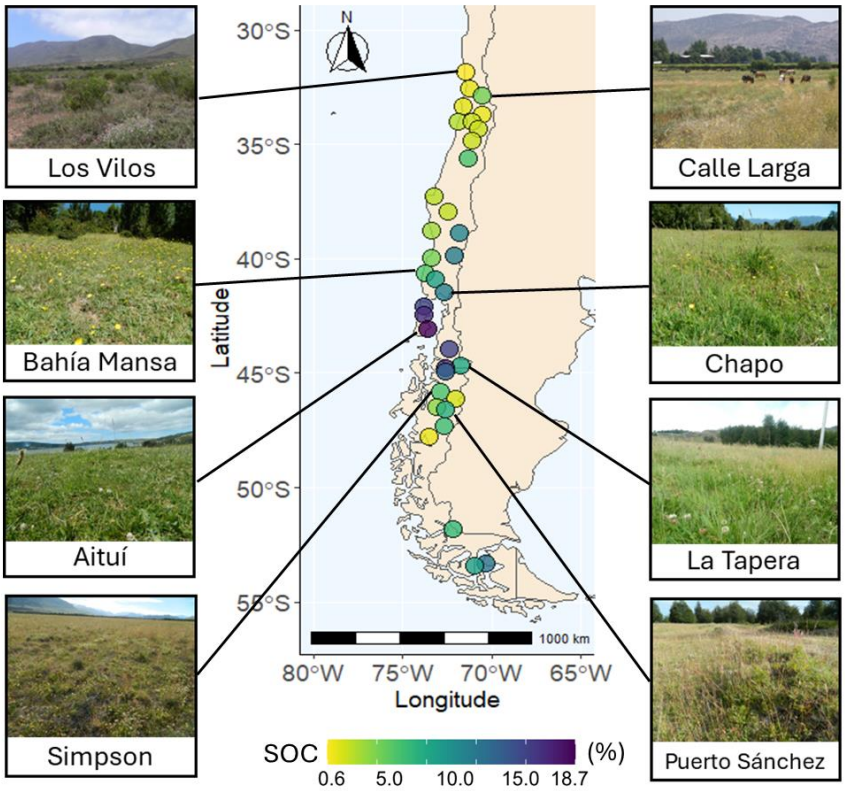



**Figure 1-3.** Map of the 35 sampling sites, with photographs (by Manuel Casanova) of eight sites. Soil organic carbon content (SOC) is indicated by color.

Overall, this geoclimatic gradient is very well suited to address the knowledge gaps that were identified in Chapter 1.5. The manageable number of samples allows for comprehensive in depth laboratory

analyses. At the same time, constraining the vegetation dynamics and the type of OM input permits a focus on soil-internal SOC transformation processes. While gradients per se do not offer spatially explicit insights, the breadth of conditions that this gradient covers allows for insights that are relevant at the macroscale. Some numbers help to illustrate the spatial relevance of this gradient. The global land cover of grasslands depends on the exact definition, but estimates range between 20 to 25 % of total land area (Suttie et al., 2005; Klein Goldewijk et al., 2017; FAO, 2021). The ten WRB soil groups that are covered by the gradient account for 59 % of the global land area (Encyclopaedia Britannica, 2024) and hold 52 % of the SOC stock (0 – 30 cm depth; FAO-UNESCO, 2018). Overall, the broad pedoclimatic settings of grasslands that this Chilean gradient covers can be found globally (Figure 1-4).

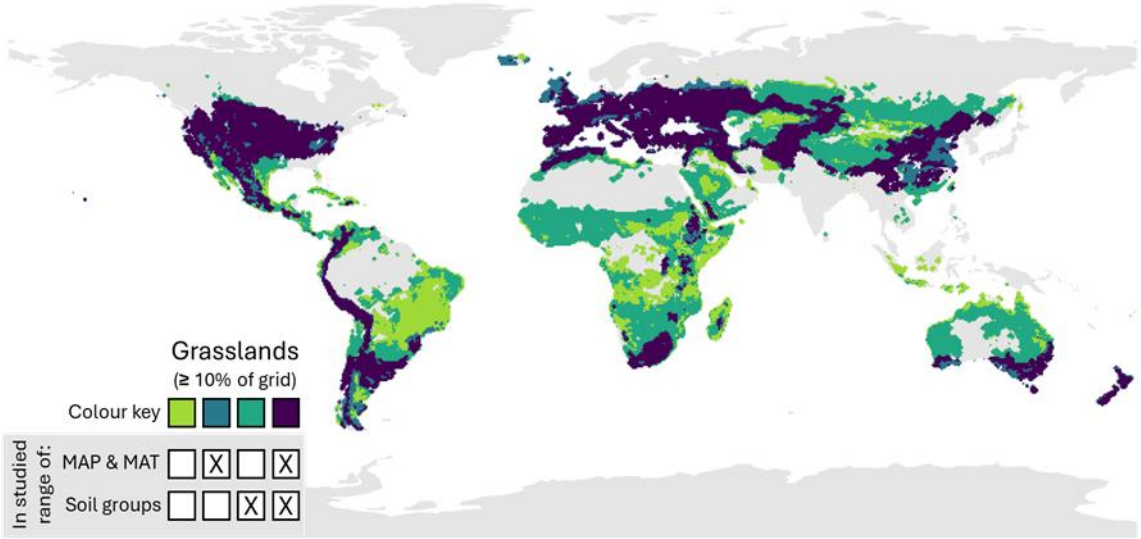



**Figure 1-4**. The global distribution of grasslands in the range of the pedoclimatic conditions that are covered by the Chilean gradient. The "X"s in the boxes of the legend indicate that grasslands cover ≥ 10 % of the area of grid cells which were: (i) neither dominated by climate nor soil groups that were included in the gradient; dominated by (ii) the included climate range or (iii) the included soil groups; or which were dominated by (iv) included climate range and included soil groups. Please note that this map has a coarse resolution (5 arc minutes, *i.e.* 9.3 x 9.3 km at the Equator) and only serves for illustrative purposes. For more details see Appendix Chapter 1.

## 1.8 Thesis outline

This thesis focuses on three selected mechanisms that link different fundamental factors of the SOC cycle (Figure 1-5). The mechanisms of interest are (1) the role of stable microaggregates as an SOC reservoir, (2) the role of competitive exclusion and substrate specialization as a determinant of bacterial community composition and (3) the role of microbial community composition as a driver of biomass-

specific C metabolism. The aim of this thesis is to investigate whether these mechanisms translate from the mesoscale to the macroscale. The insights into the relevance of the selected mechanisms at the macroscale will help to better inform the macroscale representation of the SOC cycle. Three sets of research questions will be addressed in three chapters. In the following, I will present the detailed research questions and hypotheses of Chapters 2 to 4.

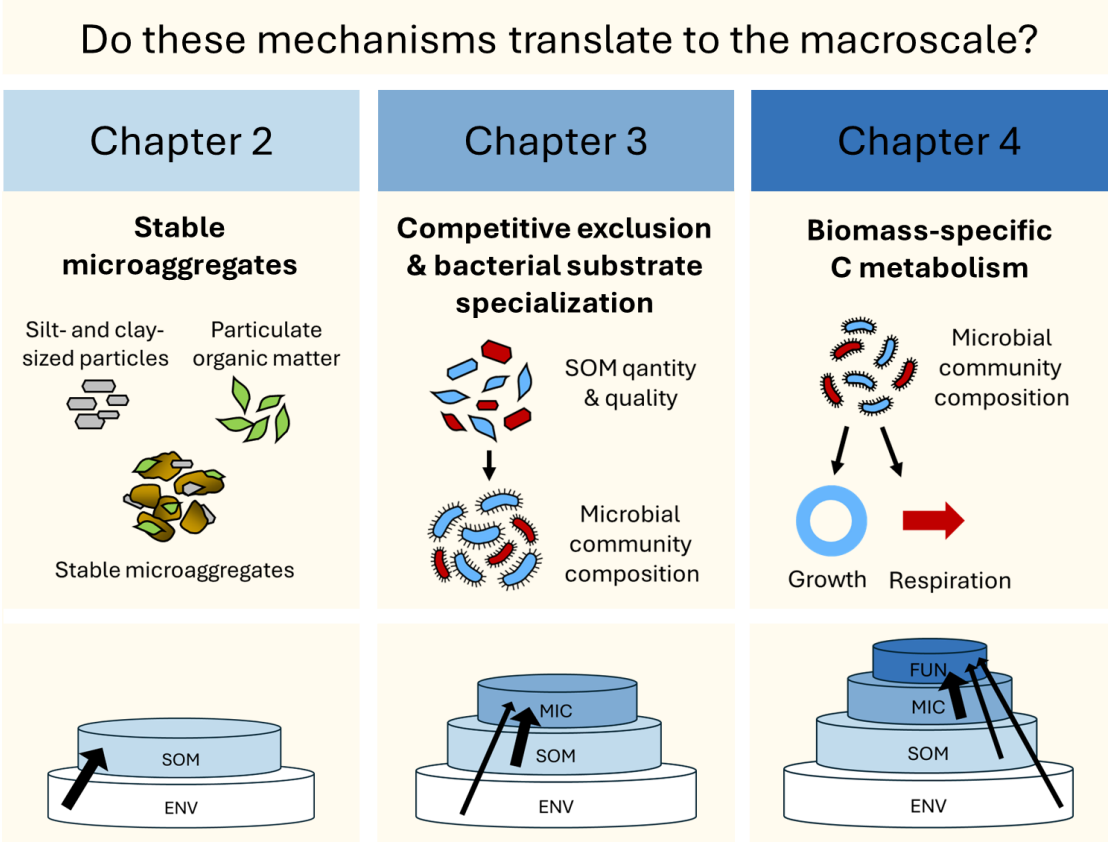



**Figure 1-5.** Conceptual summary of Chapters 2 to 4. The bottom of the figure shows the factors of the SOC cycle that will be linked in the respective chapters. ENV = environmental setting; SOM = soil organic matter quantity and quality; MIC = microbial community composition; FUN = microbial functions.

**In Chapter 2, we investigate whether stable microaggregates – which are a mechanism of spatial organization of soil – constitute a quantitatively relevant SOC fraction at the macroscale.** For this, we apply the Zimmermann fractionation scheme which separates SOC into particulate organic matter (POM), silt- and clay-sized particles (S+C) and stable microaggregates. We further investigate whether stable microaggregates are predictable from environmental conditions, and therefore merit representation at the macroscale. **We ask: Under which geoclimatic conditions do stable microaggregates constitute a relevant SOC fraction? What are the dominant geoclimatic drivers of the amount of SOC associated with stable microaggregates? How consistent are chemical characteristics of SOC in stable microaggregate across contrasting geoclimatic settings?**

H1.1: Stable microaggregates constitute a major SOC fraction with geoclimatic patterns that are distinct from the POM and S+C fractions.
H1.2: The quantitative role of stable microaggregates is more pronounced in chemically reactive soils and in soils in which OM accumulates.
H1.3: The association of SOC with stable microaggregates is limited by the amount of OM (and the biological drivers thereof) in low-SOC soils, whereas it is limited by the amount of reactive minerals in high-SOC soils.
H1.4: The OM associated with the stable microaggregate fraction is more decomposed in Andosols than in other soil types.

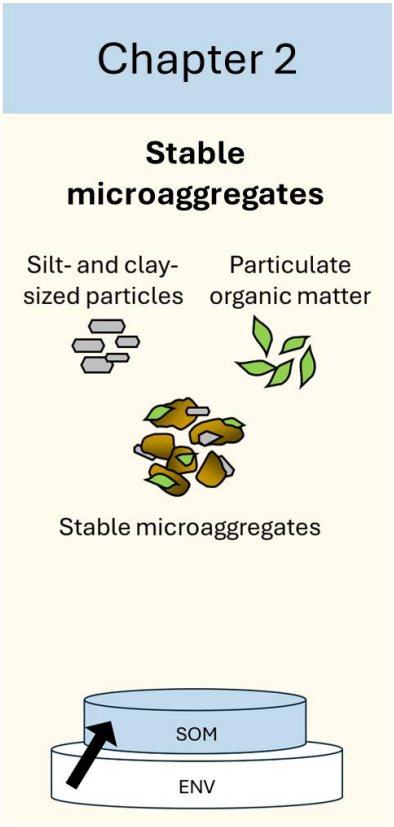
**In Chapter 3, we investigate whether there are scalable links between SOM quantity and quality and soil bacterial community composition at the macroscale.** For this, we measure qualitative characteristics of SOM and characterize bacterial community composition with sequencing of the 16S rRNA gene. **We ask: What role do SOM quantity and SOM quality play in shaping soil bacterial community composition at the large scale? Do competitive exclusion and bacterial substrate specialization translate to the macroscale as mechanisms of bacterial community assembly?**
H2.1: Bacterial alpha diversity is negatively linked to SOM quantity, because competitive exclusion is a dominant mechanism that shapes community composition along a continuum of SOM quantity.
H2.2: Bacterial alpha diversity is higher in systems with a lower degree of bulk SOM decomposition, because decomposition lowers the qualitative diversity of SOM and soil bacteria exhibit substrate specialization.

**In Chapter 4, we investigate how soil microbial traits and functions that are relevant for the SOC cycle relate to the environmental setting, SOM and the soil microbial community composition at the macroscale.** For this, we measure microbial growth and respiration rates, carbon use efficiency (CUE) and potential extracellular enzyme activities. **We ask: Which aspects of soil microbial C metabolism are related to the environmental setting vs. microbial community composition? Do metabolic mechanisms allow for improved predictions of macroscale microbial traits and functions through knowledge about the microbial community composition?**
H3.1: Microbial biomass and absolute process rates are linked most directly to the environmental setting because the environment dictates the frame for microbial activity.
H3.2: Biomass-specific process rates, CUE and microbial C:N ratio are linked to the microbial community composition, because they reflect microbial properties.
H3.3: Biomass-specific growth and respiration rates are linked to different features of the microbial community composition because they are physiologically decoupled.

The work for Chapter 2 has been conducted in the laboratory of the soil resources group at ETH Zurich. Nuno Bischofsberger and Annina Maier assisted in the laboratory work. Microbial community characterization in Chapters 3 and 4 was done at the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) in Zurich in collaboration Aline Frossard and Xingguo Han. The isotopic method for quantification of microbial growth and carbon use efficiency was conducted at the Centre for Microbiology and Environmental Systems Science at the University of Vienna in collaboration with Joerg Schnecker. Cornelia Rottensteiner assisted in the laboratory work in Vienna. Chapters 2 to 4 are intended to be published in peer-reviewed journals. Chapter 2 is currently under revision, Chapter 3 is in preparation and Chapter 4 is ready for submission. All three chapters consist of a topic-specific introduction, material and methods section, result section, discussion and conclusion.

In Chapter 5, I will provide a synthesis of the findings of all three chapters. For this, I will guide the reader through a comprehensive comparison of the SOC dynamics of two contrasting systems along the gradient. In addition, I will summarize and discuss the main findings of Chapters 2 to 4. I will place the main conclusions of each chapter in the context of the macroscale SOC cycle, and I will propose future research directions that build on my insights. Lastly, I will briefly discuss how the investigated gradient fits into the terrestrial SOC cycle. I will conclude with an outlook about future work along this gradient.

# 2. The role of climate, mineralogy and stable micro-aggregates for soil carbon dynamics along a geoclimatic gradient




This chapter has been submitted as:

Wasner D, Abramoff R, Griepentrog M, Zagal Venegas E, Boeckx P, Doetterl S. The role of climate, mineralogy and stable micro-aggregates for soil carbon dynamics along a geoclimatic gradient. Under revision for Global Biogeochemical Cycles.

## Abstract

Organic matter accumulation in soil is understood as the result of the dynamics between mineral-associated (often more decomposed, microbial derived) organic matter and free particulate (often less decomposed, plant derived) organic matter. However, on regional to global scales, the patterns and drivers behind main SOC fractions are not well understood and remain poorly linked to the pedogenetic variation across soil types. Here, we separated soil organic carbon (SOC) associated with silt- and clay-sized particles (S+C), stable microaggregates (>63 µm, SA) and free particulate organic matter (POM) from a diverse range of grassland topsoils sampled along a geoclimatic gradient. The relative contribution of the two predominantly mineral-associated fractions (S+C & SA) differed significantly across the gradient while free POM was never the dominant SOC fraction. Stable microaggregates emerged as the major SOC fraction in carbon-rich soils. The degree of decomposition of carbon in stable microaggregates was consistently between that of the S+C and POM fractions and did not change along the investigated gradient. In contrast, the carbon associated with the S+C fraction was less microbially decomposed in carbon-rich soils than in carbon-poor soils. The amount of SOC in the S+C fraction was positively correlated to pedogenic oxide contents and finer texture, whereas the amount of SOC associated with stable microaggregates was positively correlated to pedogenic oxide contents and negatively to temperature. We present a conceptual summary of our findings, which integrates the role of stable microaggregates with other major SOC fractions and illustrates their changing importance across (soil) environmental gradients.

## 2.1 Introduction

Soil organic carbon (SOC) comprises the largest terrestrial carbon stock (Canadell, 2021). At regional to global scales, climate, land use and soil geochemistry have been identified as major controls on SOC dynamics since they can drive carbon input, stabilization as well as persistence in soil (Sebastian Doetterl et al., 2015; Rasmussen et al., 2018; Heckman et al., 2022; Slessarev et al., 2022). At such large scales, climate emerges as a dominant control of bulk SOC because it acts in two ways: on short timescales it affects biological processes such as organic matter input and decomposition, and SOC has been found to be most persistent in wet and cool systems (Heckman et al., 2022). On the long term, it shapes the reactivity of the soil mineral phase via weathering processes. The reactivity of the soil mineral phase is in turn not just a function of climate alone, but of soil development and the resistance of soil parent material to weathering. In dry and alkaline systems, calcium can play an important role in SOC stabilization (Rasmussen et al., 2018). In settings where parent material, age and humid climate allow for an abundance of poorly crystalline minerals, Al and Fe hydroxides play a pronounced role in

SOC stabilization (Doetterl et al., 2018; Rasmussen et al., 2018; Slessarev et al., 2022). As a consequence of this complexity, not all processes that govern bulk SOC act strictly linear across the entire spectrum of soils (Yu et al., 2021). For instance, the abrupt switch from alkaline to acidic soil pH along the global continuum of water balance (Slessarev et al., 2016) represents a threshold in the relationship between climate and soil development, that in turn causes abrupt thresholds in soil chemistry and SOC stabilization (Chadwick and Chorover, 2001; Rasmussen et al., 2018). Changes in the type of vegetation and land use can also cause thresholds in SOC dynamics between systems. Mechanisms of SOC stabilization can for example be affected by contrasting quality of litter input in a transition from grasslands to woody systems, or by disturbance regimes such as tillage (J Six et al., 2000; Heckman et al., 2022).

Our understanding of SOC dynamics is further complicated by the fact that bulk SOC is not a homogeneous entity. Rather, bulk SOC consists of individual carbon compounds with a large variety of inherent chemical properties and different degrees and types of association with the soil mineral phase (von Lützow et al., 2008; Simpson and Simpson, 2012; Lehmann and Kleber, 2015). Consequently, not all organic carbon in a soil follows the same dynamics. This complexity is addressed by studying SOC in functionally distinct soil fractions, which can be isolated by physical and chemical fractionation schemes. There, the contrast between mineral-associated SOC vs. free particulate organic matter (POM) in soils has been recognized as an important distinction with respect to formation dynamics (Cotrufo et al., 2015; Sokol and Bradford, 2019; Lavallee et al., 2020), SOC turnover rates (Heckman et al., 2022) and SOC vulnerability to changing environmental conditions (Spycher et al., 1983; Viscarra Rossel et al., 2019; Luo et al., 2020; Lugato et al., 2021; Rocci et al., 2021). Broadly, POM has been characterized to consist of coarse plant derived organic material with a high activation energy and a high C:N ratio, whereas mineral-associated SOC primarily consists of low molecular weight compounds of plant and microbial origin, with lower activation energies and low C:N ratio (Lavallee et al., 2020). Experimentally, the processes of formation and accumulation of SOC in these two fractions have been mostly investigated for specific environmental settings (Cotrufo et al., 2015; Sokol and Bradford, 2019). However, patterns and controls of SOC fractions at large scales as well as the functional implications for SOC turnover have remained critically understudied for many soil types (Kögel-Knabner and Amelung, 2021). Recent studies and meta-analyses have started to close this knowledge gap by studying the links between the amount of SOC in these fractions and the corresponding drivers with large scale datasets (Cotrufo et al., 2019; Rocci et al., 2021; Heckman et al., 2022). The contrasting turnover of mineral-associated SOC vs. POM is now the basis for the implementation of SOC dynamics in models that represent soil-atmosphere C exchange (Zhang et al., 2021; Abramoff et al., 2022). This is a big improvement from representing SOC as one homogenous pool or as several conceptual - but unmeasurable - pools.

However, most state of the art conceptual models of SOC dynamics focus on the juxtaposition of mineral-associated fractions and particulate organic matter after a prior dispersion of physical structures such as stable microaggregates (Lavallee et al., 2020; Cotrufo et al., 2021; Zhang et al., 2021; Georgiou et al., 2022; Sokol et al., 2022b). This choice is often made because the importance of aggregates as a distinct stabilization mechanism, specifically through occlusion of SOC, remains unclear and is hard to determine. Therefore, stable microaggregates are often dismissed as distinct SOC fractions, and their relevance across larger scales remains understudied. However, several studies have demonstrated that stable microaggregates and the occluded light fraction play an important role for SOC dynamics across larger scales. Poeplau and Don (2013) showed that stable microaggregates reacted sensitively to land use change across Europe, Wiesmeier et al. (2016) successfully incorporated stable microaggregates into the RothC soil carbon model to project regional SOC dynamics, and Heckman et al. (2022) demonstrated in a global analysis that the occluded light fraction follows distinct environmental dynamics.

Conceptually, stable microaggregates are soil structures which consist of diverse mineral and organic materials that are bound together over long timescales, and that resist dispersion under natural conditions (Tisdall and Oades, 1982; Totsche et al., 2018). When aggregates are dispersed, they release both mineral-associated SOC as well as POM, which has prompted the idea that they may not be important to consider as a distinct SOC fraction at larger scales (Lavallee et al., 2020). However, at a mechanistic level, the distinct microenvironments created by physical occlusion of mineral- and non-mineral-associated organic matter in microaggregates may contribute to SOC protection (Dungait et al., 2012; Abramoff et al., 2018; Totsche et al., 2018; Heckman et al., 2022) since they restrict or alter the access of microbial decomposers to potential food and energy sources. Ways by which microaggregates are thought to protect SOC are physical separation of OM from the decomposers by occlusion and hydrophobicity, reduction of the diffusion of extracellular enzymes to their substrate and creation of micro conditions unfavorable for decomposers, such as low oxygen levels (von Lützow et al., 2006; Totsche et al., 2018). In support of this idea, SOC in stable microaggregates has a slower turnover than free POM. This has been shown many times and with various approaches, mostly characterizing the occluded light fraction after aggregate disruption. The evidence encompasses observations that the disruption of aggregates resulted in higher mineralization rates (e.g. Mueller et al., 2012, 2014, references in von Lützow et al., 2006), as well as observations that the occluded light fraction had C:N ratios and $^{13}C$ signatures similar to free POM but significantly older $^{14}C$ ages (e.g. Heckman et al., 2022; Rasmussen et al., 2005; Schrumpf et al., 2013). Further, studies that used $^{13}C$ signatures after shifts between C3 and C4 plant cultivation could demonstrate that SOC in stable microaggregates has a slower turnover (e.g. John et al., 2005).

A long-standing concept for the formation of stable microaggregates is the aggregate hierarchy model, of which there exist numerous small variations as reviewed in Six et al. (2004). Slightly modified from Oades (1984), Golchin et al. (1997) described the concept as follows: free POM gets colonized by microbial decomposers, and these decomposers transform POM into microbial decomposition products and mucilage. These microbial products function as a glue and lead to mineral coating, forming macroaggregates (typically defined as > 250 µm). When the core of these macroaggregates is decomposed, they collapse into smaller and more stable microaggregates (typically defined as 20 – 250 µm). It has been shown that the stability of aggregates increases with smaller size (Dungait et al., 2012). However, the aggregate hierarchy model may not apply to the same extent to all types of soils. The dominance of an alternative process of stable microaggregate formation has in particular been proposed for soil types with high contents of Al and Fe oxy-hydroxides (such as Ultisols or Oxisols) or short-range order minerals (such as Andosols) (J. Six et al., 2000; Huygens et al., 2005; Matus et al., 2014; Totsche et al., 2018). In such soils, coprecipitation and adsorption processes form microaggregates directly in a "bottom-up" order (Edwards and Bremner, 1967; Huygens et al., 2005; Lehmann et al., 2007; Hernandez-Soriano et al., 2018; Totsche et al., 2018), or smaller OM-metal composites act as a glue that attracts OM as well as denser mineral particles to form very stable microaggregates (Asano and Wagai, 2014; Wagai et al., 2020). In either case, it is undisputed that several components and processes can affect microaggregate formation: Mineral components, organic matter, biological activity and drying-wetting cycles (Six et al., 2004; Totsche et al., 2018). Al and Fe oxy-hydroxides and clay minerals can act either as surface coating or as precipitation nuclei, POM can act as a formation nuclei and smaller microbially derived OM compounds such as decomposition products and mucilage can act as gluing agents.

The formation process of stable microaggregates indicates that this fraction exhibits dynamics that are distinct from free POM and dispersed mineral-associated SOC. The exclusion of stable microaggregates as an SOC fraction from conceptual models might therefore represent a critical blind spot in our understanding of SOC dynamics. However, to our knowledge, there has been no systematic large-scale study with coherent methodology that evaluated intact stable microaggregates as an SOC fraction across soil types or geoclimatic gradients. This leaves a potentially crucial knowledge gap in our understanding of the distribution of stable microaggregates across large scales: Under which geoclimatic conditions do stable microaggregates constitute a relevant SOC fraction? What are the dominant geoclimatic drivers of the amount of SOC associated with stable microaggregates? And how consistent are chemical characteristics of stable microaggregate SOC across contrasting geoclimatic settings? To address these questions, we investigated the large-scale distribution of SOC among the soil fractions of 35 A-horizon topsoils, sampled across a 2300 km large geoclimatic gradient in Chile with a maximized range of climatic soil physicochemical properties (Figure 2-1, Table S2-1). We expected that

(1) stable microaggregates would constitute a major SOC fraction with geoclimatic patterns that are distinct from the POM and S+C fractions. Given the important role of OM and reactive minerals as ingredients of aggregate formation, we hypothesized that the quantitative role of stable microaggregates is more pronounced in chemically reactive soils and in soils in which OM accumulates. (2) We further expected a shift in the dominant controls of the amount of SOC in the stable microaggregate fraction. We hypothesized that the association of SOC with stable microaggregates should be limited by the amount of OM (and the biological drivers thereof) in low SOC soils, whereas it should be limited by the amount of reactive minerals in high SOC soils. (3) Lastly, we expected that the formation dynamics of the stable microaggregate fraction would vary across soil systems. We hypothesized that the OM associated with the SA fraction would be more decomposed in Andosols ("bottom-up" aggregate formation) than in other soil types (aggregate formation following the "aggregate hierarchy model").

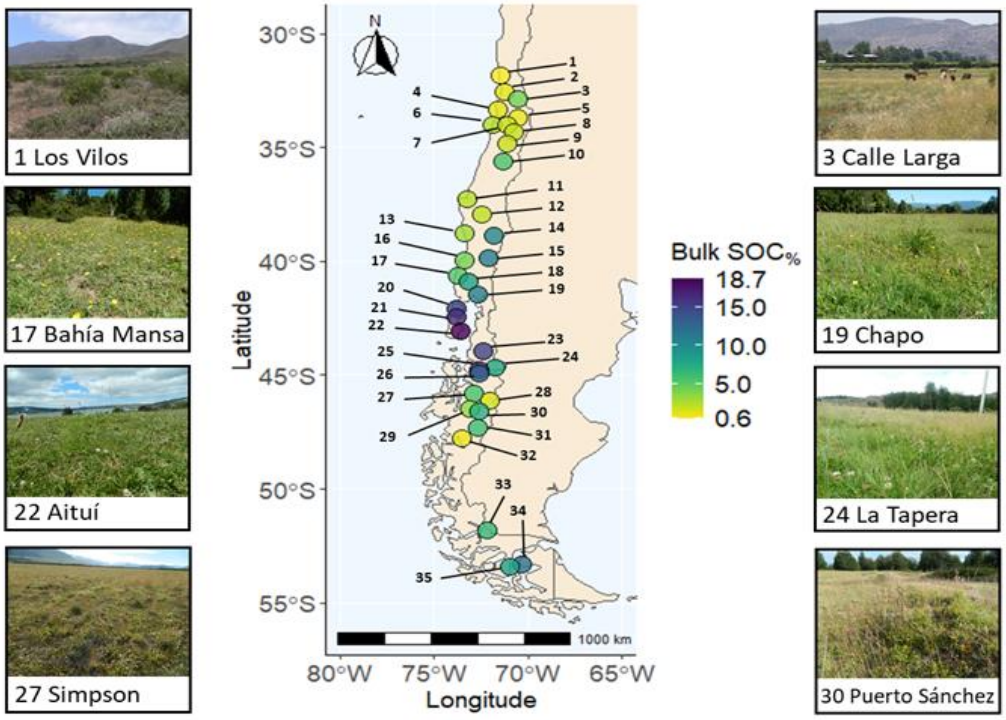



**Figure 2-1.** Map of the 35 sampling sites. Bulk $SOC_{\%}$ contents are indicated by color. Numbers refer to site IDs as used in Tables S1 and S2. In addition, photographs of 8 sites are shown with respective site IDs and site names.

## 2.2 Material & Methods

### 2.2.1 Site selection and sampling

A total of 35 A-horizon topsoils (0 - 10 cm) under extensively managed and natural grassland were sampled in the summer seasons of 2017 and 2018 across a geoclimatic gradient in Chile (Figure 2-1). The gradient was chosen due to the (i) large variability in climatic conditions that drive C input via biomass production and conditions for microbial decomposition of C, (ii) similar vegetation coverage (extensive grasslands) and (iii) the large variation in geochemical soil conditions that govern fertility and the mineral-associated SOC stabilization potential. To increase the identifiability of large scale geo-climatic controls on SOC fractions, we constrained our analysis to a coherent type of organic matter (OM) input and land use - extensive rangeland and natural grassland - ranging from cold tundra to warm (semi-)arid steppe, while excluding climatic extremes (cold and hot desert environments). Further criteria for site selection were carbonate free soil conditions (null HCl reaction) at $pH_{CaCl2} < 7.0$. We sampled soils from 18 sites (Table S2-2) that overlapped with a previous study which demonstrated that the interaction of climate and soil geochemistry can drive bulk SOC dynamics ( Doetterl et al., 2015). We added 17 new sites to expand the covered physicochemical and climatic conditions. This study was focused on topsoil, because geoclimatic contrasts are most strongly pronounced in topsoil. Topsoil is the part of soil where the majority of active C cycling happens in grasslands (i.e. where the majority of C input enters the soil system), where climate affects soil microbial decomposition most strongly and directly, and where biogeochemical alteration of the soil matrix is advanced furthest.

Climate classification following Köppen-Geiger for all sites was done with the R-package 'kgc' (Bryant et al., 2017). Mean annual temperature, precipitation and temperature seasonality (MAT, MAP & TempSeas) were taken from WorldClim Version 2 (Fick and Hijmans, 2017). Mean annual potential evapotranspiration (PET) was taken from Trabucco & Zomer (2018). Both datasets average monthly climate data from 1970 – 2000, at a resolution of 30 seconds. The gradient covers a MAT range from 3.0 to 17.1 °C, a MAP range from 217 to 2440 mm, a PET range between 715 and 1852 mm and Köppen-Geiger climate zones ranging from arid steppe (Bsk) in the north to polar tundra (ET) in the south. Most sites (30) are located in the temperate climate zone, representing climates with cold and warm summers, with or without dry season (Cfb, Cfc, Csb, Csc). Basic climatic site characteristics are summarized in Table S2-2. Dominant vegetation, landscape position and soil moisture regime were either characterized on site or supplemented from previous surveys (detailed information and references in Tables S1 and S2). Vegetation comprised in all cases species that dominate in natural grasslands and prairies, and landforms were mostly alluvial and marine terraces with flat topography (Table S2-1). Net primary productivity (NPP) estimates were obtained from a MODIS-based dataset at

500 m grid size, averaged from 2000 to 2018 (Running and Zhao, 2019). NPP ranged from 72 to 1764 g C $m^{-2}$ $yr^{-1}$. The analyzed gradient covers 10 WRB Reference soil groups (Table S2-1).

The selection of sites was aided by information from previous soil surveys by CIREN (Centro de Información de Recursos Naturales). At each of the well-described sites, a polyvinyl chloride (PVC) tube (height: 35 cm; diameter: 90 mm diameter) was vertically inserted into soil in close vicinity to the original CIREN soil profiles. The tube was then extracted by excavation of the surrounding soil, and both ends of the tubes were sealed for transport to the University of Concepción. The original sampling design of the CIREN monitored sites is described in detail in Doetterl et al. (2015). All collected soil samples were then frozen at field moisture at -20°C and were stored and shipped in this condition. Upon arrival in the laboratory in Switzerland, samples were thawed, sieved to < 2 mm, and frozen again at -20°C until further analysis. The fine soil (< 2mm) was used for further in-depth analysis of relationships between geochemical and climatic variables to SOC storage and assessment of SOC stabilization mechanisms.

### 2.2.2 Workflow summary

We separated three distinct and conceptually relevant SOC fractions: Unprotected SOC associated with (plant-derived) free particulate organic matter (POM); SOC associated with silt- and clay-sized particles (S+C, <63 µm) and SOC associated with stable microaggregates > 63 µm (SA) (Figure 2-2a). In a second step, the qualitative composition of the OM of fractions was analyzed to assess the degree of decomposition within each fraction and thereby gain further insight into the qualitative and quantitative composition of SOC and fractions across soil types. For this, we quantified the degree of decomposition (decomposition index, DI) of the three separated fractions following a principal component analysis (PCA) of fraction-specific C:N ratios, $\delta^{13}C$ and $\delta^{15}N$ values and relative chemical composition by DRIFT spectroscopy (Figure 2-2b, Figure S2-1). In a third step, to investigate which soil geochemical and climatic factors control total SOC amounts in bulk soil and fractions, we measured 22 climatic, pedogenic and physicochemical parameters that are known to affect long term soil SOC dynamics (Figure 2-2c, see selected variables in Table S2-3). Soils along the gradient were analyzed for those physicochemical properties (pH, texture, contents of pedogenic oxides and available base cations etc.) that relate to the geochemical variation in parent material and the degree of soil weathering that varies across soil types and soil development. Lastly, using rotated PCA, all explanatory variables were condensed into distinct environmental dimensions, each representing key aspects of geo-climatic controls on SOC. Then, we used regression analyses to assess the relationship of the rotated components (RCs) to the isolated SOC fractions' quantity and associated OM quality along the soil gradient. We used the identified environmental RCs as predictors, and $C_{TOT}$ in the three isolated

fractions as well as bulk $SOC_{\%}$ as dependent variables. Figure 2-2 shows a visual summary of the workflow.

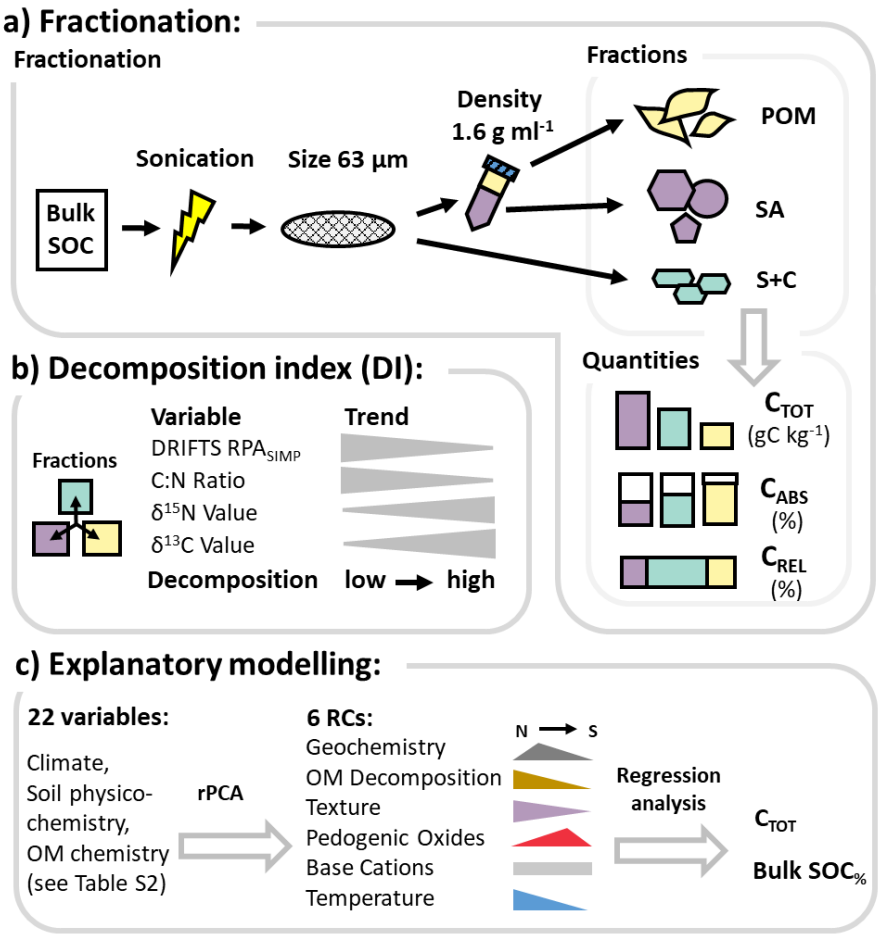



**Figure 2-2.** Schematic overview of quantitative and qualitative SOC characterization and data analysis. a) Isolation of three soil fractions; free particulate organic matter (POM), stable microaggregates > 63 µm (SA) and silt- and clay-sized particles (S+C). Quantification of the total amount of SOC per unit soil ($C_{TOT}$), fraction-specific C content ($C_{ABS}$) and relative contribution to bulk SOC ($C_{REL}$). b) To quantify the degree of microbial transformation of OM, a decomposition index (DI) was derived from principal component analysis (PCA) with fraction-specific measurements of decomposition indicators. c) Regression analysis to investigate controls of SOC fractions. We first applied rotated PCA (rPCA) to reduce 22 environmental variables (listed in Table S2-3) to six rotated components (RCs) for regression analysis.

### 2.2.3 Soil fractionation

We used a widely accepted protocol (Zimmermann et al., 2007) for combined size and density SOC fractionation, with a modified sodium polytungstate (SPT) density fractionation step (Griepentrog and Schmidt, 2013). We defined POM as particles > 63 µm and lighter than 1.6 g $cm^{-3}$, S+C as particles of all densities < 63 µm, and SA as particles > 63 µm and denser than 1.6 g $cm^{-3}$. Note that POM, as defined here, may also encompass organic matter fragments that are coated with mineral material, but are large enough to nevertheless maintain a density lower than 1.6 g $cm^{-3}$ (Wagai et al., 2009). First, air dried soils were dispersed by ultrasonication (15 g dry weight in 100 ml nanopure water) with an energy input of 24 J $ml^{-1}$ (50 Hz, 33.5 W), corresponding to an applied sonication energy of 160 J $g^{-1}$ dry soil. The sonicated soil was then wet sieved with nanopure water over a 63 µm mesh size aperture sieve until the throughflow became clear. The isolated fraction < 63 µm (S+C) was subsequently dried at 50 °C and stored for further analysis. Note that the S+C fraction is a mix of diverse free organo-mineral associations and stable silt- and clay-sized aggregates (Totsche et al., 2018; Lavallee et al., 2020; Schweizer, 2022), and that it can theoretically also contain remnants of fine organic matter (< 63 µm) that is not associated with mineral particles. The fraction > 63 µm was air dried (modified from Zimmermann et al., 2007), and a subsample of 8 g was further separated by density fractionation. For this, the subsample was first mixed in 40 ml of 1.6 g $cm^{-3}$ SPT using a vortex for 10 seconds, and then centrifuged (Sigma 3-16 KL, 15 min, 3000 xg). The floating light POM fraction was decanted onto 0.7 µm mesh size Whatman glass filters (grade GF/F), to first remove the SPT and then clean the fraction with 100 ml nanopure water with the suction of a vacuum pump, before finally being dried at 50 °C. The heavy fraction which remained as a residual after the decantation - referred to as SA (in original Zimmerman protocol referred to as “sand and stable aggregates”) - was washed five times by addition of 40 ml nanopure water, mixing on a Vortex, centrifugation (settings as above) and decantation, before being dried at 50 °C. For one selected soil sample of low, intermediate, and high clay content, fractionation was done in triplicate to evaluate the reproducibility of the procedure. The reproducibility of the method was satisfactory, with standard deviations of $Mass_{REL}$ relative to respective means at 12.0 % for POM, 4.4 % for S+C and 3.6 % for SA, averaged across the replicated soils (see Table S2-4). Standard deviation was highest in the low clay soil. The fractionation procedure resulted in an average recovery of 99.2 % soil mass (± 3.8 % S.D.), and an average total recovery of 99.1 % SOC (± 12.0 % S.D.). Detailed values for all samples are shown in Table S2-4. Note that the strength of the sonication treatment may affect absolute numbers such as the 4 % SOC threshold at which $SA_{TOT}$ exceeds $S+C_{TOT}$, while it should not affect the patterns of contribution of the C content in each fraction to bulk SOC (Poeplau and Don, 2014; Poeplau et al., 2018).

## 2.2.4 Organic matter characterization

*SOC contents and relative contribution of fractions to bulk soil*

To quantify bulk $SOC_{\%}$ and absolute fraction-specific C contents ($POM_{ABS}$/S+$C_{ABS}$/$SA_{ABS}$), C (and N) of bulk soil and the three isolated fractions were quantified on an EA-IRMS (FlashSmart Elemental Analyzer coupled with a Finnigan DELTAplusXP, Thermo Fisher Scientific, Waltham, Massachusetts, USA).

Total amounts of SOC in each fraction per unit bulk soil were calculated following equation (1).

$$Frac_{TOT} = \frac{Frac_{ABS} * Mass_{REL}}{10} \quad (1)$$

with $Frac_{TOT}$ representing the total amount of SOC in each fraction ($POM_{TOT}$/S+$C_{TOT}$/$SA_{TOT}$) [g C $kg^{-1}$], $Frac_{ABS}$ representing fraction-specific C content [%] and $Mass_{REL}$ representing the relative mass contribution of each fraction to bulk soil [%]. Note that the sand-sized fraction contains both stable microaggregates as well as sand-sized primary particles. Calculating a $C_{ABS}$ value for the stable microaggregate fraction without accounting for the mass of primary particle sand content could thus lead to pronounced underestimation of $SA_{ABS}$. Values for $SA_{ABS}$ were therefore corrected for primary particle sand content following equation (2).

$$Sand\ corrected\ SA_{ABS} = \frac{SA_{ABS} * Mass_{REL}\ of\ SA}{Mass_{REL}\ of\ SA - Mass_{SAND}} \quad (2)$$

with $Mass_{REL}$ of SA being the relative mass contribution of the SA fraction to bulk soil, and $Mass_{SAND}$ being the relative mass contribution of primary sand particles to bulk soil (see method section *Soil physicochemistry*). We acknowledge that the underlying assumption that stable microaggregates contain no sand particles at all is likely an oversimplification and might lead to an overestimation of $SA_{ABS}$ values. However, since stable microaggregates are mainly composed of silt- and clay-sized particles, this error can be assumed to be small (Oades, 1984; Golchin et al., 1997).

The contribution of the C content in each fraction to bulk SOC ($POM_{REL}$/S+$C_{REL}$/$SA_{REL}$) was calculated based on the relative contribution of each fraction to bulk SOC content and expressed as a percentage.

*Qualitative SOC characterization*

Our study system of grasslands features similar plant OM input quality ( Doetterl et al., 2015) along the gradient. The C:N ratios as well as and $^{13}C/^{12}C$ and $^{15}N/^{14}N$ stable isotope ratios of the three isolated fractions were measured using EA-IRMS (FlashSmart Elemental Analyzser coupled with a Finnigan DELTAplusXP, Thermo Fisher Scientific, Waltham, Massachusetts, USA) and interpreted as indicators of progressing microbial transformation of organic matter (OM) (Dijkstra et al., 2006; Lerch et al., 2011). C:N ratios of OM were calculated based on molar C and N contents, and stable isotope ratios of C and N are expressed relative to the Vienna Pee Dee Belemnite (VPDB) and to the atmospheric $N_2$ standards, respectively.

*Assessing the degree of decomposition of SOC*

To obtain further proxies for the degree of decomposition of plant-derived OM towards more microbially-processed OM, two additional methods were applied: (i) An assessment of the relative chemical composition of OM by diffuse reflectance infrared fourier transform spectroscopy (DRIFTS) (Demyan et al., 2012), and (ii) an assessment of thermal stability by Rock-Eval pyrolysis (Sebag et al., 2016) (Figure 2-2b).

DRIFTS was performed on bulk soil (n = 35) as well as on the three isolated fractions. For six samples of the POM fraction, we did not have enough sample material to perform DRIFTS (thus n = 29 for POM). Absorbance values were obtained on milled subsamples (< 50 µm). Samples were scanned in duplicate in the mid-infrared region (7500 – 600 $cm^{-1}$) at a resolution of 2 $cm^{-1}$, using a Fourier transform IR (FT-IR) spectrometer with a high-throughput screening extension (HTS-XT) (Bruker Optics, Vertex 70, Germany). Normalization of spectra against a gold background (Infragold NIR-MIR Reflectance Coating, Labsphere), as well as correction for atmospheric $CO_2$ and $H_2O$ were done in OPUS spectrometer software (Bruker Optics, Germany) via averaging of 32 co-added scans per sample. All subsequent spectra processing was done using the R-packages 'simplerspec' (Baumann, 2020) and 'prospectr' (Stevens and Ramirez-Lopez, 2020). To correct for light scatter, spectra were resampled to a range of 4000 – 600 $cm^{-1}$ with duplicates averaged and normalized using the normal variate method available within the R packages. Then, based on published information (Parikh et al., 2014), the following six wavenumber ranges were assigned to three types of functional groups: aliphatic C-H (anti)symmetric stretches: 2950 – 2910 and 2866 – 2836 $cm^{-1}$; aromatic C=C stretches: 1540 – 1524 and 1520 – 1510 $cm^{-1}$; carboxylic acid C=O stretch and carboxylate C-O asymmetric stretch: 1734 – 1718 and 1650 – 1636 $cm^{-1}$. Peak areas of the wavenumber ranges were integrated with a local baseline correction (Demyan et al., 2012). Following previous studies (Ryals et al., 2014; Fissore et al., 2017; Mainka et al., 2022), the assigned peak areas were interpreted as follows: aliphatic assignments as indicators of simple plant

derived OM (SIMP), aromatic assignments indicating complex plant derived OM (COMP) and carboxylic assignments - which are indicative of oxidized OM - as microbially transformed OM (MBIO). For each of the three functional groups, the peak area of a group was divided by the summed peak area of all groups to obtain its relative peak area (RPA) as a fraction of the total ($RPA_{SIMP}$, $RPA_{COMP}$ and $RPA_{MBIO}$ as % of total). As an indicator for the state of OM decomposition, we interpreted $RPA_{SIMP}$ decreasing relative to other compounds with progressing decomposition (Ryals et al., 2014; Fissore et al., 2017; S. J. Hall et al., 2018).

In parallel to the above, but on bulk soil only, Rock-Eval pyrolysis was performed to create a second and independent proxy for the degree of OM transformation related to its decomposition and stabilization (Sebag et al., 2016). For this, we used 0.1 g milled bulk soil samples, and conducted the following analyses on a pyrolyzer (Vinci Technologies, Rock-Eval 6, France). The applied protocol consisted of two phases: a pyrolysis in an inert $N_2$ atmosphere starting at a temperature of 200 °C until 650 °C, and a pyrolysis in an oxidized atmosphere between 400 °C and 850 °C, both with a heating rate of 25 °C $min^{-1}$. Subsequently, the Rock-Eval I-Index for the degree of biological transformation of OM was calculated (Sebag et al., 2016). Briefly, areas under defined segments of the S2 curve (*i.e.*, the hydrocarbons that form during thermal pyrolysis) were used following equation (3).

$$I_{ind} = log_{10} ( (A1 + A2) / A3) \qquad (3)$$

with $I_{ind}$ being the I-Index, A1 being the segment between 200 - 340 °C (*i.e.,* labile biopolymers, low intrinsic stability), A2 340 - 400 °C (resistant biopolymers, intermediate intrinsic stability) and A3 400 - 460 °C (immature geopolymers, high intrinsic stability).

For descriptive statistics summarizing all quantitative and qualitative SOC data of the fractions, see Table S2-5.

## 2.2.5 Soil physicochemistry

*Soil physical characterization*

Soil texture was determined via laser diffraction using a particle size analyzer (PSA) (LS 13 320, Beckman Coulter, USA). We determined soil texture in two different ways: First, as soil primary texture, namely the textural composition of primary particles after breaking up all physical and chemical structures that can aggregate soil. Primary texture was used to calculate the sand correction of $SA_{ABS}$ (see previous section). Second, as soil secondary texture, namely the textural composition of soil without disintegrating stable microaggregates and oxy-hydroxide concretions which are structural and stable components of natural soils (Totsche et al., 2018). Secondary texture was used to represent soil texture

effects in the explanatory models (see section 2.2.7 Statistical analysis). For the determination of soil primary texture, air dried soil was sonicated (100 J $ml^{-1}$, 50 Hz, 65 W), shaken for 18 h with 10 % Na-hexametaphosphate and treated for 96 h with 30 % $H_2O_2$ at 70 °C. Before and after $H_2O_2$ treatment, supernatant was removed by decanting after centrifugation (10 min 1700 x g). For the determination of soil secondary texture, fresh soil was shaken for 3 h with 10 % Na-hexametaphosphate to dissolve macroaggregates. Resulting primary and secondary particles were quantified with the PSA. Particle size contributions were calculated as percent of total particle volume, and size classification followed the WRB system (IUSS Working Group WRB, 2015): clay < 2 µm, 2 µm < silt < 63 µm, 63 µm < sand < 2000 µm. For the sand correction of $SA_{ABS}$, the primary particles were wet sieved over a 63 µm mesh size aperture sieve until the throughflow became clear. Then, after oven drying at 50 °C the mass contribution of the primary particles > 63 µm (i.e., sand) to total soil mass was calculated as a percentage ($Mass_{SAND}$). For one selected soil sample of low, intermediate, and high secondary sand content, $Mass_{SAND}$ was evaluated in triplicate. The reproducibility of the method was satisfactory, with a standard deviation of $Mass_{SAND}$ relative to the mean at 2.8 %, averaged across the replicated soils. Standard deviation of $Mass_{SAND}$ was highest in the soil with high secondary sand content.

*Soil chemical characterization*

As a measure of soil acidity and the potential effect it has on the mobility of elements, soil pH was determined in 0.01 M $CaCl_2$ solution (soil:solution ratio of 1:5). After shaking for 10 min, the samples were left to rest for 24 h and shortly shaken before measurement in suspension using a pH meter (713 pH Meter, Metrohm, Switzerland).

Total element content of Al and Fe in bulk soil was assessed by digestion of 1 g aliquots with an aqua regia acid solution ($HCl:HNO_3:H_2O$, 3:1:1, v:v:v, 2.5 h at 120 °C). After filtration through Whatman 41 filter papers Al and Fe were quantified using inductively coupled plasma optical emission spectrometry (5100 ICP-OES, Agilent Technologies, USA). As a measure of rock-derived base cation availability, exchangeable Ca, K and Mg were extracted with 0.1 M $BaCl_2$ (soil:solution ratio of 1:6.5) for 2 h on a horizontal shaker, filtered through Whatman 42 filter papers and quantified with ICP-OES (Hendershot and Duquette, 1986).

## 2.2.6 Soil mineralogy and weathering

To assess differences in Si contents due to weathering processes and intrinsic geochemical differences of soil parent material, total Si content of bulk soil was measured using energy dispersive X-ray fluorescence (ED-XRF) spectrometry. For this, milled soil was mixed with Licowax (Fluxana, Germany)

at a ratio of 4:1, and formed into pellets which were measured with a spectrometer (Spectro Analytical Instruments, Spectro XEPOS, Germany). Then, we calculated ratios of primary texture clay relative to Si ($clay_{prim}$:Si) and total Fe relative to Si ($Fe_{Aq.Reg}$:Si) as proxies for geochemical reactivity of the mineral matrix, based on the rationale that weathering processes can lead to enrichment of clay and Fe relative to Si (Amelung et al., 2018). To obtain the contents of metal phases that are relevant for mineral SOC stabilization in bulk soil, we conducted a sequential extraction of Al, Fe and Mn oxy-hydroxides using sodium-pyrophosphate (PP) to extract organo-metallic complexes (Bascomb, 1968, modified), and subsequently ammonium oxalate (OXA) to obtain poorly crystalline metal-oxides and short-range order minerals ((Courchesne and Turmel, 2007; Doetterl et al., 2015), modified). We acknowledge that PP- and OXA-soluble metals may also include Al- and Fe-organic complexes that are not necessarily metal oxy-hydroxides, but for simplicity we will refer to the sum of PP- and OXA-soluble metals as "pedogenic oxides". In brief, we used 0.5 g of milled material of the S+C and SA fractions, and extracted the material for 16 h on a horizontal shaker with a sodium-pyrophosphate solution at pH 10 (consisting of 0.1 M $Na_4P_2O_7$ x 10 $H_2O$ and 0.5 M $Na_2SO_4$), at a soil:solution ratio of 1:40. Subsequently, the vials were centrifuged (Sigma 3-16 KL, 10 min, 1700 xg), the supernatant was decanted, filtered through Whatman 41 filter paper, diluted to 50 ml with nanopure water, and stored at 4° C until measurement. The remaining soil residue was extracted after shaking for 2 h on a horizontal shaker with a 0.2 M ammonium oxalate solution at pH 3 at a soil:solution ratio of 1:40 (consisting of 0.2 M $(NH_4)_2C_2O_4$ x $H_2O$ and 0.2 M $C_2H_2O_4$ x 2 $H_2O$, mixed at a ratio of 1.31:1), and in the dark to prevent photodegradation of the extractant. Subsequently, the solution was filtered, diluted and stored as in the previous step. In both extracted solutions (PP and OXA), Al, Fe and Mn were quantified using ICP-OES (5100 ICP-OES, Agilent Technologies, USA). Elemental contents of PP- and OXA-soluble metals for bulk soil were estimated by summing up the respective contents in S+C and SA, weighted by the mass contribution of each fraction to bulk.

### 2.2.7 Statistical analysis

All statistical analysis and creation of the map were performed in R-Studio using R version 4.1.1 (RStudio Team, 2020), stepwise regression and random forest (RF) models were computed using the R-package 'caret' (Kuhn, 2008), Breusch-Pagan tests were done using the R-package 'lmtest' (Zeileis and Hothorn, 2002). PCAs were performed with the R-package 'psych' (Revelle, 2022), and Cochran's C tests for outlying variance were performed using the package 'outliers'.

*Patterns of SOC along the gradient*

Linear regression was used to test for relationships between all dependent variables (total amounts of SOC in each fraction per unit soil, contents and the relative contribution of SOC in each fraction to bulk SOC), as well as relationships between the degree of OM decomposition to bulk SOC and fractions. In order to check whether data needed to be log-transformed to meet the assumptions for linear regression, homoscedasticity (Breusch-Pagan test), normal distribution of residuals (Shapiro-Wilk test) and absence of residual patterns (visual check) were evaluated. In case of violation of either criterion, the dependent variables were log-transformed. If the intercepts were not significant, models for $C_{TOT}$ and $C_{ABS}$ were forced through zero. Note that in the case of $SA_{TOT}$ and $S+C_{TOT}$, the assumption of homoscedasticity was violated to a minor degree at high bulk $SOC_{\%}$ values (Table S2-7). However, we kept the untransformed models because transformation produced strongly skewed patterns in the residuals. Similarly, in the models for $SA_{ABS}$ and $S+C_{ABS}$ log-transformation would have introduced residual patterns and was thus rejected.

*Deriving the decomposition index (DI)*

To assess the degree of OM decomposition for the three fractions, we used a principal component analysis (PCA) approach. This approach links the three individual soil fractions with fraction-specific information on OM decomposition (Figure 2-2a, Figure S2-1). For this, we conducted an unrotated PCA, including scaled data on the relative contribution of simple, less decomposed plant derived OM ($RPA_{SIMP}$) to fraction OM, together with fraction specific information on C:N ratio, $\delta^{13}C$ and $\delta^{15}N$ which are generally strongly affected by OM decomposition (Fernandez et al., 2003; Dijkstra et al., 2006; Lerch et al., 2011; Fissore et al., 2017; Mainka et al., 2022). More specifically, the C:N ratio of OM narrows in the course of microbial OM decomposition, and $\delta^{13}C$ and $\delta^{15}N$ values become more positive, as the heavy isotopes become enriched. Following Cochran's C test for outlying variance, two $RPA_{SIMP}$ values of one S+C and one POM sample were excluded as outliers. The result of this analysis showed that the first principal component (PC1) explains 50 % of the variation and can be interpreted as an axis from lower towards higher levels of SOC decomposition within fractions (Figure S2-1). This PC1 is loaded with decreasing C:N ratio, decreasing $RPA_{SIMP}$, increasing enrichment of $^{15}N$, and, to a lesser extent, increasing enrichment of $^{13}C$. We subsequently used the scores of PC1 as a decomposition index (DI) and analyzed the pattern of fraction specific DIs as described in the previous section. We conducted one-way ANOVA with a subsequent Tukey HSD post-hoc test to test for significant differences in DI between the three fractions.

*Explanatory modeling - Grouping and identifying predictors*

To investigate which soil geochemical and climatic factors control SOC stocks in bulk soil and fractions, we selected a set of soil parameters and qualitative SOC variables that are known to affect long term SOC dynamics, as well as a set of climatic parameters, in total 22 variables (see above for details, and find selected variables in Table S2-3). For a mechanistic interpretation of included variables, we performed dimension reduction using rotated principal component analysis (rPCA) with varimax rotation to reduce the set of these predictor variables (Haaf et al., 2021). Rotation of the PCA reduces the number of variables that correlate with individual rotated components (RCs) and thus facilitates the interpretability of the RCs, albeit at the cost of orthogonality (Figure S2-2). We only retained RCs that met two criteria: First, an Eigenvalue greater than 1, and second, adding more than 5 % of explained variance. This resulted in six retained RCs, which were interpreted based on their dominant loadings (loading ≥ 0.5) and were subsequently used as predictor variables in the regression analysis (see below). Correlations between bulk $SOC_{\%}$ and $C_{TOT}$ in SA, S+C and POM with selected measured environmental variables are shown in Figure S2-3.

*Explanatory modeling - Model selection*

We performed stepwise regression with Monte Carlo cross validation to select the most important RCs for prediction of bulk $SOC_{\%}$ and the total amounts of SOC per unit soil in the three isolated fractions ($C_{TOT}$) (Extended S3). For this, independent variables were scaled, and the data was split 100 times into a 75 % training set and a 25 % validation set. Model performance was assessed using RMSE, and a maximum of three (out of six potential) RC predictor variables were retained in order to constrain model complexity and avoid overfitting. Using the predictors that were selected by the best stepwise regression models, and to account for potential non-linearity and interactions of controls on the dependents, we consecutively built for each dependent a (i) simple linear regression model (SLM), a (ii) linear model allowing for pairwise interaction between all predictors (ILM), and a (iii) random forest model (RF) to allow for non-linearities (with 100 trees). We performed this step with log-transformed dependent variables only, as transformation generally improved residual distribution and RMSE. The variable importance (VI) of each significant predictor was quantified as the absolute value of the t-statistic, which is calculated by dividing the coefficient of a predictor by its standard error (James et al., 2013). In order to improve comparability of the contribution of predictor variables between different models, we normalized each VI relative to the VIs of all variables in the respective model. This relative variable importance (rVI) was calculated by dividing each individual VI by the sum of all VIs in the respective model. Homoscedasticity and normal distribution of the residuals were tested as described

above for SLM and ILMs. For RFs, we applied cross validation as in the stepwise regression, with 100 trees in each iteration.

Finally, we selected the best performing of the three model structures for each dependent variable based on RMSE, distribution of the residuals and the rationale of parsimony (Table S2-7). More specifically, when model performance was similar (RMSE in 10 % range), we favored and interpreted simpler model structures to avoid overinterpretation of our data. For example, if the RMSE of a SLM was in a range of 10 % of the respective ILM or RF, we interpreted the SLM. Since all models had log-transformed dependent variables, we additionally calculated a back-transformed RMSE (btRMSE) and a relative RMSE (rRMSE) for an easier interpretation of the data. For calculation of the btRMSE, we computed residuals between the untransformed observations and back-transformed log-predictions, following equation (4).

$$btRMSE = \sqrt{[\Sigma (O_i - e^{P_i})^2 / df]} \qquad (4)$$

with $O_i$ being the observed value for the $i^{th}$ observation, $P_i$ being the predicted value for the $i^{th}$ observation, and df being the degrees of freedom of the respective model. The rRMSE was calculated as the percentage of the btRMSE relative to the untransformed range of observations, following equation (5).

$$rRMSE = (btRMSE / (O_{max} - O_{min})) * 100 \qquad (5)$$

with $O_{max}$ and $O_{min}$ being maximum and minimum of the respective untransformed dependent observation.

## 2.3 Results

### 2.3.1 Shifting importance of SOC fractions along soil gradients

Across all sites, we observed a gradient of SOC in the bulk soil from 0.6 to 18.7 %(Figure 2-3, Table S2-6). While all three investigated fractions contained more SOC in soils with high bulk $SOC_{\%}$ than in soils with low bulk $SOC_{\%}$, the difference of total SOC contained in each fraction (measured as g C $kg^{-1}$ soil) was largest for the SA fraction ($SA_{TOT}$) followed by $S+C_{TOT}$ and $POM_{TOT}$. $POM_{TOT}$ ranged from 0.9 to 48.9 g C $kg^{-1}$, $S+C_{TOT}$ ranged from 2.0 to 57.7 g C $kg^{-1}$ and $SA_{TOT}$ from 1.9 to 159.9 g C $kg^{-1}$.

The fraction-specific C content ($C_{ABS}$) in POM showed no relationship with bulk $SOC_{\%}$ but remained high (13.0 - 37.8 %) across all sites. In contrast, $C_{ABS}$ of the mineral-associated fractions strongly increased

with bulk $SOC_{\%}$, from 1.5 to 24.6 % ($S+C_{ABS}$) and from 1.6 to 37.7 % ($SA_{ABS}$). $SA_{ABS}$ changed more strongly with bulk $SOC_{\%}$ than $S+C_{ABS}$ (Figure 2-3b, Table S2-6) and the fraction-specific C content of the SA fraction became similar to the POM fraction in high bulk $SOC_{\%}$ soils.

The POM fraction never dominated bulk SOC, with the relative contribution of POM to bulk SOC ($POM_{REL}$) ranging between 3.5 - 38.8 %. Variations of $S+C_{REL}$ and $SA_{REL}$ are much larger, ranging from low to high levels of bulk $SOC_{\%}$ between 7.1 – 80.9 % and 15.3 – 97.0 % contribution to total SOC, respectively (Figure 2-3c, Table S2-5). Notably, the relative contribution of SA to SOC exceeded that of S+C in many soils of the gradient. In soils with low bulk $SOC_{\%}$, S+C was the quantitatively most important SOC fraction, and in high bulk $SOC_{\%}$ soils (> 4 % bulk $SOC_{\%}$) stable microaggregates were the most important SOC fraction.

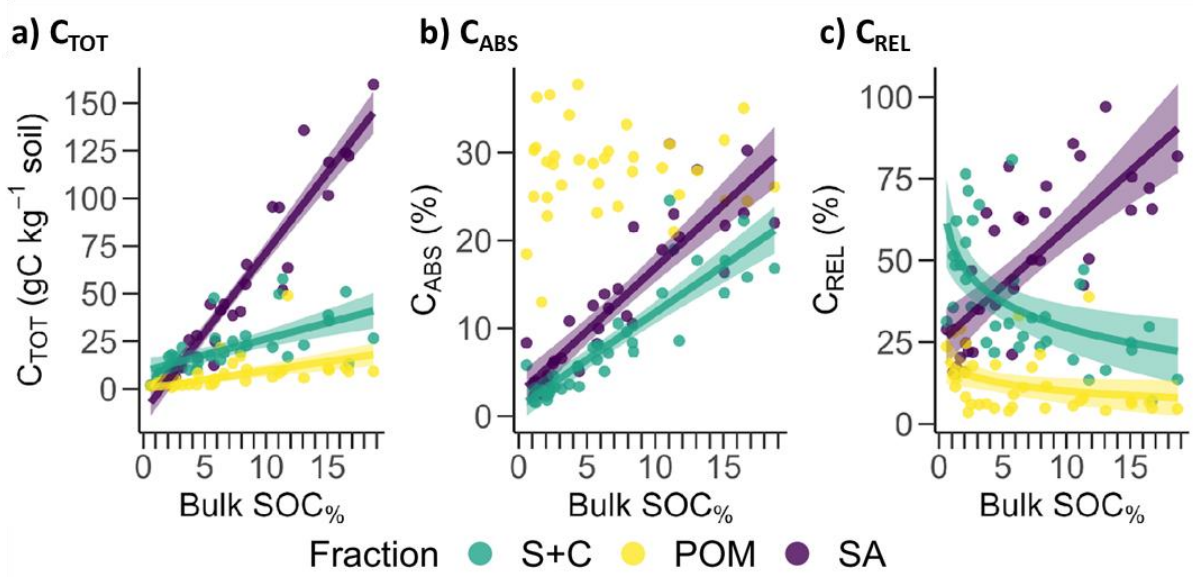



**Figure 2-3.** The quantitative development of fractions SA, S+C and POM with bulk soil organic carbon (bulk $SOC_{\%}$). SA = stable microaggregates (>63 µm), S+C = silt- and clay-sized particles, POM = free particulate organic matter. a) Total amount of SOC ($C_{TOT}$) in each of the three fractions. b) Fraction-specific C content ($C_{ABS}$) in each fraction (for SA sand corrected). c) Quantitative importance (or relative contribution, $C_{REL}$) of the fractions to bulk $SOC_{\%}$. Measured data points and fitted significant models with 95 % confidence intervals are shown. For model descriptions, see Table S2-6.

## 2.3.2 The decomposition index (DI) of the OM in the SOC fractions

The decomposition index (DI) is unitless and ranges from 3.32 to -2.55, where high values indicate a high degree of decomposition (relative to the dataset), and vice versa. DI values differed significantly between fractions (one-way ANOVA, $F(2,93) = 39.07$, $p < 0.05$; all Tukey HSD post-hoc pairwise

comparisons $p < 0.05$). POM was generally the least decomposed, SA was intermediate and S+C was the fraction with the most decomposed OM (Figure 2-4a). Of the three fractions, S+C had the most variable DI, decreasing significantly with higher bulk $SOC_{\%}$ (Figure 2-4b, Table S2-7). The DI in the POM and SA fractions was less variable without a significant pattern along the bulk $SOC_{\%}$ gradient (Figure 2-4b, Table S2-7). Further, bulk OM was considerably more decomposed in soils with low SOC content than in soils with high SOC content (Figure S2-4).

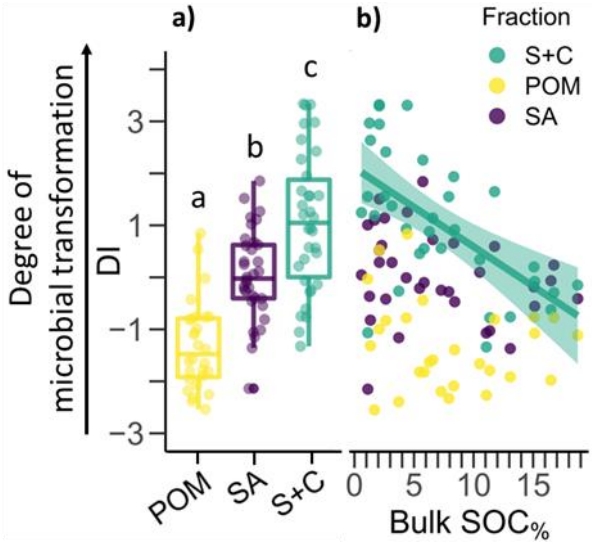



**Figure 2-4.** The decomposition index (DI) of the organic matter in fractions SA, S+C and POM. SA = stable microaggregates >63 µm), S+C = silt- and clay-sized particles, POM = free particulate organic matter. a) The DI (based on chemical characterization by DRIFT spectroscopy, C:N ratio and stable isotope ratios) of the three fractions (boxplots: center line, median; box limits, 25th and 75th percentiles; whiskers, 1.5x interquartile range; points, outliers). Letters indicate significant differences. b) Development of the DI with bulk $SOC_{\%}$. Measured data points and the fitted model with 95 % confidence intervals are shown. Only the pattern for S+C is significant at $p < 0.05$. For model descriptions, see Table S2-6.

## 2.3.3 Soil geochemical and climatic drivers of SOC

The gradient spans $pH_{CaCl2}$ values from 4.1 to 6.7, clay and sand contents from 1 to 21 % and 16 to 93 %, respectively, and pedogenic oxide contents varying up to 40-fold (Al from 0.5 to 26.5 g kg-1, and Fe from 0.7 to 20.1 g kg-1) (Table S2-3). Rotated PCA for dimension reduction yielded six rotated components (RCs, Table S2-8), which together accounted for 70 % of the variation in the dataset. 'Geochemistry' (RC1) is loaded with variables indicative of geochemical reactivity as a product of

bedrock chemistry and the degree of soil weathering (total Fe and Al, Fe:Si, Clay:Si). 'OM Decomposition' (RC2) reflects the dimension along which bulk OM becomes increasingly transformed and decomposed and is an indicator for bulk soil OM quality (assessed with DRIFT spectroscopy and Rock-Eval pyrolysis). 'Texture' (RC3) is an indicator for soil structure, positively loaded with "secondary" (see methods for explanation) silt and clay content, and negatively with secondary sand content. 'Pedogenic Oxides' (RC4) is an indicator for soil mineralogy, positively loaded with the sum of organically complexed and poorly crystalline Al- and Fe-oxides and incorporating the long term effects of climatic action via weathering. 'Base cations' (RC5) is positively loaded with plant-available Ca and Mg and serves as a fertility indicator. Lastly, 'Temperature' (RC6) is positively loaded with MAT and NPP and serves as a climate indicator. Note that the effects of important factors for SOC stabilization such as soil pH and the annual water balance (MAP - PET) are represented indirectly in the RCs OM decomposition (RC2), pedogenic oxides (RC4) and temperature (RC6), which they in parts shape and control (Table S2-8).

All regression models performed well and were able to predict patterns of the SOC fractions along the gradient ($R^2$ = 0.46 to 0.81, rRMSE = 12.5 to 17.9 %, Figure 2-5 & Table S2-6). Linear regression models always outperformed random forest models (Table S2-6), indicating the absence of major non-linear relationships (i.e. thresholds) in this dataset. Bulk $SOC_{\%}$ was negatively correlated to temperature (RC6) and positively to pedogenic oxide contents (RC 4) (Figure 2-5). $POM_{TOT}$ was positively correlated with available base cations (RC5) and negatively correlated to temperature (RC6). The $S+C_{TOT}$ was positively correlated with pedogenic oxide contents (RC4) and with (finer) texture (RC3), and weakly with (decreasing) temperature (RC6). $SA_{TOT}$ was positively correlated with pedogenic oxides (RC4) and negatively with temperature (RC6) and the degree of OM decomposition in bulk soil (RC2). Notably, geochemistry (RC1) does not feature in any of the explanatory models (Figure 2-5) as direct control, but acts as a secondary control, strongly correlated to several key predictors included in the models (RCs OM decomposition, texture, temperature, see Figure S2-2).

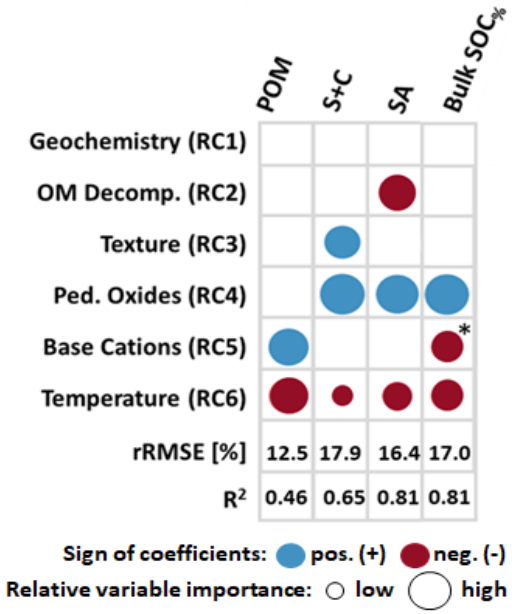



**Figure 2-5.** Links between climatic and soil physicochemical properties and total amount of SOC ($C_{TOT}$) in the soil fractions. Significant predictors (sorted by Eigenvalue, Table S2-8) of $C_{TOT}$ in free particulate organic matter (POM), silt- and clay-sized particles (S+C) and stable microaggregates (SA)and bulk $SOC_{\%}$. Circle sizes indicate the relative variable importance (rVI). * = Negative interaction term base cations (RC5) x pedogenic oxides (RC4). OM Decomp. = Organic matter Decomposition; Ped. Oxides = Pedogenic Oxides.

## 2.4 Discussion

### 2.4.1 Stable microaggregates are a key SOC fraction

Mineral-associated SOC fractions (S+C and SA) consistently dominated over POM and were key to explain the increase of SOC content along the investigated geoclimatic gradient (Figure 2-3). As hypothesized, stable microaggregates constituted an important SOC fraction and were the main SOC fraction in high-SOC soils. Importantly, $SA_{TOT}$ reached up to 160 g C $kg^{-1}$ soil, which was 2.5 and 3 times higher than the highest values for $S+C_{TOT}$ and $POM_{TOT}$, respectively. Furthermore, $SA_{TOT}$ was linked positively with pedogenic oxide contents (RC4) and negatively with temperature (RC6) and the degree of bulk OM decomposition (RC2) (Figure 2-5). This supports the hypothesis that the quantitative role of stable microaggregates is largest in chemically reactive soils and in soils in which OM accumulates (Figure 2-6a). We did not detect a universal upper limit of SOC content associated with the mineral fractions along this gradient, not withholding that such limit may exist. Instead of reaching an asymptote at a particular bulk $SOC_{\%}$ value, fraction-specific C content ($C_{ABS}$) of the S+C and SA fractions

increased across the entire bulk $SOC_{\%}$ range (Figure 2-3b). We explain this observation in two ways. First, we acknowledge that many soils (with andic properties and with generally high reactive surface area of the mineral phase) featured in the central region of this gradient have higher pedogenic oxide contents than other soil types and therefore represent conditions of extraordinarily high stabilization potential (Beare et al., 2014; Matus et al., 2014). Second, only a small fraction of SOC was present as free POM across the entire range of climatic and physicochemical conditions (Figure 2-3c). Instead, the majority of OM formed associations with the mineral matrix. This lack of accumulation of free POM is relevant because it indicates that even OM which accumulates due to climatic constraints on decomposition gets associated with the mineral matrix in the form of stable microaggregates.

Note that due to differences in OM input and disturbance regimes, the dominant role of stable microaggregates in temperate grasslands can not necessarily be directly transferred to different systems. Stable microaggregates hold less SOC in arable soils as compared to grassland soils (Poeplau and Don, 2013; Wiesmeier et al., 2016; Antony et al., 2022). Similarly, stable microaggregates also contain a smaller fraction of total SOC in forest soils, where free POM plays a proportionally larger role (Poeplau and Don, 2013; Guidi et al., 2014; Antony et al., 2022). Lastly, the importance of stable microaggregates as an SOC fraction is known to decrease with soil depth, where the silt- and clay-sized fraction plays a proportionally larger role (Schrumpf et al., 2013; Poeplau et al., 2017; Antony et al., 2022). Nevertheless, although stable microaggregates are not always as dominant as in grassland systems, they consistently constitute a considerable SOC fraction across various land uses. However, the current debate about carbon saturation thresholds in soils focuses mainly on the fraction of fully dispersed silt- and clay-sized particles (Cotrufo et al., 2019; Georgiou et al., 2022; Begill et al., 2023). Based on the importance of stable microaggregates as SOC fractions demonstrated in this study (Figure 2-6a) and across literature, we argue that the potential role of microaggregates should not be ignored in the context of SOC stabilization.

### 2.4.2 The amount of SOC in POM, S+C and SA follows distinct geoclimatic patterns

In line with our second hypothesis, all three investigated SOC fractions were related to distinct sets of environmental predictors (Figure 2-5). Temperature (RC6) was related negatively to all SOC fractions, but to a varying extent: Direct climatic proxies were most important to explain $POM_{TOT}$, less important for $SA_{TOT}$ and of least importance for $S+C_{TOT}$. We explain this discrepancy with the varying timescales at which climate can affect SOC.

At shorter timescales, climate affects biological processes. At low ambient temperature (measured range of MAT down to 3.0 °C), corresponding to high bulk $SOC_{\%}$ environments along this gradient, microbial decomposition may be more temperature constrained than plant growth and NPP. In such conditions, carbon can still be sequestered by plants, but is not efficiently decomposed by soil microbial communities (Wiesmeier et al., 2019; García-Palacios et al., 2021; Rocci et al., 2021). The consequence is a build-up of bulk SOC with a generally lower degree of decomposition (Figure 2-3 & 6, Figure S2-4) in colder environments. This interpretation is in line with observations in other studies across different land-uses, which found that low MAT leads to higher amounts of SOC as POM (Wagai et al., 2008; Rocci et al., 2021; Heckman et al., 2022). In addition, water availability, which is partially included in the variable temperature (RC6) (through MAP-PET, Table S2-8) might also be limiting for biological activity in dry and warm climates (Manzoni et al., 2012a; La Pierre et al., 2016; Zhang and Xi, 2021). For a more detailed discussion on the controls of $POM_{TOT}$, see Supplementary Text S2-1.

At longer timescales, climate acts as a driver of pedogenesis. This was supported by the fact that water balance (calculated as MAP - PET) and NPP correlated positively with RC4 “Pedogenic oxides” (Table S2-8), indicating the role of moisture and organic inputs as drivers of mineral weathering (Rasmussen et al., 2018; Slessarev et al., 2022). Higher soil moisture can favor the formation of pedogenic oxides via increased weathering rates, and thus facilitates the formation of organo-mineral associations. We therefore argue that the long-term effects of climate on SOC dynamics were mainly encapsulated in the content of pedogenic oxides (RC4), while the short-term effects were mainly reflected by temperature (RC6). However, a complete separation of these mechanisms is difficult if not impossible in nature as they occur in parallel and in interaction with each other (Figure S2-2, Supplementary Text S2-2).

While climate may affect OM turnover by microbes as well as pedogenesis, soil mineralogic properties were important proxies to predict the amount of SOC in mineral-associated SOC fractions. Pedogenic oxide content (RC4) was linked with the mineral-related fractions (Figure 2-5) and along with texture it was the dominant predictor of $S+C_{TOT}$. Soils with more pedogenic oxides and finer texture (Supplementary Text S2-3) contained a larger amount of SOC in the fraction of silt- and clay-sized particles. The S+C fraction dominated bulk $SOC_{\%}$ in low-SOC soils (Figure 2-3c), and this fraction was most directly linked to mineralogical properties (Figure 2-5); we therefore concluded that mineralogic properties were the dominant controls of bulk SOC in the low-SOC soils of this gradient that were unfavorable for aggregation (Figure 2-6a). In contrast to $S+C_{TOT}$, $SA_{TOT}$ was equally well explained by pedogenic oxide content (RC4) as well as temperature (RC4) and the degree of bulk OM decomposition (RC2). This highlights that aggregation is a consequence of mineralogical properties as well as of biological processes that are controlled by climate. The SA fraction dominated bulk $SOC_{\%}$ in high-SOC soils (Figure 2-3c), and this fraction was linked to mineralogical properties as well as directly to climate

(Figure 2-5). Therefore, we found that climate drives bulk SOC content most strongly in high-SOC soils, where biological processes contributed to SOC accumulation (Figure 2-6a).

### 2.4.3 Controls and drivers behind stable microaggregates

We expected that the controls on the amount of SOC associated with stable microaggregates would shift along the gradient. Based on the expectation that aggregation would be limited by the amount of OM (and the biological drivers thereof) in low-SOC soils, we had hypothesized a higher importance of climatic conditions in low-SOC soils. In contrast, based on the expectation that reactive minerals would limit aggregation in high-SOC soils, we had expected a higher importance of mineralogical conditions in such systems.

On the one hand, POM forms the nucleus of aggregate formation according to the hierarchy model of aggregate formation (Oades, 1984; Golchin et al., 1997). This is supported by our observation that the formation of microaggregates was partially driven by climate-controlled accumulation of POM (Figure 2-5), which was in line with this theory and with previous findings (Regelink et al., 2015; Kopittke et al., 2018; Poeplau et al., 2018; Totsche et al., 2018; Witzgall et al., 2021). The inclusion of water availability in RC6 (temperature) could represent the role that frequent wetting and drying cycles play for aggregate formation and stability (Six et al., 2004; Totsche et al., 2018; Edlinger et al., 2023). On the other hand, reactive soil minerals constitute surface coating in the initial steps of hierarchical aggregate formation (Oades, 1984; Golchin et al., 1997). It has been repeatedly observed that pedogenetic soil conditions - particularly the presence of pedogenic Al and Fe oxyhydroxides - promote the formation of microaggregates (Huygens et al., 2005; Virto et al., 2008; von Lützow et al., 2008; Regelink et al., 2015; Rasmussen et al., 2018; Wagai et al., 2020).

However, in contrast to our hypothesis of shifting controls on aggregation, statistical analyses indicated that no non-linear relationships between geoclimatic predictors and the amount of SOC in the SA fraction existed along the investigated gradient (Table S2-7). Rather, links between aggregation and the environment were consistently linear along the entire gradient without thresholds of controls, not withholding that such thresholds may exist in other environments that are not represented here. We explain this observation with the fact that many high SOC soils were soils with andic properties which contain large amounts of short-range order minerals that are highly reactive for aggregate formation (Garrido and Matus, 2012; Asano and Wagai, 2014; Matus et al., 2014; Rasmussen et al., 2018). However, Cambisols and Leptosols with high bulk $SOC_{\%}$ of 5.5 to 11.8 % also followed the same pattern suggesting the general tendency of soils to promote aggregation as SOC builds up irrespective of soil mineralogy. This further indicates that the (bio)climatic and geochemical controls on the formation of stable microaggregates were not mutually exclusive, but rather reinforced each other. Independent of

what limits aggregation in the first place, once formed, microaggregates occlude OM, slow down or even inhibit decomposition (Grandy and Neff, 2008; Mueller et al., 2012a; Heckman et al., 2022) (Grandy & Neff, 2008; Heckman et al., 2022; Mueller, Schlund, et al., 2012) and thereby likely contribute to accumulate further SOC (Kleber et al., 2007; Mueller et al., 2012a; Vogel et al., 2014).

### 2.4.4 Distinct OM quality in different SOC fractions

The chemical characteristics of OM as well as the fraction-specific C content ($C_{ABS}$) of the stable microaggregate fraction were consistently between those of particulate organic matter and the silt- and clay-sized fraction (Figure 2-4a, Figure 2-3b). Aggregate formation as in the aggregate hierarchy theory requires the occlusion of POM as an aggregation nucleus, whereas "bottom-up" aggregation starts through coprecipitation and adsorption processes of small OM-metal composites, which are assumed to be microbially transformed (Edwards and Bremner, 1967; Totsche et al., 2018; Wagai et al., 2020). As a consequence, we hypothesized that a prevalence of "bottom-up" aggregate formation would result in a larger degree of microbial transformation of OM (i.e. a higher decomposition index, DI) especially in the stable microaggregates of soils with high Al and Fe oxy-hydroxide contents. However, the DI of OM associated with the SA fraction did not change significantly along the investigated gradient (Figure 2-4b). In contrast, the DI of the S+C fraction decreased significantly with increasing bulk $SOC_{\%}$ (Figure 2-4b). It is therefore likely that shifts of OM stabilization processes in the S+C fraction masked shifts in aggregate formation mechanisms, as we will elaborate in the following section.

The S+C fraction is a mix of small silt- and clay-sized aggregates and diverse free associations of OM and minerals (Totsche et al., 2018; Lavallee et al., 2020; Schweizer, 2022). Importantly, minerals preferentially associate with decomposition products and microbially transformed OM (Kaiser et al., 1997; Grandy and Neff, 2008; Sollins et al., 2009; Kallenbach et al., 2016; Heckman et al., 2022). This is generally in line with the observation that the OM in the S+C fraction was on average more microbially transformed (i.e. a higher DI) than the particulate OM fraction (POM) (Figure 2-4a). However, the decrease of the DI in the S+C fraction along the bulk $SOC_{\%}$ gradient indicated a shift in the type of OM that associated with this fraction. In high-SOC soils with postulated "bottom-up" aggregate formation (e.g. Andosols), the OM in the S+C fraction was less microbially transformed than in low-SOC soils. We explain this with the fact that OM stabilization can follow different microbial pre-processing pathways, either *in vivo* or *ex vivo* (Liang et al., 2017). The *in vivo* pathway describes OM stabilization that follows full microbial transformation of the OM (i.e. microbial uptake of substrate, subsequent breakdown through microbial catabolism and reassembly into chemically different molecules by microbial anabolism). In contrast, the *ex vivo* pathway describes OM stabilization that bypasses microbial uptake

and therefore full microbial catabolism and anabolism. Examples of the *ex vivo* pathway are direct stabilization of soluble plant-derived OM or stabilization of structural plant-derived OM that was only modified by extracellular enzymes. The decrease of the DI in the S+C fraction along the bulk $SOC_{\%}$ gradient could therefore indicate a shift from *in vivo* to *ex vivo* dominated OM stabilization (Grandy and Neff, 2008; Sanderman et al., 2014; Angst et al., 2017; Liang et al., 2017; Mikutta et al., 2019; Cotrufo et al., 2022). More specifically, the shift could stem from increased occlusion of small (< 63 µm) plant material in high SOC soils as compared to low SOC soils where microbial transformation of OM input is more advanced. Regardless of the underlying mechanism (see Supplementary Text S2-4 for a more detailed discussion), this shift from *in vivo* to *ex vivo* stabilization and the shift from aggregate formation as in the aggregate hierarchy theory to "bottom-up" aggregation in high SOC soils could have canceled each other out in terms of OM quality. In soils dominated by "bottom up" aggregation, the OM in the smaller building blocks was relatively undecomposed due to a prevalence of *ex vivo* stabilization (Figure 2-6b). As a result, the DI of the stable microaggregate fraction remained unchanged along the entire gradient (Figure 2-6a).

### 2.4.5 Reasons for less OM decomposition in stable microaggregates

The overall lower DI (i.e. less decomposed OM) in the stable microaggregate fraction as compared to the S+C fraction (Figure 2-4a) could support the notion that microaggregates function as an efficient mechanism to protect OM from microbial decomposition (Dungait et al., 2012; Abramoff et al., 2018; Totsche et al., 2018; Heckman et al., 2022). This would be in line with findings of significantly older $^{14}C$ ages in occluded OM as compared to free POM (e.g. Heckman et al., 2022; Rasmussen et al., 2005; Schrumpf et al., 2013). However, it is important to note that persistence of OM associated with stable microaggregates must not necessarily stem from physical occlusion. It has also been hypothesized that chemically recalcitrant carbon could preferentially initiate aggregation (Wagai et al., 2009). In three high-SOC soils, $C_{ABS}$ in the SA fraction almost reached 30 % (Figure 2-3b). This suggests that the aggregate-associated OM in these soils may partially lack physical occlusion as the volume ratio of OM to mineral matrix reaches high values (Wagai et al., 2009). However, $C_{ABS}$ values in this fraction were likely overestimated due to the applied procedure for sand correction (section *Soil physical characterization*). Regardless of which process ultimately caused the observed patterns of OM quality, this study shows that (1) stable microaggregates contain OM that is distinct from POM and the OM in the S+C fraction, and that (2) this pattern is largely consistent across very diverse soil systems (Figure 2-6).

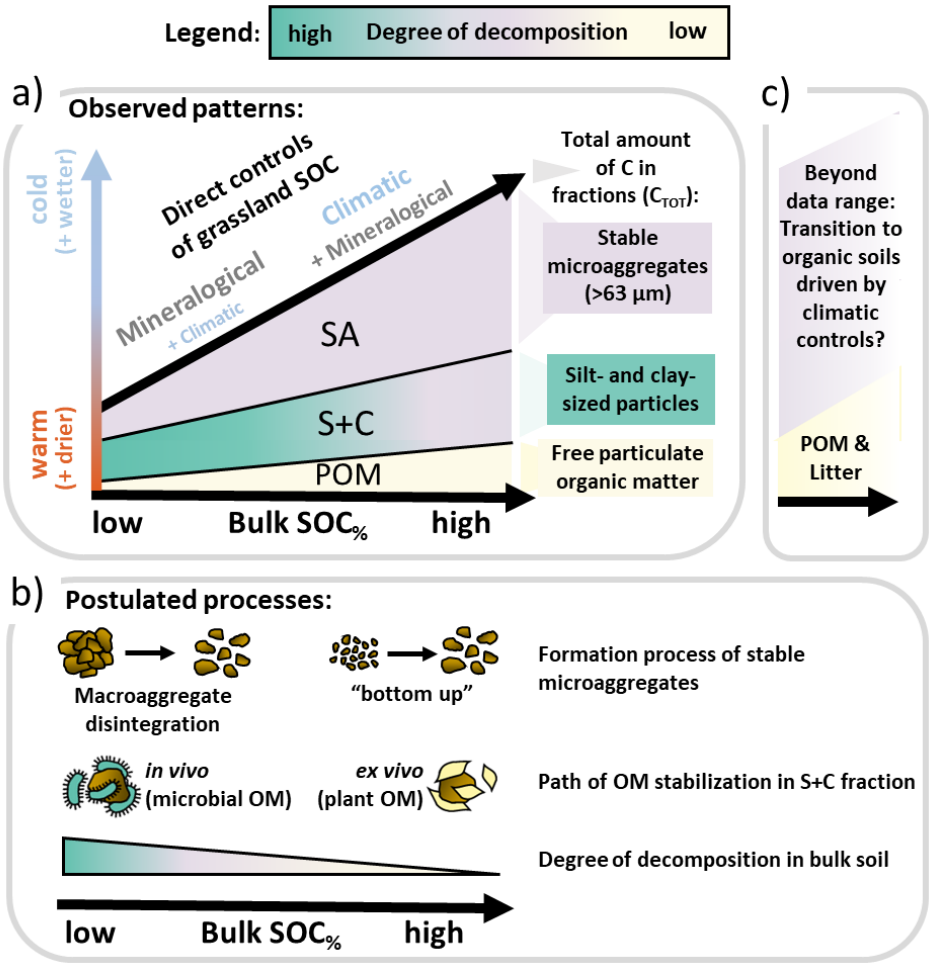



**Figure 2-6.** Conceptual summary of the controls and processes that drive bulk $SOC_{\%}$ across geoclimatic conditions in temperate grasslands. a) Bulk $SOC_{\%}$ increases with $C_{TOT}$ in free particulate organic matter (POM), silt- and clay-sized particles (S+C) and stable microaggregates (SA), where the area under the bold black arrow represents the amount of SOC in the fractions. In low-SOC soils, the S+C fraction dominates bulk $SOC_{\%}$, and in high-SOC soils the SA fraction dominates bulk $SOC_{\%}$. Low-SOC soils are present in warm climates (left y-axis), with limitations in mineral stabilization potential as direct control of bulk $SOC_{\%}$ (lower left corner). High-SOC systems are present in cold climates, where climatic constraints on decomposition become increasingly important as direct control (upper right corner). b) In low-SOC systems, stable microaggregate formation seemed to follow the "aggregate hierarchy model" concept via macroaggregates, whereas "bottom-up" aggregation of smaller particles seemed to dominate in high-SOC systems. In low-SOC systems, the dominant pathway of organic matter association with the S+C fraction seemed to be *in vivo*, whereas in high-SOC systems it seemed to be *ex vivo*. c) The observed patterns are constrained to the investigated climatic range (right side) and grassland topsoils.

## 2.5 Conclusion

The results of our study imply that stable microaggregates are an important piece in the puzzle to understand SOC dynamics at the global scale. Mineral-associated SOC constituted the majority of soil C across a wide range of grasslands, spanning a 30-fold variation in bulk SOC content. In particular, this study presents clear evidence that stable microaggregates are an important SOC fraction in mineral topsoils of grasslands with high SOC content. Moreover, stable microaggregates follow their own dynamics of formation and alteration, which are distinct from POM or non-aggregated mineral fractions. Soils form stable microaggregates when sufficient amounts of reactive minerals and OM input are present. In this study, the amount of SOC in the stable microaggregate fractions was linked to a distinct set of environmental variables, predictable across soil types and geo-climatic regions. Common geochemical and climatic proxy variables, namely extractable Al- and Fe- oxy-hydroxide contents and annual climatic means (MAT and water balance) were identified as suitable proxies for the two main ingredients (OM and reactive minerals) and processes of aggregate formation. In contrast, the amount of SOC in the free POM fraction was more strongly linked to climate, and the amount of SOC associated with silt- and clay- sized particles was mainly linked to soil mineralogical variables. Across the range of investigated soils, the OM in stable microaggregates was generally less decomposed than the OM associated with free silt- and clay-sized particles, despite potential mechanistic differences in aggregate formation between Andosols and soils without volcanic influence.

We conclude that SOC does not just accumulate either as undecomposed free particulate organic matter or as decomposed mineral-associated organic matter. Rather, a substantial proportion of SOC in the investigated soils accumulated in the form of stable microaggregates. Given that the stable microaggregate fraction (1) varies in quantitative importance across (soil) environmental gradients, (2) follows distinct dynamics across soil physicochemical and climatic conditions and (3) constitutes a qualitatively - and likely functionally - distinct SOC fraction, there is an urgent need to better understand the functional implications of stable microaggregates at large scales. We argue that particularly the distinction between free POM and OM associated with stable microaggregates may have important implications. SOC stabilization may not be accurately represented without explicitly considering the role of stable microaggregates. The drivers of microaggregation are not yet fully resolved causally across soil types and large spatial gradients. For mechanistic insights into the formation of stable microaggregates, their potential to protect and stabilize SOC, and their vulnerability under changing climatic conditions, stable microaggregates need to be considered explicitly in more experimental studies across large scales.

## 2.6 Acknowledgements

We thank Manuel Casanova for soil sampling and photographs of the sites, Katherine Rebolledo for sample shipment, Nuno Bischofsberger and Annina Maier for assistance with the laboratory work. We thank Rota Wagai and two anonymous reviewers for their constructive comments on the manuscript. Financial support in Chile was provided by Fondecyt 1121138 and 1161492, and in Switzerland by ETH Zurich.

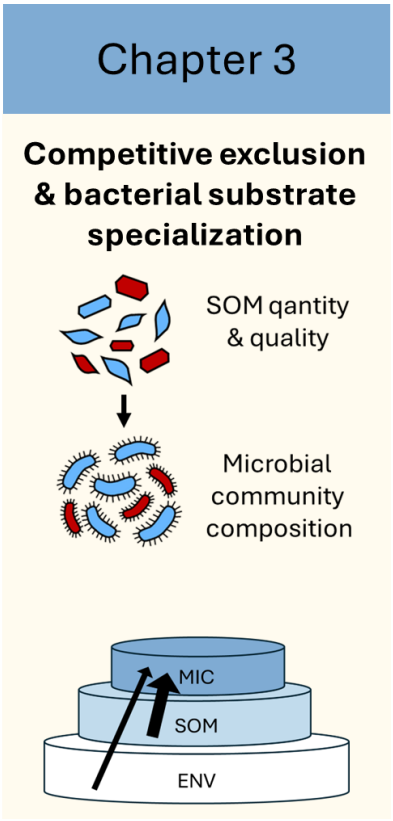
## Data availability statement

All data used in this study are available within the paper or in a permanent open access online repository of the ETH Zurich Research Collection under the following DOI: https://doi.org/10.3929/ethz-b-000649389

## Code availability

The R code used and produced for statistical analysis in this study is deposited in a permanent open access online repository of the ETH Zurich Research Collection under the following DOI: https://doi.org/10.3929/ethz-b-000649389

# 3. Quantity over quality: the effects of soil organic matter on soil bacterial diversity along a geoclimatic gradient




This chapter is in preparation for submission as:

Wasner D, Han X, Schnecker J, Frossard A, Zagal Venegas E, Doetterl S. Quantity over quality: the effects of soil organic matter on soil bacterial diversity along a geoclimatic gradient.

## Abstract

To understand the controls over soil biogeochemical cycles along environmental gradients, it is necessary to enhance our understanding of the factors that determine soil bacterial community composition at larger scales. It is well established that the quantity of soil organic matter (SOM) is an important driver of soil bacterial community composition from the regional to global scale. In addition to SOM quantity, qualitative properties of SOM have been hypothesized to affect soil bacterial community composition through the mechanism of bacterial substrate specialization. However, up to date it remains unresolved to what extent SOM quality affects soil bacterial community composition at large scales. In this study, we investigated how SOM quantity and quality structure the soil bacterial communities along a biogeochemical gradient of grassland soils. We measured relative abundance patterns of soil bacteria using 16S rRNA barcoding and characterized SOM quantity and quality with a range of different methods. We could explain a large fraction of bacterial communities (up to 59.6 % of 16S rRNA reads) in soils with high SOM quantity (along the continuum between 0.6 to 18.7 % SOC) and low soil pH (along the continuum between pH 4.1 to 6.7). Soils with high SOM quantity and low soil pH had a lower bacterial alpha diversity. SOM quality did not have a pronounced effect on soil bacterial community composition. This suggests that bacterial substrate specialization does not shape soil bacterial community composition at the investigated scale.

## 3.1 Introduction

Soil bacteria fulfill a wide range of crucial ecosystem functions and are known to differ greatly across climatic, ecological and pedological gradients (Fierer, 2017; Delgado-Baquerizo et al., 2018). It is therefore crucial to better understand the drivers of soil bacterial community composition to link environment and ecosystem functioning. One tool to study the distribution of soil bacteria across space are correlational studies along environmental gradients. Based on such studies, soil pH is understood to be the strongest driver of soil bacterial community composition at regional to global scales. Below neutral soil pH as well as in strongly alkaline soils, alpha diversity generally decreases with increasing acidity (Lauber et al., 2009; Rousk et al., 2010; Griffiths et al., 2011; Zhou et al., 2016; Bahram et al., 2018).

Soil organic matter (SOM) quantity is the second strongest determinant of soil bacterial community composition at regional to global scales (Delgado-Baquerizo et al., 2016; Bahram et al., 2018; Bastida et al., 2021). However, in contrast to soil pH, the effects of SOM quantity on soil bacterial community composition are biome-dependent. Maestre et al. (2015) and Delgado-Baquerizo et al. (2016) reported positive links between bacterial alpha diversity and SOM content for global dryland soils and soils of

the southern hemisphere, respectively. In contrast, a survey of high latitudinal soils (Siciliano et al., 2014) and a regional study across Scotland (Delgado-Baquerizo et al., 2017) found decreased bacterial alpha diversity in conditions of higher substrate availability. Mechanistically, several processes could link bacterial community composition with SOM quantity. One proposed mechanism is competitive exclusion, where taxa that are adapted to high substrate conditions (copiotrophic taxa) outcompete taxa that are adapted to low substrate conditions (oligotrophic taxa). This mechanism is supported by observations of shifts from oligotrophic to copiotrophic bacteria in response to substrate addition in experiments (Fierer et al., 2007; Langenheder and Prosser, 2008; Ng et al., 2014; Bhatnagar et al., 2018; Geyer and Barrett, 2019; Stone et al., 2023). Conceptually and in experiments, adaptation to high substrate conditions – and consequently competitive exclusion of oligotrophic taxa by copiotrophic taxa - are often connected to the potential of organisms to grow fast (Fierer et al., 2007; Ho et al., 2017). Microbial adaptations for fast growth are for instance large numbers of low-affinity transporters for substrate uptake, trade-offs in the enzymatic toolbox and increased numbers of ribosomal RNA operons (Trivedi et al., 2013; Roller et al., 2016; Ho et al., 2017). The oligotrophic-copiotrophic continuum forms the backbone of several trait-based concepts to classify the functioning of soil bacteria (Ho et al., 2017; Malik et al., 2020). The stress gradient framework expands the idea of competitive exclusion and proposes that positive interactions could also play a role in the link between alpha diversity and substrate conditions. According to this concept, positive interaction (such as specialization and facilitation) takes place between taxa in substrate-limited soils, thereby increasing alpha diversity in such systems (Bastida et al., 2021). In addition, intermediary effects of SOM have been used to partially explain links between alpha diversity and temperature at the continental to global scale. For example, several studies have proposed that increased ecosystem productivity could drive higher alpha diversity in warmer systems (Delgado-Baquerizo et al., 2016b; Zhou et al., 2016).

Importantly, bulk SOM is not a homogeneous entity. Rather, the physicochemical characteristics of SOM can vary across a range of qualitative properties. For instance, SOM quality can range from low to high oxidation state, from low to high C:N ratio, and from small to large molecules (von Lützow et al., 2008; Simpson and Simpson, 2012; Lehmann and Kleber, 2015). Continuously, a diverse range of OM compounds enters the soil (Sokol et al., 2019) and gets transformed through microbial decomposition (Grandy and Neff, 2008; Roth et al., 2019). Microbial turnover thereby converts plant-derived OM into microbial compounds, which are generally smaller and have a lower C:N ratio than plant material. Selective microbial turnover can enrich less palatable (i.e., more “recalcitrant”) compounds in the soil. Overall, the molecular diversity of OM decreases in the course of microbial transformation (Roth et al., 2015; Hoffland et al., 2020; Davenport et al., 2023; Jones et al., 2023). The physicochemical characteristics of SOM in turn strongly affect its fate and turnover. Microbially transformed OM preferentially associates with soil minerals, where it can be physically and chemically protected from

further decomposition (Grandy and Neff, 2008; Sollins et al., 2009; Kallenbach et al., 2016; Totsche et al., 2018; Heckman et al., 2022).

As a consequence of the complex mechanisms that are involved in SOM turnover and stabilization, climatic decomposition constraints and soil mineralogic properties can lead to a diverse range of SOM quality across soil horizons, soil types, and soil development stages (Mainka et al., 2022; S. Stoner et al., 2023; Wasner et al., 2023). SOM quality may in turn affect microbial community composition and nutrient acquisition strategies. Substrate specialization in bacteria has been shown among soil crust bacteria (Baran et al., 2015), and different taxa of soil bacteria varied significantly in the diversity of substrates they could use (Y. Wang et al., 2022). Further, it has been shown that more bacterial taxa are able to perform competitively well under excess of chemically labile substrates as compared to excess of chemically complex substrates (Goldfarb et al., 2011). In situ, soil bacterial communities degrade the litter native to their environment more efficiently than chemically different litter that stems from other systems (Wallenstein et al., 2013). Only a subset of bacteria possesses an enzymatic toolbox to degrade chemically more complex (structural) substrates (Trivedi et al., 2013; Berlemont and Martiny, 2015). Consequently, SOM quality has been proposed to be a main driver of soil bacterial community composition (Fierer et al., 2007). However, to the knowledge of the authors, the potential of SOM quality to shape community composition in the presence of confounding environmental factors such as climate or soil biogeochemical characteristics has only rarely been assessed. Even when this has been attempted, SOM quality was only approximated by bulk C:N ratio (Cederlund et al., 2014; Delgado-Baquerizo et al., 2016b, 2017; Bahram et al., 2018) or soil physicochemical fractionation (Szoboszlay et al., 2017). It is therefore an open question, how strongly SOM quantity versus SOM quality shape soil bacterial community composition across pedo-climatic gradients.

In this study, we aimed to investigate the role of SOM quantity versus SOM quality in shaping soil bacterial community composition across a large spatial and environmentally contrasting gradient. We sampled a set of 35 grassland topsoils (0 – 10 cm) along a 2300 km north-south transect in Chile ranging from warm arid steppe to cold tundra climate, thereby covering almost the full climatic range in which natural grasslands occur. We expected that soil bacterial diversity and community composition would be linked to SOM quantity as well as SOM quality. Experimental and regional studies have found competitive exclusion to be a dominant mechanism to shape bacterial diversity in resource rich conditions. Based on this, we hypothesized (1) to find a negative link between SOM quantity and alpha diversity. Further, decomposition has been shown to lower the qualitative diversity of SOM, and soil bacteria are known to exhibit substrate specialization in experimental conditions. We therefore hypothesized (2) to find higher alpha diversity in systems with a lower degree of bulk SOM decomposition. In order to investigate links between abundance patterns of taxonomic units and biogeochemical variables, we conducted Illumina MiSeq barcoding of a region of the 16S rRNA. We

further measured relevant physicochemical properties of the soils, compiled climatic data of the sites and provided a detailed characterization of SOM quality. For this qualitative SOM characterization, we employed RockEval pyrolysis, isotopic measurements, diffuse reflectance infrared spectroscopy as well as UV-Vis spectroscopy.

## 3.2 Material and Methods

### 3.2.1 Soil sampling

The A-horizons of 35 topsoils (0 - 10 cm) under extensive and natural grassland were sampled in the summer seasons of 2017 and 2018 across a large scale in Chile, covering 10 World Reference Base soil orders. The soils along this geoclimatic gradient span soil organic carbon (SOC) contents from 0.6 to 18.7 % (Figure 3-1). The gradient was chosen because it represents a large variability in climatic and geochemical soil conditions as well as organic matter (OM) properties, while at the same time it allows to constrain the analysis to a coherent type of land cover systems (grassland-dominated biomes). A further criterion for site selection was a carbonate free soil condition (null HCl reaction), resulting in a $pH_{CaCl2}$ from 4.1 to 6.7. After sampling, the samples were sieved to 2 mm, and kept frozen at -20 °C until further processing.

### 3.2.2 Climate classification

Climate classification of the sites was done following Köppen-Geiger with the R-package 'kgc' (Bryant et al., 2017). Mean annual temperature (MAT) and precipitation (MAP) were taken from WorldClim Version 2 (Fick and Hijmans, 2017). Mean annual potential evapotranspiration (PET) was taken from (Trabucco and Zomer, 2018). Both datasets average monthly climate data from 1970 to 2000, at a spatial resolution of 30 arc seconds. Water balance was calculated monthly as MAP minus PET. The gradient covers a MAT range from 3.0 to 17.1 °C, a range in water balance from -1382 to 1704 mm and Köppen-Geiger climate zones ranging from arid steppe (Bsk) at lower latitudes in the Chilean north to polar tundra (ET) in the Chilean south. Most sites (30) are in the temperate climate zone, under climates with cold and warm summers, with or without dry season (Cfb, Cfc, Csb, Csc). Dominant vegetation, landscape position and soil moisture regime were either characterized on site or supplemented from previous surveys (detailed information in Chapter 2).

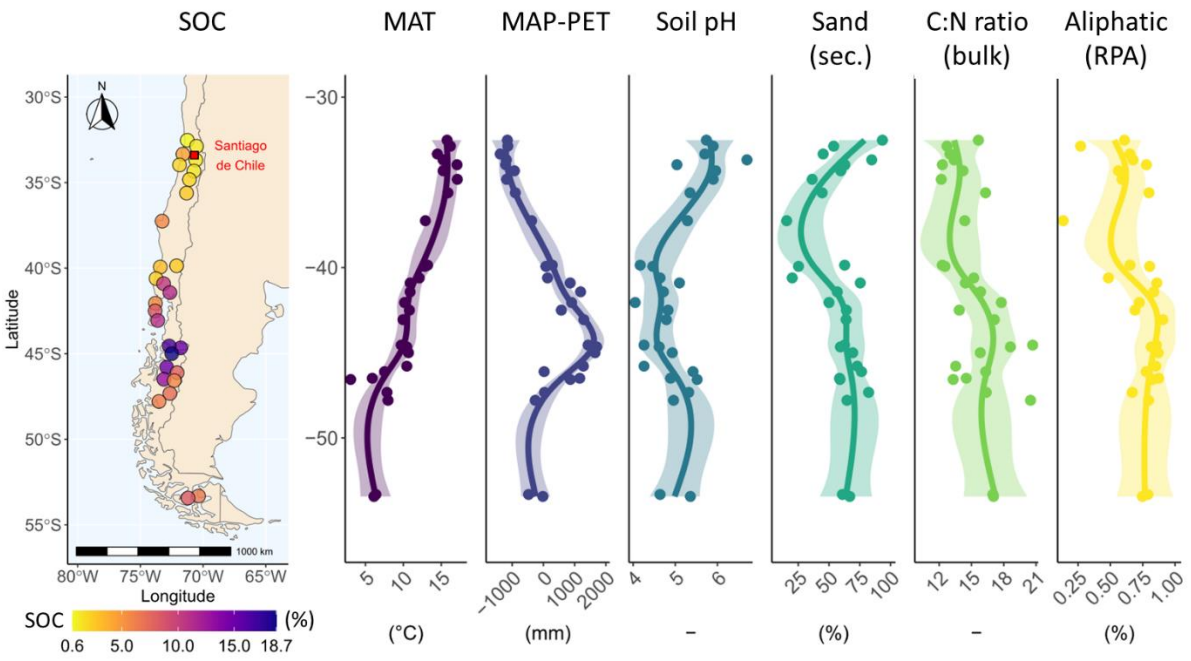



**Figure 3-1.** Map of the investigated sites (n = 35) across Chile. On the map, SOC contents are color-coded. In addition, two climatic variables (mean annual temperature (MAT); water balance (MAP-PET)), two soil physicochemical variables (soil pH; Secondary and content (sand sec.) and two descriptors of organic matter quality (C:N ratio of bulk soil, relative peak area (RPA) of aliphatic groups based on infrared spectroscopy) are shown along the gradient. Points show measured values (x-axes) and the location (latitude) at the same scale as labeled on the map (y-axis). Lines are for visual aid only (representing fits of general additive models with k = 7, with standard error of the fits).

## 3.2.3 Measurement of biogeochemical variables

*Soil pH and texture*

As a measure for soil acidity, an important criteria for soil biogeochemistry, soil pH was determined in 0.01 M $CaCl_2$ solution (fresh soil:solution ratio of 1:5). After 10 min of shaking, the samples were left to rest for 24 h and again shortly shaken before pH was measured in suspension using a pH meter (713 pH Meter, Metrohm, Switzerland).

Soil texture was determined via laser diffraction using a particle size analyzer (PSA) (LS 13 320, Beckman Coulter, USA). Importantly, in our study we determined soil texture as soil secondary texture, namely the textural composition of soil without disintegrating stable microaggregates and oxy-hydroxide concretions. These stable units are structural components of natural soils (Totsche et al., 2018), and therefore strongly influence the physicochemistry of the soil microbial environment. This is why further statistical analyses were focused on the sand fraction (as the fraction dominated by a mix of secondary concretions and stable aggregates + primary sand particles) versus the clay fraction. Briefly, fresh soil

was shaken for 3 h with 10 % Na-hexametaphosphate to dissolve macroaggregates, and the resulting particles were quantified with the PSA. Particle size contributions were calculated as percent of total particle volume, and size classification followed the WRB system (IUSS Working Group WRB, 2015): clay < 2 µm, 2 µm < silt < 63 µm, 63 µm < sand < 2000 µm.

*Quantity and quality of bulk SOM*

Soil organic carbon (SOC) and total soil nitrogen (TN) contents were quantified via total combustion using a CN analyzer (Vario MAX Cube, Elementar GmbH, Germany). Bulk C:N ratios were calculated using the molar ratio of SOC:TN. Four different approaches were used to characterize the degree of transformation and stabilization of bulk SOM: (i) RockEval pyrolysis, (ii) diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS), (iii) the relative proportion of organic carbon (OC) in the particulate organic matter fraction (POM), and (iv) the shift in the $^{13}C/^{12}C$ isotope ratio between OC in POM and bulk SOC.

(i) RockEval pyrolysis was done based on Sebag et al. (2016) to assess thermostability of SOM. Briefly, milled soil samples were pyrolyzed on a Rock-Eval 6 device (Vinci Technologies, France), first in an $N_2$ atmosphere between 200 to 650 °C, and second in an oxidized atmosphere between 400 and 850 °C, both with a heating rate of 25 °C $min^{-1}$. Subsequently, the Rock-Eval I-Index for the degree of biological transformation of SOM was calculated exactly as introduced in Sebag et al. (2016). The higher the I-index value, the less biologically transformed is the bulk SOM.

(ii) DRIFTS was done to assess the relative dominance of functional groups in SOM. Briefly, milled material (< 50 µm) was scanned in duplicate in the mid-infrared region (7500 to 600 $cm^{-1}$) at a resolution of 2 $cm^{-1}$, using a Fourier transform IR (FT-IR) spectrometer with a high-throughput screening extension (HTS-XT) (Bruker Optics, Vertex 70, Germany). Spectra were normalized against a gold background (Infragold NIR-MIR Reflectance Coating, Labsphere) and corrected for atmospheric $CO_2$ and $H_2O$ in the OPUS spectrometer software (Bruker Optics, Germany), and 32 co-added scans per sample were averaged. All subsequent spectra processing was done using the R-packages "simplerspec" (Baumann, 2020) and "prospectr" (Stevens and Ramirez-Lopez, 2020). For correction of light scatter, spectra were resampled to a range of 4000 to 600 $cm^{-1}$ with duplicates averaged and normalized using the normal variate method. Based on published information (Parikh et al., 2014), six wavenumber ranges were assigned to three types of functional groups: aliphatic C-H (anti)symmetric stretches: 2950 – 2910 and 2866 – 2836 $cm^{-1}$; aromatic C=C stretches: 1540 – 1524 and 1520 – 1510 $cm^{-1}$; carboxylic acid C=O stretch and carboxylate C-O asymmetric stretch: 1734 – 1718 and 1650 – 1636 $cm^{-1}$. Peak areas of the wavenumber ranges were integrated with a local baseline correction and for each of the three functional groups, the peak area was divided by the summed peak area of all three

groups to obtain relative peak areas (RPA). Increasing RPA carboxylic and decreasing RPA aliphatic values indicate a progressing degree of SOM decomposition (Ryals et al., 2014; Mainka et al., 2022).
(iii) To obtain the relative proportion of OC in the POM fraction as a measure for the amount of potentially labile accumulating SOM, soil physical fractionation was conducted, as described in detail in Chapter 2. Briefly, a combined size and density SOC fractionation was done based on Zimmermann et al. (2007). We defined POM as particles > 63 µm and lighter than 1.6 g $cm^{-3}$. First, air dried soils were dispersed by ultrasonication (15 g dry weight in 100 ml nanopure water) with an energy input of 24 J $ml^{-1}$ (50 Hz, 33.5 W). The sonicated soil was then wet sieved with nanopure water over a 63 µm mesh size aperture sieve until the throughflow became clear. The fraction > 63 µm was air-dried, and a subsample of 8 g was further separated by density fractionation. For this, the subsample was first mixed in 40 ml of 1.6 g $cm^{-3}$ SPT using a vortex for 10 seconds, and then centrifuged (Sigma 3-16 KL, 15 min, 3000 xg). The floating light POM fraction was decanted onto 0.45 µm mesh size Whatman glass filters (grade GF/F), to first remove the SPT and then clean the fraction with 100 ml nanopure water with the suction of a vacuum pump, before finally being oven-dried at 50 °C. The heavy fraction which remained as a residual after the decantation (stable microaggregates) was washed five times by addition of 40 ml nanopure water, mixing on a Vortex, centrifugation (settings as above) and decantation, before being oven-dried at 50 °C. Organic carbon content and $^{13}C/^{12}C$ isotope ratios of all fractions were quantified on an EA-IRMS (FlashSmart Elemental Analyzser coupled with a Finnigan DELTAplusXP, Thermo Fisher Scientific, USA), and the contribution of C in the POM fraction to bulk SOC was calculated based on the POM organic carbon content and the relative mass contribution of this fraction to bulk soil. This value is expressed as a percentage and referred to as “OC in POM fraction”. The higher this value, the more untransformed SOM has accumulated freely (and thus physically unprotected) in the soil.
(iv) To quantify the shift in the $^{13}C/^{12}C$ isotope ratio between POM and bulk SOC, the $^{13}C/^{12}C$ isotope ratio of bulk soil was quantified using EA-IRMS as described above, and both $^{13}C/^{12}C$ isotope ratios were expressed as $\delta^{13}C$ values relative to the Vienna Pee Dee Belemnite (VPDB). Subsequently, the shift (referred to as 13C shift) was calculated as the ratio of POM-$\delta^{13}C$ over bulk-$\delta^{13}C$. The higher this value is, the more microbially transformed the bulk SOM relative to POM input, because microbial decomposition leads to an accumulation of $^{13}C$ relative to $^{12}C$ (Dijkstra et al., 2006; Lerch et al., 2011).

*Quantity and quality of extractable SOM*

We used KCl-extractable SOM as a proxy for microbially available SOM. For this, we extracted soils with 1 M KCl (fresh soil:solution ratio of 1:15). After 1 h on a horizontal shaker, the extracts were filtered through Whatman 42 filter paper and concentrations of organic C (KCl-OC) and total N (KCl-N) were determined with a TOC-L Analyzer, coupled with a TNM-L unit (Shimadzu, Switzerland). Using a Tecan Infinite M200 microplate reader (Tecan GmbH, Austria) we then quantified nitrate ($NO_3^-$) and ammonia

($NH_4^+$) in the KCl extracts colorimetrically, based on the $VCl_3$/Griess reaction and the Berthelot reaction, respectively (Hood-Nowotny et al., 2010). We considered the sum of $NO_3^-$-N and $NH_4^+$-N as inorganic N (KCl-$N_i$), and estimated the concentration of extractable organic N (KCl-$N_o$) by subtracting KCl-$N_i$ from KCl-N. The C:N ratios of SOM in the extractable pool were subsequently calculated using the molar ratio of KCl-OC:KCl-$N_o$. To characterize the chemical properties of the extractable SOM, ultraviolet-visible (UV-Vis) spectroscopy was performed (Li and Hur, 2017). In brief, the pH of the extracts was adjusted to a value of 2, and the UV absorptivity between 200 to 600 nm was measured at a temperature of 20 °C, using a Cary 60UV-Spectrophotometer (Varian, USA) with a cell length of 1 cm. All absorbances were blank corrected. The aromaticity of extractable SOM (KCl Aromaticity) was calculated as the molar absorptivity at 260 nm, based on the Beer-Lambert law and expressed as L $mol^{-1}$ $cm^{-1}$ (Dilling and Kaiser, 2002). Higher values indicate that a higher proportion of the SOM has aromatic groups. The degree of humification of extractable SOM (KCl Humification) was estimated based on the ratio of absorbance values at 300 nm over 400 nm (Claret et al., 2003; Li and Hur, 2017), where higher values indicate lower degrees of humification. The molecular weight of extractable SOM (KCl Molecular weight) was estimated based on the ratio of absorbance values at 250 nm over 365 nm (Peuravuori and Pihlaja, 1997; Li and Hur, 2017), where higher values are linked to lower SOM molecular weight.

### 3.2.4 Bacterial abundance and community composition

*DNA extraction, target amplification and sequencing*

Total DNA was extracted from the bulk soil of the 35 sites in triplicates, resulting in 105 samples. The quantity of extracted soil was adjusted according to the SOC content, in order to maximize extraction efficiency: for soils with SOC < 3.0 %, each extraction was of 500 mg fresh soil , for soils with 3.0 % < SOC < 7.0 %, each extraction was of 300 mg, and for soils with SOC > 7 %, each extraction was of 250 mg. DNA was extracted with the DNeasy® PowerSoil® Pro Kit (Quiagen, Germany), following the manufacturer manual. The DNA content of each aliquot was quantified with Nanodrop (Thermo Fisher Scientific, Germany), and the V3-V4 region of the prokaryotic (bacterial and archaeal) small-subunit (16S) rRNA gene was amplified with the primers 341F (CCT AYG GGD BGC WSC AG) and 806R (GGA CTA CNV GGG THT CTA AT) as described in Frey et al. (2016). Polymerase chain reaction (PCR) amplification with tailed primers was done on a Veriti™ 96-Well Fast Thermal Cycler (Applied Biosystems, USA) using 20 ng DNA per reaction, following the protocol modified from Frey et al. (2016): an initial denaturation at 95 °C for 2 min, 36 cycles of denaturation at 94 °C for 40 s, annealing at 58 °C for 40 s and elongation at 72 °C for 1 min followed by a final elongation at 72 °C for 10 min. The amplicons were purified with Agencourt AMPure XP beads (Beckman Coulter, Berea, CA) and were quantified fluorometrically in microplates with the QuantiFluor® ONE dsDNA System (Promega, USA) following the manufacturer

manual, using a Tecan Infinite M200 microplate reader (Tecan GmbH, Austria). Purified amplicons were subsequently sent to the Genome Quebec Innovation Center (Montreal, Canada) for barcoding using the Fluidigm Access Array technology and 250 base pairs paired-end sequencing running for 518 cycles on the Illumina MiSeq v3 platform (Illumina Inc., USA). Raw sequences were deposited in the NCBI database under the accession number PRJNA1066703.

*Sequence processing with the dada2 pipeline*

Sequence processing was performed in R Studio (Version 2022.2.3.492) using the dada2 workflow (Callahan et al., 2016). Primers were trimmed off the left sides of the reads, and the right sides of the reads were truncated to 245 and 235 bases for forward and reverse reads, respectively, to improve matching quality while maintaining an overlapping region of 15 base pairs. PhiX genomes were removed. For sample inference, all samples were pseudo-pooled. Forward and reverse reads were merged with a minimum overlap of 12 base pairs, allowing for 1 mismatch. Chimeras were removed with default settings. Across all 105 samples, a total number of 3051949 reads were retained, on average 29066 per sample. A summary of the dada2 workflow is given in Table S3-1. Taxonomy was assigned by using the Silva nr99 v138.1 database (Quast et al., 2012) as a training set, and with a minimum bootstrap confidence for assigning a taxon set to 60 %. Subsequently 281 amplicon sequencing variants (ASVs) that were not bacteria were removed from the dataset.

*Rarefying of reads*

The minimum number of reads across all 105 samples was 17188. Rarefying was done by iterative random subsampling without replacement (100 iterations) to 17188 reads, implemented by the function “rrarefy” in the R-package “vegan” (Oksanen et al., 2022). All subsequent analysis was conducted on rarefied data. Rarefying is accepted for subsequent biodiversity analysis, (Cameron et al., 2021; Schloss, 2023), but is not suitable for differential abundance analysis in experimental contexts, due to variance inflation (McMurdie and Holmes, 2014). To avoid confusion, we explicitly point out that the problem of variance inflation only affects group-wise regression analysis in the context of differential abundance analysis, but not correlation analysis as presented in this study (see section *Taxa-specific distributions assessed with relative read abundance*).

*Sequence processing of rarefied reads to derive taxonomic units*

The rarefied community data of all 105 samples was filtered to keep only reads that are desirable for subsequent linkage between taxon-specific abundance patterns and biogeochemical variables

(Supplementary Text S3-1). The applied workflow for sequence filtering and processing included four goals; (1) generate taxonomic units at the highest possible taxonomic resolution, while (2) excluding very rare (i.e. not ubiquitous) taxa from analysis, (3) removing redundant abundance information and (4) minimizing unnecessary loss of abundance information. To achieve these four goals, we applied a workflow in three steps (Figure S3-1). In step 1, we aggregated the reads of assigned ASVs at five levels of taxonomic resolution (in the order from low to high taxonomic resolution: Phylum, Class, Order, Family and Genus). At each level (plus at the ASV level), the aggregated data was filtered for ubiquitous taxa. We defined ubiquitous taxa as taxa that were present in more than half of the 105 samples. The remaining taxa (that were present in less than half of the 105 samples) were defined as rare taxa. With this filtering step, we removed the reads of rare taxa and retained all reads that belonged to taxa that were ubiquitous at the respective levels of taxonomic resolution. In step 2, we aggregated the reads of rare taxa within each lower-resolution taxon and removed the reads of the respective lower-resolution taxon from further analysis. This aggregation of rare taxa was done by subtracting the sum of reads of all ubiquitous taxa within a lower-resolution taxon from the reads of the respective lower-resolution taxon (Figure S3-1). The resulting aggregated units of rare taxa reads were then referred to as "ΣRare-groups". This procedure of aggregating rare taxa into ΣRare-groups allowed to test related rare taxa for coherent patterns at lower taxonomic resolution, where they jointly fulfilled the criterion of ubiquity. Taxa belonging to unassigned lower-resolution taxa (e.g. genera within an unassigned family) were excluded from this procedure, because they were not clearly relatable to specific lower-resolution taxa. Thus, taxonomic units were only retained at a taxonomic resolution lower than ASV, if either (i) they represented aggregated rare taxa (ΣRare-groups), if (ii) they were unassigned (and thus taxonomic "dead ends'' without clearly relatable higher-resolution taxa), or if (iii) they did not contain any ubiquitous taxa at a higher taxonomic resolution. This approach removes redundant read information and therefore prevents multiple testing in downstream analysis. In step 3, the newly aggregated groups of rare taxa (ΣRare-groups) were in turn filtered by the same criterion of ubiquity as described in step 1. Through this processing, we obtained 347 taxonomic units that included the sum of the reads of all members of a taxonomic group, and 90 taxonomic units that included just the sum of the reads of the rare members of a taxonomic group (i.e., ΣRare-groups) (

Figure S3-2a). Taken together, the data was aggregated into 436 taxonomic units at the highest taxonomic resolution at which they were assigned and at which they fulfilled the criterion of ubiquity. Each taxonomic unit is unique in the dataset (i.e. no overlap of ASVs with other taxonomic units). For more detail on the sequence processing, please see Supplementary Text S3-1. We subsequently averaged the triplicate of each site by taking the mean abundance of each retained taxonomic unit, and all taxa abundances were expressed as relative abundance values (%).

### 3.2.5 Statistical analysis

All statistical analyses were performed in R Studio (Version 2022.2.3.492). All correlation matrices were done with the R-package “corrplot” (Wei and Simko, 2021), and the figure of the taxonomic tree in the supplementary material was created with the R-package “collapsibleTree” (Khan, 2018).

*Simplification of biogeochemical variables*

Measured biogeochemical variables were tested for outliers with Rosner´s test, using the R-package “EnvStats” (Millard, 2013), and significant outliers which were obviously caused by laboratory artifacts were replaced with the mean value of the respective variable (excluding significant outliers). This was done to avoid outlier based biases in downstream regression and coefficient analysis. The rotated principal component analysis (rPCA) which was applied in subsequent analysis cannot process missing values, which is why simple removal of outliers insufficient and replacement was necessary. Ranges of the biogeochemical variables are shown in Table S3-2. The biogeochemical variables were in parts highly autocorrelated (Figure S3-3a). To reduce autocorrelation and improve interpretability of the downstream analysis, we therefore used the R-package “psych” (Revelle, 2022) to perform dimension reduction with rPCA, using varimax rotation on scaled and centered variables. Rotation of the PCA reduces the number of variables that correlate with individual rotated components (RCs) and thus facilitates the interpretability of the RCs. The 8 RCs with the largest Eigenvalues were retained, in order to keep 80 % of the variance of the dataset. The retained RCs were interpreted based on their dominant loadings (loading ≥ 0.5, see Table S3-3), minimizing autocorrelation (Figure S3-3b). These 8 RCs were subsequently used as predicting variables in the downstream regression and correlation analyses.

*Drivers of bacterial diversity*

To investigate links between the environment and bacterial diversity, we conducted linear regression analysis. Diversity calculations were done at the ASV level. ASV singletons were removed from the rarefied triplicates (by setting their counts to zero), and mean community composition for each site was calculated based on the remaining counts (which accounted for 96.4 to 99.1 % of relative read abundance). Zeros were considered as zeros for the calculation of means. For alpha diversity, we calculated bacterial richness (observed ASVs) and Shannon diversity index with the R-package “vegan” (Oksanen et al., 2022). A model to predict bacterial richness and Shannon indices was built using a cross-validated stepwise regression approach to select the most explanatory RCs, implemented in the R-package “caret” (Kuhn, 2008). Briefly, the independent variables were scaled, and the data was split 100 times into a 75 % training set and a 25 % validation set. Model performance was assessed using RMSE, and a maximum of four (out of eight potential) RC predictor variables were retained to constrain

model complexity and avoid overfitting. The distribution of the residuals was assessed with the Shapiro-Wilk and the Breusch-Pagan tests, and the explanatory content of each variable (expressed as $R^2$) was extracted with the R-package "relaimpo" (Grömping, 2006). For beta diversity, permutational multivariate analysis of variance (PERMANOVA) was conducted using a Bray-Curtis dissimilarity matrix of the bacterial community data. The analysis was conducted with 1000 permutations and implemented with the adonis2 function of the R-package "vegan" (Oksanen et al., 2022).

*Taxa-specific distributions assessed with relative read abundance*

The relative read abundance of a taxonomic unit can change along a gradient either because absolute read abundance of this taxonomic unit changes (taxa-inherent trends), or because the absolute read abundances of other taxa change (Morton et al., 2017; Props et al., 2017; Alteio et al., 2021). We therefore considered relative read abundance as a measure for the dominance of the investigated taxonomic unit. This definition acknowledges that relative abundance trends do not necessarily stem from taxa-inherent properties alone, while it allows to interpret the valuable information that can be derived from relative abundance patterns. Different bacterial taxa can contain differing numbers of 16S operons (Bonk et al., 2018; Alteio et al., 2021), and therefore read counts do not directly translate into cell counts. For this study, we assume that 16S operon numbers are consistent within taxonomic units across the gradient. To the knowledge of the authors, this assumption is quietly inherent in all gradient studies of microbial communities. We therefore subsequently refer to relative read abundance as "relative abundance".

To investigate the covariation of taxonomic units and biogeochemical conditions along the gradient, we conducted taxa-specific correlation analysis. In brief, we conducted Pearson correlation analysis between relative abundance of each individual taxonomic unit that remained after filtering (see Figure S3-1 for detail) and the biogeochemical RCs. To reduce type I errors (false positives), we conducted a false discovery rate (FDR) correction on the correlation results. The FDR correction was implemented following an adjusted Benjamini-Hochberg correction (described as the "BL" approach in (Korthauer et al., 2019) across all 3496 (436 taxonomic units x 8 RCs) correlations. Specifically, p-values were corrected by the conservative Benjamini-Hochberg correction, before being multiplied by an estimate of the proportion of true null hypotheses ($\pi_0$) based on the uncorrected p-values. Following (Korthauer et al., 2019), the $\pi_0$ estimates were obtained with the R-package "swfdr" (Leek et al., 2022), using ubiquity (i.e. the percentage of sites at which taxa were present) as a covariate to adjust $\pi_0$, rendering adjustment of taxa with more observations less conservative. Correlations with adjusted p-values ≤ 0.1 were retained for discussion. This means that the retained correlations will contain less than 10 % false positives. Relative abundances of the correlated taxonomic units were summed up for each soil to obtain the percentage of explainable relative abundance in each soil. We conducted cross-validated

stepwise regression (as described in section *Drivers of bacterial diversity*) on this sum in order to describe in which soils much or little relative abundance was explainable with independent soil and environmental variables. We conducted a power analysis using the R-package "pwr" (Champely, 2020) to estimate the probability of overlooking existing links between biogeochemistry and patterns of taxonomic units in the correlation analysis. For a probability of type II errors (false negatives) of 0.2 (i.e., a statistical power of 0.8) and a p-value of 0.1, the lowest reliably detectable Pearson correlation coefficient in a single test was 0.41. However, accounting for FDR correction (using the highest unadjusted p-value that remained after discarding all correlations with an adjusted p-value > 0.1, i.e. p-value 0.0067), the lowest reliably detectable Pearson correlation coefficient was 0.55.

## 3.3 Results

### 3.3.1 Simplification of biogeochemical variables

Among the key soil and climatic variables selected with relevance for microbial processes, the gradient spans a variation of SOC from 6 to 187 g $kg^{-1}$, soil pH values between 4.1 and 6.7, (secondary) particles in the sand size between 16 to 93 % and (secondary) clay-sized particles between 1 and 9 %. SOM quality variables ranged from 12.1 to 20.7 for the bulk C:N ratio, 4 to 39 % OC in POM, and 42 to 91 % aliphatic RPA. The simplification of the 19 measured biogeochemical variables by rotated principal component analysis (rPCA) resulted in eight rotated components (RCs) free of autocorrelation (Figure S3-3). The ranges of all 19 measured biogeochemical variables are shown in Table S3-2, and the loadings assigned with the 8 RCs are shown in Table S3-3. RC1 is primarily associated with SOM quantity representing total as well as available (i.e. KCl-extractable) organic carbon and nitrogen in the bulk soil. This dimension is also positively associated with the balance of precipitation and evapotranspiration (MAP-PET) (Table S3-3), because wetter systems along the gradient accumulate more SOM (Chapter 2). Five RCs (2, 4-8) represent different aspects of organic matter quality ("SOM quality A to E") that were captured with different approaches to characterize SOM chemistry, namely UV-Vis spectroscopy (RC2), DRIFTS (RC5), the shift in the $^{13}C/^{12}C$ isotope ratio between POM and bulk SOC (RC6), and the bulk C:N ratio (RC7). RC4, referred to as "SOM quality B (POM)", accounted for the majority of correlations between taxa-specific abundance patterns and SOM quality variables (see below). This RC was mainly loaded with the relative proportion of OC in the free POM fraction, and with the aromaticity in the KCl-extractable pool. Notably, SOM quality B (POM) was also negatively correlated with the precipitation and evapotranspiration balance (MAP-PET) (Table S3-3). This is because in wet and cool soil systems, an increasing fraction of SOM is at least partially protected from microbial decomposition by association with stable microaggregates (Chapter 2). The variable SOM quality B (POM) consequently

reflected systems in which parts of SOM were present as free (unoccluded) coarse plant litter that is potentially available for decomposition. The main soil physicochemical dimensions are captured by two distinct RCs, in which RC3 represents soil texture and RC8 soil pH.

### 3.3.2 Drivers of bacterial diversity

Bacterial alpha diversity was lowest in the soils at the wet and cool extremes of the investigated gradient (northern Patagonia), and highest in the warmer and drier region of central Chile (Figure 3-2).

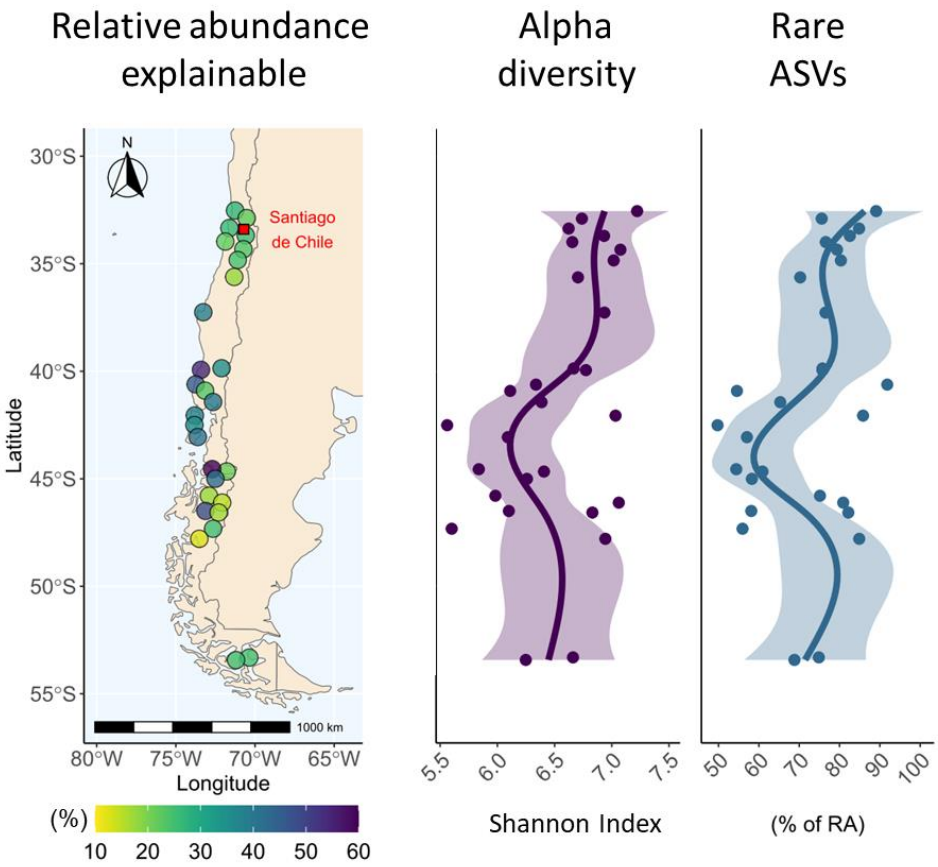



**Figure 3-2.** Results of the microbial analysis along the Chilean gradient. The color code on the map shows for each site the percentage of relative abundance that can be explained with biogeochemical variables. In addition, alpha diversity (calculated as the Shannon index) and the relative abundance (RA) of rare amplicon sequencing variants (ASVs) are shown along the gradient. Rare ASVs are defined as ASVs that were present in less than half of the 105 samples. Points represent measured data points, and the y-axes represent latitude at the same scale as labeled on the map. Lines are for visual aid only (fits of general additive models with k = 7, with standard error of the fits).

Both bacterial richness ($p < 0.001$, d.f. = 32, adj. $R^2 = 0.38$) as well as Shannon diversity ($p < 0.05$, d.f. = 32, adj. $R^2 = 0.17$) showed a negative link with SOM quantity. Richness showed a strong positive link with soil pH, and Shannon diversity a weaker negative link with SOM quality E (C:N ratio) (Figure 3-3). Beta diversity (investigated as Bray-Curtis dissimilarity) was weakly linked to six out of eight RCs, without a clearly dominant predictor emerging. Note that an analysis of patterns between soil mass and alpha diversity among the triplicates confirmed that the negative link between alpha diversity and SOM quantity is not an artifact of SOC-based soil mass adjustment in the DNA extraction step (Supplementary Text S3-2 and Figure S3-4).

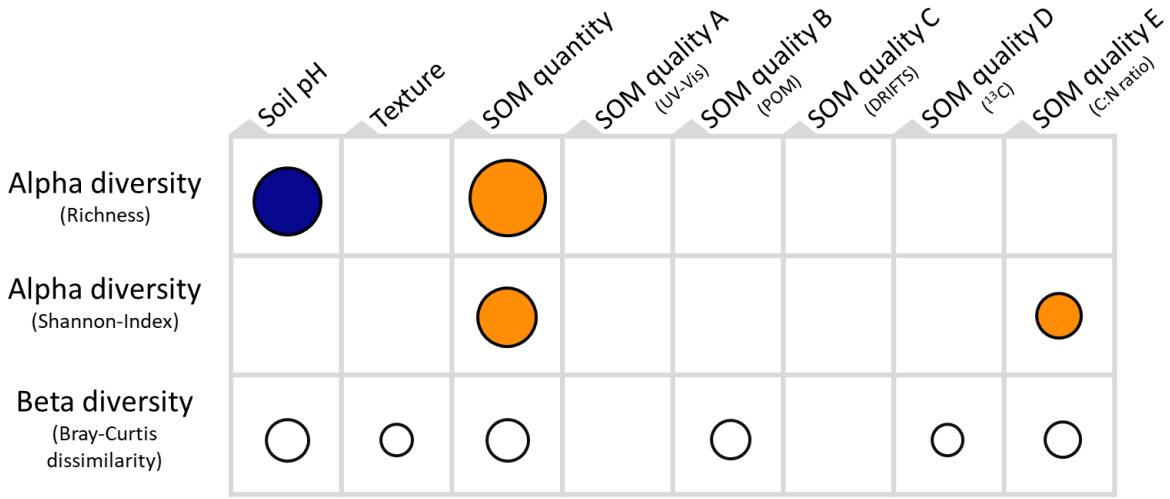



**Figure 3-3.** Significant predictors of alpha and beta diversity. Circle sizes indicate the proportion of variance explained by the individual predictors (unadjusted $R^2$, ranging from 0.04 to 0.23). Light orange indicates a negative and dark blue a positive relationship. In the case of beta diversity, circles are open because for multivariate analysis of variance using distance matrices interpretable coefficients cannot be assigned. UV-Vis = ultraviolet-visible spectroscopy; POM = particulate organic matter; DRIFTS = diffuse reflectance infrared spectroscopy.

### 3.3.3 Drivers of the relative abundance of bacterial taxonomic units along the gradient

On average, the presence of every third read (29.7 % of average relative abundance) could be explained with the tested biogeochemical variables, with a range from 12.5 % to 59.6 % for specific soils (Figure 3-2). Importantly, more relative abundance could be explained in soils with high SOM quantity and acidic soil pH (cross-validated stepwise regression, $p < 0.001$, d.f. = 28, adj. $R^2 = 0.54$) than for soils with low SOM and neutral soil pH. Of the 436 taxonomic units tested, the relative abundances of 11.5 % (n

= 50) were correlated with soil pH, 5.5 % (n = 24) with SOM quantity and 6.0 % (n = 26) with different proxies of qualitative composition of SOM (Figure 3-4a). A summary of all correlation results is presented in Supplementary File S3-1. Averaged across the gradient, soil pH, SOM quantity and SOM quality could be linked with 14.4 %, 10.1 % and 4.6 % of relative abundance, respectively. The main SOM quality variable that correlated with abundance patterns was SOM quality B (POM). All correlations with this variable were positive (Supplementary File S3-1). Across all levels of taxonomic resolutions, a similar percentage of ΣRare-groups and coherent taxa showed correlations with biogeochemical variables (Figure S3-5). This shows that ΣRare-groups can form consistent functional entities to the same extent as coherent taxa. Aggregation of rare taxa into ΣRare-groups is therefore a worthwhile step in order to maximize the amount of relative abundance data that can be used for correlation analysis at the highest possible taxonomic resolution. Note that an analysis of patterns between soil mass and relative abundances of taxonomic units among site triplicates confirmed that the positive links between relative abundance and SOM quantity were not an artifact of SOC-based soil mass adjustment in the DNA extraction step (Supplementary Text S3-2 and Figure S3-6).

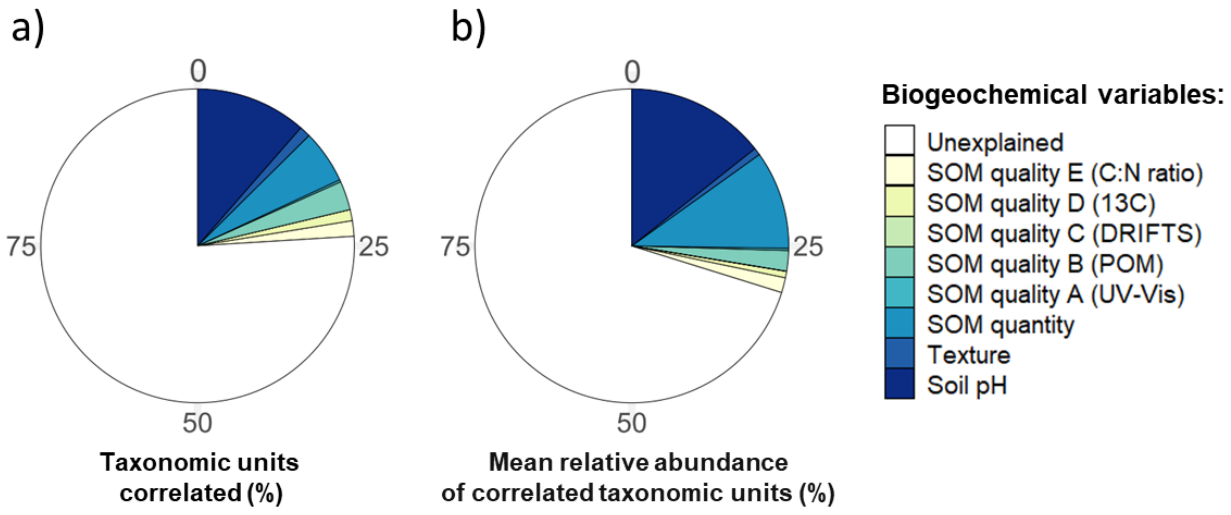



**Figure 3-4.** a) Percentage of investigated taxonomic units with relative abundances that correlate with biogeochemical variables across the dataset. b) Mean relative abundance of taxonomic units that correlate with the biogeochemical variables across the dataset. Significance threshold for all correlations: FDR-adjusted p-value ≤ 0.1. FDR = false discovery rate; UV-Vis = ultraviolet-visible spectroscopy; POM = particulate organic matter; DRIFTS = diffuse reflectance infrared spectroscopy.

## 3.4 Discussion

### 3.4.1 Soil biogeochemistry explains large parts of soil bacterial community composition

The relatively high proportion of relative abundance explainable by biogeochemical variables underlines the important role of biogeochemistry in shaping soil bacterial community composition. We could explain the largest part of bacterial community composition (up to 59.6 % of relative abundance) in soils with high SOM quantity and lower soil pH, found in the wet central region of the gradient (Figure 3-2). These soils were mostly Andosols, but also Cambisols and an Acrisol. In contrast, in the drier and more diverse soils towards the mediterranean climate of central Chile and towards the tundra systems of southern Patagonia, only a smaller fraction (as low as 12.5 % of relative abundance) of community composition could be explained with the investigated biogeochemical variables. Soils in which only a small fraction of explainable community composition was explainable encompassed a Gleysol, a Luvisol and Kastanozems as well as Mollisols, but also Cambisols and Andosols.

### 3.4.2 Alpha diversity decreases with soil pH

We found a positive relationship between richness and soil pH (Figure 3-3) in the investigated range of soil pH (moderately acidic, pH 4.1 to 6.7) which is consistent with previous findings (Lauber et al., 2009; Rousk et al., 2010; Griffiths et al., 2011; Zhou et al., 2016; Bahram et al., 2018). By removing autocorrelation between soil pH and SOM quantity through rotated component analysis (Figure S3-3 and Table S3-3), we found a direct negative effect of low soil pH on alpha diversity. Low pH represents a major stress for (soil) bacteria, which have cell internal pH values close to neutral (Lauber et al., 2009; Krulwich et al., 2011). Consequently, prevailing under acidic conditions requires a suite of specialist adaptations (Krulwich et al., 2011; Guan and Liu, 2020; Lund et al., 2020; Ramoneda et al., 2023). Low soil pH, which we predominantly found in the Andosols and Cambisols of northern Patagonia, may select for adapted specialists. In contrast, a comparatively larger number of rare microbial taxa were found at near-neutral soil pH values, where alpha diversity increased.

### 3.4.3 Alpha diversity is negatively linked to SOM quantity

As hypothesized, both measures of alpha diversity (richness and Shannon index) had a strong negative link to SOM quantity along the gradient. Low-SOM soils (found across a diversity of soil groups such as an Arenosol, a Planosol, Kastanozems, a Gleysol and a Gleysol) had a higher richness of amplicon sequencing variants (ASVs) than high-SOM soils (mostly Andisols and Cambisols) (Figure 3-2). This is,

to our knowledge, the first time that this pattern has been found across a large geoclimatic gradient. Our finding is in line with previous experimental studies (Langenheder and Prosser, 2008; Geyer and Barrett, 2019) as well as a survey of high latitudinal soils (Siciliano et al., 2014) and a regional study (Delgado-Baquerizo et al., 2017), which all observed decreased soil bacterial alpha diversity in conditions of higher substrate availability. However, this pattern could potentially be explained by two contrasting mechanisms as follows.

First, akin to macroecology, a possible underlying mechanism could be competitive exclusion. In systems with high substrate availability, taxa that are well-adapted to copiotrophic conditions could outcompete other taxa, and thus lead to lower alpha diversity. The observation of competitive exclusion across the studied gradient would have an interesting nuance - the long time scale (beyond short term substrate pulses) at which the mechanism would be required to act in order to drive bacterial community composition. The soils in this study constitute a gradient of SOM quantity in the sense of a long-term condition (and not a short-term flush of substrate). The native bacterial communities are therefore adapted to the substrate status that is measured as SOM quantity. The lower microbial diversity observed in the high-SOM Andosols and Cambisols in the central region of this gradient (Figure 3-2) would therefore imply that even in the long term, most soil bacteria are not capable of competitively coping with conditions of high substrate availability. This conclusion would be in line with the observation of limited growth-rate plasticity of taxa in response to environmental changes across ecosystems, as reported by Morrissey et al. (2019). A related mechanism that could lead to competitive exclusion in substrate rich systems are microbial interactions. The stress gradient framework suggests more positive interaction (such as specialization and facilitation) between species in substrate-limited soils (Bastida et al., 2021). Similarly, high substrate supply facilitated negative interactions which resulted in competitive exclusion in an artificial microcosm experiment (Ratzke et al., 2020). SOM quantity could further reflect effects of soil bacterial density (Schnecker et al., 2019). Soils with higher SOM quantity have higher microbial biomass. This could translate into increased proximity between bacterial cells (Raynaud and Nunan, 2014), perhaps leading to increased antagonistic interaction resulting in the exclusion of more taxa (Cordero and Datta, 2016).

The second explanation that could describe the observed patterns of microbial diversity and richness equally well would be a positive correlation between richness and actual SOM turnover (not quantified in this study). At continental to global scales, positive links between alpha diversity and temperature have been established (Delgado-Baquerizo et al., 2016b; Zhou et al., 2016). These links were partially hypothesized to be driven by increased ecosystem productivity in warmer systems. These observations caution to consider the meaning of SOM quantity for microbial diversity carefully. Does a larger amount

of SOM automatically reflect a resource bonanza in which copiotrophic bacteria can outcompete others? Or could high SOM quantity on the contrary rather imply that SOM accumulates because microbial decomposition of SOM is limited by other factors, such as climatic constraints? Along the examined geoclimatic gradient, climatic conditions shape the build-up of SOM via exerting physiological constraints on biological processes to a varying degree for vegetation and microorganisms (Chapter 2). For example, at the wet and cool extremes of the investigated gradient (northern Patagonia), annual plant growth and NPP may be less temperature constrained than microbial decomposition, so that carbon can still be sequestered by plants, but not decomposed by soil microbial communities to the same extent (Wiesmeier et al., 2019; García-Palacios et al., 2021). Moreover, the Andosols of this region might be particularly efficient at accumulating and stabilizing SOM and protecting it from microbial decomposition. In contrast, in the Kastanozems and Chernozems of warmer and drier regions (central Chile), annual SOM input by plants may be matched by microbial decomposition (Figure 3-1). This means that in some regions of the gradient, factors such as climatic constraints - and not substrate limitation - might keep microbial activity in check. In consequence, systems with a relatively high turnover of SOM inputs (i.e. where microbial decomposition activity matches the SOM input) may have higher microbial activity that supports a higher alpha diversity but also results in a smaller build-up of SOM (e.g. the C input limited, drier and warmer soils of the Chilean north). In contrast, systems in which SOM turnover is hampered by climatic constraints of microbial activity rather than by plant inputs (the cold Chilean south) may be microbially less diverse, while accumulating more SOC. In such a scenario, there would be no causal link between SOM quantity and richness, but both system properties would be consequences of the balance between SOM input and decomposition.

### 3.4.4 Alpha diversity dictates community composition along the gradient of SOM quantity

The relative abundance of 5.5 % (n = 24) of taxonomic units (which accounted on average across the gradient for 10.1 % of relative abundance) followed a trend with SOM quantity (Figure 3-4). In agreement with our hypothesis, all observed correlations between SOM quantity and relative abundance patterns were positive (Supplementary File S3-1). This could imply that taxa with clear copiotrophic traits (as opposed to clear oligotrophic traits) dominated the microbial communities along the gradient. However, we could not identify SOM quantity-related relative abundance patterns in several phyla (e.g. Gamma-Proteobacteria, Actinobacteria; see Supplementary File S3-1) that have been reported to shift relative abundance in response to substrate addition (Cleveland et al., 2007; Eilers et al., 2010; Morrissey et al., 2016; Geyer and Barrett, 2019). We therefore conclude that oligotrophic vs. copiotrophic traits per se might not necessarily result in dominance in substrate-poor

vs. substrate-rich systems, respectively. Other conditions than immediate substrate availability could limit microbial growth and activity and prevent bacteria with copiotroph traits from becoming dominant in SOM-rich systems. We argue that the observed positive links between relative abundance and SOM quantity are rather linked to the negative relationship between SOM quantity and alpha diversity (Figure 3-3). When less taxa are present, the remaining taxa are automatically more dominant in relative terms. Consequently, the fraction of individual communities that can be explained with biogeochemical variables inversely mirrors alpha diversity along this gradient (Figure 3-2). In other words, we understand community composition better in soils with a smaller fraction of rare taxa (i.e. high-SOM soils), because the drivers of rare taxa are harder to pinpoint.

### 3.4.5 Alpha diversity is only weakly linked to SOM quality

Bulk C:N ratio was highest in some Andosols, Cambisols and Leptosols of Patagonia, which were also among the soils with the largest SOM quantities (Figure 3-1), indicating that these soils contain a large fraction of less decomposed SOM. Undecomposed SOM is chemically more diverse than microbially transformed SOM (Roth et al., 2015; Davenport et al., 2023; Jones et al., 2023). We had hypothesized higher alpha diversity in systems with a lower degree of bulk SOM decomposition, based on the expectation that substrate specialization would be a strong driver of bacterial community composition. However, alpha diversity was only weakly linked to SOM quality, and Shannon diversity was lower in soils with a higher bulk C:N ratio (SOM quality E (C:N ratio), Figure 3-3). This finding contradicted our hypothesis. An experimental long-term study by Cederlund et al. (2014) neither found strong links between SOM quality and bacterial phyla, and a regional study by (Delgado-Baquerizo et al., 2017) also found a negative relationship between bulk C:N ratio and Shannon diversity. These findings are in contrast to short-term experimental studies as well as genomic and metabolic analyses, which showed a potential for strong substrate preference among soil bacteria (Goldfarb et al., 2011; Trivedi et al., 2013; Baran et al., 2015; Berlemont and Martiny, 2015; Y. Wang et al., 2022). This discrepancy indicates that bacterial substrate preference may primarily show in conditions of substrate saturation or otherwise ideal growth conditions. Substrate preference does not seem to play a major role in the assembly of *in situ* bacterial communities at the scale of bulk soil, where other environmental constraints act on bacterial physiology.

### 3.4.6 SOM quantity matters more than SOM quality for community composition

Nevertheless, certain measures of SOM quality could explain taxa-specific relative abundance patterns that SOM quantity could not explain. A similar number of taxonomic units was correlated with SOM

quality and with SOM quantity (Figure 3-4). However, the taxonomic units which correlated with SOM quality comprised only 4.6 % of average relative abundance. In contrast, the taxonomic units which correlated with SOM quantity comprised 10.1 % of average relative abundance. Therefore, SOM quality had a smaller effect on community composition than SOM quantity. Notably, the SOM quality variables jointly explained a similar amount of variability (unadjusted $R^2$ = 0.15) in the PERMANOVA of beta diversity as did soil pH, SOM quantity and soil texture together (unadjusted $R^2$ = 0.18) (Figure 3-3). Beta diversity was calculated on the ASV level without aggregation of rare ASVs, while correlation analysis was based on ubiquitous taxonomic units. Although an exact comparison of these two analyses is not possible, the strong observed contrast in explanatory content could nevertheless indicate that SOM quality affected the presence and absence of individual rare ASVs, while it related less to patterns of ubiquitous taxonomic units that were investigated in the correlation analysis.

We found a group of ASVs in the family *Micrococcaceae* that was dominant in soils with a higher C:N ratio (SOM quality E). Soil-dwelling members of the *Micrococcaceaea* have been repeatedly found in hydrocarbon polluted soils (Dastager et al., 2014), and have been shown to be able to decompose complex (aromatic) substrates (Sims et al., 1986; Storey et al., 2018). Further, two groups in the order Streptosporangiales, two groups in the family *Micromonosporaceae* and two groups in the genus *Mesorhizobium* were positively correlated with SOM quality B (POM). SOM quality B (POM) reflects systems in which free coarse plant material, potentially available for decomposition, accumulates. In agreement with this interpretation, members of the Streptosporangiales have been found to be involved in primary decomposition of plant material in soils (Otoguro et al., 2014). *Micromonosporaceae* have been found to preferentially inhabit wet soils and peat (Trujillo et al., 2014). While the pattern in our data contrasts with the apparent moisture preference, it matches with the observation that members of this family have been found to prefer systems where free undecomposed plant material accumulates. Overall, the patterns of individual taxonomic units support the interpretation that the relative abundance of parts of the soil bacterial community was linked to SOM quality. However, we did not find that SOM quality was a strong determinant of bacterial diversity or bacterial community composition along the gradient.

### 3.4.7 Limitations

Three types of limitations need to be considered in the interpretation of this study. First, a power analysis showed that even at a comparatively high type I error rate (α-value of 0.1) only strong correlative trends (Pearson correlation coefficient > 0.55) could be detected with a reliable type II error rate (statistical power of 0.8). This means that the analysis might miss correlations due to limited

statistical power. Nevertheless, this approach was necessary in order to reveal relative abundance patterns of individual taxonomic units. In addition, Pearson correlation only captures monotonic linear relationships.

Second, several environmental variables that might be relevant for soil bacterial community composition were not directly considered in this study. Such dimensions include plant diversity (Lange et al., 2015; Fierer, 2017), resource diversity (Dal Bello et al., 2021), stochasticity and dispersal limitations as well as environmental stressors (Tripathi et al., 2018; Richter-Heitmann et al., 2020).

Third, a common challenge to the interpretation of soil bacterial community composition based on simple soil DNA extraction is that a large part of bacterial communities can be inactive (Carini et al., 2016; Fierer, 2017; Couradeau et al., 2019; Camillone et al., preprint). However, we argue that this challenge is more pronounced in attempts to link community composition and microbial responses to short-term manipulation. In this study, the investigated biogeochemical variables describe the long-term stage in which bacterial communities are embedded. Thus, the distinction between active and inactive bacteria is not so relevant, since the goal is to understand the physicochemical drivers of the entire native bacterial community (integrating also inactive cells that might be active at other times). Similarly, many processes of community assembly may act on the microscale rather than on the bulk soil scale. We nevertheless show that the presence of up to every second read can be explained with bulk soil characterization.

## 3.5 Conclusions

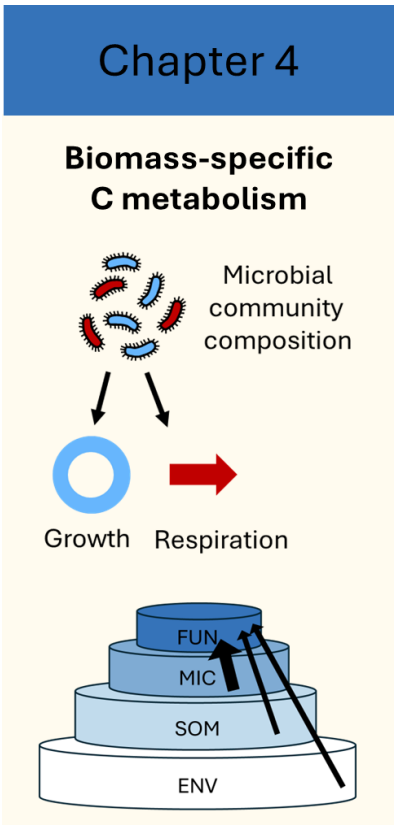
In this study, we investigated the role of SOM quantity and quality as drivers of bacterial community composition along a geoclimatic gradient of grassland soils. Bacterial alpha diversity was negatively correlated with SOM quantity. A major fraction (up to 59.6 % of 16S rRNA reads) of bacterial community composition could be explained with biogeochemical variables. This was particularly the case in soil systems with high SOM quantity (along the continuum between 0.6 to 18.7 % SOC) and low soil pH (along the continuum between pH 4.1 to 6.7), where microbial alpha diversity was lower. Such soils were located in wet and cool environments. In contrast, in drier soil systems with low SOM quantity and near-neutral soil pH, alpha diversity was higher and a smaller fraction of bacterial community composition could be explained (down to 12.5 % of 16S reads). However, the identification of the underlying mechanisms that link SOM quantity and diversity remain challenging. Experimental studies at the macroscale will be required to test whether or not alpha diversity is causally linked to SOM quantity via competitive exclusion or whether they are simply correlated, both resulting from climatic constraints.

Qualitative composition of SOM had a smaller effect on bacterial community composition than SOM quantity. This suggests that the mechanism of bacterial substrate specialization does not have a pronounced effect on soil bacterial community composition at large spatial scales. SOM quality mainly affected rare taxa. In order to detect such links more accurately, large datasets or high sequencing depth will be required, providing higher taxonomic resolution (such as family- or genus-level) to better understand soil bacterial distribution.

## 3.6 Acknowledgments

We thank Manuel Casanova for soil sampling and photographs of the sites and Katherine Rebolledo for sample shipment. We further thank Beat Stierli for lab assistance, and Luisa Minich, Annina Maier and Nuno Bischofsberger for help with the lab work. D.W. would like to thank Luisa Minich for her support. Financial support in Chile was provided by Fondecyt 1121138 and 1161492, and in Switzerland by ETH Zurich.

# 4. Environment and microbial community composition drive microbial traits and functions in the macroscale soil organic carbon cycle




This chapter is in preparation for submission as:

Wasner D, Schnecker J, Han X, Sun Y, Frossard A, Zagal Venegas E, Boeckx P, Doetterl S. Environment and microbial community composition drive microbial traits and functions in the macroscale soil organic carbon cycle.

## Abstract

Soil microbial traits and functions play a central role in soil organic carbon(SOC) dynamics. However, at large spatial scales it is unresolved whether specific environmental settings (e.g. climate, geology, soil types) (i) drive key microbial traits and functions directly, or rather indirectly through microbial community composition, and (ii) whether some microbial traits and functions are attributable to microbial community composition alone. To address this knowledge gap, we used 33 grassland topsoils (0 – 10 cm) from a north-south transect in Chile, which represents a wide range of geoclimatic conditions. First, we incubated the soils for one week in standardized conditions favorable for microbial activity and quantified a wide range of soil microbial traits and functions. For this, we measured soil microbial carbon and nitrogen, enzyme kinetics, microbial respiration and growth rates as well as carbon use efficiency (CUE). Second, we characterized climatic and physicochemical conditions as well as bacterial and fungal community composition of the soils. We then applied three different cross-validated regression approaches to investigate how strongly the measured microbial traits and functions were linked with the environmental setting vs. the microbial actors. We show that environmental factors (predominantly the amount of soil organic matter) determined patterns of microbial biomass, which in turn explained microbial respiration and growth rates. However, respiration and growth normalized for microbial biomass (*i.e.* biomass-specific) were most strongly linked to the microbial community composition instead of the environmental setting. Notably, both biomass-specific respiration and growth followed distinct trends and were related to different features of microbial community composition, which resulted in strong effects on microbial CUE. We conclude that at the investigated scale, the drivers of CUE are decoupled aspects of microbial metabolism, which is partially determined by microbial community composition. The environmental setting and the microbial community composition affect different microbial traits and functions, and therefore both factors need to be considered in the context of large scale SOC dynamics.

## 4.1 Introduction

Soil microbial activity is central to understand soil organic carbon (SOC) dynamics. Soil microbes require carbon for energy, which results in the mineralization of carbon and its respiration as $CO_2$ (in the case of aerobic metabolism). However, soil microbes also require carbon for the synthesis of biomass and extracellular products through anabolism. This carbon does not immediately leave the soil, and can potentially become stabilized over longer periods of time (decades to millennia) in the soil matrix (Cotrufo et al., 2015; Kallenbach et al., 2016; Sokol and Bradford, 2019). Unless microbial processing of SOC is hindered by direct physiological decomposition constraints (e.g. low temperatures, oxygen limitation or drought stress; Keiluweit et al., 2016; Schimel, 2018; García-Palacios et al., 2024) or

physicochemical inaccessibility (e.g. through strong organomineral associations or occlusion; Dungait et al., 2012; S. Stoner et al., 2023), all SOC ultimately undergoes microbial processing. As a consequence, microbial processes such as growth and respiration and microbial community properties such as microbial biomass carbon (MBC) and (biomass-)specific growth and respiration can be viewed as key microbial traits and functions that affect the SOC cycle.

To date, several such microbial traits and functions form the backbone of numerical models that describe organic matter turnover (Wieder et al., 2015; Chandel et al., 2023; Schimel, 2023). MBC is often used as a quantitative proxy for microbial activity because it sets an upper limit for absolute microbial process rates (Crowther et al., 2019; Chandel et al., 2023). Rates of heterotrophic respiration and microbial growth are direct measures of globally relevant processes to estimate C turnover in terrestrial ecosystems (Bond-Lamberty et al., 2018; Hashimoto et al., 2023). For example, soil microbial growth rates are estimated to be in the same order of magnitude but with diverging patterns from net primary productivity and aboveground litterfall (Gao et al., 2024), and specific growth is used to estimate microbial community turnover, which can strongly vary across systems and respond to changes (Caro et al., 2023; Camillone et al., preprint). Growth and respiration can become decoupled due to disparate physiological constraints on the processes that underlie anabolism and catabolism (Manzoni et al., 2012b; Roller and Schmidt, 2015). To capture this decoupling, different measures of (mostly partial) anabolism relative to respiration are used to quantify microbial substrate use efficiency, microbial growth efficiency (MGE) or microbial carbon use efficiency (CUE) (Hagerty et al., 2018; Geyer et al., 2019; Schimel, 2023). Such measures can be useful; for example CUE was recently found more important to explain global patterns of SOC storage than several key terrestrial carbon fluxes such as carbon input or baseline decomposition (Tao et al., 2023). Common proxies for microbial strategies are the C:N ratio of the microbial biomass (MCN), which supposedly reflects the ratio of bacteria and fungi in soil, with various functional implications (Strickland and Rousk, 2010; Malik et al., 2016), and potential activities of extracellular enzymes (PEEAs) that are involved in SOC degradation (Malik et al., 2019; Chen et al., 2020).

Microbial traits and functions must ultimately be related to the composition of the microbial community, because they describe microbial processes or properties. In the most direct scenario, microbial traits and functions are aggregated products of the entire active communities; in such a case, SOC cycling could be estimated based on the composition of a microbial community (E. K. Hall et al., 2018). However, the direct causality of this link depends on two factors: the environmental setting and scale in which microbial traits and functions are considered, and the type of the microbial trait and function. First, there are many ways by which the environmental setting, the scale and the complexity

of a soil system can affect and attenuate microbial physiology and the microbial functions that we measure (Nunan, 2017). For example, variations in microscale conditions such as substrate availability or moisture may counterbalance at bulk soil scale, and individual taxa may offset each other in terms of their traits (Pold et al., 2020; Metze et al., 2023). Direct constraints on microbial physiology such as substrate limitation may inhibit the realization of functional potential (Schimel, 2023) and ecological interactions may facilitate or inhibit microbial activity (Cordero and Datta, 2016). When such interferences take place, microbial community composition alone is not sufficient to estimate microbially mediated SOC cycling (E. K. Hall et al., 2018). In such cases, microbial traits and functions are emergent functions of the soil system. In some environmental settings, the explicit consideration and representation of microbial community composition may be obsolete to improve the description of SOC cycling, because the setting may dictate a narrow range of realized microbial activity. For example, in a soil with very little accessible substrate, respiration rates will always be low, no matter which microorganisms are present. As a consequence, it is exceedingly difficult to translate our understanding of microbial physiology from petri dishes and simple model systems into real-world soils. Second, it may also depend on the type of the microbial trait and function how closely it is linked to microbial community composition. A common distinction is made between narrow and broad microbial traits and functions (Graham et al., 2016; Osburn et al., 2021). Phylogenetically narrow traits and functions are evolutionarily constrained to few taxa, such as e.g. nitrification of methanogenesis (Rocca et al., 2015). In contrast, evolutionarily broad traits and functions like denitrification or heterotrophic respiration are widely distributed. Theory suggests that narrow traits and functions should be more strongly linked to community composition than broad functions; however, empirical evidence for this is mixed (Powell et al., 2015; Rocca et al., 2015; Graham et al., 2016; Osburn et al., 2021).

As illustrated, it may strongly depend on scale and type of trait or function whether a microbial trait or function needs to be conceptualized as a consequence of microbial community composition or as a direct consequence of the environmental setting. The scale of choice for the measurement of microbial processes in SOC cycling is most often bulk soil. Microbial traits and functions that affect the SOC cycle at the scale of bulk soil have been shown to vary with land use, climatic regime and edaphic properties (Colman and Schimel, 2013; Serna-Chavez et al., 2013; Zheng et al., 2019a). However, surprisingly few studies have attempted to directly link microbial community composition and traits and functions at large (regional to global) scales to understand patterns across pedo-climatic regions and soil types (Bier et al., 2015). Such insights are often confined to a narrow selection of soil samples from well constrained and locally confined experimental field or laboratory setups (Delgado-Baquerizo et al., 2016a; Li et al., 2019; Caro et al., 2023). Given the wide range of biogeochemical alterations that soils experience throughout their development, it is therefore unresolved whether the environmental

setting drives key microbial traits and functions at large spatial scales primarily directly, indirectly through microbial community composition, or whether some microbial traits and functions are even attributable to community composition alone.

In this study, we aimed to disentangle environmental setting and microbial community composition as drivers of key microbial traits and functions involved in the SOC cycle, at the scale of a biome. We sampled 33 mineral topsoils along a 2300 km long geoclimatic gradient in Chile, in order to capture the full range of environmental settings within the biome of temperate grasslands. We quantified 11 key microbial traits and functions that affect SOC cycling at the resolution of bulk soil after removal of direct physiological limitation to microbial activity through temperature or moisture. Inspired by the approach of Graham et al. (2016), we then applied regression analysis to investigate whether the selected microbial traits and functions are more closely linked to the environmental setting (ENV), the microbial community composition (MIC), or through additive effects of both (ENV+MIC). We hypothesized that (1) at the investigated biome scale, MBC as well as the absolute process rates would be linked most directly to ENV because the environment dictates the frame for microbial activity. We further hypothesized that (2) specific process rates, CUE and MCN would be linked to MIC, because they reflect microbial properties. Lastly, we hypothesized that (3) specific growth and respiration rates would be linked to different features of MIC because they are physiologically decoupled.

## 4.2 Material and Methods

### 4.2.1 Soil Sampling

A total of 33 A-horizon topsoils (0 - 10 cm) under grassland were sampled across a geoclimatic gradient in Chile in the summer seasons of 2017 and 2018. The gradient was constrained to a coherent type of land use and organic matter (OM) input - extensive rangeland and natural grassland - as well as carbonate free soil conditions (null HCl reaction) at $pH_{CaCl2} < 7.0$ in order to increase the identifiability of the controls on microbial traits and functions. The climate ranged from cold tundra to warm (semi-)arid steppe, excluding climatic extremes (cold and hot desert environments). Within the range of temperate grasslands, climatic and soil physicochemical contrasts were maximized to include a wide diversity of soil environments. The gradient covers a range of water balance (MAP - PET) from 1704 to -1207 mm, and a MAT range from 3.0 to 17.1 °C. The covered Köppen-Geiger climate zones range from arid steppe (Bsk) in the north to polar tundra (ET) in the south. Most sites (28) are located in the temperate climate zone, representing climates with cold and warm summers, with or without dry season (Cfb, Cfc, Csb, Csc). Basic climatic site characteristics are summarized in Table S2-2 (Appendix

Chapter 2). The gradient features 10 World Reference Base (WRB) major soil groups. It spans soils with less pronounced pedogenetic features (Leptosol, Arenosol, Cambisol), soils characterized by the influence of water (Gleysol, Planosol), humus-rich soils (Kastanozem, Chernozem) as well as soils characterized by low (Acrisol) and high (Andosol, Luvisol) mineral reactivity (Table S2-1, Appendix Chapter 2). After collection, all soil samples were frozen at field moisture at -20°C and were stored and shipped in this condition. In the laboratory in Switzerland, samples were thawed, sieved to < 2 mm, and frozen again at -20°C until further analysis.

### 4.2.2 Environmental variables

*Climate*

Mean annual temperature (MAT) and precipitation (MAP) were taken from WorldClim Version 2 (Fick and Hijmans, 2017). Mean annual potential evapotranspiration (PET) was taken from Trabucco and Zomer (2018). Both datasets have a resolution of 30 seconds and average monthly climate data from 1970 – 2000. Climate classification following Köppen-Geiger was done with the R-package 'kgc' (Bryant et al., 2017). Landscape position, soil moisture regime and dominant vegetation were either characterized on site or supplemented from previous surveys (detailed information and references in Tables S2-1 and S2-2). Landforms were mostly alluvial and marine terraces with flat topography, and vegetation comprised species that dominate in natural grasslands and prairies.

*Soil physicochemistry*

Soil pH was determined in 0.01 M CaCl2 solution (fresh soil:solution ratio of 1:5) using a pH meter (713 pH Meter, Metrohm, Switzerland).

Soil texture was determined as secondary soil texture (Totsche et al., 2018), via laser diffraction using a particle size analyzer (PSA) (LS 13 320, Beckman Coulter, USA). Fresh soil was shaken for 3 h with 10 % Na-hexametaphosphate to dissolve macroaggregates, and the resulting particles were quantified with the PSA. Particle size contributions were calculated based on particle volume, following the WRB system (IUSS Working Group WRB, 2015): clay < 2 µm, 2 µm < silt < 63 µm, 63 µm < sand < 2000 µm.

To assess weathering processes and intrinsic geochemical differences of soil parent material, total Si was measured using energy dispersive X-ray fluorescence (ED-XRF) spectrometry. Milled soil was mixed with Licowax (Fluxana, Germany) at a ratio of 4:1 and measured as pellets with a spectrometer (Spectro Analytical Instruments, Spectro XEPOS, Germany). Total content of Fe was assessed by digestion of 1 g soil aliquots with an aqua regia acid solution ($HCl:HNO_3:H_2O$, 3:1:1, v:v:v, 2.5 h at 120 °C). After filtration

through Whatman 41 filter papers Fe was quantified using inductively coupled plasma optical emission spectrometry (5100 ICP-OES, Agilent Technologies, USA). We calculated ratios of primary texture clay relative to Si ($clay_{prim}$:Si) and total Fe relative to Si (Fe:Si) as proxies for geochemical reactivity of the mineral matrix (Amelung et al., 2018).

*Soil organic matter*

Soil organic carbon (SOC) and total soil nitrogen (TN) contents were quantified via total combustion using a CN analyzer (Vario MAX Cube, Elementar GmbH, Germany), and bulk C:N ratios were calculated using the molar ratio of SOC:TN.

RockEval pyrolysis was done to assess thermostability of SOC as in Sebag et al. (2016). Milled soil samples were pyrolyzed on a Rock-Eval 6 device (Vinci Technologies, France), first in an $N_2$ atmosphere between 200 to 650 °C, and second in an oxidized atmosphere between 400 and 850 °C, both with a heating rate of 25 °C $min^{-1}$. Subsequently, the Rock-Eval I-Index was calculated exactly as in Sebag et al. (2016) to estimate the degree of biological transformation of SOC. The higher the I-index value, the less biologically transformed is the bulk SOC.

DRIFTS was done to assess the relative dominance of functional groups in SOC. Milled material (< 50 µm) was scanned in duplicate in the mid-infrared region (7500 to 600 $cm^{-1}$, resolution 2 $cm^{-1}$) using a Fourier transform IR (FT-IR) spectrometer with a high-throughput screening extension (HTS-XT) (Bruker Optics, Vertex 70, Germany). Spectra were normalized against gold (Infragold NIR-MIR Reflectance Coating, Labsphere) and corrected for atmospheric effects in the OPUS spectrometer software (Bruker Optics, Germany), and 32 co-added scans per sample were averaged. Spectra processing was done using the R-packages “simplerspec” (Baumann, 2020) and “prospectr” (Stevens and Ramirez-Lopez, 2020). To correct for light scatter, spectra were resampled to a range of 4000 to 600 $cm^{-1}$ with duplicates averaged and normalized using the normal variate method. Based on published information (Parikh et al., 2014), we assigned six wavenumber ranges to three types of functional groups: aliphatic C-H (anti)symmetric stretches: 2950 – 2910 and 2866 – 2836 $cm^{-1}$; aromatic C=C stretches: 1540 – 1524 and 1520 – 1510 $cm^{-1}$; carboxylic acid C=O stretch and carboxylate C-O asymmetric stretch: 1734 – 1718 and 1650 – 1636 $cm^{-1}$. We integrated peak areas of the wavenumber ranges with a local baseline correction and for each of the three functional groups, the peak area was divided by the summed peak area of all three groups to obtain relative peak areas (RPA). Increasing RPA carboxylic and decreasing RPA aliphatic values indicate a progressing degree of SOC decomposition (Ryals et al., 2014; Mainka et al., 2022).

### 4.2.3 Microbial community composition

*DNA extraction, target amplification and sequencing*

Total DNA in bulk soil was extracted from each of the 33 sites from 400 mg fresh topsoil samples (equilibrated to 50 % of WHC for one week at 20 °C) using the FastDNA™ SPIN Kit for Soil (MP Biomedicals), following manufacturer recommended procedures. The DNA content of each aliquot was quantified in microwell plates, using a PicoGreen assay (ThermoFisher Scientific, USA) and a microplate photometer (Tecan Infinite M200, Tecan GmbH, Austria). For bacterial 16S rRNA sequencing, the V3-V4 region of the prokaryotic (bacterial and archaeal) small-subunit (16S) rRNA gene was amplified with the primers 341F (CCT AYG GGD BGC WSC AG) and 806R (GGA CTA CNV GGG THT CTA AT) as described in Frey et al. (2016). For fungal ITS2 region sequencing, the primers ITS3 (CAH CGA TGA AGA ACG YRG) and ITS4 (TCC TSC GCT TAT TGA TAT GC) were used as described in Tedersoo et al. (2014). Polymerase chain reaction (PCR) amplification was done with tailed primers on a Veriti™ 96-Well Fast Thermal Cycler (Applied Biosystems, USA) using 30 ng DNA per reaction, following the protocol modified from Frey et al. (2016): an initial denaturation at 95 °C for 2 min, followed by 36 or 38 cycles (for 16S rRNA and ITS, respectively) of denaturation at 94 °C for 40 s, annealing at 58 °C for 40 s and elongation at 72 °C for 1 min followed by a final elongation at 72 °C for 10 min. The amplicons were purified with Agencourt AMPure XP beads (Beckman Coulter, Berea, CA) and subsequently quantified fluorometrically in microplates with the QuantiFluor® ONE dsDNA System (Promega, USA) using a microplate reader. The purified amplicons were subsequently sent to the Genome Quebec Innovation Center (Montreal, Canada) for barcoding using the Fluidigm Access Array technology and 300 base pairs paired-end sequencing on the Illumina MiSeq v3 platform (Illumina Inc., USA).

*Quantitative PCR*

In order to obtain an estimate of bacterial to fungi ratios, we quantified abundances of the bacterial 16S rRNA gene and the fungal ITS2 region by real-time quantitative PCR (qPCR) with SYBR Green dye on a QuantStudio5 Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA), according to Han et al. (2023). We used the same primers as for the amplicon sequencing. The targeted genes were 16S rRNA genes (bacteria) and ITS2 region (fungi). The qPCR reaction (10 µl) was composed of 5 µl GoTaq® qPCR Master Mix (Promega, Madison, WI, USA), 0.1 µl bovine serum albumin (BSA, 30 mg $ml^{-1}$), 1.9 µl molecular grade water, 0.5 µl of forward- and reverse-primers (10 µM), and 2 µl DNA template (2 ng $\mu l^{-1}$). The thermocycling settings were as follows: initial denaturation at 95 °C for 2 min, 40 cycles of denaturation at 95 °C for 40 s, annealing at 58°C for 40 s and elongation at 72 °C for 1 min, before a final elongation step that consisted of 95 °C for 15 s, 60 °C for 15 s and 95 °C for 15 s. Synthetic DNA

fragments were used as qPCR standards, ranging from $10^1$ to $10^8$ copies per µL, according to Han et al. (2023). Both standards and soil DNA extracts were measured in triplicate.

### 4.2.4 Microbial traits and functions relevant for SOC cycling

All microbial community traits and functions were measured at 60 % water holding capacity (WHC) and at 20 °C. Through this approach, direct temperature and moisture limitation of microbial activity was eliminated. The measured microbial communities, traits and functions reflect potential states of the investigated soils, and not necessarily their *in situ* state. While absolute values may not necessarily translate directly to *in situ* conditions, the observed patterns nevertheless allow for mechanistic interpretation. For measurements of carbon use efficiency and all properties that where measured on the same aliquots (microbial biomass carbon, microbial C:N ratio, absolute and specific growth rates and specific respiration rates), soil moisture was first set to 50 % of the WHC for one week of preincubation, before additional 10 % of the WHC were added at the start of the $^{18}O$-$H_2O$ incubation for the CUE method. For measurements of absolute respiration rates as well as potential extracellular enzyme activities, soils were directly preincubated for one week at 60 % WHC. For one selected soil of low, intermediate, and high SOC content, measurements were done in triplicate to evaluate the reproducibility of the methods.

*Microbial biomass carbon (MBC) and microbial C:N ratio (MCN)*

We used the chloroform fumigation extraction (CFE) method to quantify soil microbial C and N (modified from Vance et al., 1987). Briefly, C and N were extracted from two aliquots of each soil with 1 M KCl (fresh soil:solution ratio of 1:7.5) for 1 h on a shaker. Before extraction, one aliquot was exposed to an ethanol-free $CHCl_3$ saturated atmosphere in a dark desiccator for 48 h in order to lyse microbial cells. After fumigation, the desiccator was repeatedly evacuated to remove potential residual $CHCl_3$. After the extraction step, all samples were filtered through Whatman 40 filter paper and C and N concentrations of the KCl extracts were determined with a TOC/TN analyzer (TOC-L CPH/CPN, Shimadzu). Microbial C and N values were calculated by subtracting unfumigated from fumigated values. Values of microbial C and N as well as (unfumigated) KCl extractable C and N were converted to µg $g^{-1}$ soil and are presented without the use of an extraction coefficient. Averaged across the replicated soils, the relative standard deviation of C content was 13.2 % for the unfumigated aliquots, and 3.9 % for the fumigated aliquots. Due to procedural issues, five microbial C:N values were unrealistically high (values of 53 to 353 µg C $g^{-1}$ soil) and excluded from further analyses. In order to avoid skewed models, these values were imputed by k-nearest neighbors estimation across all 11 microbial trait and function variables, with k = 10.

*Basal heterotrophic respiration rates (respiration)*

We used an infrared gas analyzer system to estimate basal heterotrophic respiration. In detail, we weighed 10 g of each soil into airtight glass vials of 441 ml volume. The $CO_2$ concentration in the vials was measured directly after closing the vials. Measurements were done with a LI-840A infrared gas analyzer (Li-Cor Inc., Lincoln, United States) embedded into a continuous flow system that included the Flux Puppy app (v1.0.0) and was built according to (Carbone et al., 2019). Subsequently, respired $CO_2$ was allowed to build up in the vials for three consecutive intervals of 1.5 h in the dark at 20 °C. After each interval, we quantified the $CO_2$ concentration in the flow system and in the vials. $CO_2$ concentrations in the vials were corrected for the concentrations in the flow system, and the corrected $CO_2$ concentrations of the consecutive intervals of each sample were plotted against cumulative incubation time. After visual control for the absence of saturation effects, the regression slopes were taken as rates of $CO_2$ accumulation, and values were converted to respiration rates per gram dry weight of soil (µg C g $soil^{-1}$ $h^{-1}$). A reproducibility analysis of the method resulted in a relative standard deviation of 10.9 % averaged across the replicated soils.

*Carbon use efficiency (CUE), microbial growth rates (growth) and specific growth and respiration*

An approach based on $^{18}O$-$H_2O$ incorporation into DNA was used (modified after Spohn et al., 2016; Zheng et al., 2019b; Schnecker et al., 2023) to measure microbial growth and carbon use efficiency (CUE). This method might underestimate growth rates and CUE (as discussed in Geyer et al., 2019) and does not consider any products of microbial anabolism that are not captured with the CFE-method for MBC quantification, e.g. extracellular polymeric substances or extracellular enzymes (Manzoni et al., 2012b; Hagerty et al., 2018). “Microbial growth efficiency” might be a more appropriate term (Schimel et al., 2022), however we stuck with the established term CUE. We added either $^{18}O$-$H_2O$ (97 at%) or water at natural isotopic abundance to two aliquots of 400 mg soil. The added volumes ranged between 20 – 50 µl to obtain an average soil moisture of 60 % of WHC (between 55 and 66 % of WHC). After 24 h incubation at 20 °C in sealed headspace vials, we measured $CO_2$ concentrations with an infrared gas analyzer (EGM4, PP Systems). Microbial respiration rates (referred to as “respiration”) were calculated over the course of the incubation based on the difference between $CO_2$ concentrations of the $^{18}O$-$H_2O$ labeled aliquots and empty blank vials. They were converted into mass specific respiration rates (specific respiration, unit % of MBC $d^{-1}$) by division through MBC. Microbial growth rates were determined based on the incorporation of the $^{18}O$ into genomic DNA. For this, the DNA was extracted by a DNA extraction kit (FastDNA SPIN Kit for Soil, MP Biomedicals) as described above, and DNA concentrations of the extracts were determined fluorometrically with a Picogreen assay (Quant-iT

PicoGreen dsDNA Reagent, Life Technologies) following standard procedures. Subsequently, the O content and the $^{18}O$ enrichment of the purified DNA extracts were measured with a thermochemical elemental analyzer (TC/EA, Thermo Fisher) which is coupled to an isotope ratio mass spectrometer (Delta V Advantage, Thermo Fisher) via a Conflo III open split system. The amount of new DNA produced over the course of the incubation ($DNA_{new}$) was calculated as:

$$DNA_{new} = O_{DNA} * \frac{^{18}O\%_{DNA-L} - ^{18}O\%_{DNA-n.a.}}{^{18}O\%_{soil\ water}} * \frac{100}{31.21} \quad (1)$$

where $O_{DNA}$ is the total amount of DNA extracted, $^{18}O\%_{DNA\text{-}L}$ and $^{18}O\%_{DNA\text{-}n.a.}$ are the $^{18}O$ enrichments in the labeled and unlabeled aliquots respectively, and $^{18}O\%_{soil\ water}$ is the $^{18}O$ enrichment of the soil water in the labeled aliquot which was calculated based on soil water content and the volume of the added $^{18}O$-$H_2O$. The fraction term at the end of the equation encapsulates the average oxygen mass content of DNA (31.21 %). Assuming a constant soil-specific relationship between microbial C and the amount of DNA, the amount of microbial biomass produced over the course of the incubation was calculated. For this, we divided $DNA_{new}$ through DNA in the sample and multiplied this value with MBC, which was beforehand calculated by dividing microbial C through an extraction efficiency factor of 0.45 (Wu et al., 1990). The resulting absolute growth rates (referred to as “growth”) were converted into mass specific growth rates (specific growth, unit % of MBC $d^{-1}$) by division through MBC. Finally, the microbial CUE (unitless) was calculated using the following equation (Manzoni et al., 2012b):

$$CUE = \frac{C_{Growth}}{C_{Growth} + C_{Respiration}} \quad (2)$$

where $C_{Growth}$ and $C_{Respiration}$ are the measured growth rates and respiration rates expressed as ng C $g^{-1}$ soil $h^{-1}$. Averaged across the replicated soils, the relative standard deviation was 5.1 % for CUE, and 6.2 % for specific growth.

*Absolute and specific potential extracellular enzyme activities*

We used a fluorometric assay to quantify maximum catalytic capacity (Vmax) for three groups of hydrolytic enzymes (ß-Glucosidase, BG (EC 3.2.1.21); cellobiosidase, CB (EC 3.2.1.91); N-acetyl-β-glucosaminidase, NAG (EC 3.2.1.52)) that are involved in SOC degradation, targeting carbon. CB hydrolyses cellobiose from cellulose, BG releases glucose from cellobiose and other β-D-glucosides, and NAG releases glucosamine from β-1,4-linked glucosamine polymers such as chitin (Bairoch, 2000; Sinsabaugh et al., 2008). In addition, we assessed the potential activity of two extracellular oxidases

(phenoloxidase, POX; peroxidase, PEX) following Kaiser et al. (2010). In detail, we dissolved 1 g of soil in 90 ml of sodium acetate buffer (50 mM, pH 5.0) and ultrasonicated with an energy of 350 J. The resulting soil suspension was passed through a sieve of 0.3 mm mesh size and continuously stirred at 20 °C.

For each hydrolase assay, 48 replicates of 200 µl suspension were pipetted into black microplates. 50 µl of a 1:2 dilution row of substrate labeled with 4-Methylumbelliferyl (MUF) (ranging from 1.56 - 100 µM) and a blank of only sodium acetate buffer were added to sets of five aliquots of soil suspension. The MUF-labeled substrates used were MUF β-D-glucopyranoside (for BG), MUF β-D-cellobioside (for CB) and MUF N-acetyl-β-D-glucosaminide (for NAG). The remaining eight aliquots received 50 µl of a standard row of MUF (1:2 dilution row, top standard 200 µM). The microplates were incubated in the dark for 2.5 h at 20 °C and fluorescence was repeatedly measured (excitation 365 nm, emission 450 nm, automatic gain) in intervals of 30 min with a microplate photometer (Tecan Infinite M200, Tecan Trading AG). Using the soil-specific standard rows, fluorescence measurements were converted to MUF concentrations. With simple linear regression models slopes were fitted and converted to nM MUF $g^{-1}$ soil $h^{-1}$. Wells with an $R^2 < 0.6$ were removed from further analysis, and for each set of five replicates we conducted Grubb's outlier tests and excluded significant outliers. We averaged the replicates and following (Alves et al., 2021) we used the drc R-package "drc" (Ritz et al., 2015) to fit 2-parameter Michaelis Menten kinetic models to the substrate concentration rows of each soil. We retained estimates of Vmax (nM MUF $g^{-1}$ soil $h^{-1}$) that were fitted with a p-value < 0.05.

For each oxidase assay, we mixed three aliquots of 1 ml soil slurry 1:1 with 20 mM L-3,4-dihydroxyphenylalanine (DOPA) in sodium acetate buffer. The mixture was shaken for 10 min, centrifuged, and 250 ul were pipetted into transparent microplates. For PEX assays, 10 µl of 0.3 % $H_2O_2$ were added. Absorbance was measured at 450 nm at the start of the incubation and after 20 h incubation at 20 °C in the dark. With an empirically determined molar extinction coefficient of 0.62 (Kaiser et al., 2010) we converted absorbance values to DOPA concentrations and averaged the triplicates. Potential oxidase rates (nmol DOPA $g^{-1}$ soil $h^{-1}$) were calculated over time, and POX rates were subtracted from PEX rates to obtain final rates of PEX activity.

The reproducibility analysis of the method revealed a relative standard deviation of 8.6 %, 9.1 % and 11.1 % for Vmax of BG, CB and NAG respectively, and 11.2 % and 5.4 % for the potential activity rates of POX and PEX averaged across the replicated soils. Due to poor fit of the Michaelis Menten models (p-value > 0.05), four hydrolytic Vmax values were dismissed. These values were imputed by k-nearest neighbors estimation across all 11 microbial trait and function response variables, with k = 10. For the sake of clarity and brevity in the main text, and given the strong autocorrelation of Vmax of the three hydrolases and the potential rates of the two oxidases, respectively, we summed them up into two

composite variables (“hydrolases” for hydrolytic enzymes and “oxidases” for oxidative enzymes). In addition we normalized the absolute potential enzyme activity rates relative to MBC (“specific” hydrolases and “specific” oxidases). Note that the regression analysis described in section *Regression analysis* was also conducted for the absolute and specific enzyme activities of the five enzymes individually. Results of this analysis are not discussed in detail but are presented in the supplementary material.

### 4.2.5 Analysis

*Processing of microbial community data*

Sequence processing was performed in R Studio (Version 2023.06.0.421) using the dada2 workflow (Callahan et al., 2016). For 16S rRNA sequences, primers were trimmed, and the right sides of the reads were truncated to 270 and 210 bases for forward and reverse reads. For ITS sequences, we used the package “cutadapt” (Martin, 2011) to remove sequenced primers from the DNA fragments, and the right sides of the reads were truncated to 250 and 220 bases for forward and reverse reads. For both types of sequences, PhiX genomes were removed, and all respective samples were pseudo-pooled for sample inference. Forward and reverse reads were merged with a minimum overlap of 12 base pairs for 16S rRNA and 20 base pairs for ITS, in both cases allowing for 1 mismatch. Chimeras were removed with default settings. Taxonomy of bacteria was assigned using the Silva nr99 v138.1 database (Quast et al., 2012) as a training set, and taxonomy of fungi was assigned using the UNITE general FASTA database (Abarenkov et al., 2022) as a training set. In both cases, the minimum bootstrap confidence for assigning a taxonomic level was set to 60 %. Finally, all amplicon sequencing variants (ASVs) that were not bacteria or fungi, respectively, were removed from the dataset. All samples reached a reasonable coverage at the minimum sequencing depth across samples (17821 reads for 16S rRNA and 6923 reads for ITS). We therefore rarefied all samples by iterative random subsampling without replacement (100 iterations) to the respective minimum read numbers, implemented by the function rrarefy in the R-package “vegan” (Oksanen et al., 2022). Next, we aggregated amplicon sequencing variants (ASVs) at the genus level. To capture the main patterns of microbial community composition at the genus level while constraining data complexity, we filtered the most dominant genera of bacteria and fungi up to a cumulative relative read abundance of 50 % of the respective average communities. This resulted in the 12 most dominant bacterial genera, and the 14 most dominant fungal genera. We further estimated bacterial and fungal alpha diversity as ASV richness (without singletons), using the R-package “vegan” (Oksanen et al., 2022). The FUNGuild database was employed for annotating fungal trophic modes at the ASV level (Nguyen et al., 2016). Only confidence scores of 'probable' and 'highly probable' were considered in the analysis.

*Assembly of statistical predictor sets*

Based on environmental data (ENV, Table S4-1) and soil microbial community data (MIC, Table S4-3), we assembled two sets of predictors. For both sets, we used the R-package "psych" (Revelle, 2022) applied principal component analysis (PCA) on the raw input variables in order to reduce dimensions and to solve autocorrelation between variables. All PCAs were conducted with centered and scaled raw input variables. Correlation matrices of variables were visualized with the R-package "corrplot" (Wei and Simko, 2021).

For the 16 environmental variables of the ENV predictor set (Table S4-1, Figure S4-1a), we applied Rosner tests (R-package "EnvStats", (Millard, 2013) to identify in total 20 significant outliers which were obvious laboratory artifacts among 528 observations across the 16 raw input variables. PCA on data with strong outliers can result in skewed and ultimately misleading principal components. However, PCA also cannot process missing values. Therefore, all significant outliers were imputed by the k-fold nearest neighbor method (k = 10, R-package "DMwR2", (Torgo, 2016). After imputation, 4 of the 20 values remained significant outliers. We subsequently conducted rotated PCA with varimax rotation, which reduces the number of variables that correlate with individual rotated components (RCs) and thus facilitates the interpretability of the RCs. Eight RCs (free of autocorrelation, Figure S4-1b) were retained in order to reach a cumulative explained variance of > 0.8 for the environmental dataset (Table S4-2). The retained RCs were interpreted based on their dominant loadings (loading ≥ 0.5) and were subsequently used as predictor variables in the regression analysis (as the ENV predictor set).

For the MIC predictor set, we used 31 raw input variables: relative read abundances of the most dominant bacterial and fungal genera (12 and 14, respectively), relative read abundances of the sums of the remaining rare genera, bacterial and fungal richness as the Bact:Fungi ratio (Table S4-3, Figure S4-2a). For this set, we conducted unrotated PCA, because relative read abundance data is inherently autocorrelated in a way that is complex to resolve (Alteio et al., 2021) and unrotated PCA outperformed rotated PCA in terms of dimension reduction. 11 uncorrelated principal components (PCs) were retained in order to reach a cumulative explained variance of > 0.8 for the microbial community dataset (Table S4-4, Figure S4-2b). The retained PCs were used as predictor variables in the regression analysis (as the MIC predictor set).

*Regression analysis*

We conducted regression analysis to investigate whether microbial traits and functions that affect the SOC cycle would best link to the environmental setting (ENV), microbial community composition (MIC),

or an additive combination of both (ENV+MIC). We used three different regression approaches to ensure that the analysis was not subject to artifacts of model choice. In order of decreasing interpretability, we applied stepwise regression, L1 regularization (lasso) regression and nonlinear random forest regression. The three regression approaches differ in the underlying statistical approaches but have common metrics for model assessment. For each approach, cross validation is possible based on RMSE, goodness of fit can be evaluated based on $R^2$ values, and measures for relative variable importance can be calculated. To avoid overfitting, model fits (for stepwise regression) and parameter optimization (for lasso and random forest regression) were performed with Monte Carlo cross validation. Stepwise regression builds linear regression models based on iterative automated variable selection, lasso regression uses penalties to constrain model complexity, and random forest regression allows for non-linear relationships between dependent and independent variables. For detailed descriptions of the regression approaches, please see Supplementary Text S4-1. Thus, we built three regression models (ENV, MIC and ENV+MIC) for each of the targeted 11 microbial traits and functions, resulting in 33 models per statistical regression approach, and a total of 99 models. All regression analysis was performed using the R-package "caret" (Kuhn, 2008) which wraps the packages "leaps" (Lumley, 2020) for stepwise regression, "glmnet" (Friedman et al., 2010) for lasso regression and "randomForest" (Liaw and Wiener, 2002) for random forest regression. For stepwise and lasso regression, homoscedasticity was evaluated with the R-package "lmtest" (Zeileis and Hothorn, 2002).

We compared the goodness of fit of the three statistical regression approaches for each of the 33 models per statistical approach based on pairwise comparison of $R^2$ values (see Supplementary Text S4-1 for further details). We summarized the results of each of the three regression approaches for each targeted microbial trait or function by calculating the mean values and standard deviations of $R^2$.

For each model, variable importance (VI) was quantified as the absolute value of the t-statistic, which is calculated by dividing the coefficient of a predictor by its standard error (James et al., 2013). To improve comparability of the contribution of predictor variables between regression approaches, we scaled VI values (sVI) within each model between 0 and 100. For this, the absolute value of the lowest VI was assigned 0, and the absolute value of the highest VI was assigned 100. After scaling, signs of the coefficients were added to the sVI values, to facilitate interpretation. In cases where the models were not significant (p-value > 0.05), we did not consider them for further analysis. An absolute value of 100 indicates that a variable is the most important variable of a model. In contrast, an sVI of 0 in the linear models means that a variable does not feature in the respective model, whereas in the random forest models it means that a variable is the least important. Positive and negative signs reflect the coefficient signs of the respective variables in the stepwise and lasso models. We calculated mean values and

standard deviations of absolute sVI values across the three regression approaches. This allowed integration and interpretation of the model compositions resulting from the three regression approaches for all models.

## 4.3 Results

### 4.3.1 Description of the predictor datasets

The used predictor dataset representing the environmental setting (ENV) consisted of eight rotated components (RCs) (Table S4-2) which resulted from 16 edaphic and climatic variables (Table S4-1). Most importantly, ENV RC1 - "SOM quantity" (SOM being the abbreviation for soil organic matter) encapsulated total and extractable amounts of organic carbon and nitrogen, and ENV RC3 - "Texture" was positively associated with secondary sand content (Table S4-2). The predictor dataset representing the soil microbial community composition (MIC) consisted of 11 principal components (PCs) (Table S4-4) which resulted from 31 variables that described community composition and diversity (Table S4-3). Bacterial and fungal richness as well as the bacteria:fungi ratio were positively correlated (Figure S4-2a) and loaded onto MIC PC2 (Table S4-4). Notably, PC2 was also positively correlated with the relative abundance of saprotrophic fungi (Figure S4-9). The remaining components were associated with relative abundances of the dominant bacterial and fungal genera. However, autocorrelations of relative abundance data are impossible to resolve, and therefore the PCs of the MIC dataset are presented as PC1 to PC11 (instead of being named based on their dominant loadings, as the ENV RCs). Several variables were correlated between ENV and MIC (Figure S4-3). The strongest links were a positive correlation between MIC PC1 and ENV RC1 - "SOM quantity" and a negative correlation between MIC PC1 and ENV RC2 - "soil pH", as well as a positive correlation between MIC PC3 and ENV "RC5 - MAT".

### 4.3.2 Microbial traits and functions and their correlative relationships

The ranges of the 11 quantified soil microbial traits and functions are shown in Table 4-1, and values sorted by latitude are presented in Table S4-5. The correlative relationships between the 11 quantified traits and functions are shown in Figure 4-1. Microbial biomass carbon (MBC) varied 70-fold across the gradient, from 6.7 µg C $g^{-1}$ soil in a low-SOC Gleysol to 463.6 µg C $g^{-1}$ soil in a high-SOC Andosol (Table S4-5). Absolute respiration rates (respiration) varied almost 40-fold, from 0.2 µg C $g^{-1}$ $soil^{-1}$ $h^{-1}$ in a Luvisol to 7.5 µg C $g^{-1}$ soil $h^{-1}$ in an Andosol. Absolute growth rates (growth) varied 25-fold, from 0.1 µg C $g^{-1}$ soil in various soil types (Luvisol, Gleysol, Chernozem and Andosol) to 2.4 µg C $g^{-1}$ soil in an Andosol. Both rates (respiration and growth) were positively correlated with each other and with MBC (Figure 4-1), being lower in soils from the warmer and drier systems of central Chile, and higher in soils from

the cooler and wetter systems of southern Chile (Table S4-5), characterized by higher mineralogic reactivity (Chapter 2). Carbon use efficiency (CUE) ranged between 0.10 in an Andosol and 0.60 in an Arenosol, indicating large differences in the amount of carbon that is respired per unit growth. CUE was positively correlated with specific growth and negatively with specific respiration, both correlations having similar strength (Figure 4-1). Specific growth was lowest in a Luvisol (1.4 % of MBC $d^{-1}$) and highest in a Kastanozem (10.9 % of MBC $d^{-1}$). Specific respiration showed a larger variation, ranging from 2.5 % of MBC $d^{-1}$ to 68.1 % of MBC $d^{-1}$, both in Gleysols, with the maximum value being an outlier. Both specific rates were correlated with each other, but to a weaker extent than respiration and growth (Figure 4-1). The potential activities of extracellular hydrolases and oxidases were positively correlated with each other and with MBC. CUE, MCN and specific rates were not correlated with MBC or either of the absolute rates.

**Table 4-1.** Ranges of quantified microbial traits and functions across the gradient (n = 33). Abbr. = Abbreviation, Min. = Minimum, Max. = Maximum, MUF = 4-Methylumbelliferyl, DOPA = L-3,4-dihydroxyphenylalanine.

| Variable | Abbr. | Unit | Min. | Mean | Median | Max. |
|---|---|---|---|---|---|---|
| *Biomass and biomass-dependent rates:* | | | | | | |
| Microbial biomass C | MBC | µg C $g^{-1}$ soil | 6.7 | 146.6 | 124.6 | 463.6 |
| Growth | - | µg C $g^{-1}$ soil $h^{-1}$ | 0.1 | 0.7 | 0.5 | 2.4 |
| Respiration | - | µg C $g^{-1}$ soil $h^{-1}$ | 0.2 | 2.0 | 1.5 | 7.5 |
| Hydrolases | - | nmol MUF $g^{-1}$ soil $h^{-1}$ | 322.6 | 2152.6 | 1858.1 | 9514.8 |
| Oxidases | - | nmol DOPA $g^{-1}$ soil $h^{-1}$ | 1113.5 | 3700.7 | 3458.3 | 10332.9 |
| *Dimensionless variables:* | | | | | | |
| Carbon use efficiency | CUE | - | 0.10 | 0.28 | 0.24 | 0.60 |
| Microbial C:N | MCN | - | 4.8 | 12.8 | 11.1 | 39.4 |
| *Biomass-normalized rates:* | | | | | | |
| Specific growth | - | % of MBC $d^{-1}$ | 1.4 | 4.9 | 4.6 | 10.9 |
| Specific respiration | - | % of MBC $d^{-1}$ | 2.5 | 15.4 | 12.5 | 68.1 |
| Specific hydrolases | - | nmol MUF $µg^{-1}$ MBC $h^{-1}$ | 5.1 | 16.3 | 15.8 | 51.0 |
| Specific oxidases | - | nmol DOPA $µg^{-1}$ MBC $h^{-1}$ | 8.7 | 50.4 | 24.6 | 634.3 |

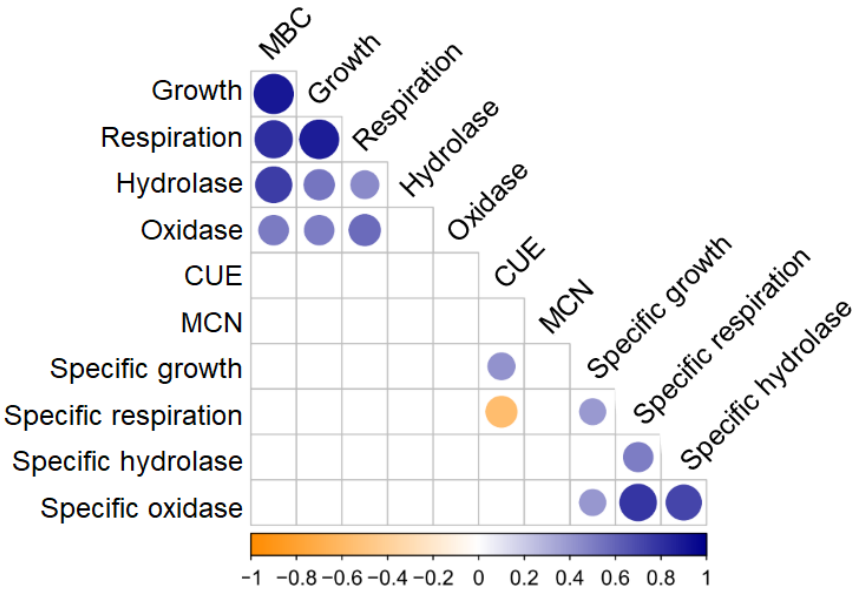



**Figure 4-1.** Correlation matrix showing significant Pearson correlations (p-value < 0.05) between the 11 investigated microbial traits and functions. Circle size and color indicate the strength of the relationships (Pearson correlation coefficient). Shown are significant correlations (Pearson correlation, $p < 0.05$). MBC = Microbial biomass carbon, CUE = Carbon use efficiency, MCN = Microbial C:N.

### 4.3.3 Predictors of microbial traits and functions

Across all regression models, the proportion of explained variation ($R^2$) varied from 0 to 0.85. On average, stepwise regression explained most variation ($R^2 = 0.43$) followed by random forest regression ($R^2 = 0.31$) and lasso regression ($R^2 = 0.27$). The $R^2$ of the different regression approaches correlated strongly (Figure S4-4), showing that all three approaches gave comparable results.

The results of the regression analysis are summarized in Figure 4-2. Detailed model descriptions are provided in Table S4-6. ENV explained a larger proportion of variation than MIC for MBC and all absolute rates (respiration, growth, hydrolases, oxidases). In contrast, MIC explained a larger proportion of variation for CUE and several specific rates (specific respiration, specific growth, specific hydrolases). The variables that were best explained were the two measures of respiration (respiration: $R^2 = 0.75$, specific respiration: $R^2 = 0.73$), CUE (mean $R^2 = 0.65$), growth ($R^2 = 0.54$) and MBC ($R^2 = 0.48$). The variables with the lowest proportion of explainable variation were MCN ($R^2 = 0.13$) and the measures of enzyme activities ($R^2 = 0.19$ to 0.29). Notably, of the individual enzyme activities, absolute activities of NAG and POX were markedly better predictable ($R^2 = 0.53$ and 0.38, respectively) than the rest (Figure S4-5), and among specific enzyme activities NAG was best predictable ($R^2 = 0.35$), with MIC (Figure S4-5). The combination of the datasets ENV+MIC did not improve explainability substantially

except for specific growth, indicating a clear distinction between which microbial traits and functions are explained by the environmental setting versus by microbial community composition.

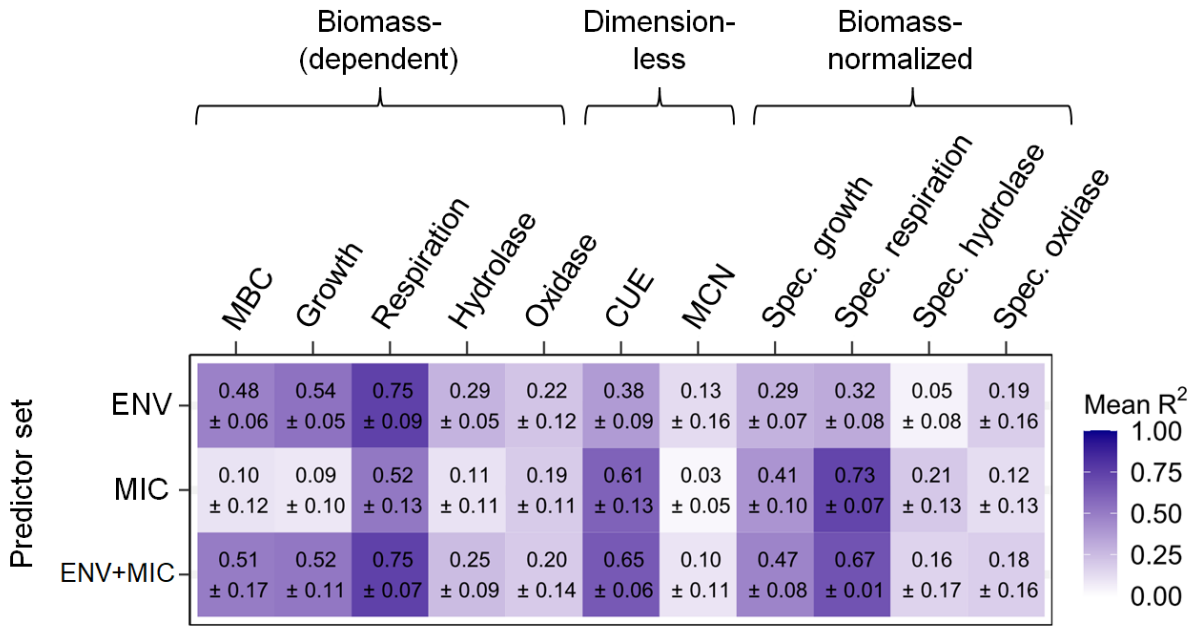



**Figure 4-2.** Proportion of the variation of 11 microbial traits and functions (x-axis) that can be explained with three different sets of predictors (y-axis). Shown are mean ± S.D. of $R^2$ across three different regression approaches (stepwise, lasso, and random forest). ENV = Environmental data, MIC = Microbial community data, MBC = Microbial biomass carbon, CUE = Carbon use efficiency, MCN = Microbial C:N, Spec. = specific.

SOM quantity was the strongest predictor (sVI = 92 to 100) for MBC and all absolute microbial functions (Figure 4-3a). In addition, growth and hydrolases (in particular the hydrolases CB and NAG, Figure S4-6) shared a weaker positive relationship with coarser texture (sVI = 37 and 25), while oxidases (in particular the oxidase PEX, Figure S4-6) had an almost equally strong negative link to soil pH (sVI = 87). CUE was higher in neutral than in acidic soils (sVI = 99), higher in systems with higher MAT (sVI = 80) and lower in soils with a wider bulk C:N ratio (sVI = 87) (Figure 4-3a). However, the MIC dataset explained more variability of CUE (Figure 4-2) through a negative link with PC1 (sVI = 100) (Figure 4-3b). Similarly, specific respiration and specific growth could both be linked to ENV variables, but were better explained with MIC variables (Figure 4-2). In regression with the ENV dataset, specific respiration was lower in neutral than in acidic soils (sVI = 99) and specific growth was higher in soils from warmer systems (sVI = 73). Moreover, both specific rates were higher in soils with coarser texture (sVI = 85 to 100) (Figure 4-3a). However, in the better performing models with MIC, the two specific rates were linked to different variables: specific respiration was strongly positively linked to PC1, while specific

growth was strongly negatively linked to PC2 (Figure 4-3b). Overall model performances for specific potential enzyme activities and MCN were weak (Figure 4-2), which was also reflected in their model structures: While specific hydrolases were linked to various components of MIC, for specific oxidases and MCN no clear predictors emerged in either the ENV or the MIC models (highest sVI = 57 and 41, respectively) (Figure 4-3, and Figure S4-6 for individual enzymes). The sVI values for the ENV+MIC models are shown in Figure S4-7.

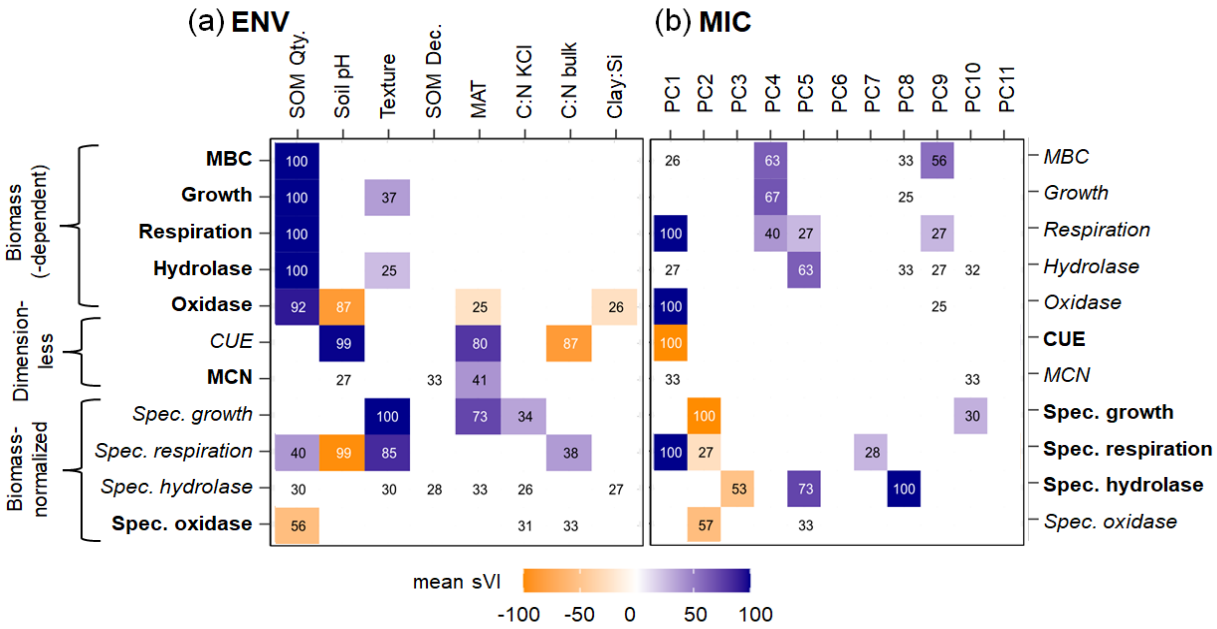



**Figure 4-3.** Scaled variable importance (sVI) ≥ 25 in the prediction of 11 microbial traits and functions (y-axis). (a) Prediction with environmental data (ENV); (b) prediction with microbial community data (MIC). Shown are mean sVI values across three different regression approaches (corresponding S.D. values are shown in Figure S4-8). Colors indicate the signs of the coefficients (negative: orange; positive: blue). Note that variables which only featured in random forest models do not have signs. The best models for each microbial trait function are highlighted by bold font. SOM Qty. = SOM quantity, SOM Dec. = SOM Decomposition, MBC = Microbial biomass carbon, CUE = Carbon use efficiency, MCN = Microbial C:N, Spec. = specific, KCl = extractable with KCl, PC = principal component.

## 4.4 Discussion

### 4.4.1 Potential extracellular enzyme activities and microbial C:N ratio remained poorly explained

Absolute rates of potential extracellular enzyme activities (PEEAs) of hydrolases and oxidases involved in C degradation were strongly correlated with MBC (Figure 4-1), as observed previously in other studies across various scales (Delgado-Baquerizo et al., 2016a; Baker and Allison, 2017; Piton et al., 2020; Osburn et al., 2021). However, along our study gradient the degree to which absolute and specific PEEAs could be explained with the available ENV and MIC predictors was low (Figure 4-2, Figure S4-6). This contradicted our hypotheses and was somewhat surprising, given that links of absolute and specific enzyme activities with community composition (Schnecker et al., 2015; Trivedi et al., 2016; Piton et al., 2020; Osburn et al., 2021) and with the environmental setting (Sinsabaugh et al., 2008; Schimel et al., 2017; Li et al., 2020; Sheng et al., 2022) have been described many times. In particular, increased investment into extracellular enzymes (i.e. higher activities of hydrolases and oxidases normalized for MBC, here referred to as "specific" activities) has been proposed to be a microbial resource acquisition trait (Malik et al., 2020) that is at a community-level trade-off with CUE (Hagerty et al., 2018; Malik et al., 2019; Zhang et al., 2023). The present study could not confirm this pattern across larger spatial scales and diverse soil types (Figure 4-1). We argue that several factors could blur patterns of PEEAs at the observed scale. (i) PEEA measurements integrate long time spans, because extracellular enzymes can stay active for a long time through sorption on minerals (Schimel et al., 2017; Sheng et al., 2022). (ii) PEEAs were measured at pH 5 and 20 °C, but pH optima and temperature responses of PEEAs vary across soils and enzyme groups, potentially skewing rate estimates along the investigated gradient (German et al., 2012; Baker and Allison, 2017; Puissant et al., 2019; Wade et al., 2021). (iii) The three assessed hydrolytic enzyme groups (BG, CB and NAG) fulfill broad functions. These enzymes alone might therefore neither be strongly relatable to bulk SOM quality, nor to specific microbial taxa (Trivedi et al., 2013; Diamond et al., 2019). Notably, the cell-wall degrading enzyme NAG was best predictable among the individual enzymes (Figure S4-6). NAG has been proposed to be more involved in microbial recycling than the plant material degrading enzymes BG and CB (Schnecker et al., 2019). (iv) Similarly, the assessed oxidative enzyme groups (POX, PEX) fulfill multiple functions; in addition to resource acquisition they can serve to degrade toxic phenols (Sinsabaugh, 2010). Ultimately, PEEAs translate between the environmental setting and the demands of the microbial community. Depending on the scale of observation, PEEAs - as measured most commonly - are therefore likely emergent functions of the interplay between environment and microbial communities, rather than aggregate functions of either environment or community alone (E. K. Hall et al., 2018). Similarly, MCN was neither linked to the environmental setting nor to microbial community composition (Figure 4-2),

although soil fungal biomass has a wider C:N ratio than soil bacterial biomass (Strickland and Rousk, 2010). We attribute this lack of relationship to potential method uncertainties associated with measurements (Hargreaves et al., 2013; D. Wang et al., 2022). Overall, we found that PEEAs and MCN, as measured commonly, neither reflected the environmental setting, nor properties that could be related to microbial community composition. Regardless of whether this was caused by method inaccuracies or an emergent nature of these properties, we therefore urge caution in the interpretation of these variables at large scales.

### 4.4.2 MBC, growth and respiration are driven by the environmental setting

In support of the first hypothesis, that microbial biomass carbon as well as the absolute process rates would be linked most directly to environmental setting parameters, SOM quantity was the single most important driver of MBC along the investigated gradient (Figure 4-3a). This confirms earlier findings on the large-scale linkage between SOM quantity and MBC (Colman and Schimel, 2013; Crowther et al., 2019; L. C. Smith et al., 2021; Gao et al., 2022). An underlying mechanism could be that soil microbes in temperate grasslands are generally limited by available carbon or nitrogen substrate, and more substrate availability can therefore sustain a higher microbial biomass (Soong et al., 2020; L. C. Smith et al., 2021). Along the investigated gradient, bulk SOM was associated to various degrees with functionally different SOM fractions (particulate organic matter, free silt- and clay-sized particles, stable microaggregates) (Chapter 2). However, bulk SOM and extractable SOM (the latter used as a proxy for bio-available SOM) were strongly correlated, as observed before (Wang et al., 2003; Delgado-Baquerizo et al., 2016a) (Table S4-2). This indicates that the variable composition of bulk SOM fractions did not affect microbially available SOM across grasslands. The strongest driver of SOM quantity was water balance (MAP – PET) (Table S4-2). The water balance can affect the balance between SOM input and decomposition (Manzoni et al., 2012a; Serna-Chavez et al., 2013; La Pierre et al., 2016; Zhang and Xi, 2021), as well as the SOM stabilization potential through weathering in the course of pedogenesis (Slessarev et al., 2022). Note that only potential direct controls of microbial activity were considered in the present study, and thus the strong influence of soil mineralogy on SOM quantity does not become evident here (Chapter 2, Doetterl et al., 2015; Rasmussen et al., 2018; Heckman et al., 2022). However, the indirect importance of parent material and mineralogy is evident by the fact that the soils with highest SOC contents – and highest MBC – were Andosols, which have a high mineral reactivity (Chapter 2, Beare et al., 2014; Matus et al., 2014) (Table S4-5).

Previous studies found that information on microbial community composition explained variability of respiration was additive to explanatory power provided by data on the environmental setting (Graham

et al., 2016; Liu et al., 2018). In this study, the variability of respiration that MIC accounted for was redundant rather than additive to SOM quantity (Figure2, Figure S4-3). We explain this finding with the thought that additive effects of community composition and environmental setting may become more likely when communities are attributable to legacy effects or stochastic assembly processes rather than to deterministic assembly by present environmental conditions (Orland et al., 2019). Possibly, systems that undergo strong environmental fluctuations or disturbances may rather allow for additive explanatory power. However, along the investigated gradient a large part of soil bacterial community composition was attributable to patterns in soil pH and SOM quantity (Chapter 3). Respiration and growth were instead most strongly linked with SOM quantity (Figure 4-3a). Several studies reported direct links between SOM quantity and respiration, bypassing microbial biomass (Wang et al., 2003; Delgado-Baquerizo et al., 2016a; L. C. Smith et al., 2021). In these studies, the authors argued that substrate availability could limit the absolute activity of microbial communities independently of their biomass. However, in the present study the strong positive correlation between respiration, growth and MBC (Figure 4-1) suggests that SOM quantity affected respiration and growth indirectly via the amount of microbial biomass (Figure 4-4a). This relationship has been frequently observed for respiration across different spatial scales (Colman and Schimel, 2013; Delgado-Baquerizo et al., 2016a; Liu et al., 2018; Osburn et al., 2021). However, it is more surprising for growth, where the inverse relationship (faster growth in systems with lower MBC) was recently reported for three soils (Caro et al., 2023). This discrepancy is likely a matter of investigated scales and illustrates the context-dependency of controls; For example, the present gradient of grassland soils encompassed a 70-fold variation in MBC (Table 4-1). Such a large difference in microbial biomass might readily override the effects that variable sizes of active fractions (i.e. percentage of microbes in a soil that are active) could have on growth. Overall, we conclude that MBC, respiration and growth at the biome scale are strongly tied to SOM quantity.

### 4.4.3 The environmental setting partially determines specific growth and specific respiration

Respiration and growth normalized for MBC (i.e. specific respiration and growth) were positively correlated (Figure 4-1), indicating that the size of the active fractions of soil microbial communities and C turnover through respiration generally co-varied along the investigated gradient. Both specific rates were higher in soils with coarser texture (Figure 4-3a), a link that was particularly pronounced in the cool and wet Andosols of central Patagonia (Table S4-5). In the case of specific growth, this effect was additive to microbial community composition (Figure S4-7, Figure S4-3). Contrary to our second hypothesis, that specific process rates, CUE and MCN would be linked to MIC variables mostly, texture

as a proxy seemed to represent mechanisms that influence specific activity independently of community composition. In our study, texture was quantified as secondary texture in which stable physical structures are preserved that can limit the access of microorganisms to SOC or influence soil pore structures (i.e. water- stable micro-aggregates). Possibly, soils rich in stable microaggregates and SOC (such as the Andosols of central Patagonia, Chapter 2) offered more viable or aerated microsites at the air-water interface, increased pore connectivity and larger probability for the contact between microbes and substrate (Don et al., 2013; Keiluweit et al., 2017; Schnecker et al., 2019; Nunan et al., 2020). This could allow a larger fraction of the microbial communities to grow and respire, thereby decoupling specific respiration and growth from biomass (L. C. Smith et al., 2021; Camillone et al., preprint). In support of this idea, growth was also higher in soils with coarser texture (Figure 4-3a). Nevertheless, finer texture is often considered to promote microbial activity at intermediate moisture conditions (Wolf et al., 2013; Angst et al., 2021b), with few exceptions (Hassink, 1994; Franzluebbers et al., 1996; Nunan et al., 2020; Witzgall et al., 2021). Importantly, the lack of this connection in other studies could be related to the fact that usually primary texture is investigated and reported, or that texture gradients are often small and limited in spatially confined studies; potentially relevant effects of variable in-situ secondary texture (Six et al., 2004; Wilpiszeski et al., 2019) may therefore be frequently overlooked.

In contrast to texture, links between specific rates and MAT, soil pH and bulk C:N ratio were not additive to microbial community composition and largely disappeared after consideration of the latter (Figure S4-3, Figure S4-7). The positive link between MAT and specific growth could potentially be explained with thermal adaptation of the communities. MAT of the sampling sites ranged from 3.0 to 17.1 °C (Table S4-1); as a consequence, soils from warm sites of central Chile were incubated at thermal conditions (20 °C) that more closely resembled their native conditions as compared to soils from cooler regions in southern Chile. According to the theory of thermal adaptation, microbial communities adjust to their native MAT in several ways, one of which may be selection of thermally adapted genotypes or taxa (Bradford, 2013; Bradford et al., 2019; Nottingham et al., 2019; Eng et al., 2023). Soil pH is an important environmental filter of microbial community composition (Lauber et al., 2009; Rousk et al., 2010), also along the investigated gradient (Chapter 2). It is therefore conceivable that soil pH indirectly affected specific respiration (and as a consequence CUE, Figure 4-3a) through the microbial community composition (Malik et al., 2018, 2019). The link between bulk C:N ratio and specific respiration is in line with a theory of stoichiometric constraints: when the C:N ratio of the substrate exceeds the C:N ratio of microbial biomass, CUE may be lower due to a relative N limitation of growth or increased overflow respiration (Manzoni et al., 2012b; Sinsabaugh et al., 2016). However, the redundancy of bulk C:N ratio and the microbial community composition (Figure S4-3, Figure S4-7) indicate that bulk C:N rather

affected CUE through the community composition. Lastly, the classification of soils into WRB major soil groups did not help to structure the influence of the environmental setting on soil microbial traits and functions, because variations within soil groups were large (Table S4-5).

### 4.4.4 Decoupled aspects of MIC drive specific respiration, specific growth and CUE

In support of the second hypothesis, respiration and growth normalized for MBC (i.e. specific respiration and growth) were overall more strongly linked to microbial community composition than to the environmental setting (Figure 4-2). Furthermore, specific respiration and specific growth were linked to different components of MIC (Figure 4-3b), which was in line with the third hypothesis, that they would be physiologically decoupled. This decoupling between specific respiration and specific growth was particularly pronounced in soil types without andic properties, which had less secondary sand content as discussed in section 4.4.3 (Table S4-5). We argue that this missing biological coupling of specific respiration and specific growth as well as the link to microbial community composition are connected and can be understood in the light of microbial physiology. Cell-specific respiration rates are evolutionary constrained, as the energy requirement of a cell depends at least partially on hardwired cellular traits (van Bodegom, 2007; Roller and Schmidt, 2015). More specifically, the potential for copiotroph behavior is likely associated with higher cell-specific respiration rates (Roller and Schmidt, 2015; Ho et al., 2017). As a consequence, specific respiration at the cell level may be a phylogenetically narrow (instead of broad) trait, rendering it a predictable aggregated trait (instead of an unpredictable emergent trait) at the bulk soil level. In the present study, seven bacterial genera were strongly correlated with PC1 (Table S4-4), which was the strongest predictor of specific respiration (Figure 4-2). Out of these, four genera dominant in soils with high specific respiration belonged to the phyla Alphaproteobacteria, Gammaproteobacteria and Actinobacteriota, which have the potential to behave copiotroph in response to substrate addition (Stone et al., 2023) (Table S4-4). In contrast to specific respiration, specific growth (or microbial turnover) is thought to be less directly driven by evolutionary hardwired obligate metabolism (Li et al., 2019; T. P. Smith et al., 2021; Stone et al., 2023). While the potential for fast growth may be evolutionarily conserved in some phyla (Stone et al., 2023), maximum growth rates are unlikely to be achieved in soil due to energy limitation (Manzoni et al., 2012b; Soong et al., 2020; Caro et al., 2023). This is supported by the observation that specific growth rates were less variable than specific respiration rates along the gradient (Table 4-1). As a consequence of hampered growth, specific growth should be decoupled from specific respiration, with a less direct link to community composition than specific respiration. This is supported by the observation that the MIC models on average explained 41 % (± 10 % S.D.) of variation for specific growth, but 73 % (± 7 % S.D.)

for specific respiration (Figure 4-2). Specific growth was most strongly linked with MIC PC2 (Figure 4-3b). MIC PC2 was associated with bacterial and fungal ASV richness, as well as the bacteria:fungi ratio and prevalence of saprotrophic taxa (Table S4-4, Figure S4-9). Our analysis links higher specific growth with lower taxonomic richness of bacteria and fungi, and with a reduced dominance of bacteria over fungi (Figure 4-3, Table S4-4). This is surprising, because it has previously been suggested that bacteria grow faster than saprotrophic fungi in soil (Rousk and Bååth, 2011; Caro et al., 2023). Irrespective of the underlying mechanisms, links between specific growth and community composition, as well as growth rate differences among bacterial taxa at submaximal growth rates have been observed before in more local studies (Li et al., 2019; Zheng et al., 2019a; Y. Chen et al., 2021; Metze et al., 2024). We propose that improved accuracy in the description of the active fractions of microbial communities (e.g. via parallel labeling of active parts of the community, (Couradeau et al., 2019; Stone et al., 2023; Metze et al., 2024) and improved characterisation of energy limitation (Kellerman et al., 2015; Davenport et al., 2023; Jones et al., 2023) could increase the resolution on proposed mechanisms. In summary, the observation that specific respiration and growth were linked to different components of the microbial communities provides evidence that microbial respiration and growth can be physiologically decoupled in soils under conditions of energy limitation (Figure 4-4b).

Interestingly, SOC substrate quality is expected to affect CUE (Gommers et al., 1988; Manzoni et al., 2012b; Roller and Schmidt, 2015; Jones et al., 2018), but we did not find this link (Figure 4-3a). Perhaps the chemical nature of available (dissolved) SOC generally causes energy-limitation (Manzoni et al., 2012b; Soong et al., 2020), irrespective of the degree of SOC decomposition at the bulk scale as measured in this study. Rather, microbial community composition was the best predictor of CUE (Figure 4-2), in line with our second hypothesis. Previous studies found that CUE increased with bacterial richness (Domeignoz-Horta et al., 2020) and bacterial dominance over fungi (Soares and Rousk, 2019). However, CUE was most strongly linked to PC1 (Figure 4-3b) which was correlated with several genera of potential copiotrophs (Table S4-4) belonging to the phyla Alphaproteobacteria, Gammaproteobacteria and Actinobacteriota; (T. P. Smith et al., 2021; Stone et al., 2023). We explain the link between CUE and microbial community composition with the fact that microbial CUE can be related to phylogeny due to physiological trade-offs between maximum growth rates and general metabolic efficiency (Saifuddin et al., 2019; Muscarella et al., 2020; T. P. Smith et al., 2021). More specifically, the potential for copiotroph behavior is likely associated with inefficient substrate use (due to higher maintenance cost) at submaximal growth rates, resulting in lower CUE as compared to oligotrophs (Roller and Schmidt, 2015; Roller et al., 2016; Ho et al., 2017). Our observations suggest that at submaximal growth rates (due to energy limitation) CUE at the investigated scale seems to be mainly driven by specific respiration, which is largely a function of community composition (Figure 4-

4). Theoretical considerations (Hagerty et al., 2018) and local case studies over soil depth (Zhang et al., 2023) and season (Simon et al., 2020; Schnecker et al., 2023) identified the same fundamental mechanism: specific respiration and specific growth are uncoupled products of microbial metabolism, and CUE is their integrated product. This also helps to understand a recent finding where CUE decreased with carbon limitation across different land uses (He et al., 2023): In situations of carbon limitation, microbial communities can downregulate growth more than respiration, thereby lowering their CUE. At the same time, the positive link between specific growth and CUE (Figure 4-1) is consistent with a concept proposed by Lipson (2015); in regimes of carbon- (i.e. energy-) limitation, increased growth rates lead to an increase of CUE.

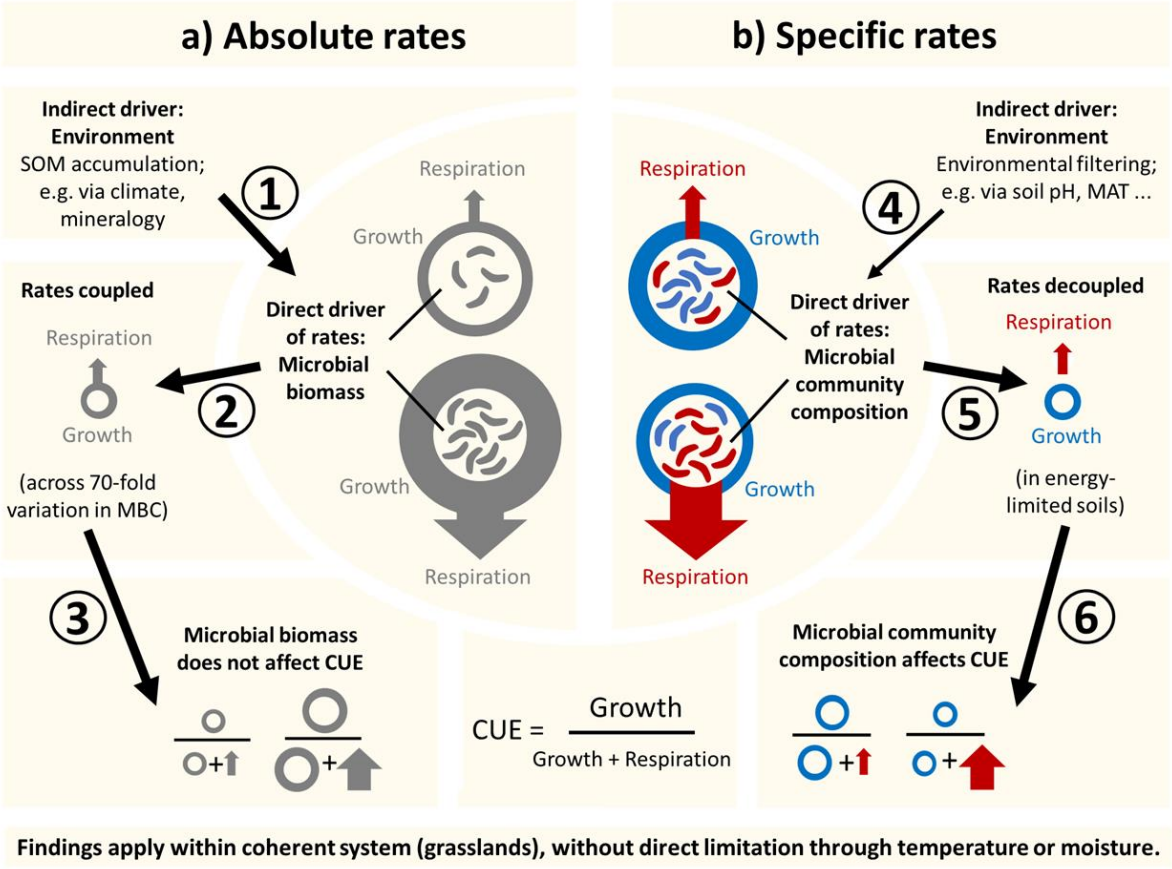



**Figure 4-4.** Summary of the findings of this study. Panel a) shows absolute rates (respiration and growth), panel b) shows specific rates (specific respiration, specific growth, *i.e.* rates normalized for MBC). At the investigated scale, the environment (SOM quantity) drives microbial biomass (1), which in turn drives coupled absolute rates of growth and respiration (2), resulting in no strong effects on CUE (3). The environment also shapes the microbial community (4), which drives decoupled specific rates of growth and respiration (5), which result in strong effects on CUE (6). MBC = microbial biomass carbon, CUE = carbon use efficiency,

The dominant (and identifiable) controls on SOC dynamics depend on scale and context (Ali et al., 2018; González-Domínguez et al., 2019; Li et al., 2020; Tian et al., 2022). The present biome-scale gradient features a large variation of environmental setting and microbial traits and functions (Table S4-1, Table 4-1) that allows insights on the controls of microbial functioning across pedo-climatic regions. At smaller scales, however, finer variation of microbial traits and functions may be of interest, which could be driven by other factors and processes that are masked at the large scale. In addition, the influence of stressors on microbial community functioning, such as moisture fluctuations, might complicate the picture by potentially overriding or amplifying the presented links (Schimel, 2023). However, we argue that the acknowledgement that physiological decoupling of specific respiration and specific growth - which takes place at cellular levels - translates to the biome-scale has the potential to strongly improve our understanding of CUE across scales. Currently, it is common practice to investigate CUE as just one variable. Based on this study, we however propose that CUE should be interpreted and represented in its components, regardless of the scale.

## 4.5 Conclusion

This study showed that the environmental setting and the microbial community composition drive different aspects of SOC cycling. Conceptualization of absolute rates as the product of microbial biomass and specific rates could help to reconcile the different roles that environmental setting and microbial community composition play in the SOC cycle. Ultimately, we conclude that disentangling microbial biomass, specific respiration rate and specific growth rate could be a promising step towards a better understanding of the SOC cycle at the large scale. At the scale of a biome (*i.e.* temperate grasslands), the environment (predominantly SOM quantity) was found to drive and dominate patterns of microbial biomass, which in turn drove growth and respiration. However, specific growth and specific respiration were most strongly linked to features of microbial community composition that followed varying trends, resulting in strong effects on CUE (Figure 4-4), and only to a limited extent to edaphic factors of the environmental setting.

The fact that specific growth, specific respiration and consequently CUE are partially predictable but decoupled properties of the microbial community composition urges us to reconsider the conceptual representation of CUE as just one value. Along the lines of Hagerty et al. (2018), we provide evidence that CUE might be more comprehensible and more valuable when it is analyzed in its components. For conceptual and numerical models, a potential way forward could be to consider specific respiration rates and specific growth rates separately. On the one hand, specific respiration rates might be a

function of relative cellular maintenance costs and, therefore, genomic indicators for the potential for copiotroph behavior could help to predict specific respiration rates. On the other hand, energy limitation and the ratio of bacteria to fungi may directly control specific growth rates (or microbial turnover) in soil systems. Further research will be required to identify more accurate measurable proxies for these factors, but the increasing amount of data for soil microbial growth rates will ultimately facilitate parametrization and validation of growth in models.

## 4.6 Acknowledgements

We thank Cornelia Rottensteiner, Annina Maier and Marco Griepentrog for their assistance with the laboratory work. We further thank Manuel Casanova for soil sampling and Katherine Rebolledo for shipment of the samples.

# 5. Synthesis

In this thesis, I used a geoclimatic soil gradient to evaluate whether physical stabilization and microbial mechanisms related to the SOC cycle translate from the mesoscale to the macroscale. The investigated mechanisms were (1) stable microaggregates as an SOC reservoir, (2) bacterial competitive exclusion and substrate specialization and (3) biomass-specific microbial C metabolism. In Chapters 2 to 4, I could show together with coauthors that stable microaggregates and biomass-specific C metabolism translate from the mesoscale to the macroscale, while we found that bacterial substrate specialization was not a relevant mechanism along the geoclimatic gradient (Figure 5-1). In the following synthesis, I will summarize the findings across all three chapters and highlight interdisciplinary links among them (Chapter 5.1). For this, I will use a comprehensive comparison of the SOC dynamics of two contrasting systems along the gradient. In addition, I will discuss the general conclusions that I have derived from each of the three chapters (Chapters 5.2 to 5.4). I will place these conclusions in the context of the current state of the art, discuss their implications for the respective research fields and propose future research directions that build on this thesis. Following these perspectives, I will briefly discuss how the investigated system aligns with the general terrestrial SOC cycle (Chapter 5.5). This will provide guidance to place the findings of this thesis in a global context. Lastly, I will conclude with a general outlook about future research with this dataset (Chapter 5.6).

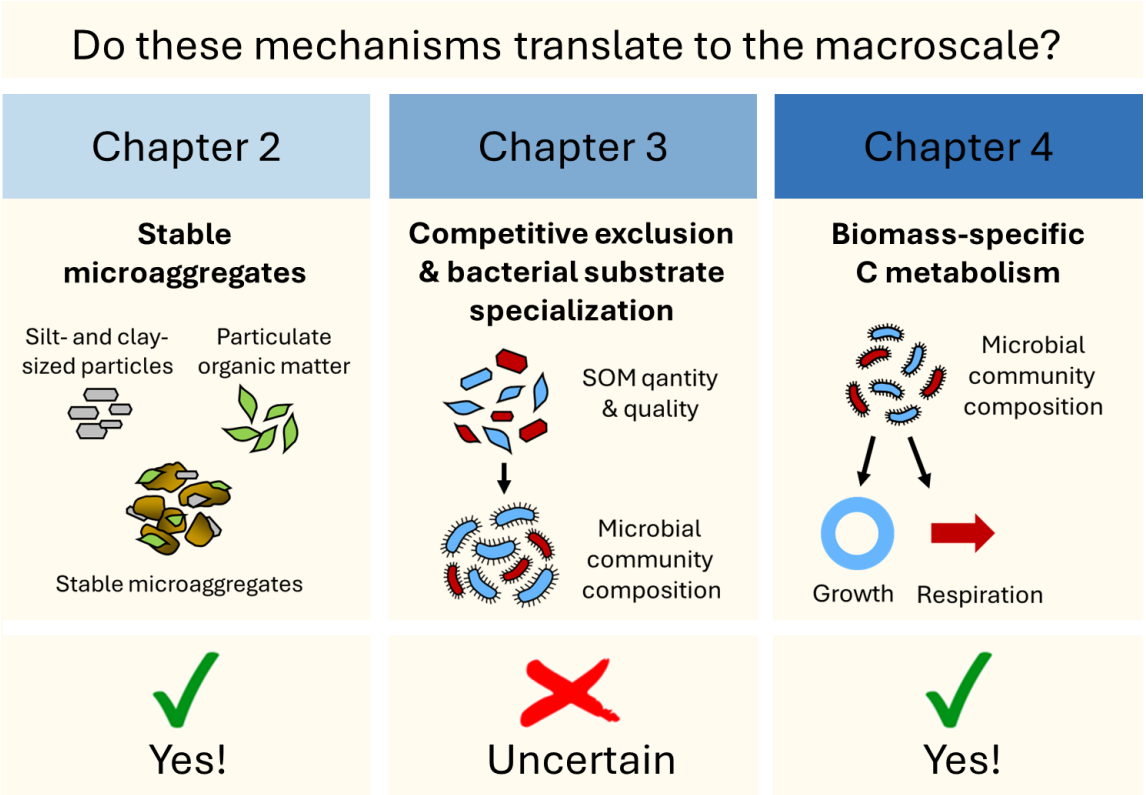


**Figure 5-1.** Conceptual summary of Chapters 2 to 4, at the bottom simplified answers to the main question: Do the investigated mechanisms translate to the macroscale?

## 5.1 Connecting the dots: soil organic carbon cycling along the gradient

In the following section I will summarize the knowledge about SOC cycling that we have gained through this thesis across diverse geoclimatic settings. The broad and interdisciplinary range of biogeochemical measurements that was conducted along this gradient allows us to paint an exciting picture of unusual clarity of macroscale SOC dynamics across temperate grasslands. To integrate the key information from Chapters 2 to 4, I will highlight two sites with contrasting geoclimatic settings (Figure 5-2): a dry and warm Arenosol which has formed on marine terraces and stabilized dunes in the north of the gradient (Los Vilos), and a Patagonian Andosol in the wet and cool central region of the gradient, which has formed on holocenic volcanic ash (Chapo). I want to emphasize that the patterns that I describe in this summary are not merely based on the comparison of these two sites that have contrasting and extreme properties. Rather, they are based on continuous patterns that have emerged across 35 sites and that have been described in detail in the respective Chapters 2 to 4. I believe that it is the consistency of the described patterns across this large range of physicochemical conditions which renders this dataset extraordinarily valuable. In the following section, values in brackets always refer to the example sites Los Vilos and Chapo, but a summary of the entire ranges and patterns is provided in Figure 5-2.

In the dry and warm soils in the north of the gradient (MAT 15.8°C, MAP-PET -1152), mineralogic reactivity was low (pedogenic oxides 1.5 g $kg^{-1}$ soil), soil pH was near neutral (5.7), and sand content was high (91.8 %). In the cool and wet soils of Patagonia (MAT 10.0°C, MAP-PET 1295), mineralogic reactivity was high (pedogenic oxides 18.1 g $kg^{-1}$ soil) and soil pH was more acidic (4.8). Most soils in Patagonia were per se less sandy (primary sand, 46.8 %), but considerable aggregation (17.1 %) caused sandy secondary textures (secondary sand, 63.9 %). These differences could partially stem from pedogenesis that might be further advanced in the wetter climate of central Patagonia. However, it is evident that climate and weathering are not the sole cause for the observed differences along the gradient. Los Vilos is an Arenosol which formed on sandy parent material; in contrast, central Patagonia is dominated by Andosols (such as Chapo) which feature highly reactive short range order minerals. Due to an autocorrelation between climatic and andic properties along the Chilean gradient, the effects of parent material and weathering were difficult the disentangle. However, it is safe to say that contrasting parent material and climate created a very strong gradient of soil physicochemical conditions. These soil physicochemical conditions caused consistent patterns of SOC dynamics.

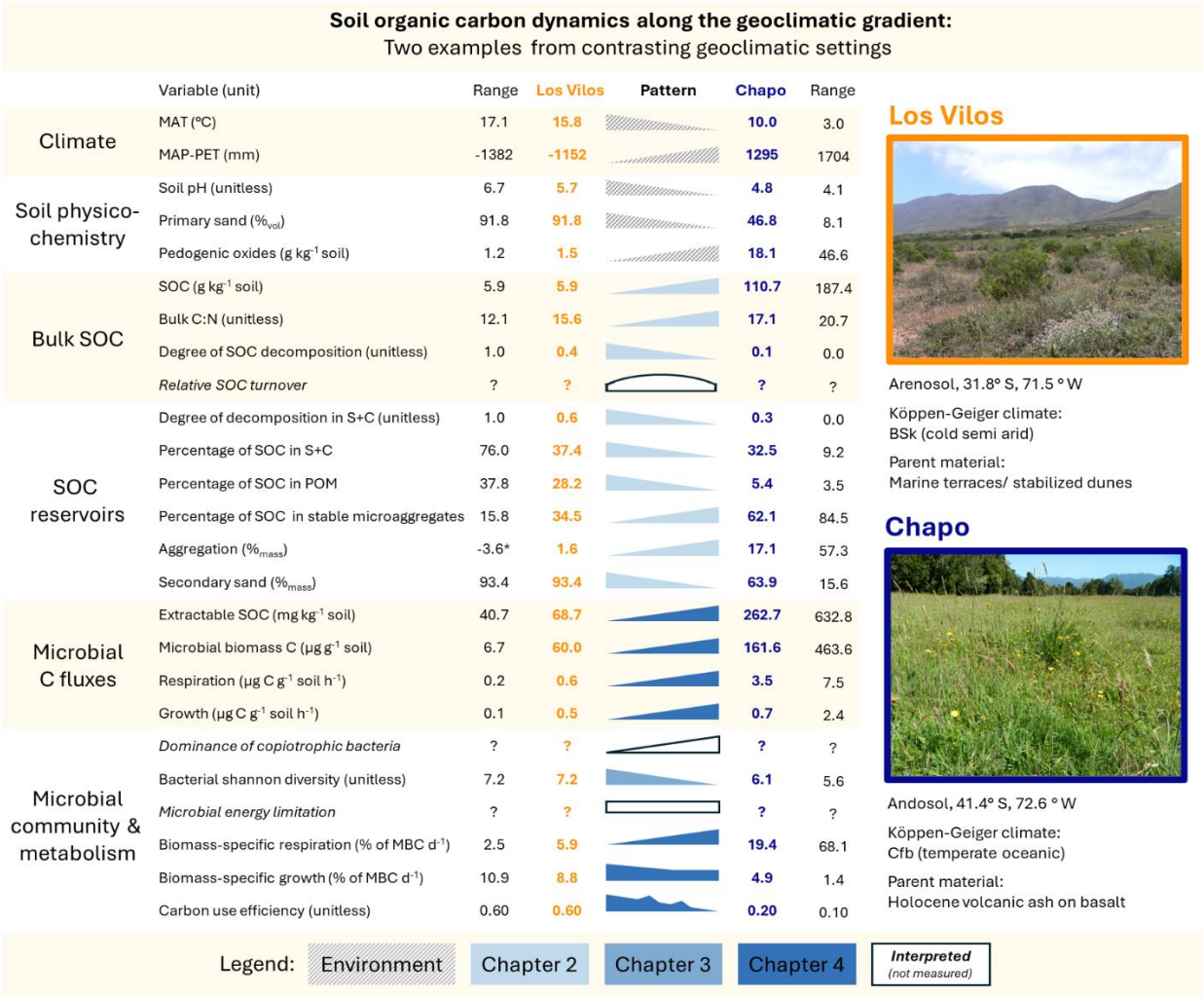



**Figure 5-2.** Conceptual summary of SOC dynamics along the gradient, highlighted with data from two sites in contrasting geoclimatic settings. Photographs by Manuel Casanova. MAT = mean annual temperature, MAP-PET = annual water balance, Primary sand = sand content after full dispersion, S+C = silt- and clay-sized fraction, POM = particulate organic matter, Secondary sand = sand content without dispersion of stable microaggregates, MBC = microbial biomass carbon.

The drier and warmer soils in the north with low mineralogical reactivity accumulated less SOC (5.9 g $kg^{-1}$ soil) than the wetter and cooler Patagonian soils with higher mineralogical reactivity (110.7 g C $kg^{-1}$ soil). The dominant SOC fraction in soils such as Los Vilos were the silt- and clay-sized particles (37.4 % of the SOC recovered in the fractionation), whereas SOC in soils such as Chapo was dominated by stable microaggregates (62.1 % of the SOC recovered in the fractionation). The amount of C associated with silt- and clay-sized particles was mainly driven by mineralogical reactivity, whereas the amount of C associated with stable microaggregates was driven by climatic as well as mineralogic properties. In soils like Los Vilos, the bulk C:N ratio was lower (Los Vilos: 15.6, Chapo: 17.1) and bulk SOM was more microbially transformed (Los Vilos: 0.4, Chapo: 0.1; index scaled between 1 and 0). Moreover, SOM associated with the silt- and clay-sized particles was more microbially transformed (Los Vilos: 0.6,

Chapo: 0.3; index scaled between 1 and 0). This indicates that *in situ* stabilization – where SOC passes through microbes – was more important in Los Vilos, whereas *ex situ* stabilization – where SOC bypasses microbial uptake – was more important in Chapo. Based on these findings, I argue that in soils like Los Vilos (warmer, drier, less reactive), the turnover of SOC relative to organic matter input is higher than in soils like Chapo (cooler, wetter, more reactive). In addition, I conclude that lower mineral reactivity (low content of pedogenic oxides and clay) limits the amount of SOC that can accumulate in Los Vilos. In low-SOC soils like Los Vilos, concentrations of extractable (and thus available) C were lower than in high-SOC soils like Chapo (68.7 and 262.7 mg $kg^{-1}$ soil, respectively). Linked to the lower available C concentration in Los Vilos was a lower microbial biomass (60.0 vs. 161.6 µg MBC $g^{-1}$ soil), which translated into lower absolute rates of microbial respiration (0.6 vs. 3.5 µg C $g^{-1}$ soil $h^{-1}$) and microbial growth (0.5 vs. 0.7 µg C $g^{-1}$ soil $h^{-1}$) at controlled temperature and moisture conditions (20 °C, 50 % WHC). I argue that as a consequence of lower substrate availability, the bacterial communities in low-SOC soils were not dominated by copiotrophic taxa, which may have made them more diverse than high-SOC soils (Shannon diversity 7.2 vs. 6.1). Due to the prevalence of oligotrophic taxa in low-SOC soils, biomass-specific respiration was lower in low-SOC soils (Los Vilos: 5.9 % of MBC $d^{-1}$) than in high-SOC soils (Chapo: 19.4 % of MBC $d^{-1}$). In high-SOC soils, copiotrophic taxa were more prevalent, which are likely to have a higher energy-demand for cell-maintenance. I argue that biomass-specific growth was less variable between the systems than specific respiration because energy limitation is likely a strong growth constraint across all soils. Nevertheless, biomass-specific growth was linked to the microbial community and to texture in such a way that it was highest in the northern soils of the gradient similar to Los Vilos (8.8 % of MBC $d^{-1}$), and lower in soils like Chapo (4.9 % of MBC $d^{-1}$). The link between texture and microbial activity suggests that aggregation may not just affect SOC availability, but perhaps also influence microbial activity through altering microscale conditions. As a consequence of the decoupled metabolic controls on biomass-specific respiration and growth, CUE tended to be highest in low-SOC soils (Los Vilos 0.60) and low in high-SOC soils (Chapo 0.2). However, this pattern was inconsistent, and microbial community composition directly was a better predictor of CUE than the environmental setting.

The high resolution and interdisciplinary characterization of this large gradient allowed to link soil physicochemistry, SOC accumulation, microbial C fluxes and microbial metabolism across a biome (temperate grasslands). Thereby, this work could (rudimentarily) account for the complexity of soil systems. By doing so, this thesis portrayed numerous links between the environmental setting, SOC quantity and quality, soil microbial community composition and microbial traits and functions. Overall, the work presented here improves our understanding of how the entire SOC cycle is directly or indirectly affected by the environmental setting. Climate and soil physicochemistry are the ultimate

controls of SOC dynamics at the macroscale (Colman and Schimel, 2013; Sebastian Doetterl et al., 2015; Rasmussen et al., 2018; Yu et al., 2021; Heckman et al., 2022). However, microbial C metabolism was more directly linked to microbial community composition. I believe that this finding adds a new and important nuance to our understanding of macroscale SOC dynamics. Despite the overarching dominance of climatic and soil physicochemical controls, neither physicochemical properties nor microbial processes alone are sufficient to fully explain all aspects of the macroscale SOC cycle. Rather, they are complementary. I believe that the clarity with which the presented gradient work shows this complementarity helps to make sense of two well-articulated but somewhat opposed perspectives: scaling from the microbial view (Schimel, 2023) and scaling from the soil physicochemical view (Baveye, 2023). More personally, this work has helped me to consolidate the microbial- and biology-focused "socialization" from my Master studies with a geographer`s physicochemical macroscale context. For me, the most important take-away from my doctoral thesis is that it opened my eyes to the importance and value of connecting perspectives and of thinking across disciplines.

## 5.2 Stable microaggregates at the macroscale: potentials and challenges

In Chapter 2, my coauthors and I investigated whether stable microaggregates constitute a relevant SOC reservoir. Are we overlooking microaggregates at the macroscale, despite knowing that aggregation is an important mechanism at the mesoscale? We could show that stable microaggregates (isolated with a density and particle size fractionation scheme, Zimmermann et al., 2007) constitute a major SOC reservoir, because they contained on average half of the SOC along the gradient. We further asked whether stable microaggregates are predictable with environmental conditions. We found that the amount of SOC in stable microaggregates was linked to climate as well as mineralogic soil properties. Stable microaggregates were quantitatively most important in wet and cool systems, and in soils with high mineralogic reactivity (particularly Andosols). Lastly, we used stable isotope analysis and DRIFT spectroscopy to investigate how consistent the chemical characteristics of stable microaggregate C were across contrasting geoclimatic settings. The degree of decomposition of C in stable microaggregates was significantly different from particulate organic matter and silt- and clay-sized fractions and did not change along the gradient. In summary, Chapter 2 demonstrated that stable microaggregates constitute a quantitatively important SOC reservoir, with environmental drivers and chemical characteristics that were distinct from particulate organic matter and silt- and clay-sized fractions.

I therefore conclude that the scaling of aggregate dynamics to the macroscale is a worthwhile endeavor. Our results and findings suggest that more macroscale models should feature aggregate turnover. One notable example of a model that already integrates aggregate dynamics is the Millennial v2 model

(Abramoff et al., 2022), which incorporates aggregate dynamics based on the AggModel by Segoli et al. (2013). In the AggModel, microaggregates form within macroaggregates, the rates of aggregate dynamics are linear, and microaggregates protect SOC (Segoli et al., 2013). Moreover, the fractionation method used in this thesis (Zimmermann et al., 2007) was initially proposed to be used for the initialization of the first-order rate model RothC. For example, Wiesmeier et al. (2016) successfully predicted SOC stocks in Bavaria (Germany) with a RothC model initialized with Zimmermann-fractions. Chapter 2 now provides powerful empirical evidence that stable microaggregates indeed constitute an SOC reservoir with consistent macroscale patterns that are coherent with mechanistic knowledge from the mesoscale. This finding therefore supports the representation of stable microaggregates at the macroscale, as implemented in Millennial v2, or more indirectly in model initialization based on the fractionation scheme described by Zimmermann et al. (2007).

Contrary to what our findings suggest, the popular simplification of separating SOC into POM and MAOM disregards the potential role of aggregates as a relevant macroscale mechanism. The POM-MAOM simplification is mostly achieved by (partial) dispersion of aggregates, followed by size fractionation, where MAOM < e.g. 63 µm < POM (Poeplau et al., 2018; Lavallee et al., 2020). The dispersion of aggregates in the POM-MAOM framework means that the aggregate-associated C ends up in the POM and MAOM fractions. The framework thereby combines free and occluded POM (which was liberated by aggregate dispersion) into one functional reservoir, as well as silt- and clay-sized particles together with remnants of dispersed aggregates that are < 63 µm. In strongly aggregated soils with high mineral reactivity (e.g. the soils in central Patagonia), this may skew the characteristics of POM and MAOM to a greater extent than in weakly aggregated soils. I am working to investigate this aspect further (see Chapter 5.6). However, the current fate of stable microaggregates at the macroscale is not solely a case of being “overlooked”, as I asked provokingly in this thesis. Rather, it is more accurate to state that stable microaggregates are disregarded knowingly from scaling. This is mainly for reasons of practicality (Lavallee et al., 2020).

Two major challenges remain which in my opinion obstruct the development of an effective macroscale representation of stable microaggregates. In the following I want to briefly discuss these two challenges. (i) A universal and unanimously accepted operational definition of stable microaggregates is missing. Different fractionation schemes consider different operational definitions, and thus arrive at different aggregate fractions. And even with such a consensus, it will be challenging to apply a consistent dispersion treatment across different laboratories or soil types, or for that reason compare “aggregates” between different studies. The method of dispersion (von Lützow et al., 2007; Poeplau et al., 2018; Just et al., 2021), detailed specifics in the sonication step (if dispersal is achieved by sonication) such as power or intervals (Cerli et al., 2012; Kaiser and Berhe, 2014; Poeplau and Don, 2014) and even aspects of sample mixing (Büks, 2023) can strongly affect the results of fractionation

procedures. Often, the description of fractionation procedures is not sufficient for exact reproduction (Poeplau and Don, 2014; Büks, 2023). This in turn creates additional uncertainties for the parameterization and calibration of models with aggregate data. Therefore, a consensus definition on a well reproducible isolation procedure for stable microaggregates – or a set of such procedures for different soil types – would greatly facilitate the scaling of aggregate dynamics. Such a method – or methods – further need to be tailored to capture the essential functions that aggregates play in the SOC cycle. This leads to the second major challenge. (ii) Although we could show that aggregates follow distinct environmental controls and have an SOC composition that is distinct from POM and the silt- and clay-sized fraction, it is still unresolved how well aggregates really protect SOC from decomposition. One might think instinctively: this is easy, $^{14}C$! However, resolving causality in the context of C age in aggregates remains a tough nut to crack. Does SOC in aggregates have a slower turnover because it is stabilized in aggregates, or does SOC that turns over slowly due to inherent chemical recalcitrancy preferentially promote aggregation (Wagai et al., 2009)? I propose that a way forward could be the use of compound-specific $^{14}C$ dating (Ascough et al., 2009; van der Voort et al., 2017; Gies et al., 2021; Grant et al., 2022). Direct comparison of the $^{14}C$ age of the same compounds with and without aggregate-association could circumvent uncertainties regarding the effect of inherent "recalcitrance" of compounds. In addition to the general aspect of stabilization, the relevance of scaling stable microaggregates also depends on whether C turnover in aggregates reacts consistently to environmental change, as has been shown for POM and MAOM (Rocci et al., 2021). Based on mechanistic knowledge about aggregate dynamics, it is reasonable to hypothesize that warming or altered moisture regimes will affect aggregate turnover (Six et al., 2004; Totsche et al., 2018; Poeplau et al., 2020b; Philippot et al., 2024). If this hypothesis can be confirmed, it might provide a strong reason to consider aggregate dynamics in macroscale simulations. To test this, a meta-analysis of aggregate dynamics in climate manipulation experiments will be necessary, similar to the approach of Rocci et al. (2021) for POM and MAOM.

All things considered, Chapter 2 demonstrated that aggregates constitute a SOC reservoir which translates to the macroscale, and that there could be a considerable gain from representing aggregates as an independent SOC reservoir. Ultimately, we need to weigh the gain derived from scaling the mechanism of aggregate turnover versus the loss that comes from not scaling it. Until the challenges which I elaborated above have been addressed, it will be helpful to understand how the lack of aggregate representation affects models. For example Laub et al. (2024) recently showed that the absence of SOC protection through aggregates leads to lower SOC turnover estimates in other pools. Once the remaining questions have been addressed, stable microaggregates will be ready for scaling, and we will no longer need to "knowingly disregard" our mechanistic knowledge at the macroscale.

## 5.3 Linking carbon and bacterial communities: a need for more sophisticated tools

In Chapter 3, we asked whether there are scalable links between SOM (including C and N) and soil bacterial community composition. To address this question, we investigated how soil bacterial community composition relates to the environment and to the quantity and qualitative characteristics of SOM. Please note that in this chapter, SOM quantity refers to a rotated component that represented the contents of total and extractable C and N. The quality of SOM was assessed based on chemical characteristics derived from spectroscopy, thermostability, stable isotope ratios and the prevalence of particulate organic matter. We found that soil pH and SOM quantity were the strongest drivers of soil bacterial community composition. More specifically, soils with high SOM quantity and low soil pH were less diverse (alpha diversity) than soils with low SOM quantity and high soil pH. However, the identification of the underlying mechanism that links SOM quantity and diversity remained challenging. With the available data, we could not evaluate whether the link between alpha diversity and SOM quantity was due to the mechanism of competitive exclusion, or due to a co-dependence of diversity and SOM quantity on SOM turnover rates. Lastly, we did not find evidence that the mechanism of bacterial substrate specialization had a pronounced effect on soil bacterial community composition. Instead, bacterial community composition only showed a weak link with the qualitative composition of SOM. Overall, these findings suggest that bacterial substrate specialization – which is a well-established mechanism at the mesoscale – does not play an important role at the macroscale which can be detected with standard methods that are used to describe qualitative properties of SOM. I therefore conclude that bacterial substrate specialization does not merit scaling as a mechanism that determines bacterial community composition.

To my knowledge, the negative link between alpha diversity and SOM quantity was nevertheless a novel finding at the macroscale. Two potential underlying mechanisms are: (i) competitive exclusion, and (ii) co-dependence of diversity and SOM quantity on SOM turnover rates. Competitive exclusion would occur if bacteria that are adjusted to high substrate conditions outcompete other bacteria (Malik et al., 2020). In contrast, a dependence of bacterial diversity on SOM turnover rates was proposed in other studies (Delgado-Baquerizo et al., 2016b; Zhou et al., 2016), although neither of these studies comprehensively elaborated or even identified potential underlying mechanisms. Nevertheless, in order to test this second hypothesis as was proposed by others, a measure of *in situ* SOM turnover would be necessary. Approaches to accurately quantify plant OM inputs into soil across gradients are complicated and very laborious. An impressive example by L. Chen et al. (2021) demonstrated that this requires empirical estimates of aboveground net primary productivity and root biomass, in combination with satellite-derived high resolution vegetation parameters. A simpler approach to get a

sense of potential *in situ* patterns of SOM turnover along the gradient could be the measurement of decomposition rates with the tea bag index (Keuskamp et al., 2013). SOM turnover estimation with this method would solely be based on mass losses (and not directly on C losses) and it would neglect input rates as well as the turnover of native SOM and labile SOM. However, it would be technically simple enough for application across this large gradient, although still requiring considerable logistical effort. A similar but much more elaborate approach could be the *in situ* addition of isotopically labelled OM, as described by (Schiedung et al., 2023a). While this approach also just quantifies turnover of the added material, and in addition needs to account for vertical losses, it directly quantifies the dynamics of substrate-derived C (and N) and even allows to trace it into different fractions. Another approach to estimate SOM turnover is based on $^{14}C$ measurements (Torn et al., 2009; Shi et al., 2020). However, different turnover rates of different carbon pools complicate this approach (Sierra et al., 2017; van der Voort et al., 2019; S. W. Stoner et al., 2023).

We implicitly assumed a decreasing molecular diversity of SOM with increasing microbial transformation based on literature (Roth et al., 2015; Hoffland et al., 2020; Davenport et al., 2023; Jones et al., 2023). However, perhaps this assumption was flawed – molecular diversity within intact macromolecules of plant litter may not necessarily be of much help to soil bacteria in search of dissolved available substrate. I now think that higher SOM turnover could theoretically increase the molecular diversity of dissolved available SOM. Possibly, low-SOC soils could (1) have a lower diversity of total bulk SOM because less plant material (which is chemically diverse, but insoluble) accumulates, while at the same time (2) a relatively larger range of intermediate decomposition products are in solution and lead to a higher diversity of available substrate. For future work, a more precise characterization of available SOM would therefore be desirable. Possible methods to achieve this could be pyrolysis combined with gas chromatography and subsequent mass spectrometry (py-GC-MS, Jones et al., 2023), liquid chromatography combined with tandem mass spectrometry (LC-MS/MS, Davenport et al., 2023), fluorescence spectroscopy combined with parallel factor analysis modelling (Fellman et al., 2008), or fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS, Kellerman et al., 2015). To my knowledge, all these methods are laborious and come with challenges and limitations in soils, but I think across strong environmental gradients they could provide an exciting avenue to further explore the link between substrate diversity and bacterial community composition. A first steppingstone could be solid-state $^{13}C$ nuclear magnetic resonance analysis in order to better evaluate the chemical composition and diversity of bulk SOM (Hall et al., 2020; Castañeda-Gómez et al., 2023; Man et al., 2023).

In Chapter 3, we found that SOM quality was mainly linked to rare bacterial taxa. However, the investigation of rare taxa was limited by sequencing depth and the size of the sequenced 16S rRNA gene fragments. Improved resolution of rare taxa abundances could therefore also provide improved

insights into the links between SOM quality and bacterial community composition. A higher sequencing depth would be achieved by pooling less samples together during the sequencing procedure, and longer gene fragments could be obtained with a different set of primers during gene amplification.
In conclusion, we did not find strong scalable links between soil bacterial community composition and SOM composition. To safely discard the existence of scalable links between the SOM quality and communities, higher resolutions of community composition and SOM composition may be necessary. A more promising link was found between bacterial diversity and SOM content. However, further work will be necessary to understand the mechanism that underlies this pattern. Nevertheless, our findings in Chapter 4 - namely, that microbial community composition affects microbial traits and functions such as CUE and biomass-specific respiration – underline the importance to better understand the mechanisms that determine microbial community composition.

## 5.4 Microbial metabolism: a promising avenue for scaling

In Chapter 4, we asked whether biomass-specific microbial metabolism translates to the macroscale. For this, we investigated how soil microbial traits and functions that are relevant for the SOC cycle are influenced by differences in the environmental setting, SOM and the soil microbial community composition. We found that microbial biomass and absolute process rates (*i.e.* rates per unit soil mass) were most directly linked to the environmental setting because the environment sets the frame for microbial activity. In accordance with our hypothesis we found that biomass-specific process rates and CUE were linked to microbial community composition. Moreover, we found that biomass-specific growth and respiration rates were linked to different members of the microbial community, indicating that they were physiologically decoupled. Overall, this demonstrates that metabolic mechanisms – which play out at the microscale – allow for improved predictions of microbial traits and functions at the macroscale.
Based on our findings, we can make suggestions for a refined representation of microbial traits and functions at the macroscale. We propose that genomic indicators for the potential of copiotrophic behavior could help to better predict specific respiration and CUE. A link between such genomic indicators (*e.g.* genome size, number of 16S operons per cell) and microbial traits has been shown repeatedly based on bacterial isolates (Saifuddin et al., 2019; Muscarella et al., 2020; T. P. Smith et al., 2021). However, this potential has not yet been unlocked in soil systems at the macroscale. In Chapter 5.6 I will briefly outline work that is underway to follow up on this idea. We further found that microbial community composition predicts specific growth only to a lesser extent. This finding is in line with previous work based on quantitative stable isotope probing (Li et al., 2019). The link between specific growth rates and microbial community composition may be weaker due to energy limitation in soil,

which constrains maximum growth rates (Soong et al., 2020; Caro et al., 2023; He et al., 2023). Therefore, a remaining challenge is to better characterize microbial energy limitation. A nested study approach (as described in Chapter 1.6) could provide a way forward. Based on chemical characterization of available SOC with py-GC-MS (Jones et al., 2023), LC-MS/MS (Davenport et al., 2023), FT-ICR-MS (Kellerman et al., 2015) or a combination of capillary electrophoresis and high performance liquid chromatography as described by van Hees et al. (2008), it should be possible to estimate a weighted mean of the degree of reduction (Manzoni et al., 2012b) or the nominal oxidation state of available SOC (Kellerman et al., 2015). As a side note, such a characterization of available SOC would go hand in hand with the determination of substrate diversity which I proposed in Chapter 5.3. If such a metric for energy limitation could be linked with specific growth rates, a simpler proxy would be desirable, for example based on fourier transformed infrared (FTIR) spectroscopy of available SOC. Another simpler measure for carbon limitation could be the degree of lignin biotransformation, which was recently linked to microbial CUE in a local warming experiment (Zhang et al., 2023).

But how can these insights be translated into models? The models reviewed in Chapter 1.4 represent microbial metabolism differently. COMISSION (Ahrens et al., 2015) uses a fixed value for CUE, and Millennial v2 (Abramoff et al., 2022) modulates CUE with temperature. MEMS (Zhang et al., 2021) and MIMICS (Wieder et al., 2014) treat CUE as a function of substrate quality. In MEMS, CUE is calculated dynamically as a function of the substrate C:N ratio. In MIMICS, different SOC pools have fixed CUE values, but functional groups of microbes differ in their diets (*i.e.* they consume different SOC pools, resulting in different CUEs for different groups of microbes). In MEND (Wang et al., 2022) and RESOM (Tang and Riley, 2014), microbial metabolism is modelled explicitly, and growth and respiration are decoupled. In both models carbon uptake, respiration, enzyme excretion and mortality are represented. In addition, MEND distinguishes between an active and an inactive pool of microbial biomass. If the link between microbial community composition, energy limitation and biomass-specific metabolism can be further refined (for instance with the approaches that I proposed above and outline in Chapter 5.6) these two models – MEND and RESOM – might offer the possibility to incorporate biomass-specific metabolism as a consequence of microbial community composition. Growth and respiration could be represented biomass-specific, informed by microbial community composition and possibly energy limitation; absolute rates could then be obtained by multiplication with microbial biomass, which is controlled by separate mechanisms that these models already feature. Overall, I believe that this is a promising avenue to improve representation of the SOC cycle at the macroscale.

## 5.5 Transferability of the gradient to other systems

I argue that the general insights of this thesis are relevant for the entire terrestrial C cycle. The range of investigated geoclimatic conditions was very broad (see Chapters 1.7 and 5.1). Based on mechanistic knowledge, there is no reason to believe that the insights on aggregation and biomass-specific microbial C metabolism would not apply beyond the investigated range of conditions. However, it is important to keep the limits of the gradient in mind when trying to transfer the magnitude and relative importance of the investigated mechanisms to other ecosystems or climatic extremes. The relative importance of individual mechanisms for SOC cycling is not just scale-dependent but can also be system-specific. In different systems, different mechanisms may play a larger role. In the following, I want to briefly discuss the limits of the investigated gradient. Which general aspects of SOC cycling will change when we transition to different ecosystems, when we move to more extreme climatic or pedogenic conditions, or when we dig deeper into subsoil? Considering the limits of the investigated gradient and the particularities of "neighboring" systems will provide guidance on how to place the findings of this thesis in a global context. Please note that the aim of this overview is to briefly give examples for the types of differences between the SOC dynamics of the study system and its "neighboring" systems (Figure 5-3). A comprehensive review of SOC cycling across systems would be beyond the scope of this chapter.

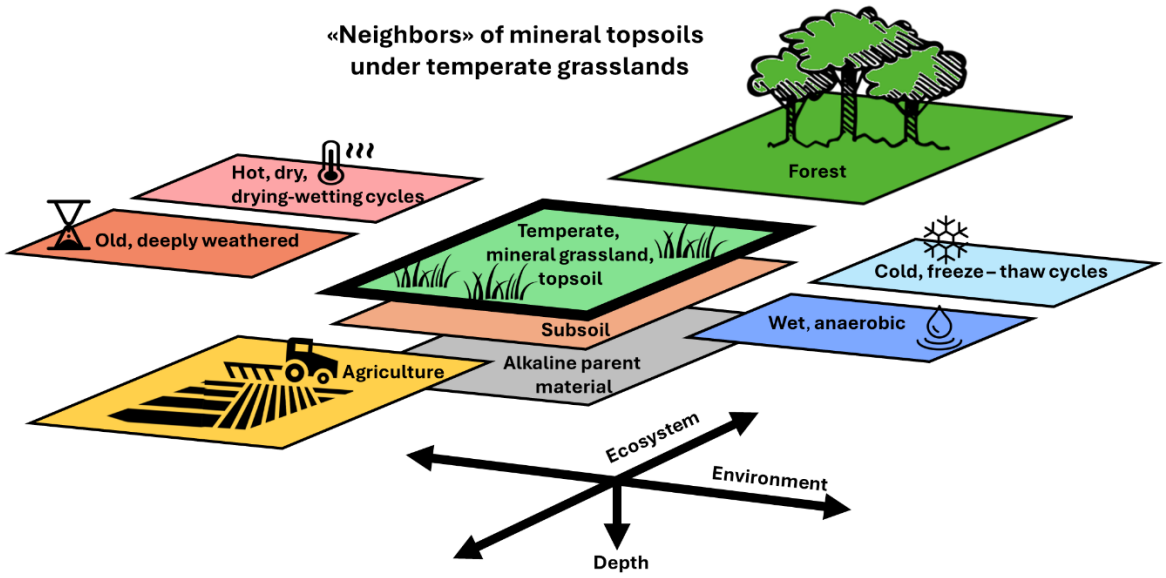



**Figure 5-3.** The "neighboring" systems of mineral topsoils in temperate grasslands. Icons were taken from thenounproject.com, detailed credits are listed in Appendix Chapter 5.

Moving to different ecosystems or land use types, a main factor that changes are the dynamics of OM input. The balance of aboveground vs. belowground OM inputs (Jobbágy and Jackson, 2000; Guo and

Gifford, 2002; Ding et al., 2021; Huang et al., 2021), as well as the chemical composition and diversity of inputs (Meier and Bowman, 2008; Mori et al., 2020; Angst et al., 2021a) depend on the type of vegetation. The entrance point as well as chemical composition of OM can have consequences for the decomposition dynamics: For example, mineral-associated SOC is more efficiently formed from belowground inputs (Sokol and Bradford, 2019; Sokol et al., 2019). In Chapters 2 to 4 of this thesis, we did not explicitly investigate how vegetation and input dynamics affected the mechanisms of interest, because within a coherent type of system, the functional differences should be minor. However, in order to transfer our insights to systems beyond temperate grasslands, differences in the dynamics of OM input likely need to be considered. For example, in agricultural soils, carbon tends to deplete over time, partially due to a depletion of OM inputs through harvest (Guo and Gifford, 2002; Sanderman et al., 2017; Wiesmeier et al., 2019). In addition to altered OM inputs, common management practices in agricultural soils (*e.g.* use of heavy machinery and tillage) can also introduce structural disturbances such as soil compaction (Batey, 2009) and disruption of aggregate formation (Six et al., 2004), which might override patterns caused by the natural physicochemical variability of soils.

In more extreme climates beyond temperate systems, direct climatic effects and system-specific mechanisms can override or modify patterns that are elsewhere continuous (Nave et al., 2021). For example, moving to very cold and wet systems beyond the gradient, or to topographic positions that accumulate water, microbial processes will become increasingly limited through metabolic constraints (Lloyd and Taylor, 1994; Freeman et al., 2001; Schipper et al., 2014). As a consequence, colder and/or wetter conditions will lead to reduced microbial decomposition, accumulation of litter and acidic soil pH. This can result in the formation of soils with thick organic layers (Histosols), soils with permanently frozen layers which can have high contents of undecomposed plant material (Cryosols) or permanently water-saturated systems where OM input exceeds decomposition (peatlands). In such systems, mineral stabilization plays a much smaller role than in the temperate mineral soils investigated in this thesis (García-Palacios et al., 2024; Hansen et al., 2024).

Moving towards the opposite climatic extreme, draught and drying-wetting cycles can govern SOC turnover. Lack of water disrupts connectivity within soils (Tecon and Or, 2017) and alters microbial functioning in the short- and in the long term (Canarini et al., 2021; Metze et al., 2023). Drying-wetting cycles further exert osmotic stress on microorganisms (Schimel, 2018; Malik and Bouskill, 2022) which can lead to pulses of $CO_2$ efflux from soils (Schimel, 2018; Manzoni et al., 2020). In dry systems, such pulses have been shown to govern SOC dynamics (Jarvis et al., 2007). Therefore, understanding the mechanisms that underly such particular events may be more informative than studying baseline soil functioning.

Old soils in (sub-)tropical regions can be deeply weathered, because long exposure to hot and wet climate leads to advanced chemical weathering (Amelung et al., 2018). Advanced weathering ultimately causes the depletion of reactive minerals, and such soils (Ferralsols, Lixisols, Acrisols) become infertile and have low carbon stabilization potential (Doetterl et al., 2018; Sayer et al., 2019; Georgiou et al., 2022; Reichenbach et al., 2023). As a consequence, a large part of biogeochemical cycling in tropical soil systems takes place in the litter layer (Sayer et al., 2024). In such systems, mineral soil as investigated in this thesis may contribute comparatively less to overall C dynamics as compared to temperate grasslands. In a way, the opposite extreme of the fast biological cycling of tropical litter layers is found when moving into the subsoil. In subsoil, mineral processes dominate over biological processes as the main control of SOC cycling (Hicks Pries et al., 2023). Subsoils are characterized by a large fraction of mineral-associated C, and low rates of C input as well as spatial constraints on the access to substrate limit microbial activity (Salomé et al., 2010; Button et al., 2022).

Alkaline soils are predominantly found on calcareous parent material (Bolan et al., 2023) and in dry regions (Slessarev et al., 2016). The investigated gradient was constrained to a soil pH below neutral, and not all patterns that were observed in this thesis necessarily extend into alkaline soils. This is because several crucial aspects of SOC dynamics can differ between acidic and alkaline soils. For example, C stabilization mechanisms change due to different soil physicochemical mechanisms (Rasmussen et al., 2018; Rowley et al., 2018), and microbial diversity and metabolic efficiency have been shown to follow different patterns in acidic and alkaline soils (Lauber et al., 2009; Malik et al., 2018; Jones et al., 2019).

From these examples, it becomes obvious that the specific features of different soil systems mandate regional tailoring of SOC dynamics when scaling from the mesoscale to the macroscale (Rasmussen et al., 2018; Heckman et al., 2021; Jungkunst et al., 2022; Wang et al., 2024). The importance of mechanisms for overall SOC dynamics needs to be weighed for each type of soil system. Understanding such differences is primarily important for the structuring, parametrization and calibration of numerical models. For example, stable microaggregates may play a smaller role in subsoils than in topsoils (Schrumpf et al., 2013; Poeplau et al., 2017), or in arable soils than in grassland soils (Poeplau and Don, 2013; Antony et al., 2022). However, while the relative importance of mechanisms may differ between systems, the fundamental mechanisms that are at work are transferrable. To stick with the example of aggregates, aggregates will probably form wherever the environmental setting allows for it, *i.e*. where organic matter accumulates in combination with sufficient mineral reactivity. Thus, system-dependence of SOC dynamics does not undermine the relevance of scaling mechanisms. Rather, it is a

strong call for more gradient studies that cover different types of soil systems! In addition, resolving the system-dependency of mechanisms will require the comparative analysis of SOC dynamics across different systems. For this, the integration of different approaches such as gradient studies, meta-analyses, databases and research networks will ultimately be imperative.

## 5.6 Research outlook along (and beyond) this geoclimatic gradient

Concluding this thesis, I want to give a brief outlook on further research that is underway along this gradient. An important aspect of this future perspective is data sharing. This gradient is now an extraordinarily well characterized study system. We plan to provide all measured data in the form of a documented dataset which can be shared with interested researchers, in order to foster further research. In combination with the metainformation provided by the regional soil surveys of CIREN (the Chilean Centro de Información de Recursos Naturales), the dataset can then also be used as a tool for researchers to inform targeted soil (and site) selection for experiments. As an additional side product of this thesis, this extensive gradient dataset might be useful for the parameterization, calibration and validation of numerical models. For example, Moritz Mainka (ETH Zurich, Switzerland) is currently using this dataset to do a sensitivity analysis of the COMISSION model (Ahrens et al., 2015), where he asks how the model reacts to variable numbers of input parameters. As another example, colleagues from Chile have already used samples from this gradient to investigate the potential of DRIFT spectroscopy to predict stable C isotopes as a proxy for SOC dynamics (Sepúlveda et al., 2021; Marcela et al., 2022).

Currently, together with colleagues I pursue three lines of investigation that follow up directly on questions from this thesis.

First, in Chapter 2 we could show that stable microaggregates have SOM characteristics (DRIFT spectroscopy data, stable isotope ratios, C:N ratio) that are distinct from particulate organic matter and silt- and clay-sized particles. However, we could not directly quantify to what extent the stable microaggregates consisted of particulate organic matter or smaller silt- and clay-sized particles. To this end, stronger physical dispersion of the isolated microaggregate fraction is in progress, coupled with SOM characterization. This will help to improve inferences about the mechanisms that underly aggregate formation along the gradient.

Second, in Chapter 4 (as well as in a global modelling study with colleagues, Gao et al., 2024) we could show that microbial metabolism affects traits and functions of microbes at the macroscale. In order to continue this promising direction of research, we collaborate with Estelle Couradeau (Pennsylvania State University, USA). During a research stay in Estelles group, I could learn the method of bioorthogonal non-canonical amino acid tagging (BONCAT) coupled with cell enumeration by flow

cytometry. Estelle and her team are working to implement this method for soil systems, and I was invited to contribute to the method development (Camillone et al., preprint). BONCAT allows to quantify the fraction of cells that are actively synthesizing proteins over the course of an incubation. This method will provide an exciting tool to determine how strongly biomass-specific respiration and growth rates are linked to the fraction of active cells. In addition, enumeration of cells together with quantitative PCR of the 16S rRNA gene (with Xingguo Han and Aline Frossard, WSL, Switzerland) will allow to test whether average 16S operon number per cell is a suitable proxy for biomass-specific respiration. This dataset could be further complemented with an in-depth characterization of the energy status of available SOM in these soils (more detailed discussion in Chapter 5.4). In summary, I hope that this work will provide a clear set of tools to improve the representation of microbial metabolism in numerical models.

Third, one important aspect of SOC dynamics that was not investigated in-depth in this thesis was the contribution of microbial necromass vs. plant material to mineral-associated SOC. In particular, a key question is: what controls at the macroscale how efficiently microbial turnover and plant inputs translate into mineral-associated SOC? While the decomposition indices in Chapter 2 (based on DRIFT spectroscopy data, stable isotope ratios, C:N ratio) provided a general insight into SOM characteristics of the SOC fractions, more sophisticated tools exist. In the master thesis of Annina Maier (ETH Zurich, Switzerland), we analyzed amino sugars in stable microaggregates and in the silt- and clay-sized fraction. However, in the course of this work, serious concerns about the accuracy of this state-of-the-art method emerged, and Marijn Van de Broek (ETH Zurich, Switzerland) will prepare a publication on this issue. While the search for a solution to this challenge continues, solid-state $^{13}C$ nuclear magnetic resonance analysis of the soils in collaboration with Myrna Simpson (University of Toronto, Canada) will provide better insights into the fate and dynamics of different SOM functional groups.

Ultimately, I believe that this gradient provides a wealth of opportunities for future research. I hope that many more insights will be derived from this dataset, from the frozen soil material that is still left – and perhaps from future sampling campaigns along the gradient that may have different but complementary foci. The insights derived from this thesis and the ongoing work demonstrate that macroscale gradient studies have strong potential as tools to evaluate whether micro- and mesoscale mechanisms translate to the macroscale.

Lastly, the work during my thesis has also sparked my interest in questions that lie beyond the capacity of this gradient. As introduced in Chapter 1.4, climate acts on SOC dynamics on variable temporal scales; short (days to months), intermediate (seasons to years) and long (decades to millennia). On each temporal scale, different mechanisms dominate. For my future work, I will be interested in turning

my attention to the integration of the temporal scales at which climate affects SOC cycling. I believe that a nested approach along altitudinal gradients could provide a potential way forward. Altitudinal gradients can control for differences in parent material and soil age, while maximizing contrasts in climate (or at least in temperature and the duration of seasons). I propose that in-depth characterization of pedogenesis in combination with detailed monitoring of microbial and vegetation dynamics along altitudinal gradients on contrasting parent material can help to better disentangle mechanisms that act in the long term via weathering vs. mechanisms that act in the intermediate term via biological dynamics. In addition, temperature and moisture manipulation experiments with soils from climatically contrasting ends of these gradients could help to embed short term microbial dynamics into such a temporally stratified picture of SOC dynamics. The resulting insights into the magnitude of SOC cycling mechanisms within one system but across temporal scales could potentially provide new answers to the question: which mechanisms do we need to scale – in the context of climate change?

# 6. References


Abarenkov, K., Zirk, A., Piirmann, T., Pöhönen, R., Ivanov, F., Nilsson, R.H., Kõljalg, U., 2022. UNITE general FASTA release for Fungi. UNITE Community. doi:10.15156/BIO /2483911

Abramoff, R.Z., Guenet, B., Zhang, H., Georgiou, K., Xu, X., Viscarra Rossel, R.A., Yuan, W., Ciais, P., 2022. Improved global-scale predictions of soil carbon stocks with Millennial Version 2. Soil Biology and Biochemistry 164, 108466. doi:10.1016/j.soilbio.2021.108466

Abramoff, R.Z., Xu, X., Hartman, M., O'Brien, S., Feng, W., Davidson, E., Finzi, A., Moorhead, D., Schimel, J., Torn, M., Mayes, M.A., 2018. The Millennial model: in search of measurable pools and transformations for modeling soil carbon in the new century. Biogeochemistry 137, 51–71. doi:10.1007/s10533-017-0409-7

Ahrens, B., Braakhekke, M.C., Guggenberger, G., Schrumpf, M., Reichstein, M., 2015. Contribution of sorption, DOC transport and microbial interactions to the 14C age of a soil organic carbon profile: Insights from a calibrated process model. Soil Biology and Biochemistry 88, 390–402. doi:10.1016/j.soilbio.2015.06.008

Ali, R.S., Kandeler, E., Marhan, S., Demyan, M.S., Ingwersen, J., Mirzaeitalarposhti, R., Rasche, F., Cadisch, G., Poll, C., 2018. Controls on microbially regulated soil organic carbon decomposition at the regional scale. Soil Biology and Biochemistry 118, 59–68. doi:10.1016/j.soilbio.2017.12.007

Alteio, L.V., Séneca, J., Canarini, A., Angel, R., Jansa, J., Guseva, K., Kaiser, C., Richter, A., Schmidt, H., 2021. A critical perspective on interpreting amplicon sequencing data in soil ecological research. Soil Biology and Biochemistry 160, 108357. doi:10.1016/j.soilbio.2021.108357

Alves, R.J.E., Callejas, I.A., Marschmann, G.L., Mooshammer, M., Singh, H.W., Whitney, B., Torn, M.S., Brodie, E.L., 2021. Kinetic Properties of Microbial Exoenzymes Vary With Soil Depth but Have Similar Temperature Sensitivities Through the Soil Profile. Frontiers in Microbiology 12, 735282. doi:10.3389/fmicb.2021.735282

Amelung, W., Blume, H.-P., Fleige, H., Horn, R., Kandeler, E., Kögel-Knabner, I., Kretzschmar, R., Stahr, K., Wilke, B.-M., 2018. Scheffer/Schachtschabel Lehrbuch der Bodenkunde. Springer Berlin Heidelberg, Berlin, Heidelberg. doi:10.1007/978-3-662-55871-3

Amelung, W., Bossio, D., de Vries, W., Kögel-Knabner, I., Lehmann, J., Amundson, R., Bol, R., Collins, C., Lal, R., Leifeld, J., Minasny, B., Pan, G., Paustian, K., Rumpel, C., Sanderman, J., van Groenigen, J.W., Mooney, S., van Wesemael, B., Wander, M., Chabbi, A., 2020. Towards a global-scale soil climate mitigation strategy. Nature Communications 11, 5427. doi:10.1038/s41467-020-18887-7

Angst, G., Mueller, K.E., Kögel-Knabner, I., Freeman, K.H., Mueller, C.W., 2017. Aggregation controls the stability of lignin and lipids in clay-sized particulate and mineral associated organic matter. Biogeochemistry 132, 307–324. doi:10.1007/s10533-017-0304-2

Angst, G., Mueller, K.E., Nierop, K.G.J., Simpson, M.J., 2021a. Plant- or microbial-derived? A review on the molecular composition of stabilized soil organic matter. Soil Biology and Biochemistry 156, 108189. doi:10.1016/j.soilbio.2021.108189

Angst, G., Pokorný, J., Mueller, C.W., Prater, I., Preusser, S., Kandeler, E., Meador, T., Straková, P., Hájek, T., van Buiten, G., Angst, Š., 2021b. Soil texture affects the coupling of litter decomposition and soil organic matter formation. Soil Biology and Biochemistry 159, 108302. doi:10.1016/j.soilbio.2021.108302

Antony, D., Collins, C.D., Clark, J.M., Sizmur, T., 2022. Soil organic matter storage in temperate lowland arable, grassland and woodland topsoil and subsoil. Soil Use and Management 38, 1532–1546. doi:10.1111/sum.12801

Asano, M., Wagai, R., 2014. Evidence of aggregate hierarchy at micro- to submicron scales in an allophanic Andisol. Geoderma 216, 62–74. doi:10.1016/j.geoderma.2013.10.005

Ascough, P.L., Bird, M.I., Brock, F., Higham, T.F.G., Meredith, W., Snape, C.E., Vane, C.H., 2009. Hydropyrolysis as a new tool for radiocarbon pre-treatment and the quantification of black carbon. Quaternary Geochronology 4, 140–147. doi:10.1016/j.quageo.2008.11.001

Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., Harend, H., Huerta-Cepas, J., Medema, M.H., Maltz, M.R., Mundra, S., Olsson, P.A., Pent, M., Põlme, S., Sunagawa, S., Ryberg, M., Tedersoo, L., Bork, P., 2018. Structure and function of the global topsoil microbiome. Nature 560, 233–237. doi:10.1038/s41586-018-0386-6

Bairoch, A., 2000. The ENZYME database in 2000. Nucleic Acids Research 28, 304–305. doi:10.1093/nar/28.1.304

Baker, N.R., Allison, S.D., 2017. Extracellular enzyme kinetics and thermodynamics along a climate gradient in southern California. Soil Biology and Biochemistry 114, 82–92. doi:10.1016/j.soilbio.2017.07.005

Baran, R., Brodie, E.L., Mayberry-Lewis, J., Hummel, E., Da Rocha, U.N., Chakraborty, R., Bowen, B.P., Karaoz, U., Cadillo-Quiroz, H., Garcia-Pichel, F., Northen, T.R., 2015. Exometabolite niche partitioning among sympatric soil bacteria. Nature Communications 6, 8289. doi:10.1038/ncomms9289

Basile-Doelsch, I., Balesdent, J., Pellerin, S., 2020. Reviews and syntheses: The mechanisms underlying carbon storage in soil. Biogeosciences 17, 5223–5242. doi:10.5194/bg-17-5223-2020

Bastian, F., Bouziri, L., Nicolardot, B., Ranjard, L., 2009. Impact of wheat straw decomposition on successional patterns of soil microbial community structure. Soil Biology and Biochemistry 41, 262–275. doi:10.1016/j.soilbio.2008.10.024

Bastida, F., Eldridge, D.J., García, C., Kenny Png, G., Bardgett, R.D., Delgado-Baquerizo, M., 2021. Soil microbial diversity–biomass relationships are driven by soil carbon content across global biomes. The ISME Journal 15, 2081–2091. doi:10.1038/s41396-021-00906-0

Bastida, F., García, C., Fierer, N., Eldridge, D.J., Bowker, M.A., Abades, S., Alfaro, F.D., Asefaw Berhe, A., Cutler, N.A., Gallardo, A., García-Velázquez, L., Hart, S.C., Hayes, P.E., Hernández, T., Hseu, Z.-Y., Jehmlich, N., Kirchmair, M., Lambers, H., Neuhauser, S., Peña-Ramírez, V.M., Pérez, C.A., Reed, S.C., Santos, F., Siebe, C., Sullivan, B.W., Trivedi, P., Vera, A., Williams, M.A., Luis Moreno, J., Delgado-Baquerizo, M., 2019. Global ecological predictors of the soil priming effect. Nature Communications 10, 3481. doi:10.1038/s41467-019-11472-7

Bastida, F., Selevsek, N., Torres, I.F., Hernández, T., García, C., 2015. Soil restoration with organic amendments: linking cellular functionality and ecosystem processes. Scientific Reports 5, 15550. doi:10.1038/srep15550

Batey, T., 2009. Soil compaction and soil management – a review. Soil Use and Management 25, 335–345. doi:10.1111/j.1475-2743.2009.00236.x

Baumann, P., 2020. Simplerspec: Soil and Plant Spectroscopic Model Building and Prediction, r package version 0.1.0.9001.

Baveye, P.C., 2023. Ecosystem-scale modelling of soil carbon dynamics: Time for a radical shift of perspective? Soil Biology and Biochemistry 184, 109112. doi:10.1016/j.soilbio.2023.109112

Baveye, P.C., Otten, W., Kravchenko, A., Balseiro-Romero, M., Beckers, É., Chalhoub, M., Darnault, C., Eickhorst, T., Garnier, P., Hapca, S., Kiranyaz, S., Monga, O., Mueller, C.W., Nunan, N., Pot, V., Schlüter, S., Schmidt, H., Vogel, H.-J., 2018. Emergent Properties of Microbial Activity in Heterogeneous Soil Microenvironments: Different Research Approaches Are Slowly Converging, Yet Major Challenges Remain. Frontiers in Microbiology 9, 1929. doi:10.3389/fmicb.2018.01929

Beare, M.H., McNeill, S.J., Curtin, D., Parfitt, R.L., Jones, H.S., Dodd, M.B., Sharp, J., 2014. Estimating the organic carbon stabilisation capacity and saturation deficit of soils: a New Zealand case study. Biogeochemistry 120, 71–87. doi:10.1007/s10533-014-9982-1

Begill, N., Don, A., Poeplau, C., 2023. No detectable upper limit of mineral-associated organic carbon in temperate agricultural soils. Global Change Biology 29, 4662–4669. doi:10.1111/gcb.16804

Beier, C., Beierkuhnlein, C., Wohlgemuth, T., Penuelas, J., Emmett, B., Körner, C., de Boeck, H., Christensen, J.H., Leuzinger, S., Janssens, I.A., Hansen, K., 2012. Precipitation manipulation

experiments – challenges and recommendations for the future. Ecology Letters 15, 899–911. doi:10.1111/j.1461-0248.2012.01793.x
Berlemont, R., Martiny, A.C., 2015. Genomic Potential for Polysaccharide Deconstruction in Bacteria. Applied and Environmental Microbiology 81, 1513–1519. doi:10.1128/AEM.03718-14
Bernard, L., Basile-Doelsch, I., Derrien, D., Fanin, N., Fontaine, S., Guenet, B., Karimi, B., Marsden, C., Maron, P., 2022. Advancing the mechanistic understanding of the priming effect on soil organic matter mineralisation. Functional Ecology 36, 1355–1377. doi:10.1111/1365-2435.14038
Bernasconi, S.M., Bauder, A., Bourdon, B., Brunner, I., Bünemann, E., Chris, I., Derungs, N., Edwards, P., Farinotti, D., Frey, B., Frossard, E., Furrer, G., Gierga, M., Göransson, H., Gülland, K., Hagedorn, F., Hajdas, I., Hindshaw, R., Ivy-Ochs, S., Jansa, J., Jonas, T., Kiczka, M., Kretzschmar, R., Lemarchand, E., Luster, J., Magnusson, J., Mitchell, E.A.D., Venterink, H.O., Plötze, M., Reynolds, B., Smittenberg, R.H., Stähli, M., Tamburini, F., Tipper, E.T., Wacker, L., Welc, M., Wiederhold, J.G., Zeyer, J., Zimmermann, S., Zumsteg, A., 2011. Chemical and Biological Gradients along the Damma Glacier Soil Chronosequence, Switzerland. Vadose Zone Journal 10, 867–883. doi:10.2136/vzj2010.0129
Bhatnagar, J.M., Peay, K.G., Treseder, K.K., 2018. Litter chemistry influences decomposition through activity of specific microbial functional guilds. Ecological Monographs 88, 429–444. doi:10.1002/ecm.1303
Bier, R.L., Bernhardt, E.S., Boot, C.M., Graham, E.B., Hall, E.K., Lennon, J.T., Nemergut, D.R., Osborne, B.B., Ruiz-González, C., Schimel, J.P., Waldrop, M.P., Wallenstein, M.D., 2015. Linking microbial community structure and microbial processes: an empirical and conceptual overview. FEMS Microbiology Ecology 91, fiv113. doi:10.1093/femsec/fiv113
Blagodatskaya, E., Kuzyakov, Y., 2013. Active microorganisms in soil: Critical review of estimation criteria and approaches. Soil Biology and Biochemistry 67, 192–211. doi:10.1016/j.soilbio.2013.08.024
Blankinship, J.C., Berhe, A.A., Crow, S.E., Druhan, J.L., Heckman, K.A., Keiluweit, M., Lawrence, C.R., Marín-Spiotta, E., Plante, A.F., Rasmussen, C., Schädel, C., Schimel, J.P., Sierra, C.A., Thompson, A., Wagai, R., Wieder, W.R., 2018. Improving understanding of soil organic matter dynamics by triangulating theories, measurements, and models. Biogeochemistry 140, 1–13. doi:10.1007/s10533-018-0478-2
Bolan, N., Srivastava, P., Rao, C.S., Satyanaraya, P.V., Anderson, G.C., Bolan, S., Nortjé, G.P., Kronenberg, R., Bardhan, S., Abbott, L.K., Zhao, H., Mehra, P., Satyanarayana, S.V., Khan, N., Wang, H., Rinklebe, J., Siddique, K.H.M., Kirkham, M.B., 2023. Distribution, characteristics and management of calcareous soils, in: Advances in Agronomy. Elsevier, pp. 81–130. doi:10.1016/bs.agron.2023.06.002
Bond-Lamberty, B., Bailey, V.L., Chen, M., Gough, C.M., Vargas, R., 2018. Globally rising soil heterotrophic respiration over recent decades. Nature 560, 80–83. doi:10.1038/s41586-018-0358-x
Bond-Lamberty, B., Ballantyne, A., Berryman, E., Fluet-Chouinard, E., Jian, J., Morris, K.A., Rey, A., Vargas, R., 2024. Twenty Years of Progress, Challenges, and Opportunities in Measuring and Understanding Soil Respiration. Journal of Geophysical Research: Biogeosciences 129, e2023JG007637. doi:10.1029/2023JG007637
Bonk, F., Popp, D., Harms, H., Centler, F., 2018. PCR-based quantification of taxa-specific abundances in microbial communities: Quantifying and avoiding common pitfalls. Journal of Microbiological Methods 153, 139–147. doi:10.1016/j.mimet.2018.09.015
Bradford, M.A., 2013. Thermal adaptation of decomposer communities in warming soils. Frontiers in Microbiology 4. doi:10.3389/fmicb.2013.00333
Bradford, M.A., McCulley, R.L., Crowther, Thomas.W., Oldfield, E.E., Wood, S.A., Fierer, N., 2019. Cross-biome patterns in soil microbial respiration predictable from evolutionary theory on thermal adaptation. Nature Ecology & Evolution 3, 223–231. doi:10.1038/s41559-018-0771-4

Bradford, M.A., Veen, G.F., Bonis, A., Bradford, E.M., Classen, A.T., Cornelissen, J.H.C., Crowther, Thomas.W., De Long, J.R., Freschet, G.T., Kardol, P., Manrubia-Freixa, M., Maynard, D.S., Newman, G.S., Logtestijn, R.S.P., Viketoft, M., Wardle, D.A., Wieder, W.R., Wood, S.A., van der Putten, W.H., 2017. A test of the hierarchical model of litter decomposition. Nature Ecology & Evolution 1, 1836–1845. doi:10.1038/s41559-017-0367-4

Bryant, C., Wheeler, N.R., Rubel, F., French, R.H., 2017. Koeppen-Geiger Climatic Zones.

Buchkowski, R.W., Bradford, M.A., Grandy, A.S., Schmitz, O.J., Wieder, W.R., 2017. Applying population and community ecology theory to advance understanding of belowground biogeochemistry. Ecology Letters 20, 231–245. doi:10.1111/ele.12712

Büks, F., 2023. Technical note: The recovery rate of free particulate organic matter from soil samples is strongly affected by the method of density fractionation. Biogeosciences 20, 1529–1535. doi:10.5194/bg-20-1529-2023

Button, E.S., Pett-Ridge, J., Murphy, D.V., Kuzyakov, Y., Chadwick, D.R., Jones, D.L., 2022. Deep-C storage: Biological, chemical and physical strategies to enhance carbon stocks in agricultural subsoils. Soil Biology and Biochemistry 170, 108697. doi:10.1016/j.soilbio.2022.108697

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13, 581–583. doi:10.1038/nmeth.3869

Cameron, E.K., Martins, I.S., Lavelle, P., Mathieu, J., Tedersoo, L., Gottschall, F., Guerra, C.A., Hines, J., Patoine, G., Siebert, J., Winter, M., Cesarz, S., Delgado-Baquerizo, M., Ferlian, O., Fierer, N., Kreft, H., Lovejoy, T.E., Montanarella, L., Orgiazzi, A., Pereira, H.M., Phillips, H.R.P., Settele, J., Wall, D.H., Eisenhauer, N., 2018. Global gaps in soil biodiversity data. Nature Ecology & Evolution 2, 1042–1043. doi:10.1038/s41559-018-0573-8

Cameron, E.S., Schmidt, P.J., Tremblay, B.J.-M., Emelko, M.B., Müller, K.M., 2021. Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. Scientific Reports 11, 22302. doi:10.1038/s41598-021-01636-1

Camillone, N.R., Bruns, M.A., Román, R., Wasner, D., Couradeau, E., 2023. Positive relationship between substrate-induced respiration rate and translationally active bacteria count in soil (preprint). Ecology. doi:10.1101/2023.12.20.572626

Camillone, N.R., Bruns, M.A., Román, R., Wasner, D., Couradeau, E., preprint. Positive relationship between substrate-induced respiration rate and translationally active bacteria count in soil. BioRxiv 2023.12.20.572626. doi:10.1101/2023.12.20.572626

Campbell, E.E., Paustian, K., 2015. Current developments in soil organic matter modeling and the expansion of model applications: a review. Environmental Research Letters 10, 123004. doi:10.1088/1748-9326/10/12/123004

Canadell, J.G., 2021. Chapter 5: Global Carbon and other Biogeochemical Cycles and Feedbacks, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Eds.)]. Cambridge University Press.

Canarini, A., Schmidt, H., Fuchslueger, L., Martin, V., Herbold, C.W., Zezula, D., Gündler, P., Hasibeder, R., Jecmenica, M., Bahn, M., Richter, A., 2021. Ecological memory of recurrent drought modifies soil processes via changes in soil microbial community. Nature Communications 12, 5308. doi:10.1038/s41467-021-25675-4

Carbone, M.S., Seyednasrollah, B., Rademacher, T.T., Basler, D., Le Moine, J.M., Beals, S., Beasley, J., Greene, A., Kelroy, J., Richardson, A.D., 2019. Flux Puppy – An open-source software application and portable system design for low-cost manual measurements of CO2 and H2O fluxes. Agricultural and Forest Meteorology 274, 1–6. doi:10.1016/j.agrformet.2019.04.012

Carini, P., Marsden, P.J., Leff, J.W., Morgan, E.E., Strickland, M.S., Fierer, N., 2016. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. Nature Microbiology 2, 16242. doi:10.1038/nmicrobiol.2016.242

Caro, T.A., McFarlin, J., Jech, S., Fierer, N., Kopf, S., 2023. Hydrogen stable isotope probing of lipids demonstrates slow rates of microbial growth in soil. Proceedings of the National Academy of Sciences 120, e2211625120. doi:10.1073/pnas.2211625120

Castañeda-Gómez, L., Lajtha, K., Bowden, R., Mohammed Jauhar, F.N., Jia, J., Feng, X., Simpson, M.J., 2023. Soil organic matter molecular composition with long-term detrital alterations is controlled by site-specific forest properties. Global Change Biology 29, 243–259. doi:10.1111/gcb.16456

Cécillon, L., Baudin, F., Chenu, C., Christensen, B.T., Franko, U., Houot, S., Kanari, E., Kätterer, T., Merbach, I., van Oort, F., Poeplau, C., Quezada, J.C., Savignac, F., Soucémarianadin, L.N., Barré, P., 2021. Partitioning soil organic carbon into its centennially stable and active fractions with machine-learning models based on Rock-Eval® thermal analysis (PARTY<sub>SOC</sub>v2.0 and PARTY<sub>SOC</sub>v2.0<sub>EU</sub>). Geoscientific Model Development 14, 3879–3898. doi:10.5194/gmd-14-3879-2021

Cederlund, H., Wessén, E., Enwall, K., Jones, C.M., Juhanson, J., Pell, M., Philippot, L., Hallin, S., 2014. Soil carbon quality and nitrogen fertilization structure bacterial communities with predictable responses of major bacterial phyla. Applied Soil Ecology 84, 62–68. doi:10.1016/j.apsoil.2014.06.003

Cerli, C., Celi, L., Kalbitz, K., Guggenberger, G., Kaiser, K., 2012. Separation of light and heavy organic matter fractions in soil — Testing for proper density cut-off and dispersion level. Geoderma 170, 403–416. doi:10.1016/j.geoderma.2011.10.009

Chadwick, O.A., Chorover, J., 2001. The chemistry of pedogenic thresholds. Geoderma 100, 321–353. doi:10.1016/S0016-7061(01)00027-1

Champely, S., 2020. pwr: Basic Functions for Power Analysis. R package version 1.3-0.

Chandel, A.K., Jiang, L., Luo, Y., 2023. Microbial Models for Simulating Soil Carbon Dynamics: A Review. Journal of Geophysical Research: Biogeosciences 128, e2023JG007436. doi:10.1029/2023JG007436

Chen, L., Fang, K., Wei, B., Qin, S., Feng, X., Hu, T., Ji, C., Yang, Y., 2021. Soil carbon persistence governed by plant input and mineral protection at regional and global scales. Ecology Letters 24, 1018–1028. doi:10.1111/ele.13723

Chen, Q.-L., Ding, J., Li, C.-Y., Yan, Z.-Z., He, J.-Z., Hu, H.-W., 2020. Microbial functional attributes, rather than taxonomic attributes, drive top soil respiration, nitrification and denitrification processes. Science of The Total Environment 734, 139479. doi:10.1016/j.scitotenv.2020.139479

Chen, Y., Neilson, J.W., Kushwaha, P., Maier, R.M., Barberán, A., 2021. Life-history strategies of soil microbial communities in an arid ecosystem. The ISME Journal 15, 649–657. doi:10.1038/s41396-020-00803-y

Claret, F., Schäfer, T., Bauer, A., Buckau, G., 2003. Generation of humic and fulvic acid from Callovo-Oxfordian clay under high alkaline conditions. Science of The Total Environment 317, 189–200. doi:10.1016/S0048-9697(03)00337-1

Cleveland, C.C., Nemergut, D.R., Schmidt, S.K., Townsend, A.R., 2007. Increases in soil respiration following labile carbon additions linked to rapid shifts in soil microbial community composition. Biogeochemistry 82, 229–240. doi:10.1007/s10533-006-9065-z

Colman, B.P., Schimel, J.P., 2013. Drivers of microbial respiration and net N mineralization at the continental scale. Soil Biology and Biochemistry 60, 65–76. doi:10.1016/j.soilbio.2013.01.003

Cordero, O.X., Datta, M.S., 2016. Microbial interactions and community assembly at microscales. Current Opinion in Microbiology 31, 227–234. doi:10.1016/j.mib.2016.03.015

Corstanje, R., Kirk, G.J.D., Pawlett, M., Read, R., Lark, R.M., 2008. Spatial variation of ammonia volatilization from soil and its scale-dependent correlation with soil properties. European Journal of Soil Science 59, 1260–1270. doi:10.1111/j.1365-2389.2008.01087.x

Costa, O.Y.A., Raaijmakers, J.M., Kuramae, E.E., 2018. Microbial Extracellular Polymeric Substances: Ecological Function and Impact on Soil Aggregation. Frontiers in Microbiology 9, 1636. doi:10.3389/fmicb.2018.01636

Cotrufo, M.F., Haddix, M.L., Kroeger, M.E., Stewart, C.E., 2022. The role of plant input physical-chemical properties, and microbial and soil chemical diversity on the formation of particulate and mineral-associated organic matter. Soil Biology and Biochemistry 168, 108648. doi:10.1016/j.soilbio.2022.108648

Cotrufo, M.F., Lavallee, J.M., Zhang, Y., Hansen, P.M., Paustian, K.H., Schipanski, M., Wallenstein, M.D., 2021. In-N-Out: A hierarchical framework to understand and predict soil carbon storage and nitrogen recycling. doi:10.1111/gcb.15782

Cotrufo, M.F., Ranalli, M.G., Haddix, M.L., Six, J., Lugato, E., 2019. Soil carbon storage informed by particulate and mineral-associated organic matter. Nature Geoscience 12, 989–994. doi:10.1038/s41561-019-0484-6

Cotrufo, M.F., Soong, J.L., Horton, A.J., Campbell, E.E., Haddix, M.L., Wall, D.H., Parton, W.J., 2015. Formation of soil organic matter via biochemical and physical pathways of litter mass loss. Nature Geoscience 8, 776–779. doi:10.1038/ngeo2520

Couradeau, E., Sasse, J., Goudeau, D., Nath, N., Hazen, T.C., Bowen, B.P., Chakraborty, R., Malmstrom, R.R., Northen, T.R., 2019. Probing the active fraction of soil microbiomes using BONCAT-FACS. Nature Communications 10, 2770. doi:10.1038/s41467-019-10542-0

Courchesne, F., Turmel, M.-C., 2007. Chapter 26: Extractable Al, Fe, Mn and Si., in: Carter, M.R., & Gregorich, E.G. (Eds.). Soil Sampling and Methods of Analysis. CRC Press.

Craine, J.M., Fierer, N., McLauchlan, K.K., 2010. Widespread coupling between the rate and temperature sensitivity of organic matter decay. Nature Geoscience 3, 854–857. doi:10.1038/ngeo1009

Creamer, C.A., Foster, A.L., Lawrence, C., McFarland, J., Schulz, M., Waldrop, M.P., 2019. Mineralogy dictates the initial mechanism of microbial necromass association. Geochimica et Cosmochimica Acta 260, 161–176. doi:10.1016/j.gca.2019.06.028

Crowther, T.W., van den Hoogen, J., Wan, J., Mayes, M.A., Keiser, A.D., Mo, L., Averill, C., Maynard, D.S., 2019. The global soil community and its influence on biogeochemistry. Science 365, eaav0550. doi:10.1126/science.aav0550

Cruz-Martínez, K., Rosling, A., Zhang, Y., Song, M., Andersen, G.L., Banfield, J.F., 2012. Effect of Rainfall-Induced Soil Geochemistry Dynamics on Grassland Soil Microbial Communities. Applied and Environmental Microbiology 78, 7587–7595. doi:10.1128/AEM.00203-12

Dal Bello, M., Lee, H., Goyal, A., Gore, J., 2021. Resource–diversity relationships in bacterial communities reflect the network structure of microbial metabolism. Nature Ecology & Evolution 5, 1424–1434. doi:10.1038/s41559-021-01535-8

Dastager, S.G., Krishnamurthi, S., Rameshkumar, N., Dharne, M., 2014. The Family Micrococcaceae, in: The Prokaryotes: Actinobacteria. Fourth Edition. (Eds: Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E., Thompson, F.). Springer.

Davenport, R., Bowen, B.P., Lynch, L.M., Kosina, S.M., Shabtai, I., Northen, T.R., Lehmann, J., 2023. Decomposition decreases molecular diversity and ecosystem similarity of soil organic matter. Proceedings of the National Academy of Sciences 120, e2303335120. doi:10.1073/pnas.2303335120

Davidson, E.A., Savage, K.E., Finzi, A.C., 2014. A big-microsite framework for soil carbon modeling. Global Change Biology 20, 3610–3620. doi:10.1111/gcb.12718

De Boeck, H.J., Vicca, S., Roy, J., Nijs, I., Milcu, A., Kreyling, J., Jentsch, A., Chabbi, A., Campioli, M., Callaghan, T., Beierkuhnlein, C., Beier, C., 2015. Global Change Experiments: Challenges and Opportunities. BioScience 65, 922–931. doi:10.1093/biosci/biv099

De Jonge, C., Guo, J., Hällberg, P., Griepentrog, M., Rifai, H., Richter, A., Ramirez, E., Zhang, X., Smittenberg, R.H., Peterse, F., Boeckx, P., Dercon, G., 2024. The impact of soil chemistry, moisture and temperature on branched and isoprenoid GDGTs in soils: A study using six

globally distributed elevation transects. Organic Geochemistry 187, 104706. doi:10.1016/j.orggeochem.2023.104706
Delgado-Baquerizo, M., Grinyer, J., Reich, P.B., Singh, B.K., 2016a. Relative importance of soil properties and microbial community for soil functionality: insights from a microbial swap experiment. Functional Ecology 30, 1862–1873. doi:10.1111/1365-2435.12674
Delgado-Baquerizo, M., Maestre, F.T., Reich, P.B., Jeffries, T.C., Gaitan, J.J., Encinar, D., Berdugo, M., Campbell, C.D., Singh, B.K., 2016. Microbial diversity drives multifunctionality in terrestrial ecosystems. Nature Communications 7, 10541. doi:10.1038/ncomms10541
Delgado-Baquerizo, M., Maestre, F.T., Reich, P.B., Trivedi, P., Osanai, Y., Liu, Y., Hamonts, K., Jeffries, T.C., Singh, B.K., 2016b. Carbon content and climate variability drive global soil bacterial diversity patterns. Ecological Monographs 86, 373–390. doi:10.1002/ecm.1216
Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D.J., Bardgett, R.D., Maestre, F.T., Singh, B.K., Fierer, N., 2018. A global atlas of the dominant bacteria found in soil. Science 359, 320–325. doi:10.1126/science.aap9516
Delgado-Baquerizo, M., Reich, P.B., Bardgett, R.D., Eldridge, D.J., Lambers, H., Wardle, D.A., Reed, S.C., Plaza, C., Png, G.K., Neuhauser, S., Berhe, A.A., Hart, S.C., Hu, H.-W., He, J.-Z., Bastida, F., Abades, S., Alfaro, F.D., Cutler, N.A., Gallardo, A., García-Velázquez, L., Hayes, P.E., Hseu, Z.-Y., Pérez, C.A., Santos, F., Siebe, C., Trivedi, P., Sullivan, B.W., Weber-Grullon, L., Williams, M.A., Fierer, N., 2020. The influence of soil age on ecosystem structure and function across biomes. Nature Communications 11, 4721. doi:10.1038/s41467-020-18451-3
Delgado-Baquerizo, M., Reich, P.B., Khachane, A.N., Campbell, C.D., Thomas, N., Freitag, T.E., Abu Al-Soud, W., Sørensen, S., Bardgett, R.D., Singh, B.K., 2017. It is elemental: soil nutrient stoichiometry drives bacterial diversity. Environmental Microbiology 19, 1176–1188. doi:10.1111/1462-2920.13642
Demyan, M.S., Rasche, F., Schulz, E., Breulmann, M., Müller, T., Cadisch, G., 2012. Use of specific peaks obtained by diffuse reflectance Fourier transform mid-infrared spectroscopy to study the composition of organic matter in a Haplic Chernozem. European Journal of Soil Science 63, 189–199. doi:10.1111/j.1365-2389.2011.01420.x
De Rosa, D., Ballabio, C., Lugato, E., Fasiolo, M., Jones, A., Panagos, P., 2024. Soil organic carbon stocks in European croplands and grasslands: How much have we lost in the past decade? Global Change Biology 30, e16992. doi:10.1111/gcb.16992
Diamond, S., Andeer, P.F., Li, Z., Crits-Christoph, A., Burstein, D., Anantharaman, K., Lane, K.R., Thomas, B.C., Pan, C., Northen, T.R., Banfield, J.F., 2019. Mediterranean grassland soil C–N compound turnover is dependent on rainfall and depth, and is mediated by genomically divergent microorganisms. Nature Microbiology 4, 1356–1367. doi:10.1038/s41564-019-0449-y
Dijkstra, P., Ishizu, A., Doucett, R., Hart, S.C., Schwartz, E., Menyailo, O.V., Hungate, B.A., 2006. 13C and 15N natural abundance of the soil microbial biomass. Soil Biology and Biochemistry 38, 3257–3266. doi:10.1016/j.soilbio.2006.04.005
Dilling, J., Kaiser, K., 2002. Estimation of the hydrophobic fraction of dissolved organic matter in water samples using UV photometry. Water Research 36, 5037–5044. doi:10.1016/S0043-1354(02)00365-2
Ding, Y., Leppälammi-Kujansuu, J., Salemaa, M., Schiestl-Aalto, P., Kulmala, L., Ukonmaanaho, L., Nöjd, P., Minkkinen, K., Makita, N., Železnik, P., Merilä, P., Helmisaari, H.-S., 2021. Distinct patterns of below- and aboveground growth phenology and litter carbon inputs along a boreal site type gradient. Forest Ecology and Management 489, 119081. doi:10.1016/j.foreco.2021.119081
Doetterl, S., Asifiwe, R.K., Baert, G., Bamba, F., Bauters, M., Boeckx, P., Bukombe, B., Cadisch, G., Cooper, M., Cizungu, L.N., Hoyt, A., Kabaseke, C., Kalbitz, K., Kidinda, L., Maier, A., Mainka, M., Mayrock, J., Muhindo, D., Mujinya, B.B., Mukotanyi, S.M., Nabahungu, L., Reichenbach, M., Rewald, B., Six, J., Stegmann, A., Summerauer, L., Unseld, R., Vanlauwe, B., Van Oost, K., Verheyen, K., Vogel, C., Wilken, F., Fiener, P., 2021. Organic matter cycling along

geochemical, geomorphic, and disturbance gradients in forest and cropland of the African Tropics – project TropSOC database version 1.0. Earth System Science Data 13, 4133–4153. doi:10.5194/essd-13-4133-2021

Doetterl, S., Berhe, A.A., Arnold, C., Bodé, S., Fiener, P., Finke, P., Fuchslueger, L., Griepentrog, M., Harden, J.W., Nadeu, E., Schnecker, J., Six, J., Trumbore, S., Van Oost, K., Vogel, C., Boeckx, P., 2018. Links among warming, carbon and microbial dynamics mediated by soil mineral weathering. Nature Geoscience 11, 589–593. doi:10.1038/s41561-018-0168-7

Doetterl, S., Berhe, A.A., Nadeu, E., Wang, Z., Sommer, M., Fiener, P., 2016. Erosion, deposition and soil carbon: A review of process-level controls, experimental tools and models to address C cycling in dynamic landscapes. Earth-Science Reviews 154, 102–122. doi:10.1016/j.earscirev.2015.12.005

Doetterl, S., Cornelis, J.-T., Six, J., Bodé, S., Opfergelt, S., Boeckx, P., Van Oost, K., 2015. Soil redistribution and weathering controlling the fate of geochemical and physical carbon stabilization mechanisms in soils of an eroding landscape. Biogeosciences 12, 1357–1371. doi:10.5194/bg-12-1357-2015

Doetterl, Sebastian, Stevens, A., Six, J., Merckx, R., Van Oost, K., Casanova Pinto, M., Casanova-Katny, A., Muñoz, C., Boudin, M., Zagal Venegas, E., Boeckx, P., 2015. Soil carbon storage controlled by interactions between geochemistry and climate. Nature Geoscience 8, 780–783. doi:10.1038/ngeo2516

Dolan, K.L., Peña, J., Allison, S.D., Martiny, J.B.H., 2017. Phylogenetic conservation of substrate use specialization in leaf litter bacteria. PLOS ONE 12, e0174472. doi:10.1371/journal.pone.0174472

Domeignoz-Horta, L.A., Pold, G., Liu, X.-J.A., Frey, S.D., Melillo, J.M., DeAngelis, K.M., 2020. Microbial diversity drives carbon use efficiency in a model soil. Nature Communications 11, 3684. doi:10.1038/s41467-020-17502-z

Don, A., Böhme, I.H., Dohrmann, A.B., Poeplau, C., Tebbe, C.C., 2017. Microbial community composition affects soil organic carbon turnover in mineral soils. Biology and Fertility of Soils 53, 445–456. doi:10.1007/s00374-017-1198-9

Don, A., Rödenbeck, C., Gleixner, G., 2013. Unexpected control of soil carbon turnover by soil carbon concentration. Environmental Chemistry Letters 11, 407–413. doi:10.1007/s10311-013-0433-3

Duan, Y.F., Grogan, P., Walker, V.K., diCenzo, G.C., 2022. Whole genome sequencing of mesorhizobia isolated from northern Canada. Canadian Journal of Microbiology 68, 661–673. doi:10.1139/cjm-2022-0102

Dungait, J.A.J., Hopkins, D.W., Gregory, A.S., Whitmore, A.P., 2012. Soil organic matter turnover is governed by accessibility not recalcitrance. Global Change Biology 18, 1781–1796. doi:10.1111/j.1365-2486.2012.02665.x

Edlinger, A., Garland, G., Banerjee, S., Degrune, F., García-Palacios, P., Herzog, C., Pescador, D.S., Romdhane, S., Ryo, M., Saghaï, A., Hallin, S., Maestre, F.T., Philippot, L., Rillig, M.C., van der Heijden, M.G.A., 2023. The impact of agricultural management on soil aggregation and carbon storage is regulated by climatic thresholds across a 3000 km European gradient. Global Change Biology 29, 3177–3192. doi:10.1111/gcb.16677

Edwards, A.P., Bremner, J.M., 1967. MICROAGGREGATES IN SOILS [1]. Journal of Soil Science 18, 64–73. doi:10.1111/j.1365-2389.1967.tb01488.x

Eilers, K.G., Lauber, C.L., Knight, R., Fierer, N., 2010. Shifts in bacterial community structure associated with inputs of low molecular weight carbon compounds to soil. Soil Biology and Biochemistry 42, 896–903. doi:10.1016/j.soilbio.2010.02.003

Encyclopaedia Britannica, G., 2024. Britannica.

Eng, A.Y., Narayanan, A., Alster, C.J., DeAngelis, K.M., 2023. Thermal adaptation of soil microbial growth traits in response to chronic warming. Applied and Environmental Microbiology 89, e00825-23. doi:10.1128/aem.00825-23

Eysenck, H.J., 1994. Systematic Reviews: Meta-analysis and its problems. BMJ 309, 789–792. doi:10.1136/bmj.309.6957.789
Falconer, R.E., Battaia, G., Schmidt, S., Baveye, P., Chenu, C., Otten, W., 2015. Microscale Heterogeneity Explains Experimental Variability and Non-Linearity in Soil Organic Matter Mineralisation. PLOS ONE 10, e0123774. doi:10.1371/journal.pone.0123774
FAO, 2021. Land use statistics and indicators statistics. Global, regional and country trends 1990–2019. FAOSTAT Analytical Brief Series No 28. Rome.
FAO-UNESCO, 2018. Soil Map of the World, digitized by ESRI. Soil climate map, USDA-NRCS, Soil Science Division, World Soil Resources, Washington D.C.
Fellman, J.B., D'Amore, D.V., Hood, E., Boone, R.D., 2008. Fluorescence characteristics and biodegradability of dissolved organic matter in forest and wetland soils from coastal temperate watersheds in southeast Alaska. Biogeochemistry 88, 169–184. doi:10.1007/s10533-008-9203-x
Fernandez, I., Mahieu, N., Cadisch, G., 2003. Carbon isotopic fractionation during decomposition of plant materials of different quality. Global Biogeochemical Cycles 17, n/a-n/a. doi:10.1029/2001GB001834
Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37, 4302–4315. doi:10.1002/joc.5086
Fierer, N., 2017. Embracing the unknown: disentangling the complexities of the soil microbiome. Nature Reviews Microbiology 15, 579–590. doi:10.1038/nrmicro.2017.87
Fierer, N., Bradford, M.A., Jackson, R.B., 2007. TOWARD AN ECOLOGICAL CLASSIFICATION OF SOIL BACTERIA. Ecology 88, 1354–1364. doi:10.1890/05-1839
Finke, P., Opolot, E., Balesdent, J., Berhe, A.A., Boeckx, P., Cornu, S., Harden, J., Hatté, C., Williams, E., Doetterl, S., 2019. Can SOC modelling be improved by accounting for pedogenesis? Geoderma 338, 513–524. doi:10.1016/j.geoderma.2018.10.018
Fissore, C., Dalzell, B.J., Berhe, A.A., Voegtle, M., Evans, M., Wu, A., 2017. Influence of topography on soil organic carbon dynamics in a Southern California grassland. CATENA 149, 140–149. doi:10.1016/j.catena.2016.09.016
Flemming, H.-C., 2016. EPS—Then and Now. Microorganisms 4, 41. doi:10.3390/microorganisms4040041
Franzluebbers, A.J., Haney, R.L., Hons, F.M., Zuberer, D.A., 1996. Active fractions of organic matter in soils with different texture. Soil Biology and Biochemistry 28, 1367–1372. doi:10.1016/S0038-0717(96)00143-5
Freeman, C., Ostle, N., Kang, H., 2001. An enzymic "latch" on a global carbon store. Nature 409, 149–149. doi:10.1038/35051650
Frey, B., Rime, T., Phillips, M., Stierli, B., Hajdas, I., Widmer, F., Hartmann, M., 2016. Microbial diversity in European alpine permafrost and active layers. FEMS Microbiology Ecology 92, fiw018. doi:10.1093/femsec/fiw018
Friedlingstein, P., O'Sullivan, M., Jones, M.W., Andrew, R.M., Bakker, D.C.E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I.T., Peters, G.P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J.G., Ciais, P., Jackson, R.B., Alin, S.R., Anthoni, P., Barbero, L., Bates, N.R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I.B.M., Cadule, P., Chamberlain, M.A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L.P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R.A., Feng, L., Ford, D.J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Houghton, R.A., Hurtt, G.C., Iida, Y., Ilyina, T., Jacobson, A.R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R.F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J.I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P.C., McKinley, G.A., Meyer, G., Morgan, E.J., Munro, D.R., Nakaoka, S.-I., Niwa, Y., O'Brien, K.M., Olsen, A., Omar, A.M., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Powis, C.M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T.M., Schwinger, J., Séférian, R., Smallman, T.L., Smith, S.M., Sospedra-Alfonso, R., Sun, Q., Sutton, A.J., Sweeney, C., Takao,

S., Tans, P.P., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G.R., van Ooijen, E., Wanninkhof, R., Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., Zheng, B., 2023. Global Carbon Budget 2023. Earth System Science Data 15, 5301–5369. doi:10.5194/essd-15-5301-2023

Friedman, J., Tibshirani, R., Hastie, T., 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software 33, 1–22. doi:https://doi.org/10.18637/jss.v033.i01

Fromin, N., Saby, N.P.A., Lensi, R., Brunet, D., Porte, B., Domenach, A.-M., Roggy, J.-C., 2013. Spatial variability of soil microbial functioning in a tropical rainforest of French Guiana using nested sampling. Geoderma 197–198, 98–107. doi:10.1016/j.geoderma.2012.12.009

Gao, D., Bai, E., Wang, S., Zong, S., Liu, Z., Fan, X., Zhao, C., Hagedorn, F., 2022. Three-dimensional mapping of carbon, nitrogen, and phosphorus in soil microbial biomass and their stoichiometry at the global scale. Global Change Biology 28, 6728–6740. doi:10.1111/gcb.16374

Gao, D., Bai, E., Wasner, D., Hagedorn, F., 2024. Global prediction of soil microbial growth rates and carbon use efficiency based on the metabolic theory of ecology. Soil Biology and Biochemistry 190, 109315. doi:10.1016/j.soilbio.2024.109315

García-Palacios, P., Bradford, M.A., Benavente-Ferraces, I., de Celis, M., Delgado-Baquerizo, M., García-Gil, J.C., Gaitán, J.J., Goñi-Urtiaga, A., Mueller, C.W., Panettieri, M., Rey, A., Sáez-Sandino, T., Schuur, E.A.G., Sokol, N.W., Tedersoo, L., Plaza, C., 2024. Dominance of particulate organic carbon in top mineral soils in cold regions. Nature Geoscience 17, 145–150. doi:10.1038/s41561-023-01354-5

García-Palacios, P., Crowther, T.W., Dacal, M., Hartley, I.P., Reinsch, S., Rinnan, R., Rousk, J., van den Hoogen, J., Ye, J.-S., Bradford, M.A., 2021. Evidence for large microbial-mediated losses of soil carbon under anthropogenic warming. Nature Reviews Earth & Environment 2, 507–517. doi:10.1038/s43017-021-00178-4

Garrido, E., Matus, F., 2012. Are organo-mineral complexes and allophane content determinant factors for the carbon level in Chilean volcanic soils? CATENA 92, 106–112. doi:10.1016/j.catena.2011.12.003

Garsia, A., Moinet, A., Vazquez, C., Creamer, R.E., Moinet, G.Y.K., 2023. The challenge of selecting an appropriate soil organic carbon simulation model: A comprehensive global review and validation assessment. Global Change Biology 29, 5760–5774. doi:10.1111/gcb.16896

Georgiou, K., Abramoff, R.Z., Harte, J., Riley, W.J., Torn, M.S., 2017. Microbial community-level regulation explains soil carbon responses to long-term litter manipulations. Nature Communications 8, 1223. doi:10.1038/s41467-017-01116-z

Georgiou, K., Jackson, R.B., Vindušková, O., Abramoff, R.Z., Ahlström, A., Feng, W., Harden, J.W., Pellegrini, A.F.A., Polley, H.W., Soong, J.L., Riley, W.J., Torn, M.S., 2022. Global stocks and capacity of mineral-associated soil organic carbon. Nature Communications 13, 3797. doi:10.1038/s41467-022-31540-9

Georgiou, K., Koven, C.D., Wieder, W.R., Hartman, M.D., Riley, W.J., Pett-Ridge, J., Bouskill, N.J., Abramoff, R.Z., Slessarev, E.W., Ahlström, A., Parton, W.J., Pellegrini, A.F.A., Pierson, D., Sulman, B.N., Zhu, Q., Jackson, R.B., 2024. Emergent temperature sensitivity of soil organic carbon driven by mineral associations. Nature Geoscience 17, 205–212. doi:10.1038/s41561-024-01384-7

Georgiou, K., Malhotra, A., Wieder, W.R., Ennis, J.H., Hartman, M.D., Sulman, B.N., Berhe, A.A., Grandy, A.S., Kyker-Snowman, E., Lajtha, K., Moore, J.A.M., Pierson, D., Jackson, R.B., 2021. Divergent controls of soil organic carbon between observations and process-based models. Biogeochemistry 156, 5–17. doi:10.1007/s10533-021-00819-2

German, D.P., Marcelo, K.R.B., Stone, M.M., Allison, S.D., 2012. The Michaelis-Menten kinetics of soil extracellular enzymes in response to temperature: a cross-latitudinal study. Global Change Biology 18, 1468–1479. doi:10.1111/j.1365-2486.2011.02615.x

Getz, W.M., Marshall, C.R., Carlson, C.J., Giuggioli, L., Ryan, S.J., Romañach, S.S., Boettiger, C., Chamberlain, S.D., Larsen, L., D'Odorico, P., O'Sullivan, D., 2018. Making ecological models adequate. Ecology Letters 21, 153–166. doi:10.1111/ele.12893

Geyer, K.M., Barrett, J.E., 2019. Unimodal productivity–diversity relationships among bacterial communities in a simple polar soil ecosystem. Environmental Microbiology 21, 2523–2532. doi:10.1111/1462-2920.14639

Geyer, K.M., Dijkstra, P., Sinsabaugh, R., Frey, S.D., 2019. Clarifying the interpretation of carbon use efficiency in soil through methods comparison. Soil Biology and Biochemistry 128, 79–88. doi:10.1016/j.soilbio.2018.09.036

Gies, H., Hagedorn, F., Lupker, M., Montluçon, D., Haghipour, N., van der Voort, T.S., Eglinton, T.I., 2021. Millennial-age glycerol dialkyl glycerol tetraethers (GDGTs) in forested mineral soils: <sup>14</sup>C-based evidence for stabilization of microbial necromass. Biogeosciences 18, 189–205. doi:10.5194/bg-18-189-2021

Golchin, A., Baldock, J.A., Oades, J.M., 1997. A Model Linking Organic Matter Decomposition, Chemistry, and Aggregate Dynamics, in: Lal, R., Kimble, J.M., Follett, R.F., Stewart, B.A. (Eds.), Soil Processes and the Carbon Cycle. CRC Press, pp. 245–266. doi:10.1201/9780203739273-17

Goldfarb, K.C., Karaoz, U., Hanson, C.A., Santee, C.A., Bradford, M.A., Treseder, K.K., Wallenstein, M.D., Brodie, E.L., 2011. Differential Growth Responses of Soil Bacterial Taxa to Carbon Substrates of Varying Chemical Recalcitrance. Frontiers in Microbiology 2. doi:10.3389/fmicb.2011.00094

Gommers, P.J.F., van Schie, B.J., van Dijken, J.P., Kuenen, J.G., 1988. Biochemical limits to microbial growth yields: An analysis of mixed substrate utilization. Biotechnology and Bioengineering 32, 86–94. doi:10.1002/bit.260320112

González-Domínguez, B., Niklaus, P.A., Studer, M.S., Hagedorn, F., Wacker, L., Haghipour, N., Zimmermann, S., Walthert, L., McIntyre, C., Abiven, S., 2019. Temperature and moisture are minor drivers of regional-scale soil organic carbon dynamics. Scientific Reports 9, 6422. doi:10.1038/s41598-019-42629-5

Graham, E.B., Knelman, J.E., Schindlbacher, A., Siciliano, S., Breulmann, M., Yannarell, A., Beman, J.M., Abell, G., Philippot, L., Prosser, J., Foulquier, A., Yuste, J.C., Glanville, H.C., Jones, D.L., Angel, R., Salminen, J., Newton, R.J., Bürgmann, H., Ingram, L.J., Hamer, U., Siljanen, H.M.P., Peltoniemi, K., Potthast, K., Bañeras, L., Hartmann, M., Banerjee, S., Yu, R.-Q., Nogaro, G., Richter, A., Koranda, M., Castle, S.C., Goberna, M., Song, B., Chatterjee, A., Nunes, O.C., Lopes, A.R., Cao, Y., Kaisermann, A., Hallin, S., Strickland, M.S., Garcia-Pausas, J., Barba, J., Kang, H., Isobe, K., Papaspyrou, S., Pastorelli, R., Lagomarsino, A., Lindström, E.S., Basiliko, N., Nemergut, D.R., 2016. Microbes as Engines of Ecosystem Function: When Does Community Structure Enhance Predictions of Ecosystem Processes? Frontiers in Microbiology 7. doi:10.3389/fmicb.2016.00214

Grandy, A.S., Neff, J.C., 2008. Molecular C dynamics downstream: The biochemical decomposition sequence and its impact on soil organic matter structure and function. Science of The Total Environment 404, 297–307. doi:10.1016/j.scitotenv.2007.11.013

Grant, K.E., Galy, V.V., Haghipour, N., Eglinton, T.I., Derry, L.A., 2022. Persistence of old soil carbon under changing climate: The role of mineral-organic matter interactions. Chemical Geology 587, 120629. doi:10.1016/j.chemgeo.2021.120629

Greco, T., Zangrillo, A., Biondi-Zoccai, G., Landoni, G., 2013. Meta-analysis: pitfalls and hints. Heart Lung Vessel 219–225. doi:pmid:24364016

Griepentrog, M., Schmidt, M.W.I., 2013. Discrepancies in utilization of density fractionation along with ultrasonic dispersion to obtain distinct pools of soil organic matter. Journal of Plant Nutrition and Soil Science 176, 500–504. doi:10.1002/jpln.201200469

Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., Whiteley, A.S., 2011. The bacterial biogeography of British soils: Mapping soil bacteria. Environmental Microbiology 13, 1642–1654. doi:10.1111/j.1462-2920.2011.02480.x

Grömping, U., 2006. Relative Importance for Linear Regression in R: The Package relaimpo. Journal of Statistical Software 17, 1–27.
Grunwald, S., 2022. Artificial intelligence and soil carbon modeling demystified: power, potentials, and perils. Carbon Footprints 1, 6. doi:10.20517/cf.2022.03
Guan, N., Liu, L., 2020. Microbial response to acid stress: mechanisms and applications. Applied Microbiology and Biotechnology 104, 51–65. doi:10.1007/s00253-019-10226-1
Gubler, A., Gross, T., Hug, A.-S., Moll-Mielewczik, J., Müller, M., Rehbein, K., Schwab, P., Wächter, D., Zimmermann, R., Meuli, R.G., 2022. Die Nationale Boden- beobachtung 2021. [object Object]. doi:10.34776/AS128G
Guidi, C., Magid, J., Rodeghiero, M., Gianelle, D., Vesterdal, L., 2014. Effects of forest expansion on mountain grassland: changes within soil organic carbon fractions. Plant and Soil 385, 373–387. doi:10.1007/s11104-014-2315-2
Guo, L.B., Gifford, R.M., 2002. Soil carbon stocks and land use change: a meta analysis. Global Change Biology 8, 345–360. doi:10.1046/j.1354-1013.2002.00486.x
Guo, X., Gao, Q., Yuan, M., Wang, G., Zhou, X., Feng, J., Shi, Z., Hale, L., Wu, Linwei, Zhou, A., Tian, R., Liu, F., Wu, B., Chen, L., Jung, C.G., Niu, S., Li, D., Xu, X., Jiang, L., Escalas, A., Wu, Liyou, He, Z., Van Nostrand, J.D., Ning, D., Liu, X., Yang, Y., Schuur, Edward.A.G., Konstantinidis, K.T., Cole, J.R., Penton, C.R., Luo, Y., Tiedje, J.M., Zhou, J., 2020. Gene-informed decomposition model predicts lower soil carbon loss due to persistent microbial adaptation to warming. Nature Communications 11, 4897. doi:10.1038/s41467-020-18706-z
Hagerty, S.B., Allison, S.D., Schimel, J.P., 2018. Evaluating soil microbial carbon use efficiency explicitly as a function of cellular processes: implications for measurements and models. Biogeochemistry 140, 269–283. doi:10.1007/s10533-018-0489-z
Hall, E.K., Bernhardt, E.S., Bier, R.L., Bradford, M.A., Boot, C.M., Cotner, J.B., del Giorgio, P.A., Evans, S.E., Graham, E.B., Jones, S.E., Lennon, J.T., Locey, K.J., Nemergut, D., Osborne, B.B., Rocca, J.D., Schimel, J.P., Waldrop, M.P., Wallenstein, M.D., 2018. Understanding how microbiomes influence the systems they inhabit. Nature Microbiology 3, 977–982. doi:10.1038/s41564-018-0201-z
Hall, S.J., Berhe, A.A., Thompson, A., 2018. Order from disorder: do soil organic matter composition and turnover co-vary with iron phase crystallinity? Biogeochemistry 140, 93–110. doi:10.1007/s10533-018-0476-4
Hall, S.J., Ye, C., Weintraub, S.R., Hockaday, W.C., 2020. Molecular trade-offs in soil organic carbon composition at continental scale. Nature Geoscience 13, 687–692. doi:10.1038/s41561-020-0634-x
Han, X., Beck, K., Bürgmann, H., Frey, B., Stierli, B., Frossard, A., 2023. Synthetic oligonucleotides as quantitative PCR standards for quantifying microbial genes. Frontiers in Microbiology 14, 1279041. doi:10.3389/fmicb.2023.1279041
Hansen, P.M., Even, R., King, A.E., Lavallee, J., Schipanski, M., Cotrufo, M.F., 2024. Distinct, direct and climate-mediated environmental controls on global particulate and mineral-associated organic carbon storage. Global Change Biology 30, e17080. doi:10.1111/gcb.17080
Hargreaves, S.K., Roberto, A.A., Hofmockel, K.S., 2013. Reaction- and sample-specific inhibition affect standardization of qPCR assays of soil bacterial communities. Soil Biology and Biochemistry 59, 89–97. doi:10.1016/j.soilbio.2013.01.007
Hashimoto, S., Ito, A., Nishina, K., 2023. Divergent data-driven estimates of global soil respiration. Communications Earth & Environment 4, 460. doi:10.1038/s43247-023-01136-2
Hassink, J., 1994. Effect of soil texture on the size of the microbial biomass and on the amount of c and n mineralized per unit of microbial biomass in dutch grassland soils. Soil Biology and Biochemistry 26, 1573–1581. doi:10.1016/0038-0717(94)90100-7
He, P., Zhang, Y., Shen, Q., Ling, N., Nan, Z., 2023. Microbial carbon use efficiency in different ecosystems: A meta-analysis based on a biogeochemical equilibrium model. Global Change Biology 29, 4758–4774. doi:10.1111/gcb.16861

Heckman, K., Hicks Pries, C.E., Lawrence, C.R., Rasmussen, C., Crow, S.E., Hoyt, A.M., Fromm, S.F., Shi, Z., Stoner, S., McGrath, C., Beem-Miller, J., Berhe, A.A., Blankinship, J.C., Keiluweit, M., Marín-Spiotta, E., Monroe, J.G., Plante, A.F., Schimel, J., Sierra, C.A., Thompson, A., Wagai, R., 2022. Beyond bulk: Density fractions explain heterogeneity in global soil carbon abundance and persistence. Global Change Biology 28, 1178–1196. doi:10.1111/gcb.16023

Heckman, K.A., Nave, L.E., Bowman, M., Gallo, A., Hatten, J.A., Matosziuk, L.M., Possinger, A.R., SanClements, M., Strahm, B.D., Weiglein, T.L., Rasmussen, C., Swanston, C.W., 2021. Divergent controls on carbon concentration and persistence between forests and grasslands of the conterminous US. Biogeochemistry 156, 41–56. doi:10.1007/s10533-020-00725-z

Hendershot, W.H., Duquette, M., 1986. A Simple Barium Chloride Method for Determining Cation Exchange Capacity and Exchangeable Cations. Soil Science Society of America Journal 50, 605–608. doi:10.2136/sssaj1986.03615995005000030013x

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLOS ONE 12, e0169748. doi:10.1371/journal.pone.0169748

Hernandez-Soriano, M.C., Dalal, R.C., Warren, F.J., Wang, P., Green, K., Tobin, M.J., Menzies, N.W., Kopittke, P.M., 2018. Soil Organic Carbon Stabilization: Mapping Carbon Speciation from Intact Microaggregates. Environmental Science & Technology 52, 12275–12284. doi:10.1021/acs.est.8b03095

Hicks Pries, C.E., Ryals, R., Zhu, B., Min, K., Cooper, A., Goldsmith, S., Pett-Ridge, J., Torn, M., Berhe, A.A., 2023. The Deep Soil Organic Carbon Response to Global Change. Annual Review of Ecology, Evolution, and Systematics 54, 375–401. doi:10.1146/annurev-ecolsys-102320-085332

Hijmans R, 2022. raster: Geographic Data Analysis and Modeling. R package version 3.5-29.

Hinckley, E.S., Wieder, W., Fierer, N., Paul, E., 2014. Digging Into the World Beneath Our Feet: Bridging Across Scales in the Age of Global Change. Eos, Transactions American Geophysical Union 95, 96–97. doi:10.1002/2014EO110004

Ho, A., Lonardo, D.P.D., Bodelier, P.L.E., 2017. Revisiting life strategy concepts in environmental microbial ecology. FEMS Microbiology Ecology fix006. doi:10.1093/femsec/fix006

Hoffland, E., Kuyper, T.W., Comans, R.N.J., Creamer, R.E., 2020. Eco-functionality of organic matter in soils. Plant and Soil 455, 1–22. doi:10.1007/s11104-020-04651-9

Hood-Nowotny, R., Umana, N.H.-N., Inselbacher, E., Oswald- Lachouani, P., Wanek, W., 2010. Alternative Methods for Measuring Inorganic, Organic, and Total Dissolved Nitrogen in Soil. Soil Science Society of America Journal 74, 1018–1027. doi:10.2136/sssaj2009.0389

Huang, J., Liu, W., Yang, S., Yang, L., Peng, Z., Deng, M., Xu, S., Zhang, B., Ahirwal, J., Liu, L., 2021. Plant carbon inputs through shoot, root, and mycorrhizal pathways affect soil organic carbon turnover differently. Soil Biology and Biochemistry 160, 108322. doi:10.1016/j.soilbio.2021.108322

Huygens, D., Boeckx, P., Van Cleemput, O., Oyarzún, C., Godoy, R., 2005. Aggregate and soil organic carbon dynamics in South Chilean Andisols. Biogeosciences 2, 159–174. doi:10.5194/bg-2-159-2005

IUSS Working Group WRB, 2015. World Reference Base for Soil Resources 2014, Update 2015. International Soil Classification System for Naming Soils and Creating Legends for Soil Maps. World Soil Resources Reports No. 106.

Jackson, R.B., Lajtha, K., Crow, S.E., Hugelius, G., Kramer, M.G., Piñeiro, G., 2017. The Ecology of Soil Carbon: Pools, Vulnerabilities, and Biotic and Abiotic Controls. Annual Review of Ecology, Evolution, and Systematics 48, 419–445. doi:10.1146/annurev-ecolsys-112414-054234

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. R. An introduction to statistical learning. Springer New York.

Jarvis, P., Rey, A., Petsikos, C., Wingate, L., Rayment, M., Pereira, J., Banza, J., David, J., Miglietta, F., Borghetti, M., Manca, G., Valentini, R., 2007. Drying and wetting of Mediterranean soils stimulates decomposition and carbon dioxide emission: the "Birch effect." Tree Physiology 27, 929–940. doi:10.1093/treephys/27.7.929

Jenkinson, D.S., Andrew, S.P.S., Lynch, J.M., Gross, M.J., Tinker, P.B., 1990. The Turnover of Organic Carbon and Nitrogen in Soil. Philosophical Transactions: Biological Sciences 329.

Jenny, H., 1941. Factors of soil formation: a system of quantitative pedology, Unabridged, unaltered republ., new foreword. ed, Dover books on earth sciences. Dover Publ, New York.

Jian, J., Vargas, R., Anderson-Teixeira, K.J., Stell, E., Herrmann, V., Horn, M., Kholod, N., Manzon, J., Marchesi, R., Paredes, D., Bond-Lamberty, B.P., 2021. A Global Database of Soil Respiration Data, Version 5.0. ORNL DAAC, Oak Ridge, Tennessee, USA. doi:https://doi.org/10.3334/ORNLDAAC/1827

Jobbágy, E.G., Jackson, R.B., 2000. THE VERTICAL DISTRIBUTION OF SOIL ORGANIC CARBON AND ITS RELATION TO CLIMATE AND VEGETATION. Ecological Applications 10, 423–436. doi:10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2

John, B., Yamashita, T., Ludwig, B., Flessa, H., 2005. Storage of organic carbon in aggregate and density fractions of silty soils under different types of land use. Geoderma 128, 63–79. doi:10.1016/j.geoderma.2004.12.013

Johnson, D.R., Goldschmidt, F., Lilja, E.E., Ackermann, M., 2012. Metabolic specialization and the assembly of microbial communities. The ISME Journal 6, 1985–1991. doi:10.1038/ismej.2012.46

Jones, A.R., Dalal, R.C., Gupta, V.V.S.R., Schmidt, S., Allen, D.E., Jacobsen, G.E., Bird, M., Grandy, A.S., Sanderman, J., 2023. Molecular complexity and diversity of persistent soil organic matter. Soil Biology and Biochemistry 184, 109061. doi:10.1016/j.soilbio.2023.109061

Jones, C.D., 2021. Numerical modeling of the global climate and carbon cycle system. In: Letcher, M.T. (Editor): Climate Change, Third Edition, in: Climate Change, Third Edition. Elsevier, pp. 67–91. doi:10.1016/B978-0-12-821575-3.00004-9

Jones, D.L., Cooledge, E.C., Hoyle, F.C., Griffiths, R.I., Murphy, D.V., 2019. pH and exchangeable aluminum are major regulators of microbial energy flow and carbon use efficiency in soil microbial communities. Soil Biology and Biochemistry 138, 107584. doi:10.1016/j.soilbio.2019.107584

Jones, D.L., Hill, P.W., Smith, A.R., Farrell, M., Ge, T., Banning, N.C., Murphy, D.V., 2018. Role of substrate supply on microbial carbon use efficiency and its role in interpreting soil microbial community-level physiological profiles (CLPP). Soil Biology and Biochemistry 123, 1–6. doi:10.1016/j.soilbio.2018.04.014

Jungkunst, H.F., Göpel, J., Horvath, T., Ott, S., Brunn, M., 2022. Global soil organic carbon–climate interactions: Why scales matter. WIREs Climate Change 13, e780. doi:10.1002/wcc.780

Just, C., Poeplau, C., Don, A., van Wesemael, B., Kögel-Knabner, I., Wiesmeier, M., 2021. A Simple Approach to Isolate Slow and Fast Cycling Organic Carbon Fractions in Central European Soils—Importance of Dispersion Method. Frontiers in Soil Science 1, 692583. doi:10.3389/fsoil.2021.692583

Kaiser, C., Franklin, O., Dieckmann, U., Richter, A., 2014. Microbial community dynamics alleviate stoichiometric constraints during litter decay. Ecology Letters 17, 680–690. doi:10.1111/ele.12269

Kaiser, C., Koranda, M., Kitzler, B., Fuchslueger, L., Schnecker, J., Schweiger, P., Rasche, F., Zechmeister-Boltenstern, S., Sessitsch, A., Richter, A., 2010. Belowground carbon allocation by trees drives seasonal patterns of extracellular enzyme activities by altering microbial community composition in a beech forest soil. New Phytologist 187, 843–858. doi:10.1111/j.1469-8137.2010.03321.x

Kaiser, K., Guggenberger, G., Haumaier, L., Zech, W., 1997. Dissolved organic matter sorption on sub soils and minerals studied by $^{13}$C-NMR and DRIFT spectroscopy. European Journal of Soil Science 48, 301–310. doi:10.1111/j.1365-2389.1997.tb00550.x

Kaiser, K., Kalbitz, K., 2012. Cycling downwards – dissolved organic matter in soils. Soil Biology and Biochemistry 52, 29–32. doi:10.1016/j.soilbio.2012.04.002
Kaiser, M., Berhe, A.A., 2014. How does sonication affect the mineral and organic constituents of soil aggregates?—A review. Journal of Plant Nutrition and Soil Science 177, 479–495. doi:10.1002/jpln.201300339
Kalbitz, K., Kaiser, K., 2008. Contribution of dissolved organic matter to carbon storage in forest mineral soils. Journal of Plant Nutrition and Soil Science 171, 52–60. doi:10.1002/jpln.200700043
Kallenbach, C.M., Frey, S.D., Grandy, A.S., 2016. Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls. Nature Communications 7, 13630. doi:10.1038/ncomms13630
Keiluweit, M., Nico, P.S., Kleber, M., Fendorf, S., 2016. Are oxygen limitations under recognized regulators of organic carbon turnover in upland soils? Biogeochemistry 127, 157–171. doi:10.1007/s10533-015-0180-6
Keiluweit, M., Wanzek, T., Kleber, M., Nico, P., Fendorf, S., 2017. Anaerobic microsites have an unaccounted role in soil carbon stabilization. Nature Communications 8, 1771. doi:10.1038/s41467-017-01406-6
Kellerman, A.M., Kothawala, D.N., Dittmar, T., Tranvik, L.J., 2015. Persistence of dissolved organic matter in lakes related to its molecular characteristics. Nature Geoscience 8, 454–457. doi:10.1038/ngeo2440
Keuskamp, J.A., Dingemans, B.J.J., Lehtinen, T., Sarneel, J.M., Hefting, M.M., 2013. Tea Bag Index: a novel approach to collect uniform decomposition data across ecosystems. Methods in Ecology and Evolution 4, 1070–1075. doi:10.1111/2041-210X.12097
Khan, A., 2018. collapsibleTree: Interactive Collapsible Tree Diagrams using “D3.js”. R package version 0.1.7.
Kindler, R., Siemens, J., Kaiser, K., Walmsley, D.C., Bernhofer, C., Buchmann, N., Cellier, P., Eugster, W., Gleixner, G., Grũnwald, T., Heim, A., Ibrom, A., Jones, S.K., Jones, M., Klumpp, K., Kutsch, W., Larsen, K.S., Lehuger, S., Loubet, B., Mckenzie, R., Moors, E., Osborne, B., Pilegaard, K., Rebmann, C., Saunders, M., Schmidt, M.W.I., Schrumpf, M., Seyfferth, J., Skiba, U., Soussana, J., Sutton, M.A., Tefs, C., Vowinckel, B., Zeeman, M.J., Kaupenjohann, M., 2011. Dissolved carbon leaching from soil is a crucial component of the net ecosystem carbon balance. Global Change Biology 17, 1167–1185. doi:10.1111/j.1365-2486.2010.02282.x
Kleber, M., Bourg, I.C., Coward, E.K., Hansel, C.M., Myneni, S.C.B., Nunan, N., 2021. Dynamic interactions at the mineral–organic matter interface. Nature Reviews Earth & Environment 2, 402–421. doi:10.1038/s43017-021-00162-y
Kleber, M., Eusterhues, K., Keiluweit, M., Mikutta, C., Mikutta, R., Nico, P.S., 2015. Mineral–Organic Associations: Formation, Properties, and Relevance in Soil Environments, in: Advances in Agronomy. Elsevier, pp. 1–140. doi:10.1016/bs.agron.2014.10.005
Kleber, M., Sollins, P., Sutton, R., 2007. A conceptual model of organo-mineral interactions in soils: self-assembly of organic molecular fragments into zonal structures on mineral surfaces. Biogeochemistry 85, 9–24. doi:10.1007/s10533-007-9103-5
Klein Goldewijk, K., Beusen, A., Doelman, J., Stehfest, E., 2017. Anthropogenic land use estimates for the Holocene – HYDE 3.2. Earth System Science Data 9, 927–953. doi:10.5194/essd-9-927-2017
Knapp, A.K., Carroll, C.J.W., Fahey, T.J., 2014. Patterns and Controls of Terrestrial Primary Production in a Changing World, in: Monson, R.K. (Ed.), Ecology and the Environment. Springer New York, New York, NY, pp. 205–246. doi:10.1007/978-1-4614-7501-9_2
Kögel-Knabner, I., Amelung, W., 2021. Soil organic matter in major pedogenic soil groups. Geoderma 384, 114785. doi:10.1016/j.geoderma.2020.114785
König, S., Vogel, H.-J., Harms, H., Worrich, A., 2020. Physical, Chemical and Biological Effects on Soil Bacterial Dynamics in Microscale Models. Frontiers in Ecology and Evolution 8, 53. doi:10.3389/fevo.2020.00053

Kopittke, P.M., Hernandez-Soriano, M.C., Dalal, R.C., Finn, D., Menzies, N.W., Hoeschen, C., Mueller, C.W., 2018. Nitrogen-rich microbial products provide new organo-mineral associations for the stabilization of soil organic matter. Global Change Biology 24, 1762–1770. doi:10.1111/gcb.14009

Korthauer, K., Kimes, P.K., Duvallet, C., Reyes, A., Subramanian, A., Teng, M., Shukla, C., Alm, E.J., Hicks, S.C., 2019. A practical guide to methods controlling false discoveries in computational biology. Genome Biology 20, 118. doi:10.1186/s13059-019-1716-1

Kramer, M.G., Sanderman, J., Chadwick, O.A., Chorover, J., Vitousek, P.M., 2012. Long-term carbon storage through retention of dissolved aromatic acids by reactive particles in soil. Global Change Biology 18, 2594–2605. doi:10.1111/j.1365-2486.2012.02681.x

Krause, S., Le Roux, X., Niklaus, P.A., Van Bodegom, P.M., Lennon, J.T., Bertilsson, S., Grossart, H.-P., Philippot, L., Bodelier, P.L.E., 2014. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. Frontiers in Microbiology 5. doi:10.3389/fmicb.2014.00251

Krulwich, T.A., Sachs, G., Padan, E., 2011. Molecular aspects of bacterial pH sensing and homeostasis. Nature Reviews Microbiology 9, 330–343. doi:10.1038/nrmicro2549

Kuhn, M., 2008. Caret package.

La Pierre, K.J., Blumenthal, D.M., Brown, C.S., Klein, J.A., Smith, M.D., 2016. Drivers of Variation in Aboveground Net Primary Productivity and Plant Community Composition Differ Across a Broad Precipitation Gradient. Ecosystems 19, 521–533. doi:10.1007/s10021-015-9949-7

Labouyrie, M., Ballabio, C., Romero, F., Panagos, P., Jones, A., Schmid, M.W., Mikryukov, V., Dulya, O., Tedersoo, L., Bahram, M., Lugato, E., van der Heijden, M.G.A., Orgiazzi, A., 2023. Patterns in soil microbial diversity across Europe. Nature Communications 14, 3311. doi:10.1038/s41467-023-37937-4

Lange, M., Eisenhauer, N., Sierra, C.A., Bessler, H., Engels, C., Griffiths, R.I., Mellado-Vázquez, P.G., Malik, A.A., Roy, J., Scheu, S., Steinbeiss, S., Thomson, B.C., Trumbore, S.E., Gleixner, G., 2015. Plant diversity increases soil microbial activity and soil carbon storage. Nature Communications 6, 6707. doi:10.1038/ncomms7707

Langenheder, S., Prosser, J.I., 2008. Resource availability influences the diversity of a functional group of heterotrophic soil bacteria. Environmental Microbiology 10, 2245–2256. doi:10.1111/j.1462-2920.2008.01647.x

Laub, M., Blagodatsky, S., Van de Broek, M., Schlichenmaier, S., Kunlanit, B., Six, J., Vityakon, P., Cadisch, G., 2024. SAMM version 1.0: a numerical model for microbial- mediated soil aggregate formation. Geoscientific Model Development 17, 931–956. doi:10.5194/gmd-17-931-2024

Lauber, C.L., Hamady, M., Knight, R., Fierer, N., 2009. Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale. Applied and Environmental Microbiology 75, 5111–5120. doi:10.1128/AEM.00335-09

Lavallee, J.M., Soong, J.L., Cotrufo, M.F., 2020. Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century. Global Change Biology 26, 261–273. doi:10.1111/gcb.14859

Lawrence, C.R., Beem-Miller, J., Hoyt, A.M., Monroe, G., Sierra, C.A., Stoner, S., Heckman, K., Blankinship, J.C., Crow, S.E., McNicol, G., Trumbore, S., Levine, P.A., Vindušková, O., Todd-Brown, K., Rasmussen, C., Hicks Pries, C.E., Schädel, C., McFarlane, K., Doetterl, S., Hatté, C., He, Y., Treat, C., Harden, J.W., Torn, M.S., Estop-Aragonés, C., Asefaw Berhe, A., Keiluweit, M., Della Rosa Kuhnen, Á., Marin-Spiotta, E., Plante, A.F., Thompson, A., Shi, Z., Schimel, J.P., Vaughn, L.J.S., von Fromm, S.F., Wagai, R., 2020. An open-source database for the synthesis of soil radiocarbon data: International Soil Radiocarbon Database (ISRaD) version 1.0. Earth System Science Data 12, 61–76. doi:10.5194/essd-12-61-2020

Le Noë, J., Manzoni, S., Abramoff, R., Bölscher, T., Bruni, E., Cardinael, R., Ciais, P., Chenu, C., Clivot, H., Derrien, D., Ferchaud, F., Garnier, P., Goll, D., Lashermes, G., Martin, M., Rasse, D., Rees, F., Sainte-Marie, J., Salmon, E., Schiedung, M., Schimel, J., Wieder, W., Abiven, S., Barré, P.,

Cécillon, L., Guenet, B., 2023. Soil organic carbon models need independent time-series validation for reliable prediction. Communications Earth & Environment 4, 158. doi:10.1038/s43247-023-00830-5
Leek, J.T., Jager, L., Boca, S.M., Konopka, T., 2022. swfdr: Estimation of the science-wise false discovery rate and the false discovery rate conditional on covariates. R package version 1.22.0.
Lehmann, J., Hansel, C.M., Kaiser, C., Kleber, M., Maher, K., Manzoni, S., Nunan, N., Reichstein, M., Schimel, J.P., Torn, M.S., Wieder, W.R., Kögel-Knabner, I., 2020. Persistence of soil organic carbon caused by functional complexity. Nature Geoscience 13, 529–534. doi:10.1038/s41561-020-0612-3
Lehmann, J., Kinyangi, J., Solomon, D., 2007. Organic matter stabilization in soil microaggregates: implications from spatial heterogeneity of organic carbon contents and carbon forms. Biogeochemistry 85, 45–57. doi:10.1007/s10533-007-9105-3
Lehmann, J., Kleber, M., 2015. The contentious nature of soil organic matter. Nature 528, 60–68. doi:10.1038/nature16069
Lerch, T.Z., Nunan, N., Dignac, M.-F., Chenu, C., Mariotti, A., 2011. Variations in microbial isotopic fractionation during soil organic matter decomposition. Biogeochemistry 106, 5–21. doi:10.1007/s10533-010-9432-7
Li, J., Mau, R.L., Dijkstra, P., Koch, B.J., Schwartz, E., Liu, X.-J.A., Morrissey, E.M., Blazewicz, S.J., Pett-Ridge, J., Stone, B.W., Hayer, M., Hungate, B.A., 2019. Predictive genomic traits for bacterial growth in culture versus actual growth in soil. The ISME Journal 13, 2162–2172. doi:10.1038/s41396-019-0422-z
Li, P., Hur, J., 2017. Utilization of UV-Vis spectroscopy and related data analyses for dissolved organic matter (DOM) studies: A review. Critical Reviews in Environmental Science and Technology 47, 131–154. doi:10.1080/10643389.2017.1309186
Li, Q., Chang, J., Li, L., Lin, X., Li, Y., 2023. Research progress of nano-scale secondary ion mass spectrometry (NanoSIMS) in soil science: Evolution, applications, and challenges. Science of The Total Environment 905, 167257. doi:10.1016/j.scitotenv.2023.167257
Li, Q., Chen, J., Feng, J., Wu, J., Zhang, Q., Jia, W., Lin, Q., Cheng, X., 2020. How do Biotic and Abiotic Factors Regulate Soil Enzyme Activities at Plot and Microplot Scales Under Afforestation? Ecosystems 23, 1408–1422. doi:10.1007/s10021-019-00477-4
Liang, C., Schimel, J.P., Jastrow, J.D., 2017. The importance of anabolism in microbial control over soil carbon storage. Nature Microbiology 2, 17105. doi:10.1038/nmicrobiol.2017.105
Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2, 18–22.
Lipson, D.A., 2015. The complex relationship between microbial growth rate and yield and its implications for ecosystem processes. Frontiers in Microbiology 6. doi:10.3389/fmicb.2015.00615
Liu, X.-J.A., Finley, B.K., Mau, R.L., Schwartz, E., Dijkstra, P., Bowker, M.A., Hungate, B.A., 2020. The soil priming effect: Consistent across ecosystems, elusive mechanisms. Soil Biology and Biochemistry 140, 107617. doi:10.1016/j.soilbio.2019.107617
Liu, Y.-R., Delgado-Baquerizo, M., Wang, J.-T., Hu, H.-W., Yang, Z., He, J.-Z., 2018. New insights into the role of microbial community composition in driving soil respiration rates. Soil Biology and Biochemistry 118, 35–41. doi:10.1016/j.soilbio.2017.12.003
Lloyd, J., Taylor, J.A., 1994. On the Temperature Dependence of Soil Respiration. Functional Ecology 8, 315. doi:10.2307/2389824
Lugato, E., Lavallee, J.M., Haddix, M.L., Panagos, P., Cotrufo, M.F., 2021. Different climate sensitivity of particulate and mineral-associated soil organic matter. Nature Geoscience 14, 295–300. doi:10.1038/s41561-021-00744-x
Lumley, T., 2020. based on Fortran code by Alan Miller. leaps: Regression Subset Selection. R package version 3.1.
Lund, P.A., De Biase, D., Liran, O., Scheler, O., Mira, N.P., Cetecioglu, Z., Fernández, E.N., Bover-Cid, S., Hall, R., Sauer, M., O'Byrne, C., 2020. Understanding How Microorganisms Respond to

Acid pH Is Central to Their Control and Successful Exploitation. Frontiers in Microbiology 11, 556140. doi:10.3389/fmicb.2020.556140
Luo, Y., Ahlström, A., Allison, S.D., Batjes, N.H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E.A., Finzi, A., Georgiou, K., Guenet, B., Hararuk, O., Harden, J.W., He, Y., Hopkins, F., Jiang, L., Koven, C., Jackson, R.B., Jones, C.D., Lara, M.J., Liang, J., McGuire, A.D., Parton, W., Peng, C., Randerson, J.T., Salazar, A., Sierra, C.A., Smith, M.J., Tian, H., Todd-Brown, K.E.O., Torn, M., Groenigen, K.J., Wang, Y.P., West, T.O., Wei, Y., Wieder, W.R., Xia, J., Xu, Xia, Xu, Xiaofeng, Zhou, T., 2016. Toward more realistic projections of soil carbon dynamics by Earth system models. Global Biogeochemical Cycles 30, 40–56. doi:10.1002/2015GB005239
Luo, Z., Viscarra Rossel, R.A., Shi, Z., 2020. Distinct controls over the temporal dynamics of soil carbon fractions after land use change. Global Change Biology 26, 4614–4625. doi:10.1111/gcb.15157
Lynch, M.D.J., Neufeld, J.D., 2015. Ecology and exploration of the rare biosphere. Nature Reviews Microbiology 13, 217–229. doi:10.1038/nrmicro3400
Maestre, F.T., Delgado-Baquerizo, M., Jeffries, T.C., Eldridge, D.J., Ochoa, V., Gozalo, B., Quero, J.L., García-Gómez, M., Gallardo, A., Ulrich, W., Bowker, M.A., Arredondo, T., Barraza-Zepeda, C., Bran, D., Florentino, A., Gaitán, J., Gutiérrez, J.R., Huber-Sannwald, E., Jankju, M., Mau, R.L., Miriti, M., Naseri, K., Ospina, A., Stavi, I., Wang, D., Woods, N.N., Yuan, X., Zaady, E., Singh, B.K., 2015. Increasing aridity reduces soil microbial diversity and abundance in global drylands. Proceedings of the National Academy of Sciences 112, 15684–15689. doi:10.1073/pnas.1516684112
Mainka, M., Summerauer, L., Wasner, D., Garland, G., Griepentrog, M., Berhe, A.A., Doetterl, S., 2022. Soil geochemistry as a driver of soil organic matter composition: insights from a soil chronosequence. Biogeosciences 19, 1675–1689. doi:10.5194/bg-19-1675-2022
Malik, A.A., Bouskill, N.J., 2022. Drought impacts on microbial trait distribution and feedback to soil carbon cycling. Functional Ecology 36, 1442–1456. doi:10.1111/1365-2435.14010
Malik, A.A., Chowdhury, S., Schlager, V., Oliver, A., Puissant, J., Vazquez, P.G.M., Jehmlich, N., von Bergen, M., Griffiths, R.I., Gleixner, G., 2016. Soil Fungal:Bacterial Ratios Are Linked to Altered Carbon Cycling. Frontiers in Microbiology 7. doi:10.3389/fmicb.2016.01247
Malik, A.A., Martiny, J.B.H., Brodie, E.L., Martiny, A.C., Treseder, K.K., Allison, S.D., 2020. Defining trait-based microbial strategies with consequences for soil carbon cycling under climate change. The ISME Journal 14, 1–9. doi:10.1038/s41396-019-0510-0
Malik, A.A., Puissant, J., Buckeridge, K.M., Goodall, T., Jehmlich, N., Chowdhury, S., Gweon, H.S., Peyton, J.M., Mason, K.E., van Agtmaal, M., Blaud, A., Clark, I.M., Whitaker, J., Pywell, R.F., Ostle, N., Gleixner, G., Griffiths, R.I., 2018. Land use driven change in soil pH affects microbial carbon cycling processes. Nature Communications 9, 3591. doi:10.1038/s41467-018-05980-1
Malik, A.A., Puissant, J., Goodall, T., Allison, S.D., Griffiths, R.I., 2019. Soil microbial communities with greater investment in resource acquisition have lower growth yield. Soil Biology and Biochemistry 132, 36–39. doi:10.1016/j.soilbio.2019.01.025
Man, M., Gregorich, E.G., Beare, M.H., Ellert, B.H., Simpson, M.J., 2023. Distinct dynamics of plant- and microbial-derived soil organic matter in relation to varying climate and soil properties in temperate agroecosystems. Geochimica et Cosmochimica Acta 361, 276–287. doi:10.1016/j.gca.2023.10.008
Manzoni, S., Chakrawal, A., Fischer, T., Schimel, J.P., Porporato, A., Vico, G., 2020. Rainfall intensification increases the contribution of rewetting pulses to soil heterotrophic respiration. Biogeosciences 17, 4007–4023. doi:10.5194/bg-17-4007-2020
Manzoni, S., Schimel, J.P., Porporato, A., 2012a. Responses of soil microbial communities to water stress: results from a meta-analysis. Ecology 93, 930–938. doi:10.1890/11-0026.1
Manzoni, S., Taylor, P., Richter, A., Porporato, A., Ågren, G.I., 2012b. Environmental and stoichiometric controls on microbial carbon-use efficiency in soils. New Phytologist 196, 79–91. doi:10.1111/j.1469-8137.2012.04225.x

Marcela, H., de los Ángeles Sepulveda, M., Muñoz, C., Casanova, M., Wasner, D., Bodé, S., Doetterl, S., Boeckx, P., Zagal, E., 2022. Predicting Soil Organic Carbon Mineralization Rates Using δ13C, Assessed by Near-Infrared Spectroscopy, in Depth Profiles Under Permanent Grassland Along a Latitudinal Transect in Chile. Journal of Soil Science and Plant Nutrition 22, 2105–2117. doi:10.1007/s42729-022-00797-w

Marschmann, G.L., Pagel, H., Kügler, P., Streck, T., 2019. Equifinality, sloppiness, and emergent structures of mechanistic soil biogeochemical models. Environmental Modelling & Software 122, 104518. doi:10.1016/j.envsoft.2019.104518

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal 17, 10–12. doi:https://doi.org/10.14806/ej.17.1.200

Martín, M.A., Martínez, F.S.J., Giráldez, J.V., Pachepsky, Y., Vogel, H., 2021. Editorial for the special issue on "Advances in soil scaling: Theories, techniques and applications." European Journal of Soil Science 72, 491–494. doi:10.1111/ejss.13063

Martiny, J.B.H., Jones, S.E., Lennon, J.T., Martiny, A.C., 2015. Microbiomes in light of traits: A phylogenetic perspective. Science 350, aac9323. doi:10.1126/science.aac9323

Matus, F., Rumpel, C., Neculman, R., Panichini, M., Mora, M.L., 2014. Soil carbon storage and stabilisation in andic soils: A review. CATENA 120, 102–110. doi:10.1016/j.catena.2014.04.008

Matus, F.J., 2021. Fine silt and clay content is the main factor defining maximal C and N accumulations in soils: a meta-analysis. Scientific Reports 11, 6438. doi:10.1038/s41598-021-84821-6

Mbé, B., Monga, O., Pot, V., Otten, W., Hecht, F., Raynaud, X., Nunan, N., Chenu, C., Baveye, P.C., Garnier, P., 2022. Scenario modelling of carbon mineralization in 3D soil architecture at the microscale: Toward an accessibility coefficient of organic matter for bacteria. European Journal of Soil Science 73, e13144. doi:10.1111/ejss.13144

McMurdie, P.J., Holmes, S., 2014. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. PLoS Computational Biology 10, e1003531. doi:10.1371/journal.pcbi.1003531

Meier, C.L., Bowman, W.D., 2008. Links between plant litter chemistry, species diversity, and below-ground ecosystem function. Proceedings of the National Academy of Sciences 105, 19780–19785. doi:10.1073/pnas.0805600105

Mengist, W., Soromessa, T., Legese, G., 2020. Method for conducting systematic literature review and meta-analysis for environmental science research. MethodsX 7, 100777. doi:10.1016/j.mex.2019.100777

Metze, D., Schnecker, J., Canarini, A., Fuchslueger, L., Koch, B.J., Stone, B.W., Hungate, B.A., Hausmann, B., Schmidt, H., Schaumberger, A., Bahn, M., Kaiser, C., Richter, A., 2023. Microbial growth under drought is confined to distinct taxa and modified by potential future climate conditions. Nature Communications 14, 5895. doi:10.1038/s41467-023-41524-y

Metze, D., Schnecker, J., de Carlan, C.L.N., Bhattarai, B., Verbruggen, E., Ostonen, I., Janssens, I.A., Sigurdsson, B.D., Hausmann, B., Kaiser, C., Richter, A., 2024. Soil warming increases the number of growing bacterial taxa but not their growth rates. Science Advances 10, eadk6295. doi:10.1126/sciadv.adk6295

Meyer, H., Pebesma, E., 2022. Machine learning-based global maps of ecological variables and the challenge of assessing them. Nature Communications 13, 2208. doi:10.1038/s41467-022-29838-9

Mikutta, R., Turner, S., Schippers, A., Gentsch, N., Meyer-Stüve, S., Condron, L.M., Peltzer, D.A., Richardson, S.J., Eger, A., Hempel, G., Kaiser, K., Klotzbücher, T., Guggenberger, G., 2019. Microbial and abiotic controls on mineral-associated organic matter in soil profiles along an ecosystem gradient. Scientific Reports 9, 10294. doi:10.1038/s41598-019-46501-4

Millard, S.P., 2013. EnvStats: An R Package for Environmental Statistics. Springer, New York.

Mori, A.S., Cornelissen, J.H.C., Fujii, S., Okada, K., Isbell, F., 2020. A meta-analysis on decomposition quantifies afterlife effects of plant diversity as a global change driver. Nature Communications 11, 4547. doi:10.1038/s41467-020-18296-w

Morrissey, E.M., Mau, R.L., Hayer, M., Liu, X.-J.A., Schwartz, E., Dijkstra, P., Koch, B.J., Allen, K., Blazewicz, S.J., Hofmockel, K., Pett-Ridge, J., Hungate, B.A., 2019. Evolutionary history constrains microbial traits across environmental variation. Nature Ecology & Evolution 3, 1064–1069. doi:10.1038/s41559-019-0918-y

Morrissey, E.M., Mau, R.L., Schwartz, E., Caporaso, J.G., Dijkstra, P., van Gestel, N., Koch, B.J., Liu, C.M., Hayer, M., McHugh, T.A., Marks, J.C., Price, L.B., Hungate, B.A., 2016. Phylogenetic organization of bacterial activity. The ISME Journal 10, 2336–2340. doi:10.1038/ismej.2016.28

Morton, J.T., Sanders, J., Quinn, R.A., McDonald, D., Gonzalez, A., Vázquez-Baeza, Y., Navas-Molina, J.A., Song, S.J., Metcalf, J.L., Hyde, E.R., Lladser, M., Dorrestein, P.C., Knight, R., 2017. Balance Trees Reveal Microbial Niche Differentiation. MSystems 2, e00162-16. doi:10.1128/mSystems.00162-16

Mueller, C.W., Gutsch, M., Kothieringer, K., Leifeld, J., Rethemeyer, J., Brueggemann, N., Kögel-Knabner, I., 2014. Bioavailability and isotopic composition of CO2 released from incubated soil organic matter fractions. Soil Biology and Biochemistry 69, 168–178. doi:10.1016/j.soilbio.2013.11.006

Mueller, C.W., Kölbl, A., Hoeschen, C., Hillion, F., Heister, K., Herrmann, A.M., Kögel-Knabner, I., 2012a. Submicron scale imaging of soil organic matter dynamics using NanoSIMS – From single particles to intact aggregates. Organic Geochemistry 42, 1476–1488. doi:10.1016/j.orggeochem.2011.06.003

Mueller, C.W., Schlund, S., Prietzel, J., Kögel-Knabner, I., Gutsch, M., 2012b. Soil Aggregate Destruction by Ultrasonication Increases Soil Organic Matter Mineralization and Mobility. Soil Science Society of America Journal 76, 1634–1643. doi:10.2136/sssaj2011.0186

Muscarella, M.E., Howey, X.M., Lennon, J.T., 2020. Trait-based approach to bacterial growth efficiency. Environmental Microbiology 22, 3494–3504. doi:10.1111/1462-2920.15120

Naipal, V., Ciais, P., Wang, Y., Lauerwald, R., Guenet, B., Van Oost, K., 2018. Global soil organic carbon removal by water erosion under climate change and land use change during AD 1850–2005. Biogeosciences 15, 4459–4480. doi:10.5194/bg-15-4459-2018

Nakhavali, M., Lauerwald, R., Regnier, P., Guenet, B., Chadburn, S., Friedlingstein, P., 2021. Leaching of dissolved organic carbon from mineral soils plays a significant role in the terrestrial carbon balance. Global Change Biology 27, 1083–1096. doi:10.1111/gcb.15460

Nave, L.E., Bowman, M., Gallo, A., Hatten, J.A., Heckman, K.A., Matosziuk, L., Possinger, A.R., SanClements, M., Sanderman, J., Strahm, B.D., Weiglein, T.L., Swanston, C.W., 2021. Patterns and predictors of soil organic carbon storage across a continental-scale network. Biogeochemistry 156, 75–96. doi:10.1007/s10533-020-00745-9

Ng, E.-L., Patti, A.F., Rose, M.T., Schefe, C.R., Wilkinson, K., Smernik, R.J., Cavagnaro, T.R., 2014. Does the chemical nature of soil carbon drive the structure and functioning of soil microbial communities? Soil Biology and Biochemistry 70, 54–61. doi:10.1016/j.soilbio.2013.12.004

Nguyen, N.H., Song, Z., Bates, S.T., Branco, S., Tedersoo, L., Menke, J., Schilling, J.S., Kennedy, P.G., 2016. FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. Fungal Ecology 20, 241–248. doi:10.1016/j.funeco.2015.06.006

Nottingham, A.T., Bååth, E., Reischke, S., Salinas, N., Meir, P., 2019. Adaptation of soil microbial growth to temperature: Using a tropical elevation gradient to predict future changes. Global Change Biology 25, 827–838. doi:10.1111/gcb.14502

Nunan, N., 2017. The microbial habitat in soil: Scale, heterogeneity and functional consequences. Journal of Plant Nutrition and Soil Science 180, 425–429. doi:10.1002/jpln.201700184

Nunan, N., Schmidt, H., Raynaud, X., 2020. The ecology of heterogeneity: soil bacterial communities and C dynamics. Phil. Trans. R. Soc. B 375. doi:http://dx.doi.org/10.1098/rstb.2019.0249

Oades, J.M., 1984. Soil organic matter and structural stability: mechanisms and implications for management. Plant and Soil 76.

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O`Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D.,

Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M., Lahti, L., McGlinn, D., Oullette, M., Ribeiro Cunha, E., Smith, T., Stier, A., Ter Braak, C., Weedon, J., 2022. vegan: Community Ecology Package. R package version 2.6-4.

Orland, C., Emilson, E.J.S., Basiliko, N., Mykytczuk, N.C.S., Gunn, J.M., Tanentzap, A.J., 2019. Microbiome functioning depends on individual and interactive effects of the environment and community structure. The ISME Journal 13, 1–11. doi:10.1038/s41396-018-0230-x

O'Rourke, S.M., Angers, D.A., Holden, N.M., McBratney, A.B., 2015. Soil organic carbon across scales. Global Change Biology 21, 3561–3574. doi:10.1111/gcb.12959

Ortega-Ramírez, P., Pot, V., Laville, P., Schlüter, S., Amor-Quiroz, D.A., Hadjar, D., Mazurier, A., Lacoste, M., Caurel, C., Pouteau, V., Chenu, C., Basile-Doelsch, I., Henault, C., Garnier, P., 2023. Pore distances of particulate organic matter predict N 2 O emissions from intact soil at moist conditions. Geoderma 429, 116224. doi:10.1016/j.geoderma.2022.116224

Osburn, E.D., Badgley, B.D., Strahm, B.D., Aylward, F.O., Barrett, J.E., 2021. Emergent properties of microbial communities drive accelerated biogeochemical cycling in disturbed temperate forests. Ecology 102, e03553. doi:10.1002/ecy.3553

Otoguro, M., Yamamura, H., Quintana, E.T., 2014. The Family Streptosporangiaceae, in: The Prokaryotes: Actinobacteria. Fourth Edition. (Eds: Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E., Thompson, F.). Springer, Berlin, Heidelberg.

Pachepsky, Y., Hill, R.L., 2017. Scale and scaling in soils. Geoderma 287, 4–30. doi:10.1016/j.geoderma.2016.08.017

Parikh, S.J., Goyne, K.W., Margenot, A.J., Mukome, F.N.D., Calderón, F.J., 2014. Soil Chemical Insights Provided through Vibrational Spectroscopy, in: Advances in Agronomy. Elsevier, pp. 1–148. doi:10.1016/B978-0-12-800132-5.00001-8

Parton, W.J., Schimel, D.S., Cole, C.V., Ojima, D.S., 1987. Analysis of Factors Controlling Soil Organic Matter Levels in Great Plains Grasslands. Soil Science Society of America Journal 51, 1173–1179. doi:10.2136/sssaj1987.03615995005100050015x

Paustian, K., Larson, E., Kent, J., Marx, E., Swan, A., 2019. Soil C Sequestration as a Biological Negative Emission Strategy. Frontiers in Climate 1, 8. doi:10.3389/fclim.2019.00008

Paustian, K., Lehmann, J., Ogle, S., Reay, D., Robertson, G.P., Smith, P., 2016. Climate-smart soils. Nature 532, 49–57. doi:10.1038/nature17174

Pedersen, L.L., Smets, B.F., Dechesne, A., 2015. Measuring biogeochemical heterogeneity at the micro scale in soils and sediments. Soil Biology and Biochemistry 90, 122–138. doi:10.1016/j.soilbio.2015.08.003

Pepe-Ranney, C., Campbell, A.N., Koechli, C.N., Berthrong, S., Buckley, D.H., 2016. Unearthing the Ecology of Soil Microorganisms Using a High Resolution DNA-SIP Approach to Explore Cellulose and Xylose Metabolism in Soil. Frontiers in Microbiology 7. doi:10.3389/fmicb.2016.00703

Peuravuori, J., Pihlaja, K., 1997. Molecular size distribution and spectroscopic properties of aquatic humic substances. Analytica Chimica Acta 337, 133–149. doi:10.1016/S0003-2670(96)00412-6

Philippot, L., Chenu, C., Kappler, A., Rillig, M.C., Fierer, N., 2024. The interplay between microbial communities and soil properties. Nature Reviews Microbiology 22, 226–239. doi:10.1038/s41579-023-00980-5

Piton, G., Foulquier, A., Martínez-García, L.B., Legay, N., Hedlund, K., Martins da Silva, P., Nascimento, E., Reis, F., Sousa, J.P., De Deyn, G.B., Clement, J.C., 2020. Disentangling drivers of soil microbial potential enzyme activity across rain regimes: An approach based on the functional trait framework. Soil Biology and Biochemistry 148, 107881. doi:10.1016/j.soilbio.2020.107881

Poeplau, C., Don, A., 2014. Effect of ultrasonic power on soil organic carbon fractions. Journal of Plant Nutrition and Soil Science 177, 137–140. doi:10.1002/jpln.201300492

Poeplau, C., Don, A., 2013. Sensitivity of soil organic carbon stocks and fractions to different land-use changes across Europe. Geoderma 192, 189–201. doi:10.1016/j.geoderma.2012.08.003
Poeplau, C., Don, A., Six, J., Kaiser, M., Benbi, D., Chenu, C., Cotrufo, M.F., Derrien, D., Gioacchini, P., Grand, S., Gregorich, E., Griepentrog, M., Gunina, A., Haddix, M., Kuzyakov, Y., Kühnel, A., Macdonald, L.M., Soong, J., Trigalet, S., Vermeire, M.-L., Rovira, P., van Wesemael, B., Wiesmeier, M., Yeasmin, S., Yevdokimov, I., Nieder, R., 2018. Isolating organic carbon fractions with varying turnover rates in temperate agricultural soils – A comprehensive method comparison. Soil Biology and Biochemistry 125, 10–26. doi:10.1016/j.soilbio.2018.06.025
Poeplau, C., Jacobs, A., Don, A., Vos, C., Schneider, F., Wittnebel, M., Tiemeyer, B., Heidkamp, A., Prietz, R., Flessa, H., 2020a. Stocks of organic carbon in German agricultural soils—Key results of the first comprehensive inventory. Journal of Plant Nutrition and Soil Science 183, 665–681. doi:10.1002/jpln.202000113
Poeplau, C., Kätterer, T., Leblans, N.I.W., Sigurdsson, B.D., 2017. Sensitivity of soil carbon fractions and their specific stabilization mechanisms to extreme soil warming in a subarctic grassland. Global Change Biology 23, 1316–1327. doi:10.1111/gcb.13491
Poeplau, C., Sigurðsson, P., Sigurdsson, B.D., 2020b. Depletion of soil carbon and aggregation after strong warming of a subarctic Andosol under forest and grassland cover. SOIL 6, 115–129. doi:10.5194/soil-6-115-2020
Pold, G., Domeignoz-Horta, L.A., Morrison, E.W., Frey, S.D., Sistla, S.A., DeAngelis, K.M., 2020. Carbon Use Efficiency and Its Temperature Sensitivity Covary in Soil Bacteria. MBio 11, e02293-19. doi:10.1128/mBio.02293-19
Possinger, A.R., Zachman, M.J., Enders, A., Levin, B.D.A., Muller, D.A., Kourkoutis, L.F., Lehmann, J., 2020. Organo–organic and organo–mineral interfaces in soil at the nanometer scale. Nature Communications 11, 6103. doi:10.1038/s41467-020-19792-9
Powell, J.R., Welsh, A., Hallin, S., 2015. Microbial functional diversity enhances predictive models linking environmental parameters to ecosystem properties. Ecology 96, 1985–1993. doi:10.1890/14-1127.1
Props, R., Kerckhof, F.-M., Rubbens, P., De Vrieze, J., Hernandez Sanabria, E., Waegeman, W., Monsieurs, P., Hammes, F., Boon, N., 2017. Absolute quantification of microbial taxon abundances. The ISME Journal 11, 584–587. doi:10.1038/ismej.2016.117
Puissant, J., Jones, B., Goodall, T., Mang, D., Blaud, A., Gweon, H.S., Malik, A., Jones, D.L., Clark, I.M., Hirsch, P.R., Griffiths, R., 2019. The pH optimum of soil exoenzymes adapt to long term changes in soil pH. Soil Biology and Biochemistry 138, 107601. doi:10.1016/j.soilbio.2019.107601
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 41, D590–D596. doi:10.1093/nar/gks1219
Raich, J.W., Potter, C.S., Bhagawati, D., 2002. Interannual variability in global soil respiration, 1980-94: GLOBAL SOIL RESPIRATION 1980-94. Global Change Biology 8, 800–812. doi:10.1046/j.1365-2486.2002.00511.x
Ramoneda, J., Stallard-Olivera, E., Hoffert, M., Winfrey, C.C., Stadler, M., Niño-García, J.P., Fierer, N., 2023. Building a genome-based understanding of bacterial pH preferences. SCIENCE ADVANCES.
Ranjard, L., Richaume, A., 2001. Quantitative and qualitative microscale distribution of bacteria in soil. Research in Microbiology 152, 707–716. doi:10.1016/S0923-2508(01)01251-7
Rasmussen, C., Heckman, K., Wieder, W.R., Keiluweit, M., Lawrence, C.R., Berhe, A.A., Blankinship, J.C., Crow, S.E., Druhan, J.L., Hicks Pries, C.E., Marin-Spiotta, E., Plante, A.F., Schädel, C., Schimel, J.P., Sierra, C.A., Thompson, A., Wagai, R., 2018. Beyond clay: towards an improved set of variables for predicting soil organic matter content. Biogeochemistry 137, 297–306. doi:10.1007/s10533-018-0424-3

Rasmussen, C., Torn, M.S., Southard, R.J., 2005. Mineral Assemblage and Aggregates Control Carbon Dynamics in a California Conifer Forest. Soil Science Society of America Journal 69, 1711–1721. doi:10.2136/sssaj2005.0040

Ratzke, C., Barrere, J., Gore, J., 2020. Strength of species interactions determines biodiversity and stability in microbial communities. Nature Ecology & Evolution 4, 376–383. doi:10.1038/s41559-020-1099-4

Raynaud, X., Nunan, N., 2014. Spatial Ecology of Bacteria at the Microscale in Soil. PLoS ONE 9, e87217. doi:10.1371/journal.pone.0087217

Regelink, I.C., Stoof, C.R., Rousseva, S., Weng, L., Lair, G.J., Kram, P., Nikolaidis, N.P., Kercheva, M., Banwart, S., Comans, R.N.J., 2015. Linkages between aggregate formation, porosity and soil chemical properties. Geoderma 247–248, 24–37. doi:10.1016/j.geoderma.2015.01.022

Reichenbach, M., Fiener, P., Hoyt, A., Trumbore, S., Six, J., Doetterl, S., 2023. Soil carbon stocks in stable tropical landforms are dominated by geochemical controls and not by land use. Global Change Biology 29, 2591–2607. doi:10.1111/gcb.16622

Revelle, W., 2022. psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. R package version 2.2.5.

Richter-Heitmann, T., Hofner, B., Krah, F.-S., Sikorski, J., Wüst, P.K., Bunk, B., Huang, S., Regan, K.M., Berner, D., Boeddinghaus, R.S., Marhan, S., Prati, D., Kandeler, E., Overmann, J., Friedrich, M.W., 2020. Stochastic Dispersal Rather Than Deterministic Selection Explains the Spatio-Temporal Distribution of Soil Bacteria in a Temperate Grassland. Frontiers in Microbiology 11, 1391. doi:10.3389/fmicb.2020.01391

Ritz, C., Baty, F., Streibig, J.C., Gerhard, D., 2015. Dose-response analysis using R. PLoS One. doi:10.1371/journal.pone.0146021

Rocca, J.D., Hall, E.K., Lennon, J.T., Evans, S.E., Waldrop, M.P., Cotner, J.B., Nemergut, D.R., Graham, E.B., Wallenstein, M.D., 2015. Relationships between protein-encoding gene abundance and corresponding process are commonly assumed yet rarely observed. The ISME Journal 9, 1693–1699. doi:10.1038/ismej.2014.252

Rocci, K.S., Lavallee, J.M., Stewart, C.E., Cotrufo, M.F., 2021. Soil organic carbon response to global environmental change depends on its distribution between mineral-associated and particulate organic matter: A meta-analysis. Science of The Total Environment 793, 148569. doi:10.1016/j.scitotenv.2021.148569

Roller, B.R.K., Schmidt, T.M., 2015. The physiology and ecological implications of efficient growth. The ISME Journal 9, 1481–1487. doi:10.1038/ismej.2014.235

Roller, B.R.K., Stoddard, S.F., Schmidt, T.M., 2016. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. Nature Microbiology 1, 16160. doi:10.1038/nmicrobiol.2016.160

Roth, V.-N., Dittmar, T., Gaupp, R., Gleixner, G., 2015. The Molecular Composition of Dissolved Organic Matter in Forest Soils as a Function of pH and Temperature. PLOS ONE 10, e0119188. doi:10.1371/journal.pone.0119188

Roth, V.-N., Lange, M., Simon, C., Hertkorn, N., Bucher, S., Goodall, T., Griffiths, R.I., Mellado-Vázquez, P.G., Mommer, L., Oram, N.J., Weigelt, A., Dittmar, T., Gleixner, G., 2019. Persistence of dissolved organic matter explained by molecular changes during its passage through soil. Nature Geoscience 12, 755–761. doi:10.1038/s41561-019-0417-4

Rousk, J., Bååth, E., 2011. Growth of saprotrophic fungi and bacteria in soil: Growth of saprotrophic fungi and bacteria in soil. FEMS Microbiology Ecology 78, 17–30. doi:10.1111/j.1574-6941.2011.01106.x

Rousk, J., Bååth, E., Brookes, P.C., Lauber, C.L., Lozupone, C., Caporaso, J.G., Knight, R., Fierer, N., 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. The ISME Journal 4, 1340–1351. doi:10.1038/ismej.2010.58

Rowley, M.C., Grand, S., Verrecchia, É.P., 2018. Calcium-mediated stabilisation of soil organic carbon. Biogeochemistry 137, 27–49. doi:10.1007/s10533-017-0410-1

RStudio Team, 2020. RStudio: Integrated Development for R.

Running, S.W., Zhao, M., 2019. User's Guide Daily GPP and Annual NPP (MOD17A2H/A3H) and Year-end Gap-Filled (MOD17A2HGF/A3HGF) Products NASA Earth Observing System MODIS Land Algorithm (For Collection 6), Version 4.2.

Russell, J.B., Cook, G.M., 1995. Energetics of Bacterial Growth: Balance of Anabolic and Catabolic Reactions. MICROBIOL. REV. 59.

Ryals, R., Kaiser, M., Torn, M.S., Berhe, A.A., Silver, W.L., 2014. Impacts of organic matter amendments on carbon and nitrogen dynamics in grassland soils. Soil Biology and Biochemistry 68, 52–61. doi:10.1016/j.soilbio.2013.09.011

Saifuddin, M., Bhatnagar, J.M., Segrè, D., Finzi, A.C., 2019. Microbial carbon use efficiency predicted from genome-scale metabolic models. Nature Communications 10, 3568. doi:10.1038/s41467-019-11488-z

Salomé, C., Nunan, N., Pouteau, V., Lerch, T.Z., Chenu, C., 2010. Carbon dynamics in topsoil and in subsoil may be controlled by different regulatory mechanisms. Global Change Biology 16, 416–426. doi:10.1111/j.1365-2486.2009.01884.x

Sanderman, J., Amundson, R., 2008. A comparative study of dissolved organic carbon transport and stabilization in California forest and grassland soils. Biogeochemistry 89, 309–327. doi:10.1007/s10533-008-9221-8

Sanderman, J., Baldock, J.A., Dangal, S.R.S., Ludwig, S., Potter, S., Rivard, C., Savage, K., 2021. Soil organic carbon fractions in the Great Plains of the United States: an application of mid-infrared spectroscopy. Biogeochemistry 156, 97–114. doi:10.1007/s10533-021-00755-1

Sanderman, J., Hengl, T., Fiske, G.J., 2017. Soil carbon debt of 12,000 years of human land use. Proceedings of the National Academy of Sciences 114, 9575–9580. doi:10.1073/pnas.1706103114

Sanderman, J., Maddern, T., Baldock, J., 2014. Similar composition but differential stability of mineral retained organic matter across four classes of clay minerals. Biogeochemistry 121, 409–424. doi:10.1007/s10533-014-0009-8

Sayer, E.J., Leitman, S.F., Wright, S.J., Rodtassana, C., Vincent, A.G., Bréchet, L.M., Castro, B., Lopez, O., Wallwork, A., Tanner, E.V.J., 2024. Tropical forest above-ground productivity is maintained by nutrients cycled in litter. Journal of Ecology 1365-2745.14251. doi:10.1111/1365-2745.14251

Sayer, E.J., Lopez-Sangil, L., Crawford, J.A., Bréchet, L.M., Birkett, A.J., Baxendale, C., Castro, B., Rodtassana, C., Garnett, M.H., Weiss, L., Schmidt, M.W.I., 2019. Tropical forest soil carbon stocks do not increase despite 15 years of doubled litter inputs. Scientific Reports 9, 18030. doi:10.1038/s41598-019-54487-2

Schaetzl, R.J., 2013. Catenas and Soils. In: Shroder, J, Pope, G A (Ed.). Treatise on Geomorphology., in: Treatise on Geomorphology. Academic Press, San Diega, CA, pp. 145–158. doi:10.1016/B978-0-12-374739-6.00074-9

Schiedung, M., Bellè, S.-L., Hoeschen, C., Schweizer, S.A., Abiven, S., 2023a. Enhanced loss but limited mobility of pyrogenic and organic matter in continuous permafrost-affected forest soils. Soil Biology and Biochemistry 178, 108959. doi:10.1016/j.soilbio.2023.108959

Schiedung, M., Don, A., Beare, M.H., Abiven, S., 2023b. Soil carbon losses due to priming moderated by adaptation and legacy effects. Nature Geoscience 16, 909–914. doi:10.1038/s41561-023-01275-3

Schimel, J., 2023. Modeling ecosystem-scale carbon dynamics in soil: The microbial dimension. Soil Biology and Biochemistry 178, 108948. doi:10.1016/j.soilbio.2023.108948

Schimel, J., Becerra, C.A., Blankinship, J., 2017. Estimating decay dynamics for enzyme activities in soils from different ecosystems. Soil Biology and Biochemistry 114, 5–11. doi:10.1016/j.soilbio.2017.06.023

Schimel, J., Weintraub, M.N., Moorhead, D., 2022. Estimating microbial carbon use efficiency in soil: Isotope-based and enzyme-based methods measure fundamentally different aspects of microbial resource use. Soil Biology and Biochemistry 169, 108677. doi:10.1016/j.soilbio.2022.108677

Schimel, J.P., 2018. Life in Dry Soils: Effects of Drought on Soil Microbial Communities and Processes. Annual Review of Ecology, Evolution, and Systematics 49, 409–432. doi:10.1146/annurev-ecolsys-110617-062614

Schipper, L.A., Hobbs, J.K., Rutledge, S., Arcus, V.L., 2014. Thermodynamic theory explains the temperature optima of soil microbial processes and high $Q_{10}$ values at low temperatures. Global Change Biology 20, 3578–3586. doi:10.1111/gcb.12596

Schloss, P.D., 2023. Rarefaction is currently the best approach to control for uneven sequencing effort in amplicon sequence analyses (preprint). Microbiology. doi:10.1101/2023.06.23.546313

Schnecker, J., Baldaszti, L., Gündler, P., Pleitner, M., Sandén, T., Simon, E., Spiegel, F., Spiegel, H., Urbina Malo, C., Zechmeister-Boltenstern, S., Richter, A., 2023. Seasonal dynamics of soil microbial growth, respiration, biomass, and carbon use efficiency in temperate soils. Geoderma 440, 116693. doi:10.1016/j.geoderma.2023.116693

Schnecker, J., Bowles, T., Hobbie, E.A., Smith, R.G., Grandy, A.S., 2019. Substrate quality and concentration control decomposition and microbial strategies in a model soil system. Biogeochemistry 144, 47–59. doi:10.1007/s10533-019-00571-8

Schnecker, J., Wild, B., Takriti, M., Eloy Alves, R.J., Gentsch, N., Gittel, A., Hofer, A., Klaus, K., Knoltsch, A., Lashchinskiy, N., Mikutta, R., Richter, A., 2015. Microbial community composition shapes enzyme patterns in topsoil and subsoil horizons along a latitudinal transect in Western Siberia. Soil Biology and Biochemistry 83, 106–115. doi:10.1016/j.soilbio.2015.01.016

Schrumpf, M., Kaiser, K., Guggenberger, G., Persson, T., Kögel-Knabner, I., Schulze, E.-D., 2013. Storage and stability of organic carbon in soils as related to depth, occlusion within aggregates, and attachment to minerals. Biogeosciences 10, 1675–1691. doi:10.5194/bg-10-1675-2013

Schwaerzel, K., Seidling, W., Hansen, K., Strich, S., Lorenz, M., 2022. Part I: Objectives, Strategy and Implementation of ICP Forests. Version 2022-2. In: UNECE ICP Forests Programme Coordinating Centre (ed.): Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Thünen Institute of Forest Ecosystems, Eberswalde, Germany,.

Schweizer, S.A., 2022. Perspectives from the Fritz-Scheffer Awardee 2021: Soil organic matter storage and functions determined by patchy and piled-up arrangements at the microscale. Journal of Plant Nutrition and Soil Science jpln.202200217. doi:10.1002/jpln.202200217

Sebag, D., Verrecchia, E.P., Cécillon, L., Adatte, T., Albrecht, R., Aubert, M., Bureau, F., Cailleau, G., Copard, Y., Decaens, T., Disnar, J.-R., Hetényi, M., Nyilas, T., Trombino, L., 2016. Dynamics of soil organic matter based on new Rock-Eval indices. Geoderma 284, 185–203. doi:10.1016/j.geoderma.2016.08.025

Segoli, M., De Gryze, S., Dou, F., Lee, J., Post, W.M., Denef, K., Six, J., 2013. AggModel: A soil organic matter model with measurable pools for use in incubation studies. Ecological Modelling 263, 1–9. doi:10.1016/j.ecolmodel.2013.04.010

Sepúlveda, M. de los Á., Hidalgo, M., Araya, J., Casanova, M., Muñoz, C., Doetterl, S., Wasner, D., Colpaert, B., Bodé, S., Boeckx, P., Zagal, E., 2021. Near-infrared spectroscopy: Alternative method for assessment of stable carbon isotopes in various soil profiles in Chile. Geoderma Regional 25, e00397. doi:10.1016/j.geodrs.2021.e00397

Serna-Chavez, H.M., Fierer, N., van Bodegom, P.M., 2013. Global drivers and patterns of microbial abundance in soil. Global Ecology and Biogeography 22, 1162–1172. doi:10.1111/geb.12070

Sheng, Y., Dong, H., Coffin, E., Myrold, D., Kleber, M., 2022. The Important Role of Enzyme Adsorbing Capacity of Soil Minerals in Regulating β-Glucosidase Activity. Geophysical Research Letters 49, e2021GL097556. doi:10.1029/2021GL097556

Shi, Z., Allison, S.D., He, Y., Levine, P.A., Hoyt, A.M., Beem-Miller, J., Zhu, Q., Wieder, W.R., Trumbore, S., Randerson, J.T., 2020. The age distribution of global soil carbon inferred from radiocarbon measurements. Nature Geoscience 13, 555–559. doi:10.1038/s41561-020-0596-z

Shi, Z., Crowell, S., Luo, Y., Moore, B., 2018. Model structures amplify uncertainty in predicted soil carbon responses to climate change. Nature Communications 9, 2171. doi:10.1038/s41467-018-04526-9

Siciliano, S.D., Palmer, A.S., Winsley, T., Lamb, E., Bissett, A., Brown, M.V., van Dorst, J., Ji, M., Ferrari, B.C., Grogan, P., Chu, H., Snape, I., 2014. Soil fertility is associated with fungal and bacterial richness, whereas pH is associated with community composition in polar soil microbial communities. Soil Biology and Biochemistry 78, 10–20. doi:10.1016/j.soilbio.2014.07.005

Sierra, C.A., Ahrens, B., Bolinder, M.A., Braakhekke, M.C., von Fromm, S., Kätterer, T., Luo, Z., Parvin, N., Wang, G., 2024. Carbon sequestration in the subsoil and the time required to stabilize carbon for climate change mitigation. Global Change Biology 30, e17153. doi:10.1111/gcb.17153

Sierra, C.A., Müller, M., Metzler, H., Manzoni, S., Trumbore, S.E., 2017. The muddle of ages, turnover, transit, and residence times in the carbon cycle. Global Change Biology 23, 1763–1773. doi:10.1111/gcb.13556

Sierra, C.A., Trumbore, S.E., Davidson, E.A., Vicca, S., Janssens, I., 2015. Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture. Journal of Advances in Modeling Earth Systems 7, 335–356. doi:10.1002/2014MS000358

Simon, E., Canarini, A., Martin, V., Séneca, J., Böckle, T., Reinthaler, D., Pötsch, E.M., Piepho, H.-P., Bahn, M., Wanek, W., Richter, A., 2020. Microbial growth and carbon use efficiency show seasonal responses in a multifactorial climate change experiment. Communications Biology 3, 584. doi:10.1038/s42003-020-01317-1

Simpson, M.J., Simpson, A.J., 2012. The Chemical Ecology of Soil Organic Matter Molecular Constituents. Journal of Chemical Ecology 38, 768–784. doi:10.1007/s10886-012-0122-x

Sims, G.K., Sommers, L.E., Konopka, A., 1986. Degradation of Pyridine by *Micrococcus luteus* Isolated from Soil. Applied and Environmental Microbiology 51, 963–968. doi:10.1128/aem.51.5.963-968.1986

Sinsabaugh, R.L., 2010. Phenol oxidase, peroxidase and organic matter dynamics of soil. Soil Biology and Biochemistry 42, 391–404. doi:10.1016/j.soilbio.2009.10.014

Sinsabaugh, R.L., Belnap, J., Findlay, S.G., Shah, J.J.F., Hill, B.H., Kuehn, K.A., Kuske, C.R., Litvak, M.E., Martinez, N.G., Moorhead, D.L., Warnock, D.D., 2014. Extracellular enzyme kinetics scale with resource availability. Biogeochemistry 121, 287–304. doi:10.1007/s10533-014-0030-y

Sinsabaugh, R.L., Lauber, C.L., Weintraub, M.N., Ahmed, B., Allison, S.D., Crenshaw, C., Contosta, A.R., Cusack, D., Frey, S., Gallo, M.E., Gartner, T.B., Hobbie, S.E., Holland, K., Keeler, B.L., Powers, J.S., Stursova, M., Takacs-Vesbach, C., Waldrop, M.P., Wallenstein, M.D., Zak, D.R., Zeglin, L.H., 2008. Stoichiometry of soil enzyme activity at global scale: Stoichiometry of soil enzyme activity. Ecology Letters 11, 1252–1264. doi:10.1111/j.1461-0248.2008.01245.x

Sinsabaugh, R.L., Turner, B.L., Talbot, J.M., Waring, B.G., Powers, J.S., Kuske, C.R., Moorhead, D.L., Follstad Shah, J.J., 2016. Stoichiometry of microbial carbon use efficiency in soils. Ecological Monographs 86, 172–189. doi:10.1890/15-2110.1

Sistla, S.A., Rastetter, E.B., Schimel, J.P., 2014. Responses of a tundra system to warming using SCAMPS: a stoichiometrically coupled, acclimating microbe–plant–soil model. Ecological Monographs 84, 151–170. doi:10.1890/12-2119.1

Six, J., Bossuyt, H., Degryze, S., Denef, K., 2004. A history of research on the link between (micro)aggregates, soil biota, and soil organic matter dynamics. Soil and Tillage Research 79, 7–31. doi:10.1016/j.still.2004.03.008

Six, J, Elliott, E.T., Paustian, K., 2000. Soil macroaggregate turnover and microaggregate formation: a mechanism for C sequestration under no-tillage agriculture. Soil Biology and Biochemistry 32, 2099–2103. doi:10.1016/S0038-0717(00)00179-6

Six, J., Elliott, E.T., Paustian, K., 2000. Soil Structure and Soil Organic Matter II. A Normalized Stability Index and the Effect of Mineralogy. Soil Science Society of America Journal 64, 1042–1049. doi:10.2136/sssaj2000.6431042x

Slessarev, E.W., Chadwick, O.A., Sokol, N.W., Nuccio, E.E., Pett-Ridge, J., 2022. Rock weathering controls the potential for soil carbon storage at a continental scale. Biogeochemistry 157, 1–13. doi:10.1007/s10533-021-00859-8

Slessarev, E.W., Lin, Y., Bingham, N.L., Johnson, J.E., Dai, Y., Schimel, J.P., Chadwick, O.A., 2016. Water balance creates a threshold in soil pH at the global scale. Nature 540, 567–569. doi:10.1038/nature20139

Smercina, D.N., Bailey, V.L., Hofmockel, K.S., 2021. Micro on a macroscale: relating microbial-scale soil processes to global ecosystem function. FEMS Microbiology Ecology 97, fiab091. doi:10.1093/femsec/fiab091

Smith, L.C., Orgiazzi, A., Eisenhauer, N., Cesarz, S., Lochner, A., Jones, A., Bastida, F., Patoine, G., Reitz, T., Buscot, F., Rillig, M.C., Heintz-Buschart, A., Lehmann, A., Guerra, C.A., 2021. Large-scale drivers of relationships between soil microbial properties and organic carbon across Europe. Global Ecology and Biogeography 30, 2070–2083. doi:10.1111/geb.13371

Smith, T.P., Clegg, T., Bell, T., Pawar, S., 2021. Systematic variation in the temperature dependence of bacterial carbon use efficiency. Ecology Letters 24, 2123–2133. doi:10.1111/ele.13840

Soares, M., Rousk, J., 2019. Microbial growth and carbon use efficiency in soil: Links to fungal-bacterial dominance, SOC-quality and stoichiometry. Soil Biology and Biochemistry 131, 195–205. doi:10.1016/j.soilbio.2019.01.010

Sokol, N.W., Bradford, M.A., 2019. Microbial formation of stable soil carbon is more efficient from belowground than aboveground input. Nature Geoscience 12, 46–53. doi:10.1038/s41561-018-0258-6

Sokol, N.W., Sanderman, J., Bradford, M.A., 2019. Pathways of mineral-associated soil organic matter formation: Integrating the role of plant carbon source, chemistry, and point of entry. Global Change Biology 25, 12–24. doi:10.1111/gcb.14482

Sokol, N.W., Slessarev, E., Marschmann, G.L., Nicolas, A., Blazewicz, S.J., Brodie, E.L., Firestone, M.K., Foley, M.M., Hestrin, R., Hungate, B.A., Koch, B.J., Stone, B.W., Sullivan, M.B., Zablocki, O., LLNL Soil Microbiome Consortium, Trubl, G., McFarlane, K., Stuart, R., Nuccio, E., Weber, P., Jiao, Y., Zavarin, M., Kimbrel, J., Morrison, K., Adhikari, D., Bhattacharaya, A., Nico, P., Tang, J., Didonato, N., Paša-Tolić, L., Greenlon, A., Sieradzki, E.T., Dijkstra, P., Schwartz, E., Sachdeva, R., Banfield, J., Pett-Ridge, J., 2022a. Life and death in the soil microbiome: how ecological processes influence biogeochemistry. Nature Reviews Microbiology 20, 415–430. doi:10.1038/s41579-022-00695-z

Sokol, N.W., Whalen, E.D., Jilling, A., Kallenbach, C., Pett-Ridge, J., Georgiou, K., 2022b. Global distribution, formation and fate of mineral-associated soil organic matter under a changing climate: A trait-based perspective. Functional Ecology 36, 1411–1429. doi:10.1111/1365-2435.14040

Sollins, P., Kramer, M.G., Swanston, C., Lajtha, K., Filley, T., Aufdenkampe, A.K., Wagai, R., Bowden, R.D., 2009. Sequential density fractionation across soils of contrasting mineralogy: evidence for both microbial- and mineral-controlled soil organic matter stabilization. Biogeochemistry 96, 209–231. doi:10.1007/s10533-009-9359-z

Sommer, M., Schlichting, E., 1997. Archetypes of catenas in respect to matter — a concept for structuring and grouping catenas. Geoderma 76, 1–33. doi:10.1016/S0016-7061(96)00095-X

Soong, J.L., Fuchslueger, L., Marañon-Jimenez, S., Torn, M.S., Janssens, I.A., Penuelas, J., Richter, A., 2020. Microbial carbon limitation: The need for integrating microorganisms into our understanding of ecosystem carbon cycling. Global Change Biology 26, 1953–1961. doi:10.1111/gcb.14962

Spohn, M., Klaus, K., Wanek, W., Richter, A., 2016. Microbial carbon use efficiency and biomass turnover times depending on soil depth – Implications for carbon cycling. Soil Biology and Biochemistry 96, 74–81. doi:10.1016/j.soilbio.2016.01.016

Spycher, G., Sollins, P., Rose, S., 1983. Carbon and nitrogen in the light fraction of a forest soil: vertical distribution and seasonal patterns. Soil Science 135, 79–87.

Stevens, A., Ramirez-Lopez, L., 2020. Prospectr: Miscellaneous Func- tions for Processing and Sample Selection of Spectroscopic Data, r package version 0.2.0.

Stone, B.W.G., Dijkstra, P., Finley, B.K., Fitzpatrick, R., Foley, M.M., Hayer, M., Hofmockel, K.S., Koch, B.J., Li, J., Liu, X.J.A., Martinez, A., Mau, R.L., Marks, J., Monsaint-Queeney, V., Morrissey, E.M., Propster, J., Pett-Ridge, J., Purcell, A.M., Schwartz, E., Hungate, B.A., 2023. Life history strategies among soil bacteria—dichotomy for few, continuum for many. The ISME Journal 17, 611–619. doi:10.1038/s41396-022-01354-0

Stoner, S., Trumbore, S.E., González-Pérez, J.A., Schrumpf, M., Sierra, C.A., Hoyt, A.M., Chadwick, O., Doetterl, S., 2023. Relating mineral–organic matter stabilization mechanisms to carbon quality and age distributions using ramped thermal analysis. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 381, 20230139. doi:10.1098/rsta.2023.0139

Stoner, S.W., Schrumpf, M., Hoyt, A., Sierra, C.A., Doetterl, S., Galy, V., Trumbore, S., 2023. How well does ramped thermal oxidation quantify the age distribution of soil carbon? Assessing thermal stability of physically and chemically fractionated soil organic matter. Biogeosciences 20, 3151–3163. doi:10.5194/bg-20-3151-2023

Storey, S., Ashaari, M.M., Clipson, N., Doyle, E., de Menezes, A.B., 2018. Opportunistic Bacteria Dominate the Soil Microbiome Response to Phenanthrene in a Microcosm-Based Study. Frontiers in Microbiology 9, 2815. doi:10.3389/fmicb.2018.02815

Strickland, M.S., Rousk, J., 2010. Considering fungal:bacterial dominance in soils – Methods, controls, and ecosystem implications. Soil Biology and Biochemistry 42, 1385–1395. doi:10.1016/j.soilbio.2010.05.007

Sulman, B.N., Moore, J.A.M., Abramoff, R., Averill, C., Kivlin, S., Georgiou, K., Sridhar, B., Hartman, M.D., Wang, G., Wieder, W.R., Bradford, M.A., Luo, Y., Mayes, M.A., Morrison, E., Riley, W.J., Salazar, A., Schimel, J.P., Tang, J., Classen, A.T., 2018. Multiple models and experiments underscore large uncertainty in soil carbon dynamics. Biogeochemistry 141, 109–123. doi:10.1007/s10533-018-0509-z

Sulman, B.N., Phillips, R.P., Oishi, A.C., Shevliakova, E., Pacala, S.W., 2014. Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO2. Nature Climate Change 4, 1099–1102. doi:10.1038/nclimate2436

Suttie, J.M., Reynolds, S.G., Batello, C., 2005. Grasslands of the World, FAO.

Szoboszlay, M., Dohrmann, A.B., Poeplau, C., Don, A., Tebbe, C.C., 2017. Impact of land-use change and soil organic carbon quality on microbial diversity in soils across Europe. FEMS Microbiology Ecology 93. doi:10.1093/femsec/fix146

Tang, J., Riley, W.J., 2015. Weaker soil carbon–climate feedbacks resulting from microbial and abiotic interactions. Nature Climate Change 5, 56–60. doi:10.1038/nclimate2438

Tao, F., Huang, Y., Hungate, B.A., Manzoni, S., Frey, S.D., Schmidt, M.W.I., Reichstein, M., Carvalhais, N., Ciais, P., Jiang, L., Lehmann, J., Wang, Y.-P., Houlton, B.Z., Ahrens, B., Mishra, U., Hugelius, G., Hocking, T.D., Lu, X., Shi, Z., Viatkin, K., Vargas, R., Yigini, Y., Omuto, C., Malik, A.A., Peralta, G., Cuevas-Corona, R., Di Paolo, L.E., Luotto, I., Liao, C., Liang, Y.-S., Saynes, V.S., Huang, X., Luo, Y., 2023. Microbial carbon use efficiency promotes global soil carbon storage. Nature 618, 981–985. doi:10.1038/s41586-023-06042-3

Tatsukami, Y., Nambu, M., Morisaka, H., Kuroda, K., Ueda, M., 2013. Disclosure of the differences of Mesorhizobium loti under the free-living and symbiotic conditions by comparative proteome analysis without bacteroid isolation. BMC Microbiology 13, 180. doi:10.1186/1471-2180-13-180

Tecon, R., Or, D., 2017. Biophysical processes supporting the diversity of microbial life in soil. FEMS Microbiology Reviews 41, 599–623. doi:10.1093/femsre/fux039

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N.S., Wijesundera, R., Ruiz, L.V., Vasco-Palacios, A.M., Thu, P.Q., Suija, A., Smith, M.E., Sharp, C., Saluveer, E., Saitta, A., Rosas, M., Riit, T., Ratkowsky, D., Pritsch, K., Põldmaa, K., Piepenbring, M., Phosri, C., Peterson, M., Parts, K., Pärtel, K., Otsing, E., Nouhra, E., Njouonkou, A.L., Nilsson, R.H., Morgado, L.N.,

Mayor, J., May, T.W., Majuakim, L., Lodge, D.J., Lee, S.S., Larsson, K.-H., Kohout, P., Hosaka, K., Hiiesalu, I., Henkel, T.W., Harend, H., Guo, L., Greslebin, A., Grelet, G., Geml, J., Gates, G., Dunstan, W., Dunk, C., Drenkhan, R., Dearnaley, J., De Kesel, A., Dang, T., Chen, X., Buegger, F., Brearley, F.Q., Bonito, G., Anslan, S., Abell, S., Abarenkov, K., 2014. Global diversity and geography of soil fungi. Science 346, 1256688. doi:10.1126/science.1256688

Tennekes, M., 2018. tmap: Thematic Maps in R. doi:10.18637/jss.v084.i06

Tian, H., Lu, C., Yang, J., Banger, K., Huntzinger, D.N., Schwalm, C.R., Michalak, A.M., Cook, R., Ciais, P., Hayes, D., Huang, M., Ito, A., Jain, A.K., Lei, H., Mao, J., Pan, S., Post, W.M., Peng, S., Poulter, B., Ren, W., Ricciuto, D., Schaefer, K., Shi, X., Tao, B., Wang, W., Wei, Y., Yang, Q., Zhang, B., Zeng, N., 2015. Global patterns and controls of soil organic carbon dynamics as simulated by multiple terrestrial biosphere models: Current status and future directions. Global Biogeochemical Cycles 29, 775–792. doi:10.1002/2014GB005021

Tian, P., Zhao, X., Liu, S., Sun, Z., Jing, Y., Wang, Q., 2022. Soil microbial respiration in forest ecosystems along a north-south transect of eastern China: Evidence from laboratory experiments. CATENA 211, 105980. doi:10.1016/j.catena.2021.105980

Tisdall, J.M., Oades, J.M., 1982. Organic matter and water-stable aggregates in soils. Journal of Soil Science 33, 141–163. doi:10.1111/j.1365-2389.1982.tb01755.x

Torgo, L., 2016. Data Mining with R, learning with case studies, 2nd edition. Chapman and Hall/CRC.

Torn, M.S., Chabbi, A., Crill, P., Hanson, P.J., Janssens, I.A., Luo, Y., Pries, C.H., Rumpel, C., Schmidt, M.W.I., Six, J., Schrumpf, M., Zhu, B., 2015. A call for international soil experiment networks for studying, predicting, and managing global change impacts. SOIL 1, 575–582. doi:10.5194/soil-1-575-2015

Torn, M.S., Swanston, C.W., Castanha, C., Trumbore, S.E., 2009. Storage and Turnover of Organic Matter in Soil, in: Senesi, N., Xing, B., Huang, P.M. (Eds.), Biophysico-Chemical Processes Involving Natural Nonliving Organic Matter in Environmental Systems. Wiley, pp. 219–272. doi:10.1002/9780470494950.ch6

Totsche, K.U., Amelung, W., Gerzabek, M.H., Guggenberger, G., Klumpp, E., Knief, C., Lehndorff, E., Mikutta, R., Peth, S., Prechtel, A., Ray, N., Kögel-Knabner, I., 2018. Microaggregates in soils. Journal of Plant Nutrition and Soil Science 181, 104–136. doi:10.1002/jpln.201600451

Trabucco, A., Zomer, R.J., 2018. Global Aridity Index and Potential Evapo-Transpiration (ET0) Climate Database v2 10.

Tripathi, B.M., Stegen, J.C., Kim, M., Dong, K., Adams, J.M., Lee, Y.K., 2018. Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. The ISME Journal 12, 1072–1083. doi:10.1038/s41396-018-0082-4

Trivedi, P., Anderson, I.C., Singh, B.K., 2013. Microbial modulators of soil carbon storage: integrating genomic and metabolic knowledge for global prediction. Trends in Microbiology 21, 641–651. doi:10.1016/j.tim.2013.09.005

Trivedi, P., Delgado-Baquerizo, M., Trivedi, C., Hu, H., Anderson, I.C., Jeffries, T.C., Zhou, J., Singh, B.K., 2016. Microbial regulation of the soil carbon cycle: evidence from gene–enzyme relationships. The ISME Journal 10, 2593–2604. doi:10.1038/ismej.2016.65

Trujillo, M.E., Hong, K., Genilloud, O., 2014. The Family Micromonosporaceae, in: The Prokaryotes: Actinobacteria. Fourth Edition. (Eds: Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E., Thompson, F.). Springer, Berlin, Heidelberg.

Trumbore, S., 2006. Carbon respired by terrestrial ecosystems – recent progress and challenges. Global Change Biology 12, 141–153. doi:10.1111/j.1365-2486.2006.01067.x

Trumbore, S., 2000. AGE OF SOIL ORGANIC MATTER AND SOIL RESPIRATION: RADIOCARBON CONSTRAINTS ON BELOWGROUND C DYNAMICS. Ecological Applications 10, 399–411. doi:10.1890/1051-0761(2000)010[0399:AOSOMA]2.0.CO;2

van Bodegom, P., 2007. Microbial Maintenance: A Critical Review on Its Quantification. Microbial Ecology 53, 513–523. doi:10.1007/s00248-006-9049-5

van der Voort, T.S., Mannu, U., Hagedorn, F., McIntyre, C., Walthert, L., Schleppi, P., Haghipour, N., Eglinton, T.I., 2019. Dynamics of deep soil carbon – insights from <sup>14</sup>C

time series across a climatic gradient. Biogeosciences 16, 3233–3246. doi:10.5194/bg-16-3233-2019
van der Voort, T.S., Zell, C.I., Hagedorn, F., Feng, X., McIntyre, C.P., Haghipour, N., Graf Pannatier, E., Eglinton, T.I., 2017. Diverse Soil Carbon Dynamics Expressed at the Molecular Level. Geophysical Research Letters 44. doi:10.1002/2017GL076188
van Hees, P.A.W., Johansson, E., Jones, D.L., 2008. Dynamics of simple carbon compounds in two forest soils as revealed by soil solution concentrations and biodegradation kinetics. Plant and Soil 310, 11–23. doi:10.1007/s11104-008-9623-3
Vance, E.D., Brookes, P.C., Jenkinson, D.S., 1987. An extraction method for measuring soil microbial biomass C. Soil Biology and Biochemistry 19, 703–707. doi:10.1016/0038-0717(87)90052-6
Virto, I., Barré, P., Chenu, C., 2008. Microaggregation and organic matter storage at the silt-size scale. Geoderma 146, 326–335. doi:10.1016/j.geoderma.2008.05.021
Viscarra Rossel, R.A., Lee, J., Behrens, T., Luo, Z., Baldock, J., Richards, A., 2019. Continental-scale soil carbon composition and vulnerability modulated by regional environmental controls. Nature Geoscience 12, 547–552. doi:10.1038/s41561-019-0373-z
Vogel, C., Mueller, C.W., Höschen, C., Buegger, F., Heister, K., Schulz, S., Schloter, M., Kögel-Knabner, I., 2014. Submicron structures provide preferential spots for carbon and nitrogen sequestration in soils. Nature Communications 5, 2947. doi:10.1038/ncomms3947
von Fromm, S.F., Hoyt, A.M., Lange, M., Acquah, G.E., Aynekulu, E., Berhe, A.A., Haefele, S.M., McGrath, S.P., Shepherd, K.D., Sila, A.M., Six, J., Towett, E.K., Trumbore, S.E., Vågen, T.-G., Weullow, E., Winowiecki, L.A., Doetterl, S., 2021. Continental-scale controls on soil organic carbon across sub-Saharan Africa. SOIL 7, 305–332. doi:10.5194/soil-7-305-2021
von Lützow, M., Kögel-Knabner, I., Ekschmitt, K., Flessa, H., Guggenberger, G., Matzner, E., Marschner, B., 2007. SOM fractionation methods: Relevance to functional pools and to stabilization mechanisms. Soil Biology and Biochemistry 39, 2183–2207. doi:10.1016/j.soilbio.2007.03.007
von Lützow, M., Kögel-Knabner, I., Ekschmitt, K., Matzner, E., Guggenberger, G., Marschner, B., Flessa, H., 2006. Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions – a review. European Journal of Soil Science 57, 426–445. doi:10.1111/j.1365-2389.2006.00809.x
von Lützow, M., Kögel-Knabner, I., Ludwig, B., Matzner, E., Flessa, H., Ekschmitt, K., Guggenberger, G., Marschner, B., Kalbitz, K., 2008. Stabilization mechanisms of organic matter in four temperate soils: Development and application of a conceptual model. Journal of Plant Nutrition and Soil Science 171, 111–124. doi:10.1002/jpln.200700047
Vos, M., Wolf, A.B., Jennings, S.J., Kowalchuk, G.A., 2013. Micro-scale determinants of bacterial diversity in soil. FEMS Microbiology Reviews 37, 936–954. doi:10.1111/1574-6976.12023
Wade, J., Li, C., Vollbracht, K., Hooper, D.G., Wills, S.A., Margenot, A.J., 2021. Prescribed pH for soil β-glucosidase and phosphomonoesterase do not reflect pH optima. Geoderma 401, 115161. doi:10.1016/j.geoderma.2021.115161
Wadoux, A.M.J. -C., Samuel-Rosa, A., Poggio, L., Mulder, V.L., 2020. A note on knowledge discovery and machine learning in digital soil mapping. European Journal of Soil Science 71, 133–136. doi:10.1111/ejss.12909
Wagai, R., Kajiura, M., Asano, M., 2020. Iron and aluminum association with microbially processed organic matter via meso-density aggregate formation across soils: organo-metallic glue hypothesis. SOIL 6, 597–627. doi:10.5194/soil-6-597-2020
Wagai, R., Mayer, L.M., Kitayama, K., 2009. Nature of the “occluded” low-density fraction in soil organic matter studies: A critical review. Soil Science and Plant Nutrition 55, 13–25. doi:10.1111/j.1747-0765.2008.00356.x
Wagai, R., Mayer, L.M., Kitayama, K., Knicker, H., 2008. Climate and parent material controls on organic matter storage in surface soils: A three-pool, density-separation approach. Geoderma 147, 23–33. doi:10.1016/j.geoderma.2008.07.010

Wallenstein, M.D., Haddix, M.L., Ayres, E., Steltzer, H., Magrini-Bair, K.A., Paul, E.A., 2013. Litter chemistry changes more rapidly when decomposed at home but converges during decomposition–transformation. Soil Biology and Biochemistry 57, 311–319. doi:10.1016/j.soilbio.2012.09.027

Wan, J., Crowther, T.W., 2022. Uniting the scales of microbial biogeochemistry with trait-based modelling. Functional Ecology 36, 1457–1472. doi:10.1111/1365-2435.14035

Wang, B., An, S., Liang, C., Liu, Y., Kuzyakov, Y., 2021. Microbial necromass as the source of soil organic carbon in global ecosystems. Soil Biology and Biochemistry 162, 108422. doi:10.1016/j.soilbio.2021.108422

Wang, D., Wang, S., Du, X., He, Q., Liu, Y., Wang, Z., Feng, K., Li, Y., Deng, Y., 2022. ddPCR surpasses classical qPCR technology in quantitating bacteria and fungi in the environment. Molecular Ecology Resources 22, 2587–2598. doi:10.1111/1755-0998.13644

Wang, G., Gao, Q., Yang, Y., Hobbie, S.E., Reich, P.B., Zhou, J., 2022. Soil enzymes as indicators of soil function: A step toward greater realism in microbial ecological modeling. Global Change Biology 28, 1935–1950. doi:10.1111/gcb.16036

Wang, W.J., Dalal, R.C., Moody, P.W., Smith, C.J., 2003. Relationships of soil respiration to microbial biomass, substrate availability and clay content. Soil Biology and Biochemistry 35, 273–284. doi:10.1016/S0038-0717(02)00274-2

Wang, Y., Wilhelm, R.C., Swenson, T.L., Silver, A., Andeer, P.F., Golini, A., Kosina, S.M., Bowen, B.P., Buckley, D.H., Northen, T.R., 2022. Substrate Utilization and Competitive Interactions Among Soil Bacteria Vary With Life-History Strategies. Frontiers in Microbiology 13, 914472. doi:10.3389/fmicb.2022.914472

Wang, Z., Kumar, J., Weintraub-Leff, S.R., Todd-Brown, K., Mishra, U., Sihi, D., 2024. Upscaling Soil Organic Carbon Measurements at the Continental Scale Using Multivariate Clustering Analysis and Machine Learning. Journal of Geophysical Research: Biogeosciences 129, e2023JG007702. doi:10.1029/2023JG007702

Wasner, D., Abramoff, R., Griepentrog, M., Venegas, E.Z., Boeckx, P., Dötterl, S., 2023. Systematic and predictable variation of soil carbon reservoirs along a geoclimatic gradient (preprint). In Review. doi:10.21203/rs.3.rs-2856937/v1

Wei, H., Guenet, B., Vicca, S., Nunan, N., Asard, H., AbdElgawad, H., Shen, W., Janssens, I.A., 2014. High clay content accelerates the decomposition of fresh organic matter in artificial soils. Soil Biology and Biochemistry 77, 100–108. doi:10.1016/j.soilbio.2014.06.006

Wei, T., Simko, V., 2021. R package “corrplot”: Visualization of a Correlation Matrix (Version 0.92).

Weintraub, S.R., Flores, A.N., Wieder, W.R., Sihi, D., Cagnarini, C., Gonçalves, D.R.P., Young, M.H., Li, L., Olshansky, Y., Baatz, R., Sullivan, P.L., Groffman, P.M., 2019. Leveraging Environmental Research and Observation Networks to Advance Soil Carbon Science. Journal of Geophysical Research: Biogeosciences 124, 1047–1055. doi:10.1029/2018JG004956

Whitaker, J., Ostle, N., Nottingham, A.T., Ccahuana, A., Salinas, N., Bardgett, R.D., Meir, P., McNamara, N.P., 2014. Microbial community composition explains soil respiration responses to changing carbon inputs along an A ndes-to- A mazon elevation gradient. Journal of Ecology 102, 1058–1071. doi:10.1111/1365-2745.12247

Wieder, W.R., Allison, S.D., Davidson, E.A., Georgiou, K., Hararuk, O., He, Y., Hopkins, F., Luo, Y., Smith, M.J., Sulman, B., Todd-Brown, K., Wang, Y.-P., Xia, J., Xu, X., 2015. Explicitly representing soil microbial processes in Earth system models: Soil microbes in earth system models. Global Biogeochemical Cycles 29, 1782–1800. doi:10.1002/2015GB005188

Wieder, W.R., Grandy, A.S., Kallenbach, C.M., Bonan, G.B., 2014. Integrating microbial physiology and physio-chemical principles in soils with the MIcrobial-MIneral Carbon Stabilization (MIMICS) model. Biogeosciences 11, 3899–3917. doi:10.5194/bg-11-3899-2014

Wieder, W.R., Pierson, D., Earl, S., Lajtha, K., Baer, S., Ballantyne, F., Berhe, A.A., Billings, S., Brigham, L.M., Chacon, S.S., Fraterrigo, J., Frey, S.D., Georgiou, K., de Graaff, M.-A., Grandy, A.S., Hartman, M.D., Hobbie, S.E., Johnson, C., Kaye, J., Kyker-Snowman, E., Litvak, M.E., Mack, M.C., Malhotra, A., Moore, J.A.M., Nadelhoffer, K., Rasmussen, C., Silver, W.L., Sulman, B.N.,

Walker, X., Weintraub, S., 2021. SoDaH: the SOils DAta Harmonization database, an open-sourcesynthesis of soil data from research networks, version 1.0. doi:10.5194/essd-2020-195
Wieder, W.R., Sulman, B.N., Hartman, M.D., Koven, C.D., Bradford, M.A., 2019. Arctic Soil Governs Whether Climate Change Drives Global Losses or Gains in Soil Carbon. Geophysical Research Letters 46, 14486–14495. doi:10.1029/2019GL085543
Wiesmeier, M., Poeplau, C., Sierra, C.A., Maier, H., Frühauf, C., Hübner, R., Kühnel, A., Spörlein, P., Geuß, U., Hangen, E., Schilling, B., von Lützow, M., Kögel-Knabner, I., 2016. Projected loss of soil organic carbon in temperate agricultural soils in the 21st century: effects of climate change and carbon input trends. Scientific Reports 6, 32525. doi:10.1038/srep32525
Wiesmeier, M., Urbanski, L., Hobley, E., Lang, B., von Lützow, M., Marin-Spiotta, E., van Wesemael, B., Rabot, E., Ließ, M., Garcia-Franco, N., Wollschläger, U., Vogel, H.-J., Kögel-Knabner, I., 2019. Soil organic carbon storage as a key function of soils - A review of drivers and indicators at various scales. Geoderma 333, 149–162. doi:10.1016/j.geoderma.2018.07.026
Wilpiszeski, R.L., Aufrecht, J.A., Retterer, S.T., Sullivan, M.B., Graham, D.E., Pierce, E.M., Zablocki, O.D., Palumbo, A.V., Elias, D.A., 2019. Soil Aggregate Microbial Communities: Towards Understanding Microbiome Interactions at Biologically Relevant Scales. Applied and Environmental Microbiology 85, e00324-19. doi:10.1128/AEM.00324-19
Witzgall, K., Vidal, A., Schubert, D.I., Höschen, C., Schweizer, S.A., Buegger, F., Pouteau, V., Chenu, C., Mueller, C.W., 2021. Particulate organic matter as a functional soil component for persistent soil organic carbon. Nature Communications 12, 4115. doi:10.1038/s41467-021-24192-8
Wolf, A.B., Vos, M., de Boer, W., Kowalchuk, G.A., 2013. Impact of Matric Potential and Pore Size Distribution on Growth Dynamics of Filamentous and Non-Filamentous Soil Bacteria. PLoS ONE 8, e83661. doi:10.1371/journal.pone.0083661
Wu, J., Joergensen, R.G., Pommerening, B., Chaussod, R., Brookes, P.C., 1990. Measurement of soil microbial biomass C by fumigation-extraction—an automated procedure. Soil Biology and Biochemistry 22, 1167–1169. doi:10.1016/0038-0717(90)90046-3
Xu, X., Shi, Z., Li, D., Rey, A., Ruan, H., Craine, J.M., Liang, J., Zhou, J., Luo, Y., 2016. Soil properties control decomposition of soil organic carbon: Results from data-assimilation analysis. Geoderma 262, 235–242. doi:10.1016/j.geoderma.2015.08.038
Yan, Z., Wang, Z., Fu, Z., Zhang, Y., Peng, X., Zheng, J., 2023. Microscale heterogeneity controls macroscopic soil heterotrophic respiration by regulating resource availability and environmental stress. Biogeochemistry 164, 431–449. doi:10.1007/s10533-023-01044-9
Yu, W., Huang, W., Weintraub-Leff, S.R., Hall, S.J., 2022. Where and why do particulate organic matter (POM) and mineral-associated organic matter (MAOM) differ among diverse soils? Soil Biology and Biochemistry 172, 108756. doi:10.1016/j.soilbio.2022.108756
Yu, W., Weintraub, S.R., Hall, S.J., 2021. Climatic and Geochemical Controls on Soil Carbon at the Continental Scale: Interactions and Thresholds. Global Biogeochemical Cycles 35, e2020GB006781. doi:10.1029/2020GB006781
Zeileis, A., Hothorn, T., 2002. Diagnostic Checking in Regression Relationships. R News 2, 7–10.
Zhang, C., Xi, N., 2021. Precipitation Changes Regulate Plant and Soil Microbial Biomass Via Plasticity in Plant Biomass Allocation in Grasslands: A Meta-Analysis. Frontiers in Plant Science 12, 614968. doi:10.3389/fpls.2021.614968
Zhang, Q., Qin, W., Feng, J., Li, X., Zhang, Z., He, J.-S., Schimel, J.P., Zhu, B., 2023. Whole-soil-profile warming does not change microbial carbon use efficiency in surface and deep soils. Proceedings of the National Academy of Sciences 120, e2302190120. doi:10.1073/pnas.2302190120
Zhang, Y., Lavallee, J.M., Robertson, A.D., Even, R., Ogle, S.M., Paustian, K., Cotrufo, M.F., 2021. Simulating measurable ecosystem carbon and nitrogen dynamics with the mechanistically defined MEMS 2.0 model. Biogeosciences 18, 3147–3171. doi:10.5194/bg-18-3147-2021
Zhao, Q., Callister, S.J., Thompson, A.M., Kukkadapu, R.K., Tfaily, M.M., Bramer, L.M., Qafoku, N.P., Bell, S.L., Hobbie, S.E., Seabloom, E.W., Borer, E.T., Hofmockel, K.S., 2020. Strong mineralogic control of soil organic matter composition in response to nutrient addition across diverse

grassland sites. Science of The Total Environment 736, 137839. doi:10.1016/j.scitotenv.2020.137839

Zheng, Q., Hu, Y., Zhang, S., Noll, L., Böckle, T., Dietrich, M., Herbold, C.W., Eichorst, S.A., Woebken, D., Richter, A., Wanek, W., 2019a. Soil multifunctionality is affected by the soil environment and by microbial community composition and diversity. Soil Biology and Biochemistry 136, 107521. doi:10.1016/j.soilbio.2019.107521

Zheng, Q., Hu, Y., Zhang, S., Noll, L., Böckle, T., Richter, A., Wanek, W., 2019b. Growth explains microbial carbon use efficiency across soils differing in land use and geology. Soil Biology and Biochemistry 128, 45–55. doi:10.1016/j.soilbio.2018.10.006

Zhou, J., Deng, Y., Shen, L., Wen, C., Yan, Q., Ning, D., Qin, Y., Xue, K., Wu, L., He, Z., Voordeckers, J.W., Nostrand, J.D.V., Buzzard, V., Michaletz, S.T., Enquist, B.J., Weiser, M.D., Kaspari, M., Waide, R., Yang, Y., Brown, J.H., 2016. Temperature mediates continental-scale diversity of microbes in forest soils. Nature Communications 7, 12083. doi:10.1038/ncomms12083

Zimmermann, M., Leifeld, J., Schmidt, M.W.I., Smith, P., Fuhrer, J., 2007. Measured soil organic matter fractions can be related to pools in the RothC model. European Journal of Soil Science 58, 658–667. doi:10.1111/j.1365-2389.2006.00855.x

Zuccarini, P., Sardans, J., Asensio, L., Peñuelas, J., 2023. Altered activities of extracellular soil enzymes by the interacting global environmental changes. Global Change Biology 29, 2067–2091. doi:10.1111/gcb.16604

# 7. Appendix

## 7.1 Appendix Chapter 1

**Methods for Figure 1-4.** The map of the global distribution of grasslands was based on the Hyde 3.2 dataset of land use estimates (Klein Goldewijk et al., 2017), at a resolution o 5 arc-minutes (9.3 x 9.3 km at the equator). All grids in which managed pastures and rangelands covered in sum ≥ 10 % area in the year 2005 were retained. Climatic data (MAT and MAP) was retrieved from WorldClim Version 2 (Fick and Hijmans, 2017) at a resolution of 5 arc-minutes, and a raster file with the overlap of the range of studied values (MAT 3.0 to 17.1 °C; MAP 217 to 2440 mm) was created. The global distribution of the ten studied WRB major soil groups (Acrisol, Andosol, Arenosol, Cambisol, Chernozem, Gleysol, Kastanozem, Leptosol, Luvisol, Planosol) featured in this gradient were added (Hengl et al., 2017). The raster file of soil groups was aggregated to a resolution of 5 arc-minutes using a majority filter (model). Raster processing was done with the R-package 'raster' (Hijmans, 2022), and the map was created with the R-package 'tmap' (Tennekes, 2018).

## 7.2 Appendix Chapter 2

**Supplementary Text S2-1.** The positive link between $POM_{TOT}$ and base cations (RC5) hints at a positive effect of plant available nutrients on plant productivity, and thus OM input to soil (Figure 2-3a). However, the plant debris derived $POM_{TOT}$ was also negatively correlated to temperature (RC6), which in turn was positively correlated to NPP (Table S2-8). This apparent contradiction points at multiple, in parts conflicting, environmental factors that can control the amount of plant derived, less decomposed OM in soils. First, aboveground NPP might be poorly linked to the actual input of OM into soils. Second, the fate of plant derived OM may be strongly affected by the type of input (Sokol and Bradford, 2019; Sokol et al., 2019; Cotrufo et al., 2022) (aboveground vs. belowground, soluble vs. structural). Third, the degree to which plant OM input and microbial decomposition are constrained by climate can differ across large climatic gradients, as discussed in the main text. However, base cation contents could also be a consequence of POM content, where greater POM contents could yield a greater cation exchange capacity. This interpretation is supported by the lack of a relationship between base cation contents and soil pH (Table S2-8). In this case, the direction of causality in the described relationship would be opposite, and RC6 "Temperature" would be the sole driver of POM: In cooler systems there is more POM.

**Supplementary Text S2-2.** Despite varimax rotation maximizing differences in loading between components, annual water balance (calculated as MAP - PET) and NPP also correlate positively with RC4 "Pedogenic oxides". Vice versa, the Rock-Eval I-index correlates also with RC6 "Temperature". We argue that the mixed loading of these components is indicative of the role of moisture and organic inputs as drivers of mineral weathering (Rasmussen et al., 2018) and in explaining SOC decomposition. On the other hand, NPP likely follows an optimum curve with MAT, as plant productivity becomes increasingly water limited at higher temperatures (Knapp et al., 2014). The majority of soils in this dataset comes from temperate climate zones with most soils from humid soil moisture regimes (udic, n = 18), where no strong effects of water limitation or water logging on biological activity are expected (Table S2-2). Nevertheless, the gradient also features aridic (n =1), ustic (n = 3) and xeric (n = 6) soils, where low water availability could act limiting on plant growth, and perudic soils (n = 7) where water logging could occur seasonally. Note that the annual water balance (MAP-PET) is correlated negatively with temperature (RC6), and positively with pedogenic oxides (RC4) (Table S2-8). However, the annual water balance did not constitute a major independent axis of variation in the gradient (Table S2-8) and thus does not feature as a separate explanatory variable in our regression analysis and subsequently in our conceptual framework. It is rather an indirect feature in our framework as an implicit control on some of the effects on SOC stabilization seen through temperature and pedogenic oxides. For example,

potential drought effects can limit biological activity (Sokol et al., 2022a) towards hot conditions, and temporary water logging in moist cold conditions on the temperature axis. Second, higher soil moisture can favor the formation of pedogenic oxides via increased weathering rates (Slessarev et al., 2022), and thus affects the formation of organo-mineral associations. Further, higher moisture can aid the formation of stable microaggregates, which is influenced by changing redoximorphic conditions and drying-wetting cycles (Six et al., 2004; Totsche et al., 2018). Unsurprisingly, geochemical reactivity (RC1) is positively correlated with finer texture (RC3) and to a lesser extent with warmer temperatures (RC6) (Figure S2-2). The degree of organic matter decomposition in the bulk soil (RC2) is strongly positively correlated with finer texture (RC3) and higher temperature (RC6), and to a lesser extent positively with geochemical reactivity (RC1) and negatively with pedogenic oxide contents (RC4). The link with temperature suggests that OM decomposition in the bulk soil (RC2) may also contain variation of the temperature dimension in the dataset (RC6), where in cooler (and wetter) systems bulk OM is less decomposed. The positive link between geochemical reactivity (RC1) and the degree of OM decomposition at the bulk level (RC2) has been observed before for a weathering chronosequence (Mainka et al., 2022). We further observed a negative link between pedogenic oxides (RC4) and temperature (RC6), which might be a masked effect of moisture. The cooler systems in Patagonia are wetter than the warmer and drier systems of central Chile (Table S2-8), and humic climate enhances mineral weathering (Slessarev et al., 2022). We also observed a negative interaction between pedogenic oxides (RC4) and base cations (RC5) in the predictive model for bulk SOC% (Figure 2-5). We relate this to the fact that in soil systems where base cations are abundant, they have more explanatory content for bulk SOC% than in systems where they are not abundant, coupled with the generally higher abundance of base cations in soil parent material that is also rich in iron, aluminum and manganese. This becomes evident by the results of a simple linear model (SLM) for predicting bulk SOC%, where base cations feature as a weak positive predictor (rVIPedogenic oxides = 0.50, rVITemperature = 0.29, rVIBase cations = 0.22, model description see Table S7). Thus, if base cations play a role in driving bulk SOC%, this could stem either from their role as plant nutrients (see Supplementary Text S2-1), or more directly from an involvement in SOC stabilization via cation-bridging at the higher end of the pH-range (Rasmussen et al., 2018; Rowley et al., 2018). However, please note that base cation contents could also simply be a consequence of POM content: greater POM contents yield a greater cation exchange capacity. This interpretation is supported by the lack of a relationship between base cation contents and pH (Table S2-8). Ultimately, we could not resolve the direction of causality regarding the role of base cations, and this example urges for caution when drawing causal inferences from a correlative analysis. Despite its important role for many biological and weathering related processes (Amelung et al., 2018), soil pH did not constitute a main loading on any RC as it can vary due to contrasting climatic, biological and soil geochemical ecosystem properties (Rasmussen et al., 2018). The strongest

association of pH is with organic matter decomposition (RC2), where soils with a lower degree of decomposition are more acidic (Table S2-8).

**Supplementary Text S2-3.** Other studies identified silt and clay content as a main predictor of SOC contents, usually after full dispersion of stable microaggregates (Matus, 2021). We also observed a positive relationship between $S+C_{TOT}$ and secondary texture (RC2) since in soils with more silt- and clay-sized particles, more SOC was accumulated in the respective fraction. Note that we did not find a direct effect of texture on bulk SOC%. However, our findings do not question this general relationship; as aggregation likely occludes S+C particles, this relationship may be masked in our data because for the regression analysis we used secondary texture that does not disrupt chemically and physically stable (micro)aggregates.

Given that we can show that soil mineralogy is a main driver of $SA_{TOT}$, we caution against oversimplified classification of incompletely dispersed size fractions as POM, without at least including a density separation step. Related to this point, the S+C fraction likely also contains small microaggregates and concretions < 63 µm (Virto et al., 2008). However, since we found average organic matter in S+C and SA to be significantly different in terms of decomposition state (Figure 2-3a), we argue that both fractions on average capture functionally different pools.

We observed an increasing dominance of $S+C_{TOT}$ relative to $POM_{TOT}$ in soils with higher silt and clay content (Figure 2-3a). This was likely caused by more than one mechanism. First, high surface area in finer textured soils is often the result of weathering processes which increase mineral reactivity. This leads to improved conditions for OM stabilization with minerals, thus facilitating a transfer of OM from POM to S+C. Second, the clay fraction can harbor high concentrations of soil microbial biomass (Ranjard and Richaume, 2001; Wei et al., 2014), which could translate into higher activity of microbial SOC decomposition, driving the transformation of POM into forms which can readily get stabilized with minerals. Higher microbial carbon use efficiency as observed in fine textured compared to coarse textured soils (Angst et al., 2021b) could further accelerate this transformation.

**Supplementary Text S2-4.** Carbon stabilization following full microbial transformation of OM (i.e. microbial uptake of substrate, subsequent breakdown through microbial catabolism and reassembly into chemically different molecules by microbial anabolism) can be referred to as the in vivo pathway (Liang et al., 2017). In contrast, SOC stabilization that for various reasons bypasses microbial uptake - and subsequently full microbial catabolism and anabolism -, can be referred to as the *ex vivo* pathway (Liang et al., 2017). Examples for the *ex vivo* pathway are direct stabilization of soluble plant-derived OM, or stabilization of structural plant-derived OM that only got modified by extracellular enzymes. In our dataset, a higher decomposition index (DI) corresponds to a higher proportion of microbially

transformed OM in the respective reservoir, as the patterns of the variables on which the DI is based (fraction specific C:N ratio, relative peak area of $RPA_{SIMP}$, $\delta^{13}C$ and $\delta^{15}N$) show (Figure S2-1). This implies that in soils with a high DI a larger proportion of the OM was stabilized via the *in vivo* pathway (in contrast to the *ex vivo* pathway). We therefore argue that with increasing bulk SOC% we see an increasing importance of the *ex vivo* pathway relative to the *in vivo* pathway for SOC stabilization associated with the S+C fraction. Theoretically, a large variety of mechanisms could underlie such a shift: (1) Changes in the composition of the clay mineral phase and its sorptive affinity could influence the stabilization of chemically distinct types of OM in S+C (Kaiser et al., 1997; Sollins et al., 2009; Kramer et al., 2012; Creamer et al., 2019; Sokol et al., 2019). (2) Lower decomposition rates caused by lower temperatures could lead to the relative accumulation of less-decomposed OM. Related to this, at low OM content, it could be for microbes to recycle the available substrate, which could lead to a stabilization of more processed OM in the S+C fraction. (3) Increasing SOC contents could facilitate increasing OM – OM interactions (Possinger et al., 2020; Schweizer, 2022) on mineral surfaces, which may be less selective in terms of OM chemistry. (4) It has previously been suggested that more leaching in wetter environments may favor the *ex vivo* pathway over the in vivo pathway (Sokol et al., 2019), which would also result in the pattern that we observe. (5) In soils with more aggregation (i.e., soils with higher bulk $SOC_{\%}$), it is reasonable to assume a larger contribution of small stable aggregates in the silt- and clay-sized fraction (relative to free organo-mineral associations). These small stable aggregates could in turn occlude more poorly transformed OM, and thus shift the DI of the S+C fraction. Similarly, small microaggregates of the silt- and clay-sized fractions are particularly stable in Andosols as compared to other soil types (Asano and Wagai, 2014), which could cause that in Andosols more stable microaggregates might resist dispersion. (6) Macroaggregation may be higher in SOC rich soils and thus smaller (S+C-sized) particles may have been occluded in macroaggregates before fractionation. Within macroaggregates, the organic matter associated with these S+C-particles may be protected from decomposition by a certain degree of physical occlusion in addition to chemical stabilization mechanisms. However, it remains unclear if and to what extent occlusion in macroaggregates can protect OM from decomposition on short to intermediate time scales. In conclusion, while our data does not reveal which mechanism(s) may cause the patterns in this dataset, plenty of observed and proposed mechanisms render our observation and interpretation plausible, that the OM in the silt- and clay-sized fraction is on average less microbially transformed in high-SOC soils than in low-SOC soils.

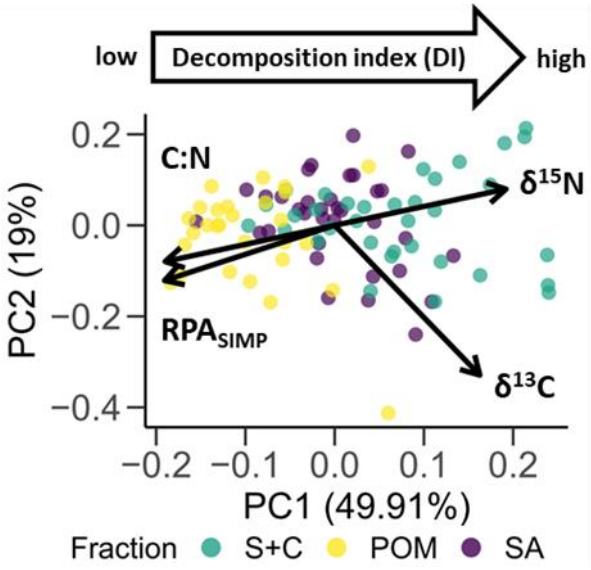



**Figure S2-1.** Results of a PCA based on fraction specific C:N ratio, relative peak area of $RPA_{SIMP}$, $\delta^{13}C$ and $\delta^{15}N$. n = 96 ($n_{SA}$ = 35, $n_{S+C}$ = 34, $n_{POM}$ = 28). PC1 explains 50 % of variation in the input variables and can be interpreted as a decomposition index (DI) as described in the methods section. The Eigenvectors of the input variables are shown as arrows: C:N ratio and $RPA_{SIMP}$ increase towards the SA fraction, while $\delta^{15}N$ and $\delta^{13}C$ values increase towards the S+C fraction.

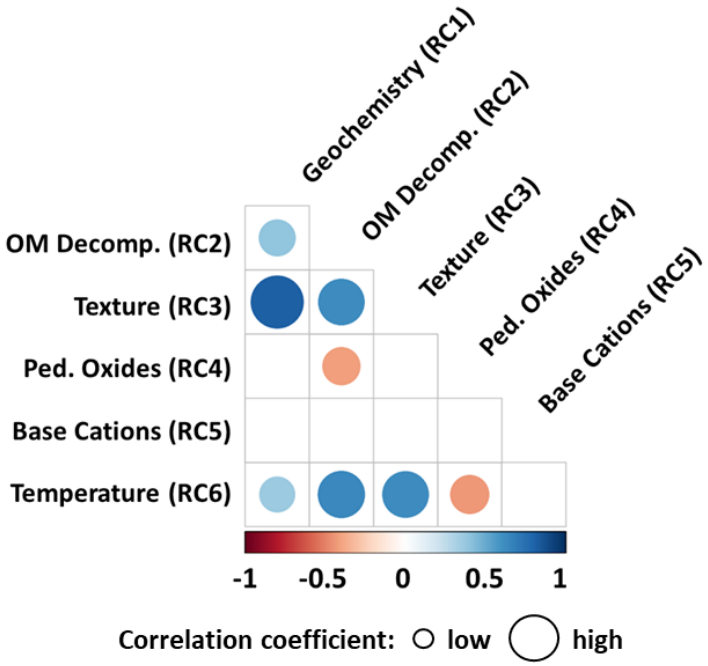



**Figure S2-2**. Correlation matrix of the six RCs retained after rPCA. Shown are significant correlations (Pearson correlation, $p < 0.05$). Circle size and color indicate the strength of the relationships (Pearson correlation coefficient).

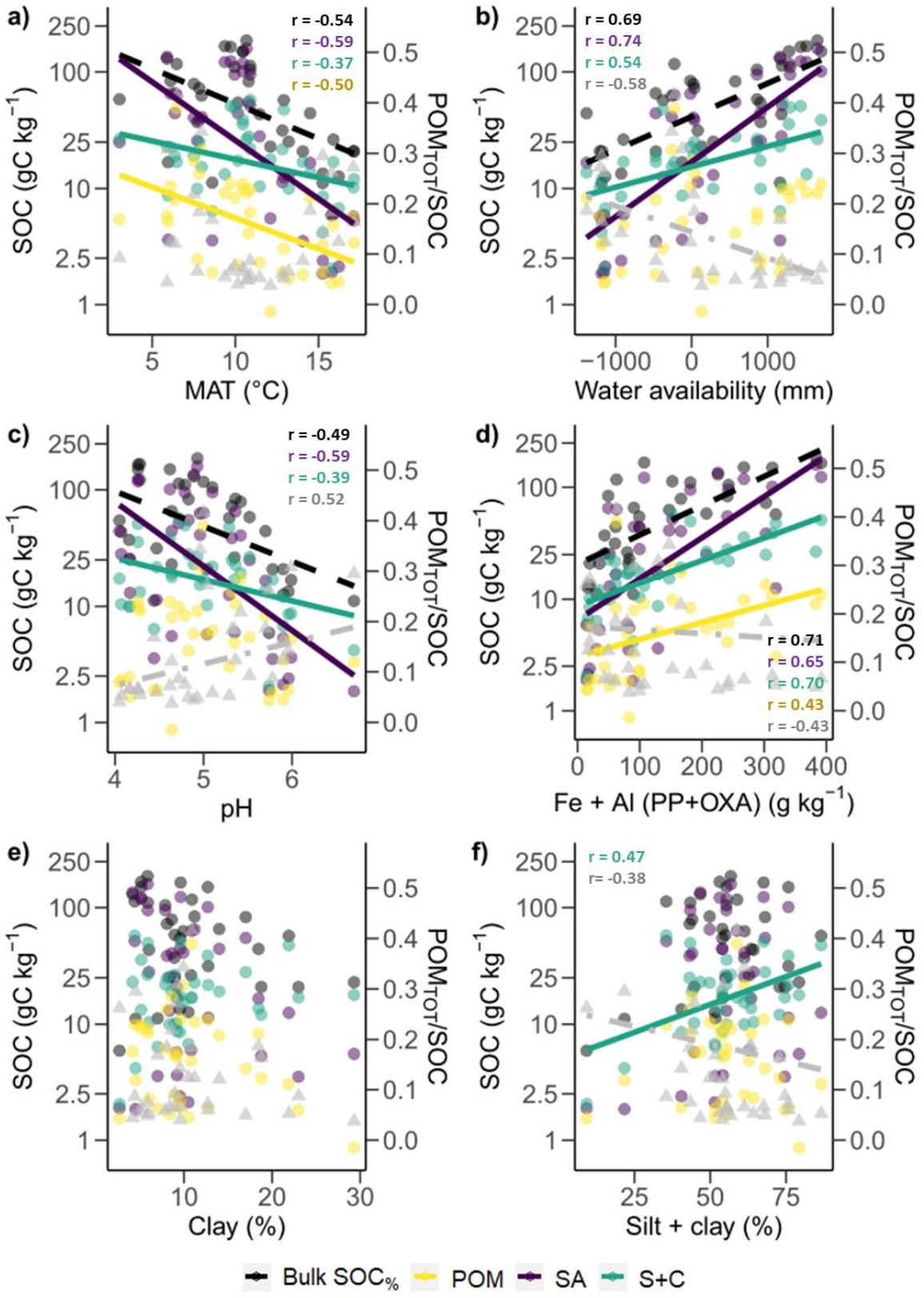



**Figure S2-3.** Correlations between bulk $SOC_{\%}$ and $C_{TOT}$ in SA, S+C and POM as well as $POM_{TOT}$/SOC with measured environmental variables. Shown are relationships with a) MAT, b) water balance (MAP – PET), c) pH value, d) the sum of Fe and Al sequentially extracted with pyrophosphate ammonium oxalate, e)

primary clay content and f) the sum of primary clay and silt content. Dots indicate datapoints of SOC amounts, triangles datapoints of $POM_{TOT}$/SOC ratio, lines indicate slopes of significant (p-values < 0.05) Pearson correlations. All correlations of SOC amounts performed with log-transformed SOC quantities. Only in the case of pedogenic oxides (panel d) transformation led to small decreases in the strength of relationship for bulk $SOC_{\%}$ and $SA_{TOT}$.

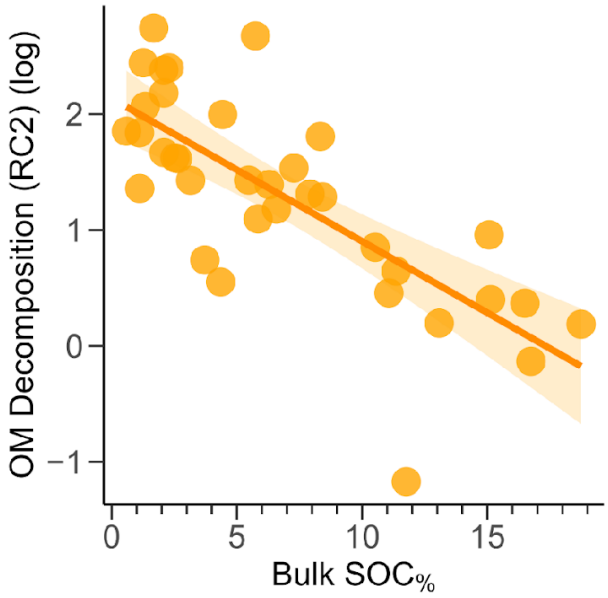



**Figure S2-4.** The degree of bulk organic matter decomposition decreases with bulk SOC content. As a measure for the degree of bulk organic matter decomposition, the rotated component RC2 "OM Decomposition" was used (decomposition index (DI) cannot be calculated for bulk soil), which strongly correlates with a decrease of aliphatic compounds ($RPA_{SIMP}$) relative to other compounds, and with a decrease of the Rock-Eval derived I-index (see Table S2-8). Log-transformed data points and fitted model with 95 % confidence intervals are shown. Because the rotated component had negative values (min. -4.69), log-transformation was done after adding + 5 to all data points. RMSE = 3.53, adj. $R^2$ = 0.55, p-value < 0.05, n = 35, df = 33.

**Table S2-1.** Geomorphology and soil information of sampling sites. Site IDs refer to the numbers as used in Figure 2-1. Literature used for metadata is indexed with numbers from 1 – 11: 1 = Aburto F., Hernández C., Pfeiffer M., Casanova M., Luzio W. (2008). Soil: a work of art of nature. Northern field guide between 30S° and 33S° In: Casanova M., Luzio W. (Eds.) The international conference and field workshops on soil classification. University of Chile, Santiago-Ovalle, Chile; 2 = CIREN (1997). Soil survey, region V of Chile. Natural Resources Information Centre, Publication N° 116 Santiago (in Spanish); 3 = CIREN (1996). Soil survey of the Metropolitan region of Chile. Natural Resources Information Centre, Publication N° 115 Santiago (in Spanish); 4 = CIREN (1996). Soil survey, region VI of Chile. Natural Resources Information Centre, Publication N° 114 Santiago (in Spanish); 5 = CIREN (1997). Soil survey of the VII region of Chile. Natural Resources Information Centre, Publication N° 121 Santiago (in Spanish); 6 = CIREN (1999). Soil survey, region VIII of Chile. Natural Resources Information Centre, Publication N° 117 Santiago (in Spanish); 7 = CIREN (2002). Soil survey, region IX of Chile. Natural Resources Information Centre, Publication N° 122 Santiago (in Spanish); 8 = CIREN (2003). Soil survey, region X of Chile. Natural Resources Information Centre, Publication N° 123 Santiago (in Spanish); 9 = CIREN (2005). Soil survey, region XI of Chile. Natural Resources Information Centre, Publication N° 130 Santiago (in Spanish); 10 = INIA (2014). Taxonomic characterization of soils from valleys with agricultural interest from Aysen Region (Patagonia Occidental-Chile.). In: Stolpe N., Hepp C. (Eds.). Bulletin INIA (Instituto de Investigaciones Agropecuarias) N° 299. Coyhaique, Chile. (in Spanish); 11 = CNR (1997). Comprehensive study of irrigation and drainage of Magellan, XII Region, vol 2 (in Spanish).

| ID | Site name | Geomorphology | WRB Reference soil groups | Profile description (depth in cm) |
|---|---|---|---|---|
| 1 | Los Vilos | Marine terraces/stabilized dunes | Eutric Arenosol | A1(0-11), A21(11-23), A22(23-38), C1(38-54), C2(54-82), C3(82-108), 2C4(108-190), 3C5(190-220), 4C6(220-275) |
| 2 | Catapilco | Marine terraces | Albic Planosol | A1(0-19), E(19-25), Bt(25-54), C(54-75+) |
| 3 | Calle Larga | Piedmont | Luvic Kastanozem / Luvic Chernozem | Ap(0-15), Bt1(15-34), Bt2(34-63), C(63-80+) |
| 4 | Bochinche | Marine terraces | Kastanozem/Chernozem | Ap(0-31), B(31-55), C(55-120) |
| 5 | Clarillo | Alluvial terrace | Kastanozem | A1(0-35), B2(35-60), IIB31(60-82), B32(82-115), B33(115-145) |
| 6 | Matanzas | Marine terraces | Chernozem / Kastanozem / Phaeozem | A1(0-20), B21(20-64), B22(64-95), C(95+) |
| 7 | Pudahuel | Piedmont | Duric Kastanozem / Duric Chernozem | A1(0-19), A3(19-34), B(34-48), C1(48-55), C2m(55-73), C3(72-120) |
| 8 | Pimpinela | Piedmont | Haplic Luvisol | A(0-40), B1(40-65), B2(65-125), B3(125+) |
| 9 | Teno | Drumlins | Duric Chernozem | A1(0-18), B1(18-40), B2(40-55), C(57+) |
| 10 | Bramaderos | High alluvial terraces (undulated) | Andosol | Ap(0-17), A12(17-48), AB(48-82), C(82-120+) |
| 11 | Arauco | Alluvial terraces | Dystric Cambisol | Ap(0-20), B1(20-64), B2(64-120+) |
| 12 | Collipulli | High remaining hillsides | Luvisol | Ap(0-15), A3(15-35), B1(35-47), Bt2(47-84), Bt3(84-150) |
| 13 | Puerto Saavedra | Alluvial terraces | Gleysol | A1(0-20), Bw1(20-40), Bw2(40-60), Bw3(60-85), Bw4(85-100), C(100+) |
| 14 | Melipeuco | Recent alluvial terrace | Andosol | A1(0-27), B1(27-55), B2(55-95), C(95-110+) |
| 15 | Choshuenco | Hillocks & hills | Dystric Cambisol | A1(0-29), Bw1(29-54), Bw2(54-86), C(86-100+) |
| 16 | Hueicoya | Hills | Humic Acrisol | A1(0-19), B1(19-33), Bt1(33-64), BC(64-90), C(90-130) |
| 17 | Bahia Mansa | Marine terraces | Dystric Cambisol | A1(0-6), A2(6-22), B1(22-44), B2(44-82), C(82-100) |
| 18 | Corte Alto | Remaining terraces | Andosol | A(0-18), Bw1(18-47), Bw2(47-77), 2C1(77-122), 2C2(122-170), 3C3(170-194) |
| 19 | Chapo | Hillocks & hills | Andosol | A1(0-25), B1(25-60), B2(60-110) |
| 20 | Chiloe Norte | Fluvioglacial terraces | Petroduric Andosol | A(0-18), Bt1(18-52), Bt2(52-90), C(90+) |
| 21 | Pachabran | Gentle rolling hills | Petroduric Histic Andosol | A1(0-6), A2(6-15), B1(15-38), B2(38-52), BC(52-75), C(75+) |
| 22 | Aitui | Rolling hills | Fulvic Andosol | A1(0-9), A2(9-23), B1(23-40), Bt(40-79), BC(79-110), C(110+) |
| 23 | La Junta | Alluvial terraces | Thaptic Andosol | A1(0-14), B1(14-27), B2(29-41), BC1(41-59), BC2(59-70), 2C(70-85), 3B(85-103) |
| 24 | La Tapera | Gentle rolling hills | Andosol | O(0-13), A1(13-29), B1(29-58), B2(58-75), 2C(75-90) |
| 25 | Puerto Cisnes | Fluvio-glacial terrace | Fulvic Andosol | A1(0-10), A2(10-29), A3(29-49), Bw(49-61), Ab(61-92), ACb(92-118), Ab(118+) |
| 26 | Puerto Aysen | Alluvial terrace | Fulvic Andosol | A(0-10), BW(10-23), AB1(23-25), AB2(35-64), Bwb(64-88), Ab(88-130) |
| 27 | Simpson | Old alluvial terraces | Dystric Cambisol | A1(0-10), A2(10-31), B1(31-68), B2(68-82), B3(82-110) |
| 28 | Cerro Castillo | Hills | Andosol | Oe(0-5), A1(0-11), A2(11-28), Bw(28-68), BC(68-80), C(80+) |
| 29 | Bahia Exploradores | Hills | Andosol | Oe(0-5), A(0-20), Bw(20-31), 2Ab(31-57) |
| 30 | Puerto Sanchez | Hillocks & hills | Andosol | O(0-6), A1(6-15), B1(15-31), B2(31-49), BC(49-67) |
| 31 | Cochrane | Rolling hills | Dystric Cambisol | A(0-9), B(9-40), 2B(40-80) |
| 32 | Caleta Tortel | Alluvial terrace | Umbric Gleysol | Oi(10-7), Oe(7-0), Bg(0-31), C1(31-53), C2(53+) |
| 33 | Ultima Esperanza | Glacial outwash | Leptosol | A(0-20), B(20-35), B1(35-50), B2(50+) |
| 34 | Porvenir | Marine terraces | Leptosol | no data |
| 35 | Aguas Frescas | Marine terraces | Leptosol | O(0-13), A(13-25), B1(25-38), B2(38-50), C(50-60) |

**Table S2-1 cont.**

| ID | Site name | Soil depth (end of B, cm) | Parent material | Ref. | Vegetation |
|---|---|---|---|---|---|
| 1 | Los Vilos | 38 | Eolian sands. | 1 | *Haplopappus, Proustia & Tessaria absinthioides + (Flourensia thurifera, Puya chilensis & Baccharis spp)* |
| 2 | Catapilco | 54 | Granite. | 2 | *Acacia caven & Avena barbata + (AMP: Bromus spp. & Vulpia sp.)* |
| 3 | Calle Larga | 63 | Sedimentary soil on andesitic rocks with varying degrees of weathering. | 2 | *Centaurea melitensis & Brassica rapa* |
| 4 | Bochinche | 55 | Sedimentary origin on a substratum of rounded gravels with abundant quartz and orthoclase. | 2 | *Baccharis linearis + ( AMP: Bromus mollis & Nasella chilensis)* |
| 5 | Clarillo | 145 | Soils of alluvial origin, with volcanic glass and mica found in high amounts in B32. | 3 | *Acacia caven & Avena barbata + (AMP: Bromus spp. & Vulpia sp.)* |
| 6 | Matanzas | 95+ | Partially weathered sandstone and marine sediments of claylite and siltstone type. | 4 | *Chaetanthera chilensis* |
| 7 | Pudahuel | 48 | Volcanic ashes of pumiceous type, redeposited by water. | 3 | *Acacia caven + (AMP: Avena silvestre, Bromus mollis, Vulpia dertonensis, Rapistrum rugosum & Raphanus sativus + Malesherbia humilis, Chaetanthera chilensis & Aira caryophyllea )* |
| 8 | Pimpinela | 125 | Soils of colluvial origin, stratified. | 4 | *Avena barbata, Brassica rapa & Raphanus sativus* |
| 9 | Teno | 55+ | Substrate consisting of partially weathered volcaniclastic breccia. | 5 | *Avena barbata & Brassica rapa* |
| 10 | Bramaderos | 120 | Sedimentary soil, volcanic origin (volcanic ash). | 5 | *Galega officinalis , Brassica rapa, Convolvulus arvensis & Hypochoeris sp.* |
| 11 | Arauco | 120 | Alluvial sediments very rich in mica and mica- and quartz-rich alluvial sediments from the metamorphic formation of the Cordillera de la Costa. | 6 | *Trifolium repens & Bromus valdivianus + (ballicas & poas + Hypochoeris)* |
| 12 | Collipulli | 150 | Soil developed from ancient volcanic ashes, resting on a substratum of fluvioglacial conglomerate, partially on a substratum made up of a fluvioglacial conglomerate, partially weathered with an andesitic-basaltic composition. | 7 | *Leontodon saxatilis + (Bromus + Hypochoeris sp.+ Trifolium repens)* |
| 13 | Puerto Saavedra | 100+ | Sedimentary soil, sandstone of alluvial origin. | 7 | *Trifolium repens, Holcus lanatus & Bromus + (Lotus, Hypochoeris, Plantago lanceolata, Ranunculus)* |
| 14 | Melipeuco | 95 | Soils of volcanic origin, on recent alluvial terrace, stratified, consisting of volcanic ashes mixed with basaltic sands. | 7 | *Dactylis glomerata + (Lolium perenne + Poa pratensis+ Hypochoeris sp.+ Lotus corniculatus + Taraxacum officinalis)* |
| 15 | Choshuenco | 86 | The soil is derived from volcanic ashes and sands of varying grain size that have been deposited on coarse sands and volcanic slag. The substratum consists of volcanic scoria and sand probably of basaltic origin. | 8 | *Holcus lanatus, Trisetum sp., Taraxacum officinale, Deschampsia antarctica + Leucanthemum vulgare* |
| 16 | Hueicoya | 90 | The parent material, mica schist, is found with varying degrees of weathering. | 8 | *Juncus balticus, Caltha appendiculata, Ranunculus sp. & Hordeum sp.* |
| 17 | Bahia Mansa | 82 | Substratum of highly weathered metamorphic rocks of the phyllite type. | 8 | *Hipochoeris incana, Taraxacum officinale, Bromus* |
| 18 | Corte Alto | 77 | Formed from Holocene volcanic ashes. | 8 | *Dactylis glomerata & Paspalum dilatatum* |
| 19 | Chapo | 110 | Formed from Holocene volcanic ash, The soil is generally found directly on a basaltic rock with which it forms a lithological discontinuity. | 8 | *Holcus lanatus, Acaena ovalifolia, Hypochoeris radicata, Agrostis capillaris, Lotus uliginosus & Elymus gayanus* |
| 20 | Chiloe Norte | 90+ | Formed from volcanic ashes. The soil substratum is an ironstone under which there is a cemented fluviogacial deposit of sands and gravels, at variable depths of between 60 and 90 cm. | 8 | *Holcus lanatus, Agrostis capillaris, Lolium perenne & Plantago lanceolata* |
| 21 | Pachabran | 75+ | Soils derived from volcanic ashes evolved in conditions of better drainage than the rest of the soils of Chiloé Island. They are older soils that occupy the highest fluvioglacial terraces close to the Coastal Mountain Range. | 8 | *Ranunculus peduncularis, Holcus lanatus, Poa pratensis, Rumex crispus* |
| 22 | Aitui | 110+ | Volcanic ash derivatives. | 8 | *Lotus uliginosus, Holcus lanatus, Hypochoeris radicata, Poa pratensis, Leucanthemum vulgare* |
| 23 | La Junta | 103 | Soils of alluvial origin, in an alluvial terrace position of the Rosselot River. | 9 | *Holcus lanatus + (Lotus + Hypochoeris + Plantago lanceolada + Agrostis tenuis)* |
| 24 | La Tapera | 75 | The substrate is grey glacial clay. | 9 | *Dactylis glomerata, Trifolium repens, Bromus valdivianus & Poa pratensis* |
| 25 | Puerto Cisnes | 118+ | Volcanic ash. | 10 | *Cotula scariosa, Paspalum dilatatum & Cynodon dactilum* |
| 26 | Puerto Aysen | 130 | Volcanic ash. | 10 | *Holcus lanatus, Deschampsia sp. & Hypochoeris radicata* |
| 27 | Simpson | 110 | The substrate consists of variable weathered sediments. | 9 | *Trifolium repens, Holcus lanatus, Hypochoeris incana & Agrostis sp.* |
| 28 | Cerro Castillo | 80+ | Volcanic ash. | 10 | *Holcus lanatus, Digitalis purpurea, Dactylis glomerata, Festuca pallecens, Lolium perenne, Taraxacum officinale, Trifolium pratense, Trifolium repens + Acaena ovalifolia* |
| 29 | Bahia Exploradores | 57 | Volcanic ash on glacial deposit. | 10 | *Dactylis glomerata, Paspalum dilatum & Trifolium repens* |
| 30 | Puerto Sanchez | 67 | Fluvioglacial deposits and metamorhic rocks. | 9 | *Acaena pinnatifida, Anemone multifida & Plantago lanceolata* |
| 31 | Cochrane | 80 | Fluvioglacial deposits. | 9 | *Agrostis spp., Achillea millefolium, Holcus lanatus & Hypochoeris incana* |
| 32 | Caleta Tortel | 31 | Alluvial deposit. | 10 | *Empetrum rubrum, Gaultheria pumila & Senecio triodon* |
| 33 | Ultima Esperanza | 50+ | Glacial outwash. | 11 | *Poa pratensis, Trisetum cumingii, Holcus lanatus & Deschampsia flexuosa. Agrostis sp., Festuca arundinacea, Hypochoeris incana (Rumex crispus + Achillea millefolium+ Dactylis glomerata)* |
| 34 | Porvenir | 120+ | Marine Terraces. | 11 | *Chiliotrichioum diffussum, Baccharis magellanicum, Poa pratensis, Festuca arundinacea, Trifolium repens & Dactylis glomerata* |
| 35 | Aguas Frescas | 60+ | Evolved from volcanic ash, on top of compacted glacial sediments. | 11 | *Hieracium aurantiacum, Trifolium repens, Agrostis spp., Festuca arundicnacea, Berberis buxifolia & Deschampsia antarctica* |

**Table S2-2.** General site information. Site IDs refer to the numbers as used in Figure 2-1.

| ID | Site name | Latitude (WGS 84) (°S) | Longitude (WGS 84) (°W) | Altitude m.a.s.l. | MAP mm | PET mm | MAT °C | TempSeas °C | NPP gC $m^{-2}$ $yr^{-1}$ | Climate Classification after Köppen-Geiger | Soil moisture regime | Described in Doetterl et al., 2015 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Los Vilos | 31.8128 | 71.5001 | 95 | 217 | 1369 | 15.8 | 2.9 | 665 | Arid, steppe, cold (BSk) | Aridic | yes |
| 2 | Catapilco | 32.5265 | 71.2608 | 116 | 321 | 1458 | 16.2 | 3.2 | 682 | Temperate, dry summer, warm summer (Csb) | Ustic | |
| 3 | Calle Larga | 32.8761 | 70.5216 | 768 | 426 | 1808 | 14.5 | 4.2 | 847 | Arid, steppe, cold (BSk) | Xeric | yes |
| 4 | Bochinche | 33.3390 | 71.6124 | 159 | 436 | 1362 | 15.3 | 3.0 | 976 | Temperate, dry summer, warm summer (Csb) | Ustic | yes |
| 5 | Clarillo | 33.6804 | 70.5857 | 665 | 514 | 1698 | 15.2 | 4.8 | 728 | Temperate, dry summer, warm summer (Csb) | Xeric | |
| 6 | Matanzas | 33.9686 | 71.8756 | 124 | 533 | 1430 | 15.9 | 3.4 | 1067 | Temperate, dry summer, warm summer (Csb) | Ustic | yes |
| 7 | Pudahuel | 33.9991 | 71.1405 | 194 | 451 | 1658 | 17.1 | 4.4 | 788 | Temperate, dry summer, warm summer (Csb) | Xeric | yes |
| 8 | Pimpinela | 34.3239 | 70.7297 | 417 | 675 | 1852 | 17.1 | 4.9 | 716 | Temperate, dry summer, warm summer (Csb) | Xeric | |
| 9 | Teno | 34.8247 | 71.0918 | 298 | 865 | 1336 | 14.1 | 4.9 | 810 | Temperate, dry summer, warm summer (Csb) | Xeric | |
| 10 | Bramaderos | 35.6133 | 71.3146 | 374 | 935 | 1312 | 12.9 | 4.7 | 931 | Temperate, dry summer, warm summer (Csb) | Xeric | |
| 11 | Arauco | 37.2534 | 73.2669 | 13 | 1426 | 1141 | 13.2 | 2.7 | 1158 | Temperate, dry summer, warm summer (Csb) | Udic | yes |
| 12 | Collipulli | 37.9206 | 72.4872 | 226 | 1301 | 1168 | 12.1 | 4.0 | 982 | Temperate, dry summer, warm summer (Csb) | Udic | |
| 13 | Puerto Saavedra | 38.7741 | 73.3897 | 3 | 1175 | 1094 | 12.8 | 2.9 | 1277 | Temperate, dry summer, warm summer (Csb) | Udic | yes |
| 14 | Melipeuco | 38.8744 | 71.8555 | 409 | 2025 | 1166 | 10.9 | 3.9 | 1042 | Temperate, no dry season, cold summer (Cfc) | Udic | |
| 15 | Choshuenco | 39.8594 | 72.1112 | 263 | 2249 | 1059 | 10.9 | 3.8 | 1078 | Temperate, dry summer, warm summer (Csb) | Udic | yes |
| 16 | Hueicoya | 39.9325 | 73.4128 | 6 | 2102 | 921 | 11.5 | 2.8 | 1764 | Temperate, no dry season, warm summer (Cfb) | Udic | yes |
| 17 | Bahia Mansa | 40.6089 | 73.7489 | 156 | 1705 | 801 | 10.1 | 2.4 | 1290 | Temperate, no dry season, warm summer (Cfb) | Udic | |
| 18 | Corte Alto | 40.9034 | 73.1540 | 118 | 1443 | 869 | 10.8 | 3.0 | 771 | Temperate, no dry season, warm summer (Cfb) | Udic | yes |
| 19 | Chapo | 41.4271 | 72.6478 | 226 | 2204 | 909 | 10.0 | 3.3 | 1044 | Temperate, no dry season, warm summer (Cfb) | Udic | |
| 20 | Chiloe Norte | 42.0530 | 73.7991 | 62 | 2440 | 736 | 10.5 | 2.5 | 1605 | Temperate, no dry season, warm summer (Cfb) | Perudic | |
| 21 | Pachabrán | 42.4219 | 73.8243 | 204 | 2167 | 733 | 9.7 | 2.5 | 1516 | Temperate, no dry season, warm summer (Cfb) | Perudic | yes |
| 22 | Aituí | 43.0579 | 73.6171 | 30 | 2411 | 757 | 10.7 | 2.6 | 1443 | Temperate, no dry season, warm summer (Cfb) | Perudic | yes |
| 23 | La Junta | 43.9639 | 72.3965 | 43 | 2145 | 865 | 10.4 | 3.7 | 605 | Temperate, no dry season, warm summer (Cfb) | Perudic | yes |
| 24 | La Tapera | 44.6585 | 71.7844 | 626 | 1032 | 1008 | 7.5 | 4.2 | 495 | Temperate, no dry season, cold summer (Cfc) | Udic | |
| 25 | Puerto Cisnes | 44.7569 | 72.6146 | 33 | 2271 | 759 | 9.4 | 3.7 | 671 | Temperate, no dry season, warm summer (Cfb) | Perudic | |
| 26 | Puerto Aysen | 44.9020 | 72.6246 | 80 | 1885 | 722 | 5.9 | 3.8 | 636 | Temperate, no dry season, warm summer (Cfb) | Perudic | |
| 27 | Simpson | 45.7864 | 72.9076 | 462 | 1576 | 715 | 3.0 | 3.7 | 73 | Polar, tundra (ET) | Udic | yes |
| 28 | Cerro Castillo | 46.1076 | 72.0736 | 640 | 818 | 942 | 6.0 | 4.2 | 499 | Temperate, dry summer, cold summer (Csc) | Udic | |
| 29 | Bahia Exploradores | 46.4954 | 73.1277 | 180 | 1642 | 734 | 7.7 | 3.7 | 319 | Polar, tundra (ET) | Udic | |
| 30 | Puerto Sánchez | 46.5682 | 72.6128 | 330 | 990 | 954 | 7.8 | 3.9 | 560 | Temperate, no dry season, cold summer (Cfc) | Udic | yes |
| 31 | Cochrane | 47.3235 | 72.6455 | 304 | 772 | 1033 | 8.0 | 4.1 | 619 | Temperate, no dry season, cold summer (Cfc) | Udic | |
| 32 | Caleta Tortel | 47.7915 | 73.5179 | 10 | 1961 | 798 | 8.6 | 3.2 | 487 | Polar, tundra (ET) | Udic | |
| 33 | Ultima Esperanza | 51.8070 | 72.1649 | 67 | 386 | 859 | 6.4 | 3.8 | 459 | Temperate, no dry season, cold summer (Cfc) | Udic | yes |
| 34 | Porvenir | 53.3148 | 70.3611 | 59 | 553 | 780 | 6.5 | 3.2 | 433 | Temperate, no dry season, cold summer (Cfc) | Perudic | yes |
| 35 | Agua Fresca | 53.4327 | 70.9886 | 58 | 703 | 720 | 6.1 | 3.1 | 447 | Temperate, no dry season, warm summer (Cfb) | Udic | yes |

**Table S2-3.** Soil physicochemical and climatic variation along the gradient. Ranges of all soil physicochemical and climatic variables that were used for the rPCA for dimension reduction.

| Variable | Unit | Min. | Mean | Median | Max. |
|---|---|---|---|---|---|
| pH ($CaCl_2$) | - | 4.05 | 5.04 | 4.96 | 6.70 |
| Sand | % | 15.6 | 59.1 | 62.6 | 93.4 |
| Silt | % | 5.6 | 36.6 | 35.1 | 72.7 |
| Clay | % | 0.8 | 4.3 | 2.5 | 21.3 |
| Al (Aq.Reg.) | $g\ kg^{-1}$ | 8.90 | 27.66 | 25.68 | 92.15 |
| Fe (Aq.Reg.) | $g\ kg^{-1}$ | 13.10 | 33.26 | 26.88 | 106.88 |
| Al (PP + OXA) | $g\ kg^{-1}$ | 0.45 | 7.84 | 4.39 | 26.44 |
| Fe (PP + OXA) | $g\ kg^{-1}$ | 0.74 | 6.05 | 5.92 | 20.11 |
| Mn (PP + OXA) | $g\ kg^{-1}$ | 0.06 | 0.29 | 0.21 | 1.15 |
| $Fe_{Aq.Reg.}$ : Si | - | 0.063 | 0.233 | 0.241 | 0.678 |
| $Clay_{prim}$ : Si | - | 0.009 | 0.052 | 0.041 | 0.180 |
| MAT | °C | 3.0 | 10.9 | 10.7 | 17.1 |
| MAP - PET | mm | -1382 | 207 | 81 | 1704 |
| NPP | $gC\ m^{-2}\ yr^{-1}$ | 73 | 842 | 771 | 1764 |
| TempSeas | °C | 2.4 | 3.6 | 3.7 | 4.9 |
| Ca ($BaCl_2$) | $g\ kg^{-1}$ | 0.22 | 2.10 | 2.21 | 6.39 |
| Mg ($BaCl_2$) | $g\ kg^{-1}$ | 0.03 | 0.31 | 0.21 | 1.43 |
| K ($BaCl_2$) | $g\ kg^{-1}$ | 0.01 | 0.24 | 0.20 | 0.73 |
| $RPA_{SIMP}$ (bulk) | % | -14.9 | 69.4 | 77.7 | 90.5 |
| $RPA_{COMP}$ (bulk) | % | 2.8 | 16.1 | 13.1 | 54.9 |
| $RPA_{MBIO}$ (bulk) | % | 5.4 | 14.6 | 10.8 | 65.4 |
| Rock-Eval I-Index | - | -0.09 | 0.18 | 0.15 | 0.41 |

**Table S2-4. Mass-balance accounting of the fractionation procedure for all samples.** $Mass_{REL}$ gives the mass percentage of each fraction relative to the mass of the fractionated bulk soil sample. The balance of the mass gives the sum of $Mass_{REL}$ of the three fractions relative to the mass of the fractionated bulk soil sample, expressed as percentage. The balance of $C_{TOT}$ gives the sum of the amount of SOC in the three fractions relative to the amount of SOC in bulk soil, expressed as percentage. For three samples, the fractionation was done in triplicates.

| | $Mass_{REL}$ (%) | | | Balance (%) | |
|---|---|---|---|---|---|
| ID | S+C | POM | SA | Mass | $C_{TOT}$ |
| 1 | 3.5 | 0.8 | 94.1 | 98.4 | 83.8 |
| 2a | 44.5 | 0.5 | 53.3 | 98.3 | 75.2 |
| 2b | 43.7 | 0.7 | 55.0 | - | - |
| 2c | 44.6 | 0.5 | 54.8 | - | - |
| 3 | 40.7 | 2.8 | 54.8 | 98.3 | 84.4 |
| 4 | 29.7 | 0.5 | 69.2 | 99.4 | 93.5 |
| 5 | 12.6 | 1.3 | 85.6 | 99.5 | 79.5 |
| 6 | 51.4 | 1.6 | 46.5 | 99.6 | 102.0 |
| 7 | 37.8 | 2.5 | 59.0 | 99.3 | 93.6 |
| 8 | 65.1 | 1.4 | 37.6 | 104.2 | 95.9 |
| 9a | 74.2 | 0.6 | 25.6 | 100.4 | 100.5 |
| 9b | 72.1 | 0.6 | 26.6 | - | - |
| 9c | 71.0 | 0.6 | 27.7 | - | - |
| 10 | 74.1 | 1.3 | 24.5 | 100.0 | 107.2 |
| 11a | 21.5 | 0.6 | 77.4 | 99.5 | 88.6 |
| 11b | 21.4 | 0.5 | 77.8 | - | - |
| 11c | 26.6 | 0.8 | 67.7 | - | - |
| 12 | 66.7 | 0.2 | 31.6 | 98.5 | 96.9 |
| 13 | 70.0 | 0.7 | 28.4 | 99.1 | 108.0 |
| 14 | 15.6 | 2.2 | 80.9 | 98.8 | 110.9 |
| 15 | 30.4 | 4.7 | 65.3 | 100.4 | 97.7 |
| 16 | 18.2 | 0.6 | 80.1 | 98.9 | 85.8 |
| 17 | 22.2 | 0.8 | 76.3 | 99.2 | 113.5 |
| 18 | 33.5 | 1.5 | 64.9 | 99.9 | 102.3 |
| 19 | 20.3 | 2.7 | 77.5 | 100.5 | 132.0 |
| 20 | 27.6 | 3.5 | 70.2 | 101.3 | 97.3 |
| 21 | 23.0 | 3.1 | 74.1 | 100.2 | 108.1 |
| 22 | 15.8 | 3.5 | 84.7 | 104.0 | 100.1 |
| 23 | 20.1 | 4.0 | 80.0 | 104.1 | 104.5 |
| 24 | 20.1 | 5.2 | 72.7 | 97.9 | 94.8 |
| 25 | 8.4 | 3.7 | 67.2 | 79.3 | 77.8 |
| 26 | 13.0 | 2.1 | 85.7 | 100.8 | 117.5 |
| 27 | 32.8 | 2.0 | 63.7 | 98.6 | 94.2 |
| 28 | 29.1 | 4.0 | 65.7 | 98.7 | 98.1 |
| 29 | 27.3 | 0.7 | 71.0 | 99.1 | 95.6 |
| 30 | 23.7 | 5.6 | 70.1 | 99.3 | 100.0 |
| 31 | 30.1 | 7.3 | 62.9 | 100.4 | 120.0 |
| 32 | 36.9 | 0.6 | 61.1 | 98.6 | 95.4 |
| 33 | 26.6 | 2.6 | 69.2 | 98.4 | 100.3 |
| 34 | 19.7 | 19.4 | 59.7 | 98.7 | 102.7 |
| 35 | 30.9 | 3.5 | 65.0 | 99.4 | 109.6 |

**Table S2-5.** Summary of quantitative and qualitative SOC data of the three isolated fractions. Shown without outliers, which were removed based on Cochran's C test. Thus, DRIFTS-values ($RPA_{SIMP}$, $RPA_{COMP}$ and $RPA_{MBIO}$) are $n_{SA}$ = 35, $n_{S+C}$ = 34, $n_{POM}$ = 28, $n_{bulk}$ = 34; all other values are n = 35.

| | | POM | | | | | SA | | | | | S+C | | | | | bulk soil | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Unit** | **Min.** | **Mean** | **Median** | **S.D.** | **Max.** | **Min.** | **Mean** | **Median** | **S.D.** | **Max.** | **Min.** | **Mean** | **Median** | **S.D.** | **Max.** | **Min.** | **Mean** | **Median** | **S.D.** | **Max.** |
| **$Mass_{REL}$** | % | 0.2 | 2.8 | 2.1 | 3.4 | 19.4 | 24.5 | 64.4 | 67.2 | 17.5 | 94.1 | 3.5 | 31.9 | 27.6 | 18.6 | 74.2 | - | - | - | - | - |
| **$C_{TOT}$** | $gC\ kg^{-1}$ | 0.9 | 7.5 | 5.4 | 8.6 | 48.9 | 1.9 | 44.9 | 28.1 | 46.1 | 159.9 | 2.0 | 20.9 | 17.3 | 13.9 | 57.7 | 5.9 | 68.1 | 57.4 | 52.8 | 187.4 |
| **$C_{ABS}$** | % | 13.0 | 28.0 | 28.7 | 5.1 | 37.8 | 2.3 | 12.3 | 10.5 | 8.6 | 31.0 | 1.5 | 8.4 | 6.4 | 6.4 | 24.6 | - | - | - | - | - |
| **$C_{REL}$** | % | 3.5 | 12.6 | 8.2 | 9.3 | 38.8 | 15.3 | 48.6 | 49.9 | 24.2 | 97.0 | 7.1 | 37.8 | 32.4 | 18.9 | 80.9 | - | - | - | - | - |
| **C:N (molar)** | - | 14.0 | 24.0 | 23.1 | 6.9 | 50.1 | 9.8 | 15.6 | 15.5 | 2.5 | 21.5 | 10.7 | 13.8 | 14.0 | 2.0 | 19.3 | 12.1 | 15.4 | 15.2 | 2.4 | 20.7 |
| **$\delta^{13}C$** | ‰ | -30.2 | -28.0 | -28.5 | 1.4 | -22.5 | -29.2 | -27.7 | -28.0 | 1.0 | -25.2 | -28.6 | -27.2 | -27.5 | 1.0 | -24.7 | -28.9 | -27.4 | -27.9 | 1.1 | -24.3 |
| **$\delta^{15}N$** | ‰ | -1.9 | 2.2 | 2.5 | 2.0 | 6.5 | -0.2 | 3.2 | 3.5 | 1.8 | 6.8 | -0.8 | 3.9 | 4.5 | 2.0 | 7.6 | 0.1 | 3.8 | 3.7 | 2.0 | 7.7 |
| **$RPA_{SIMP}$** | % | 72.6 | 87.5 | 88.5 | 4.5 | 94.5 | 65.6 | 84.4 | 86.3 | 9.4 | 99.9 | 46.5 | 74.4 | 76.6 | 12.9 | 92.4 | 13.5 | 71.8 | 77.8 | 17.7 | 90.5 |
| **$RPA_{COMP}$** | % | 0.6 | 4.8 | 4.3 | 3.4 | 19.0 | -2.8 | 9.8 | 7.1 | 9.5 | 33.9 | 2.4 | 14.0 | 12.9 | 8.9 | 36.8 | 2.8 | 15.1 | 12.3 | 11.4 | 54.9 |
| **$RPA_{MBIO}$** | % | 4.5 | 7.7 | 6.9 | 2.3 | 11.8 | -7.1 | 5.8 | 5.1 | 4.8 | 18.0 | 4.5 | 11.6 | 10.2 | 4.9 | 21.6 | 5.4 | 13.1 | 10.7 | 9.4 | 55.5 |

**Table S2-6.** Comparison of performances of the simple linear model (SLM), linear model allowing for interactions (ILM) and non-linear random forest model for each predicted SOC variable. The model with the best performance is highlighted in yellow. Transf. = transformation of the dependent variable; BP = p-value of Breusch-Pagan test for homoscedasticity of model residuals; SW = p-value of Shapiro-Wilk test for homoscedasticity of model residuals.

| | | Type | Transf. | RSME [log-scale] | btRMSE [orig. unit] | rRMSE [%] | Adj. $R^2$ | p-value | df | n | BP | SW |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_{TOT}$ | POM [gC $kg^{-1}$] | **SLM** | **log** | **0.66** | **6.00** | **12.5** | **0.46** | **< 0.05** | **31** | **35** | **0.18** | **0.53** |
| | | ILM | log | 0.61 | | | 0.52 | < 0.05 | 28 | 35 | 0.06 | 0.97 |
| | | RF | log | 0.70 | | | 0.43 | | | 35 | | |
| | S+C [gC $kg^{-1}$] | **SLM** | **log** | **0.43** | **9.94** | **17.9** | **0.65** | **< 0.05** | **31** | **35** | **0.66** | **0.05** |
| | | ILM | log | 0.44 | | | 0.63 | < 0.05 | 28 | 35 | 0.86 | 0.53 |
| | | RF | log | 0.49 | | | 0.56 | | | 35 | | |
| | SA [gC $kg^{-1}$] | **SLM** | **log** | **0.63** | **26.00** | **16.4** | **0.81** | **< 0.05** | **30** | **35** | **0.24** | **0.09** |
| | | ILM | log | 0.63 | | | 0.81 | < 0.05 | 24 | 35 | 0.47 | 0.05 |
| | | RF | log | 0.78 | | | 0.71 | | | 35 | | |
| | Bulk $SOC_{\%}$ [%] | SLM | log | 0.47 | | | 0.75 | < 0.05 | 31 | 35 | 0.20 | 0.13 |
| | | **ILM** | **log** | **0.40** | **3.08** | **17.0** | **0.81** | **< 0.05** | **28** | **35** | **0.34** | **0.63** |
| | | RF | log | 0.49 | | | 0.75 | | | 35 | | |

**Table S2-7.** Regression analysis of SOC quantities and qualities in the fractions with bulk $SOC_{\%}$. * BP-test not possible in models that have zero-intercept; **$C_{TOT}$ of POM is forced through zero and has one outlier. Transf. = transformation of the dependent variable; BP = p-value of Breusch-Pagan test for homoscedasticity of model residuals; SW = p-value of Shapiro-Wilk test for homoscedasticity of model residuals.

| Figure | Model | | Transf. | RSME | Adj. $R^2$ | p-value | df | n | BP | SW |
|---|---|---|---|---|---|---|---|---|---|---|
| Fig. 2a | $C_{TOT}$ | **POM**** | | 7.95 | 0.51 | < 0.05 | 34 | 35 | *NA** | *< 0.05* |
| | | **S+C** | | 10.89 | 0.39 | < 0.05 | 33 | 35 | *< 0.05* | *< 0.05* |
| | | **SA** | | 12.90 | 0.92 | < 0.05 | 33 | 35 | *< 0.05* | 0.12 |
| Fig. 2b | $C_{ABS}$ | **POM** | | 5.21 | -0.03 | 0.92 | 33 | 35 | 0.14 | 0.43 |
| | | **S+C** | | 3.08 | 0.77 | < 0.05 | 33 | 35 | 0.07 | *< 0.05* |
| | | **SA** | | 4.05 | 0.78 | < 0.05 | 33 | 35 | *< 0.05* | 0.06 |
| Fig. 2c | $C_{REL}$ | **POM** | log | 0.65 | 0.12 | < 0.05 | 33 | 35 | 0.46 | 0.92 |
| | | **S+C** | log | 0.40 | 0.45 | < 0.05 | 33 | 35 | 0.13 | 0.91 |
| | | **SA** | | 15.81 | 0.57 | < 0.05 | 33 | 35 | 0.47 | 0.26 |
| Fig. 3b | **DI** | **POM** | | 0.86 | 0.00 | 0.41 | 26 | 28 | 0.07 | 0.45 |
| | | **S+C** | | 1.08 | 0.35 | < 0.05 | 32 | 34 | 0.11 | 0.44 |
| | | **SA** | | 0.86 | 0.03 | 0.18 | 32 | 34 | 0.14 | 0.84 |

**Table S2-8.** Results of the rPCA for dimension reduction. The six retained rotated components (RCs) with Eigenvalue > 1 that resulted from this rPCA. The top shows Eigenvalues and individual as well as cumulative variability explained by the RCs. Loadings are shown for all 22 input variables. Interpretation and naming of the RCs was based on variables with loadings ≥ 0.5 (highlighted), and the assigned names for the RCs are given at the top. OM Decomp. = OM Decomposition; Ped. Oxides = Pedogenic Oxides.

| | RC1 | RC2 | RC3 | RC4 | RC5 | RC6 |
|---|---|---|---|---|---|---|
| | **Geochemistry** | **OM Decomp.** | **Texture** | **Ped. Oxides** | **Base Cations** | **Temperature** |
| Eigenvalue | 6.54 | 4.56 | 2.62 | 2.09 | 1.54 | 1.25 |
| Proportion var. | 0.18 | 0.14 | 0.13 | 0.09 | 0.08 | 0.07 |
| Cumulative var. | 0.18 | 0.32 | 0.45 | 0.54 | 0.63 | 0.7 |
| pH ($CaCl_2$) | -0.11 | 0.38 | -0.13 | -0.10 | 0.04 | 0.15 |
| Sand | -0.28 | -0.16 | -0.93 | 0.05 | -0.04 | -0.12 |
| Silt | 0.20 | 0.12 | 0.96 | 0.00 | 0.04 | 0.07 |
| Clay | 0.46 | 0.25 | 0.67 | -0.23 | 0.03 | 0.23 |
| Al (Aq.Reg.) | 0.87 | 0.09 | 0.19 | 0.17 | 0.07 | 0.04 |
| Fe (Aq.Reg.) | 0.92 | 0.03 | 0.24 | -0.06 | 0.10 | 0.10 |
| Al (PP + OXA) | 0.12 | -0.02 | -0.06 | 0.98 | -0.04 | -0.10 |
| Fe (PP + OXA) | 0.32 | -0.16 | -0.09 | 0.66 | 0.00 | -0.10 |
| Mn (PP + OXA) | 0.51 | 0.23 | 0.16 | -0.10 | 0.11 | -0.03 |
| $Fe_{Aq.Reg.}$ : Si | 0.90 | -0.02 | 0.15 | 0.28 | 0.06 | -0.04 |
| $Clay_{prim}$ : Si | 0.75 | 0.17 | 0.43 | 0.09 | 0.11 | 0.02 |
| MAT | 0.01 | 0.10 | 0.17 | -0.22 | -0.06 | 0.94 |
| MAP - PET | 0.11 | -0.31 | -0.03 | 0.45 | -0.06 | -0.34 |
| NPP | 0.18 | -0.07 | 0.17 | 0.24 | 0.20 | 0.53 |
| TempSeas | 0.06 | 0.19 | 0.13 | -0.04 | 0.01 | 0.01 |
| Ca ($BaCl_2$) | 0.18 | -0.19 | 0.02 | 0.06 | 0.82 | -0.04 |
| Mg ($BaCl_2$) | 0.11 | -0.10 | 0.08 | -0.09 | 0.92 | -0.02 |
| K ($BaCl_2$) | -0.17 | -0.06 | 0.00 | -0.06 | 0.38 | 0.07 |
| $RPA_{SIMP}$ (bulk) | -0.07 | -0.95 | -0.16 | 0.06 | 0.11 | -0.06 |
| $RPA_{COMP}$ (bulk) | -0.01 | 0.85 | 0.04 | -0.28 | -0.19 | 0.18 |
| $RPA_{MBIO}$ (bulk) | 0.13 | 0.86 | 0.25 | 0.17 | 0.00 | -0.07 |
| Rock-Eval I-Index | -0.31 | -0.50 | -0.26 | -0.09 | 0.10 | -0.49 |

## 7.3 Appendix Chapter 3

**File S3-1.** This file summarizes the results of the correlation analyses for all taxonomic units. Each row corresponds to one taxonomic unit. Listed are the taxonomic classifications from phylum to ASV, the actual taxonomic resolution of the taxonomic unit (column “Level”), as well as the taxonomic classification of the taxonomic unit at its respective taxonomic resolution (column “Name”). The column “Mean RA” gives the mean relative abundance of the taxonomic unit across all 35 sites. The following columns give the correlation coefficients (r) of significant correlations (FDR corrected, at a significance threshold $p < 0.1$) for each predictor variable with relative abundance. For visual aid, negative correlations are highlighted in red, and positive correlations are highlighted in blue. As provided, the taxonomic units are hierarchically sorted in alphabetic order by Phylum, Class, Order, Family, Genus, ASV. For this thesis, the file is provided as a table.

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA % | pH r | Texture r | SOM Quantity r | SOM Quality A r | SOM Quality B r | SOM Quality C r | SOM Quality D r | SOM Quality E r |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Acidobacteriota | Acidobacteriae | Acidobacteriales | Acidobacteriaceae (Subgroup 1) | Acidipila-Silvibacterium | NA | Genus | ASV394 | Acidipila-Silvibacterium | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Acidobacteriales | Acidobacteriaceae (Subgroup 1) | Edaphobacter | NA | Genus | ASV810 | Edaphobacter | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Acidobacteriales | Acidobacteriaceae (Subgroup 1) | Granulicella | NA | Genus | ASV1135 | Granulicella | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Acidobacteriales | Acidobacteriaceae (Subgroup 1) | Occallatibacter | NA | Genus | ASV501 | Occallatibacter | 0.12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Acidobacteriales | Acidobacteriaceae (Subgroup 1) | U_Genus 2 | NA | Genus | ASV1088 | U_Genus 2 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Acidobacteriales | Koribacteraceae | U_Genus 3 | NA | Family | ASV100 | Koribacteraceae | 0.37 | -0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Acidobacteriales | Unested_Family 1 | U_Genus 5 | NA | Family | ASV60 | Unested_Family 1 | 1.51 | -0.8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Bryobacterales | Bryobacteraceae | Bryobacter | ASV332 | ASV | ASV332 | ASV332 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Bryobacterales | Bryobacteraceae | Bryobacter | ASV564 | ASV | ASV564 | ASV564 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.46 |
| Acidobacteriota | Acidobacteriae | Bryobacterales | Rare Bryobacterales | U_Genus 8 | NA | Family | ASV180 | Rare Bryobacterales | 0.97 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Rare Acidobacteriae | U_Family 2 | U_Genus 6 | NA | Order | ASV60 | Rare Acidobacteriae | 0.04 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Solibacterales | Rare Solibacterales | U_Genus 9 | NA | Family | ASV56 | Rare Solibacterales | 1.07 | -0.7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Solibacterales | Solibacteraceae | Candidatus Solibacter | ASV114 | ASV | ASV114 | ASV114 | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Solibacterales | Solibacteraceae | Candidatus Solibacter | ASV224 | ASV | ASV224 | ASV224 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Solibacterales | Solibacteraceae | Candidatus Solibacter | ASV56 | ASV | ASV56 | ASV56 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | Subgroup 2 | U_Family 6 | U_Genus 10 | NA | Order | ASV245 | Subgroup 2 | 0.69 | -0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Acidobacteriae | U_Order 3 | U_Family 7 | U_Genus 11 | NA | Order | ASV6430 | U_Order 3 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Blastocatellia | 11-24 | U_Family 8 | U_Genus 12 | NA | Order | ASV790 | 11-24 | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Blastocatellia | Blastocatellales | Blastocatellaceae | JGI 0001001-H03 | NA | Genus | ASV425 | JGI 0001001-H03 | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Blastocatellia | Blastocatellales | Blastocatellaceae | Stenotrophobacter | NA | Genus | ASV490 | Stenotrophobacter | 0.07 | 0.56 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Blastocatellia | Blastocatellales | Blastocatellaceae | U_Genus 14 | NA | Genus | ASV570 | U_Genus 14 | 0.07 | 0.46 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Blastocatellia | Blastocatellales | Rare Blastocatellales | U_Genus 13 | NA | Family | ASV425 | Rare Blastocatellales | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Blastocatellia | Pyrinomonadales | U_Family 10 | U_Genus 15 | NA | Order | ASV124 | Pyrinomonadales | 0.90 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Blastocatellia | Rare Blastocatellia | U_Family 11 | U_Genus 16 | NA | Order | ASV124 | Rare Blastocatellia | 0.02 | 0.52 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Holophagae | Subgroup 7 | U_Family 12 | U_Genus 17 | NA | Order | ASV288 | Subgroup 7 | 0.50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Subgroup 5 | U_Order 6 | U_Family 14 | U_Genus 19 | NA | Class | ASV4332 | Subgroup 5 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Thermoanaerobaculia | U_Order 7 | U_Family 15 | U_Genus 20 | NA | Class | ASV2284 | Thermoanaerobaculia | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | U_Class 2 | U_Order 9 | U_Family 19 | U_Genus 26 | NA | Class | ASV37377 | U_Class 2 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Vicinamibacteria | Subgroup 17 | U_Family 16 | U_Genus 21 | NA | Order | ASV958 | Subgroup 17 | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Vicinamibacteria | Vicinamibacterales | Unested_Family 2 | U_Genus 24 | NA | Family | ASV234 | Unested_Family 2 | 1.81 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Acidobacteriota | Vicinamibacteria | Vicinamibacterales | Vicinamibacteraceae | Unested_Genus 1 | NA | Genus | ASV769 | Unested_Genus 1 | 0.30 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Acidimicrobiia | IMCC26256 | U_Family 20 | U_Genus 27 | NA | Order | ASV87 | IMCC26256 | 1.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Acidimicrobiia | Microtrichales | Iamiaceae | U_Genus 29 | NA | Family | ASV1095 | Iamiaceae | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Acidimicrobiia | Microtrichales | Ilumatobacteraceae | CL500-29 marine group | NA | Genus | ASV107 | CL500-29 marine group | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Acidimicrobiia | Microtrichales | Ilumatobacteraceae | U_Genus 32 | NA | Genus | ASV430 | U_Genus 32 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Acidimicrobiia | Microtrichales | U_Family 23 | U_Genus 33 | NA | Family | ASV337 | U_Family 23 | 0.19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Acidimicrobiia | U_Order 11 | U_Family 24 | U_Genus 34 | NA | Order | ASV276 | U_Order 11 | 0.24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Catenulisporales | Catenulisporaceae | U_Genus 35 | NA | Family | ASV945 | Catenulisporaceae | 0.04 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Corynebacteriales | Mycobacteriaceae | Mycobacterium | ASV421 | ASV | ASV421 | ASV421 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Corynebacteriales | Mycobacteriaceae | Mycobacterium | ASV689 | ASV | ASV689 | ASV689 | 0.02 | 0 | 0 | 0 | 0 | 0.46 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Corynebacteriales | Mycobacteriaceae | Mycobacterium | Rare Mycobacterium | ASV | ASV144 | Rare Mycobacterium | 1.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Corynebacteriales | Nocardiaceae | Nocardia | NA | Genus | ASV5460 | Nocardia | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | -0.48 | 0 |
| Actinobacteriota | Actinobacteria | Corynebacteriales | Nocardiaceae | Rhodococcus | NA | Genus | ASV532 | Rhodococcus | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Acidothermaceae | Acidothermus | ASV68 | ASV | ASV68 | ASV68 | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Acidothermaceae | Acidothermus | ASV82 | ASV | ASV82 | ASV82 | 0.08 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Acidothermaceae | Rare Acidothermaceae | NA | Genus | ASV68 | Rare Acidothermaceae | 0.37 | -0.7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Frankiaceae | Jatrophihabitans | NA | Genus | ASV739 | Jatrophihabitans | 0.16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Geodermatophilaceae | Blastococcus | ASV106 | ASV | ASV106 | ASV106 | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Geodermatophilaceae | Blastococcus | Rare Blastococcus | ASV | ASV106 | Rare Blastococcus | 0.21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Geodermatophilaceae | Rare Geodermatophilaceae | NA | Genus | ASV106 | Rare Geodermatophilaceae | 0.22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Nakamurellaceae | Nakamurella | ASV494 | Genus | ASV494 | Nakamurella | 0.02 | 0.49 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | Nakamurellaceae | Rare Nakamurellaceae | NA | Genus | ASV442 | Rare Nakamurellaceae | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**File S3-1 cont.**

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA | pH | Texture | SOM Quantity | SOM Quality A | SOM Quality B | SOM Quality C | SOM Quality D | SOM Quality E |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | % | r | r | r | r | r | r | r | r |
| Actinobacteriota | Actinobacteria | Frankiales | Sporichthyaceae | Unested_Genus 2 | NA | Genus | ASV3261 | Unested_Genus 2 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Frankiales | U_Family 28 | U_Genus 46 | NA | Family | ASV1383 | U_Family 28 | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Kineosporiales | Kineosporiaceae | Angustibacter | NA | Genus | ASV1543 | Angustibacter | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Kineosporiales | Kineosporiaceae | U_Genus 48 | NA | Genus | ASV2218 | U_Genus 48 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Cellulomonadaceae | Cellulomonas | NA | Genus | ASV624 | Cellulomonas | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Intrasporangiaceae | Oryzihumus | NA | Genus | ASV529 | Oryzihumus | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Intrasporangiaceae | Pedococcus-Phycicoccus | NA | Genus | ASV916 | Pedococcus-Phycicoccus | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Intrasporangiaceae | Rare Intrasporangiaceae | NA | Genus | ASV279 | Rare Intrasporangiaceae | 0.06 | 0 | -0.52 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Intrasporangiaceae | Terrabacter | NA | Genus | ASV279 | Terrabacter | 0.07 | 0 | -0.53 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Intrasporangiaceae | U_Genus 51 | NA | Genus | ASV1179 | U_Genus 51 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Microbacteriaceae | Rare Microbacteriaceae | NA | Genus | ASV280 | Rare Microbacteriaceae | 0.19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Microbacteriaceae | U_Genus 53 | NA | Genus | ASV411 | U_Genus 53 | 0.24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Acaricomes | NA | Genus | ASV14375 | Acaricomes | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Arthrobacter | NA | Genus | ASV94 | Arthrobacter | 0.23 | 0 | 0 | 0 | 0.49 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Paenarthrobacter | ASV53 | ASV | ASV53 | ASV53 | 0.12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Paenarthrobacter | Rare Paenarthrobacter | ASV | ASV53 | Rare Paenarthrobacter | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | ASV1 | ASV | ASV1 | ASV1 | 5.88 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | ASV181 | ASV | ASV181 | ASV181 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | ASV23 | ASV | ASV23 | ASV23 | 0.22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | ASV27 | ASV | ASV27 | ASV27 | 0.20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | ASV3 | ASV | ASV3 | ASV3 | 1.32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | ASV80 | ASV | ASV80 | ASV80 | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | ASV92 | ASV | ASV92 | ASV92 | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | ASV965 | ASV | ASV965 | ASV965 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Pseudarthrobacter | Rare Pseudarthrobacter | ASV | ASV1 | Rare Pseudarthrobacter | 0.32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Rare Micrococcaceae | NA | Genus | ASV1 | Rare Micrococcaceae | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Unested_Genus 3 | ASV21 | ASV | ASV21 | ASV21 | 0.22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.52 |
| Actinobacteriota | Actinobacteria | Micrococcales | Micrococcaceae | Unested_Genus 3 | Rare Unested_Genus 3 | ASV | ASV21 | Rare Unested_Genus 3 | 0.97 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.65 |
| Actinobacteriota | Actinobacteria | Micrococcales | Rare Micrococcales | U_Genus 55 | NA | Family | ASV1 | Rare Micrococcales | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micrococcales | U_Family 33 | U_Genus 58 | NA | Family | ASV10820 | U_Family 33 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micromonosporales | Micromonosporaceae | Actinoplanes | NA | Genus | ASV973 | Actinoplanes | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micromonosporales | Micromonosporaceae | Dactylosporangium | NA | Genus | ASV1063 | Dactylosporangium | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micromonosporales | Micromonosporaceae | Luedemannella | NA | Genus | ASV778 | Luedemannella | 0.11 | 0 | 0 | 0 | 0 | 0.49 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micromonosporales | Micromonosporaceae | Micromonospora | NA | Genus | ASV870 | Micromonospora | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micromonosporales | Micromonosporaceae | Unested_Genus 4 | NA | Genus | ASV426 | Unested_Genus 4 | 0.19 | 0 | 0 | 0 | 0 | 0.68 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Micromonosporales | Rare Micromonosporales | U_Genus 59 | NA | Family | ASV426 | Rare Micromonosporales | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Propionibacteriales | Nocardioidaceae | Kribbella | NA | Genus | ASV126 | Kribbella | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Propionibacteriales | Nocardioidaceae | Marmoricola | NA | Genus | ASV474 | Marmoricola | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Propionibacteriales | Nocardioidaceae | Nocardioides | NA | Genus | ASV709 | Nocardioides | 0.21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Propionibacteriales | Nocardioidaceae | Rare Nocardioidaceae | NA | Genus | ASV126 | Rare Nocardioidaceae | 0.08 | 0 | 0 | 0 | 0 | 0.49 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Pseudonocardiales | Pseudonocardiaceae | Pseudonocardia | NA | Family | ASV342 | Pseudonocardiaceae | 0.19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Pseudonocardiales | Rare Pseudonocardiales | U_Genus 62 | NA | Family | ASV342 | Rare Pseudonocardiales | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Streptomycetales | Rare Streptomycetales | U_Genus 63 | NA | Family | ASV48 | Rare Streptomycetales | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Streptomycetales | Streptomycetaceae | Streptomyces | ASV48 | ASV | ASV48 | ASV48 | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Streptomycetales | Streptomycetaceae | Streptomyces | Rare Streptomyces | ASV | ASV48 | Rare Streptomyces | 0.23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Streptomycetales | Streptomycetaceae | U_Genus 64 | NA | Genus | ASV882 | U_Genus 64 | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Streptosporangiales | Streptosporangiaceae | Rare Streptosporangiaceae | NA | Genus | ASV110 | Rare Streptosporangiaceae | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Streptosporangiales | Streptosporangiaceae | Streptosporangium | ASV110 | ASV | ASV110 | ASV110 | 0.07 | 0 | 0 | 0 | 0 | 0.6 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Streptosporangiales | Thermomonosporaceae | Actinoallomurus | NA | Genus | ASV903 | Actinoallomurus | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | Streptosporangiales | Thermomonosporaceae | U_Genus 68 | NA | Genus | ASV2256 | U_Genus 68 | 0.03 | 0 | 0 | 0 | 0 | 0.61 | 0 | 0 | 0 |
| Actinobacteriota | Actinobacteria | U_Order 14 | U_Family 39 | U_Genus 69 | NA | Order | ASV5436 | U_Order 14 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | MB-A2-108 | U_Order 15 | U_Family 40 | U_Genus 70 | NA | Class | ASV577 | MB-A2-108 | 0.16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**File S3-1 cont.**

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA % | pH r | Texture r | SOM Quantity r | SOM Quality A r | SOM Quality B r | SOM Quality C r | SOM Quality D r | SOM Quality E r |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Actinobacteriota | Thermoleophilia | Gaiellales | Gaiellaceae | U_Genus 71 | NA | Family | ASV143 | Gaiellaceae | 0.30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Thermoleophilia | Gaiellales | U_Family 43 | U_Genus 74 | NA | Family | ASV232 | U_Family 43 | 0.95 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Thermoleophilia | Solirubrobacterales | 67-14 | U_Genus 75 | NA | Family | ASV168 | 67-14 | 0.33 | 0.57 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Thermoleophilia | Solirubrobacterales | Solirubrobacteraceae | Conexibacter | NA | Genus | ASV383 | Conexibacter | 0.56 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Thermoleophilia | Solirubrobacterales | Solirubrobacteraceae | Solirubrobacter | NA | Genus | ASV1114 | Solirubrobacter | 0.21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Thermoleophilia | Solirubrobacterales | Solirubrobacteraceae | U_Genus 78 | NA | Genus | ASV587 | U_Genus 78 | 0.25 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Thermoleophilia | Solirubrobacterales | U_Family 45 | U_Genus 79 | NA | Family | ASV5293 | U_Family 45 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | Thermoleophilia | U_Order 17 | U_Family 46 | U_Genus 80 | NA | Order | ASV1139 | U_Order 17 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Actinobacteriota | U_Class 4 | U_Order 18 | U_Family 47 | U_Genus 81 | NA | Class | ASV2551 | U_Class 4 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Armatimonadota | Fimbriimonadia | Fimbriimonadales | Fimbriimonadaceae | Unested_Genus 5 | NA | Order | ASV7852 | Fimbriimonadales | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Chitinophagales | Chitinophagaceae | Edaphobaculum | NA | Genus | ASV6528 | Edaphobaculum | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Chitinophagales | Chitinophagaceae | Ferruginibacter | NA | Genus | ASV1275 | Ferruginibacter | 0.08 | 0.51 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Chitinophagales | Chitinophagaceae | Puia | NA | Genus | ASV364 | Puia | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Chitinophagales | Chitinophagaceae | Rare Chitinophagaceae | NA | Genus | ASV364 | Rare Chitinophagaceae | 0.26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Chitinophagales | Chitinophagaceae | U_Genus 88 | NA | Genus | ASV1070 | U_Genus 88 | 0.23 | 0 | 0 | 0 | 0 | 0.43 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Cytophagales | Hymenobacteraceae | Hymenobacteraceae | NA | Family | ASV1460 | Hymenobacteraceae | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Cytophagales | Microscillaceae | U_Genus 91 | NA | Family | ASV63889 | Microscillaceae | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Cytophagales | Rare Cytophagales | U_Genus 90 | NA | Family | ASV1460 | Rare Cytophagales | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Flavobacteriales | Flavobacteriaceae | Flavobacterium | NA | Genus | ASV1865 | Flavobacterium | 0.12 | 0.48 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Sphingobacteriales | Rare Sphingobacteriales | U_Genus 95 | NA | Family | ASV2160 | Rare Sphingobacteriales | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Sphingobacteriales | Sphingobacteriaceae | Mucilaginibacter | NA | Genus | ASV4224 | Mucilaginibacter | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Bacteroidia | Sphingobacteriales | Sphingobacteriaceae | Pedobacter | NA | Genus | ASV2160 | Pedobacter | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bacteroidota | Kapabacteria | U_Order 23 | U_Family 56 | U_Genus 96 | NA | Class | ASV7506 | Kapabacteria | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bdellovibrionota | Bdellovibrionia | Bdellovibrionales | Bdellovibrionaceae | Bdellovibrio | NA | Family | ASV5330 | Bdellovibrionaceae | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bdellovibrionota | Oligoflexia | 0319-6G20 | U_Family 60 | U_Genus 100 | NA | Order | ASV6718 | 0319-6G20 | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | AD3 | U_Order 27 | U_Family 62 | U_Genus 102 | NA | Class | ASV105 | AD3 | 0.66 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Anaerolineae | Anaerolineales | Anaerolineaceae | U_Genus 105 | NA | Family | ASV547 | Anaerolineaceae | 0.35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Anaerolineae | Caldilineales | Caldilineaceae | Unested_Genus 6 | NA | Family | ASV331 | Caldilineaceae | 0.17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Anaerolineae | RBG-13-54-9 | U_Family 66 | U_Genus 107 | NA | Order | ASV2329 | RBG-13-54-9 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Anaerolineae | SBR1031 | A4b | Unested_Genus 7 | NA | Genus | ASV1147 | Unested_Genus 7 | 0.19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Anaerolineae | SBR1031 | Unested_Family 3 | U_Genus 110 | NA | Family | ASV962 | Unested_Family 3 | 0.19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Chloroflexia | Chloroflexales | Roseiflexaceae | Unested_Genus 8 | NA | Genus | ASV569 | Unested_Genus 8 | 0.52 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Chloroflexia | Kallotenuales | AKIW781 | U_Genus 113 | NA | Family | ASV716 | AKIW781 | 0.23 | 0.49 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Chloroflexia | Thermomicrobiales | JG30-KF-CM45 | U_Genus 115 | NA | Family | ASV122 | JG30-KF-CM45 | 1.01 | 0.53 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Dehalococcoidia | S085 | U_Family 72 | U_Genus 118 | NA | Order | ASV727 | S085 | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | JG30-KF-CM66 | U_Order 31 | U_Family 74 | U_Genus 120 | NA | Class | ASV1960 | JG30-KF-CM66 | 0.12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | KD4-96 | U_Order 32 | U_Family 75 | U_Genus 121 | NA | Class | ASV112 | KD4-96 | 1.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | B12-WMSP1 | U_Family 76 | U_Genus 122 | NA | Order | ASV211 | B12-WMSP1 | 0.23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | C0119 | U_Family 77 | U_Genus 123 | NA | Order | ASV347 | C0119 | 0.32 | 0 | -0.54 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | Ktedonobacterales | JG30-KF-AS9 | U_Genus 124 | NA | Family | ASV366 | JG30-KF-AS9 | 0.33 | -0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae | 1959-1 | NA | Genus | ASV996 | 1959-1 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae | FCPS473 | NA | Genus | ASV167 | FCPS473 | 0.19 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae | HSB OF53-F07 | NA | Genus | ASV33 | HSB OF53-F07 | 3.32 | 0 | 0 | 0.53 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae | JG30-KF-32 | NA | Genus | ASV249 | JG30-KF-32 | 0.32 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae | Ktedonobacter | NA | Genus | ASV193 | Ktedonobacter | 0.20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae | Rare Ktedonobacteraceae | NA | Genus | ASV33 | Rare Ktedonobacteraceae | 0.23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Ktedonobacteria | Ktedonobacterales | Ktedonobacteraceae | U_Genus 129 | NA | Genus | ASV88 | U_Genus 129 | 0.70 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | OLB14 | U_Order 35 | U_Family 81 | U_Genus 130 | NA | Class | ASV1991 | OLB14 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | Rare Chloroflexi | U_Order 34 | U_Family 80 | U_Genus 128 | NA | Class | ASV33 | Rare Chloroflexi | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | TK10 | U_Order 36 | U_Family 82 | U_Genus 131 | NA | Class | ASV479 | TK10 | 0.56 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Chloroflexi | U_Class 9 | U_Order 37 | U_Family 83 | U_Genus 132 | NA | Class | ASV3799 | U_Class 9 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Cyanobacteria | Cyanobacteriia | Chloroplast | U_Family 84 | U_Genus 133 | NA | Order | ASV3451 | Chloroplast | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**File S3-1 cont.**

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA % | pH r | Texture r | SOM Quantity r | SOM Quality A r | SOM Quality B r | SOM Quality C r | SOM Quality D r | SOM Quality E r |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cyanobacteria | Sericytochromatia | U_Order 40 | U_Family 87 | U_Genus 136 | NA | Class | ASV7106 | Sericytochromatia | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Cyanobacteria | Vampirivibrionia | Obscuribacterales | Obscuribacteraceae | Unested_Genus 9 | NA | Family | ASV3130 | Obscuribacteraceae | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dependentiae | Babeliae | Babeliales | Vermiphilaceae | U_Genus 139 | NA | Class | ASV8732 | Babeliae | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Dependentiae | Rare Dependentiae | U_Order 42 | U_Family 90 | U_Genus 140 | NA | Class | ASV7445 | Rare Dependentiae | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Desulfobacterota | Desulfuromonadia | Geobacterales | Geobacteraceae | U_Genus 141 | NA | Family | ASV1785 | Geobacteraceae | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Elusimicrobiota | U_Class 13 | U_Order 45 | U_Family 94 | U_Genus 145 | NA | Phylum | ASV13201 | Elusimicrobiota | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Alicyclobacillales | Alicyclobacillaceae | Tumebacillus | NA | Family | ASV719 | Alicyclobacillaceae | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV127 | ASV | ASV127 | ASV127 | 0.06 | 0.56 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV136 | ASV | ASV136 | ASV136 | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV154 | ASV | ASV154 | ASV154 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV162 | ASV | ASV162 | ASV162 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -0.47 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV18 | ASV | ASV18 | ASV18 | 0.34 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV230 | ASV | ASV230 | ASV230 | 0.04 | 0 | 0 | 0.56 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV255 | ASV | ASV255 | ASV255 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV257 | ASV | ASV257 | ASV257 | 0.04 | 0 | 0 | 0.6 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV264 | ASV | ASV264 | ASV264 | 0.04 | 0.51 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV338 | ASV | ASV338 | ASV338 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV377 | ASV | ASV377 | ASV377 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV38 | ASV | ASV38 | ASV38 | 0.16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV381 | ASV | ASV381 | ASV381 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV390 | ASV | ASV390 | ASV390 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV46 | ASV | ASV46 | ASV46 | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV52 | ASV | ASV52 | ASV52 | 0.12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV6 | ASV | ASV6 | ASV6 | 0.81 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV61 | ASV | ASV61 | ASV61 | 0.11 | 0 | 0 | 0 | 0 | -0.6 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | ASV979 | ASV | ASV979 | ASV979 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Bacillus | Rare Bacillus | ASV | ASV6 | Rare Bacillus | 0.78 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Oceanobacillus | NA | Genus | ASV503 | Oceanobacillus | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Rare Bacillaceae | NA | Genus | ASV2 | Rare Bacillaceae | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Unested_Genus 10 | ASV2 | ASV | ASV2 | ASV2 | 1.90 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Bacillaceae | Unested_Genus 10 | Rare Unested_Genus 10 | ASV | ASV2 | Rare Unested_Genus 10 | 3.82 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | [Paenibacterium] salmoninarum group | NA | Genus | ASV15391 | [Paenibacterium] salmoninarum group | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Carnobacterium | NA | Genus | ASV305 | Carnobacterium | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Chungangia | ASV205 | ASV | ASV205 | ASV205 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Chungangia | Rare Chungangia | ASV | ASV205 | Rare Chungangia | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Domibacillus | NA | Genus | ASV696 | Domibacillus | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Filibacter | NA | Genus | ASV1724 | Filibacter | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Lysinibacillus | ASV101 | ASV | ASV101 | ASV101 | 0.08 | 0 | 0 | 0.48 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Lysinibacillus | ASV26 | ASV | ASV26 | ASV26 | 0.21 | 0 | 0 | 0.68 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Lysinibacillus | ASV268 | ASV | ASV268 | ASV268 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Lysinibacillus | ASV351 | ASV | ASV351 | ASV351 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Lysinibacillus | ASV395 | ASV | ASV395 | ASV395 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Lysinibacillus | Rare Lysinibacillus | ASV | ASV26 | Rare Lysinibacillus | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Psychrobacillus | ASV103 | ASV | ASV103 | ASV103 | 0.08 | 0 | 0 | 0.73 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Psychrobacillus | ASV37 | ASV | ASV37 | ASV37 | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Psychrobacillus | ASV51 | ASV | ASV51 | ASV51 | 0.13 | 0 | 0 | 0.65 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Psychrobacillus | Rare Psychrobacillus | ASV | ASV37 | Rare Psychrobacillus | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Rare Planococcaceae | NA | Genus | ASV14 | Rare Planococcaceae | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Rummeliibacillus | NA | Genus | ASV777 | Rummeliibacillus | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Solibacillus | ASV225 | ASV | ASV225 | ASV225 | 0.04 | 0 | 0 | 0.56 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Solibacillus | ASV455 | ASV | ASV455 | ASV455 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Solibacillus | ASV506 | ASV | ASV506 | ASV506 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**File S3-1 cont.**

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA % | pH r | Texture r | SOM Quantity r | SOM Quality A r | SOM Quality B r | SOM Quality C r | SOM Quality D r | SOM Quality E r |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Solibacillus | ASV671 | ASV | ASV671 | ASV671 | 0.02 | 0 | 0 | 0.58 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Solibacillus | ASV225 | ASV | ASV225 | Rare Solibacillus | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Sporosarcina | ASV14 | ASV | ASV14 | ASV14 | 0.39 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Sporosarcina | ASV189 | ASV | ASV189 | ASV189 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Sporosarcina | ASV28 | ASV | ASV28 | ASV28 | 0.20 | 0 | 0 | 0.57 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Sporosarcina | ASV42 | ASV | ASV42 | ASV42 | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Sporosarcina | ASV62 | ASV | ASV62 | ASV62 | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.53 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Sporosarcina | ASV14 | ASV | ASV14 | Rare Sporosarcina | 0.32 | 0 | 0 | 0.54 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Unested_Genus 11 | ASV330 | ASV | ASV330 | ASV330 | 0.03 | 0.53 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Planococcaceae | Unested_Genus 11 | ASV330 | ASV | ASV330 | Rare Unested_Genus 11 | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Bacillales | Unested_Family 4 | U_Genus 152 | ASV24 | Family | ASV24 | Unested_Family 4 | 0.44 | 0.49 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Erysipelotrichales | Erysipelotrichaceae | Turicibacter | ASV219 | ASV | ASV219 | ASV219 | 0.04 | 0 | 0 | 0.67 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Erysipelotrichales | Erysipelotrichaceae | Turicibacter | ASV219 | ASV | ASV219 | Rare Turicibacter | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Paenibacillales | Paenibacillaceae | Ammoniphilus | ASV194 | ASV | ASV194 | ASV194 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Paenibacillales | Paenibacillaceae | Ammoniphilus | ASV194 | ASV | ASV194 | Rare Ammoniphilus | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Paenibacillales | Paenibacillaceae | Cohnella | ASV1607 | Genus | ASV1607 | Cohnella | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Paenibacillales | Paenibacillaceae | Paenibacillus | ASV505 | ASV | ASV505 | ASV505 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Paenibacillales | Paenibacillaceae | Paenibacillus | ASV505 | ASV | ASV505 | Rare Paenibacillus | 0.59 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Paenibacillales | Paenibacillaceae | U_Genus 156 | ASV1289 | Genus | ASV1289 | U_Genus 156 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Rare Bacilli | U_Family 97 | U_Genus 149 | ASV2 | Order | ASV2 | Rare Bacilli | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | Thermoactinomycetales | U_Family 101 | U_Genus 157 | ASV702 | Order | ASV702 | Thermoactinomycetales | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Bacilli | U_Order 48 | U_Family 102 | U_Genus 158 | ASV2501 | Order | ASV2501 | U_Order 48 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Clostridiales | Caloramatoraceae | Fonticella | ASV4407 | Genus | ASV4407 | Fonticella | 0.08 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Clostridiales | Clostridiaceae | Clostridium sensu stricto 1 | ASV79 | Genus | ASV79 | Clostridium sensu stricto 1 | 0.30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Clostridiales | Clostridiaceae | Clostridium sensu stricto 13 | ASV30 | ASV | ASV30 | ASV30 | 0.20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Clostridiales | Clostridiaceae | Clostridium sensu stricto 13 | ASV30 | ASV | ASV30 | Rare Clostridium sensu stricto 13 | 0.02 | 0 | 0 | 0.51 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Clostridiales | Clostridiaceae | Clostridium sensu stricto 9 | ASV172 | Genus | ASV172 | Clostridium sensu stricto 9 | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Clostridiales | Clostridiaceae | Rare Clostridiaceae | ASV30 | Genus | ASV30 | Rare Clostridiaceae | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Clostridiales | Clostridiaceae | U_Genus 162 | ASV2761 | Genus | ASV2761 | U_Genus 162 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Lachnospirales | Lachnospiraceae | U_Genus 163 | ASV4148 | Family | ASV4148 | Lachnospiraceae | 0.04 | 0 | 0 | 0.64 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Oscillospirales | U_Family 105 | U_Genus 165 | ASV1335 | Order | ASV1335 | Oscillospirales | 0.06 | -0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Peptostreptococcales -Tissierellales | Peptostreptococcaceae | Paeniclostridium | NA | Genus | ASV133 | Paeniclostridium | 0.06 | 0 | 0 | 0.6 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Peptostreptococcales -Tissierellales | Peptostreptococcaceae | Rare Peptostreptococcaceae | NA | Genus | ASV16 | Rare Peptostreptococcaceae | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Peptostreptococcales -Tissierellales | Peptostreptococcaceae | Romboutsia | ASV113 | ASV | ASV113 | ASV113 | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Peptostreptococcales -Tissierellales | Peptostreptococcaceae | Romboutsia | ASV16 | ASV | ASV16 | ASV16 | 0.38 | 0 | 0 | 0.57 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Peptostreptococcales -Tissierellales | Peptostreptococcaceae | Unested_Genus 12 | ASV41 | ASV | ASV41 | ASV41 | 0.15 | 0 | 0 | 0.62 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Clostridia | Rare Clostridia | U_Family 107 | U_Genus 168 | NA | Order | ASV16 | Rare Clostridia | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | Rare Firmicutes | U_Order 47 | U_Family 98 | U_Genus 150 | NA | Class | ASV2 | Rare Firmicutes | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Firmicutes | U_Class 15 | U_Order 50 | U_Family 108 | U_Genus 169 | NA | Class | ASV21242 | U_Class 15 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Gemmatimonadota | Gemmatimonadetes | Gemmatimonadales | Gemmatimonadaceae | Gemmatimonas | NA | Genus | ASV1163 | Gemmatimonas | 0.41 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Gemmatimonadota | Gemmatimonadetes | Gemmatimonadales | Gemmatimonadaceae | Unested_Genus 13 | NA | Genus | ASV1104 | Unested_Genus 13 | 0.61 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Methylomirabilota | Methylomirabilia | Rokubacteriales | Unested_Family 5 | U_Genus 173 | NA | Family | ASV1480 | Unested_Family 5 | 0.16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Myxococcia | Myxococcales | Anaeromyxobacteraceae | U_Genus 175 | NA | Family | ASV2129 | Anaeromyxobacteraceae | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Myxococcia | Myxococcales | Myxococcaceae | Rare Myxococcaceae | NA | Genus | ASV1236 | Rare Myxococcaceae | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Myxococcia | Myxococcales | Myxococcaceae | Unested_Genus 14 | NA | Genus | ASV1236 | Unested_Genus 14 | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**File S3-1 cont.**

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA % | pH r | Texture r | SOM Quantity r | SOM Quality A r | SOM Quality B r | SOM Quality C r | SOM Quality D r | SOM Quality E r |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Myxococcota | Polyangia | Haliangiales | U_Family 114 | U_Genus 178 | NA | Order | ASV2562 | Haliangiales | 0.21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | mle1-27 | U_Family 115 | U_Genus 179 | NA | Order | ASV7539 | mle1-27 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | Polyangiales | Bltii41 | U_Genus 180 | NA | Family | ASV2485 | Bltii41 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | Polyangiales | Phaselicystidaceae | U_Genus 181 | NA | Family | ASV650 | Phaselicystidaceae | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | Polyangiales | Polyangiaceae | Aetherobacter | NA | Genus | ASV2816 | Aetherobacter | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | Polyangiales | Polyangiaceae | Pajaroellobacter | NA | Genus | ASV1717 | Pajaroellobacter | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | Polyangiales | Polyangiaceae | Rare Polyangiaceae | NA | Genus | ASV1717 | Rare Polyangiaceae | 0.02 | -0.7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | Polyangiales | Polyangiaceae | U_Genus 186 | NA | Genus | ASV4680 | U_Genus 186 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | Polyangiales | Sandaracinaceae | Unested_Genus 15 | NA | Genus | ASV2832 | Unested_Genus 15 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Polyangia | Rare Polyangia | U_Family 117 | U_Genus 183 | NA | Order | ASV650 | Rare Polyangia | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Myxococcota | Rare Myxococcota | U_Order 56 | U_Family 118 | U_Genus 184 | NA | Class | ASV650 | Rare Myxococcota | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Nitrospirota | Nitrospiria | U_Order 57 | U_Family 119 | U_Genus 188 | NA | Class | ASV615 | Nitrospiria | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Microgenomatia | Candidatus Daviesbacteria | U_Family 121 | U_Genus 190 | NA | Order | ASV4124 | Candidatus Daviesbacteria | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Microgenomatia | Rare Microgenomatia | U_Family 122 | U_Genus 191 | NA | Order | ASV4124 | Rare Microgenomatia | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0.64 | 0 |
| Patescibacteria | Paracubacteria | Candidatus Adlerbacteria | U_Family 123 | U_Genus 192 | NA | Order | ASV4608 | Candidatus Adlerbacteria | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Paracubacteria | Candidatus Kaiserbacteria | U_Family 124 | U_Genus 193 | NA | Order | ASV9005 | Candidatus Kaiserbacteria | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Paracubacteria | Candidatus Nomurabacteria | U_Family 125 | U_Genus 194 | NA | Order | ASV5969 | Candidatus Nomurabacteria | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Paracubacteria | Rare Parcubacteria | U_Family 126 | U_Genus 195 | NA | Order | ASV2073 | Rare Parcubacteria | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Paracubacteria | U_Order 61 | U_Family 127 | U_Genus 196 | NA | Order | ASV5380 | U_Order 61 | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Rare Patescibacteria | U_Order 63 | U_Family 129 | U_Genus 200 | NA | Class | ASV71 | Rare Patescibacteria | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Saccharimonadia | Saccharimonadales | LWQ8 | U_Genus 197 | NA | Family | ASV680 | LWQ8 | 0.16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Saccharimonadia | Saccharimonadales | Saccharimonadaceae | TM7a | NA | Genus | ASV71 | TM7a | 0.32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Saccharimonadia | Saccharimonadales | U_Family 130 | U_Genus 202 | NA | Family | ASV3550 | U_Family 130 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Patescibacteria | Saccharimonadia | Saccharimonadales | WWH38 | U_Genus 201 | NA | Family | ASV666 | WWH38 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | BD7-11 | U_Order 64 | U_Family 131 | U_Genus 203 | NA | Class | ASV718 | BD7-11 | 0.03 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | OM190 | U_Order 65 | U_Family 132 | U_Genus 204 | NA | Class | ASV934 | OM190 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Phycisphaerae | Phycisphaerales | Phycisphaeraceae | U_Genus 205 | NA | Family | ASV3443 | Phycisphaeraceae | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Phycisphaerae | Rare Phycisphaerae | U_Family 135 | U_Genus 210 | NA | Order | ASV251 | Rare Phycisphaerae | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Phycisphaerae | Tepidisphaerales | CPla-3 termite group | U_Genus 207 | NA | Family | ASV3429 | CPla-3 termite group | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Phycisphaerae | Tepidisphaerales | Rare Tepidisphaerales | U_Genus 209 | NA | Family | ASV251 | Rare Tepidisphaerales | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Phycisphaerae | Tepidisphaerales | WD2101 soil group | U_Genus 208 | NA | Family | ASV251 | WD2101 soil group | 2.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Gemmatales | Gemmataceae | Fimbriiglobus | NA | Genus | ASV2364 | Fimbriiglobus | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Gemmatales | Gemmataceae | Gemmata | NA | Genus | ASV180 | Gemmata | 0.23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Gemmatales | Gemmataceae | Unested_Genus 16 | NA | Genus | ASV86 | Unested_Genus 16 | 2.24 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Gemmatales | Gemmataceae | Zavarzinella | NA | Genus | ASV1042 | Zavarzinella | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Isosphaerales | Isosphaeraceae | Aquisphaera | ASV121 | ASV | ASV121 | ASV121 | 0.06 | 0 | 0 | 0.48 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Isosphaerales | Isosphaeraceae | Aquisphaera | ASV40 | ASV | ASV40 | ASV40 | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Isosphaerales | Isosphaeraceae | Aquisphaera | Rare Aquisphaera | ASV | ASV40 | Rare Aquisphaera | 0.86 | -0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Isosphaerales | Isosphaeraceae | Candidatus Nostocoida | NA | Genus | ASV876 | Candidatus Nostocoida | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Isosphaerales | Isosphaeraceae | Singulisphaera | NA | Genus | ASV371 | Singulisphaera | 0.31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Isosphaerales | Isosphaeraceae | Tundrisphaera | NA | Genus | ASV446 | Tundrisphaera | 0.29 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Isosphaerales | Isosphaeraceae | U_Genus 215 | NA | Genus | ASV213 | U_Genus 215 | 0.88 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Pirellulales | Pirellulaceae | Pir4 lineage | ASV196 | ASV | ASV196 | ASV196 | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Pirellulales | Pirellulaceae | Pir4 lineage | ASV70 | ASV | ASV70 | ASV70 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Pirellulales | Pirellulaceae | Pir4 lineage | ASV74 | ASV | ASV74 | ASV74 | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Pirellulales | Pirellulaceae | Pir4 lineage | Rare Pir4 lineage | ASV | ASV70 | Rare Pir4 lineage | 0.58 | 0.71 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**File S3-1 cont.**

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA % | pH r | Texture r | SOM Quantity r | SOM Quality A r | SOM Quality B r | SOM Quality C r | SOM Quality D r | SOM Quality E r |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Planctomycetota | Planctomycetes | Pirellulales | Pirellulaceae | Pirellula | NA | Genus | ASV2709 | Pirellula | 0.16 | 0.65 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Pirellulales | Pirellulaceae | Unested_Genus 17 | ASV85 | ASV | ASV85 | ASV85 | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Pirellulales | Pirellulaceae | Unested_Genus 17 | Rare Unested_Genus 17 | ASV | ASV85 | Rare Unested_Genus 17 | 0.87 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Planctomycetales | Gimesiaceae | U_Genus 217 | NA | Family | ASV1305 | Gimesiaceae | 0.06 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Planctomycetales | Rubinisphaeraceae | SH-PL14 | NA | Genus | ASV2975 | SH-PL14 | 0.04 | 0.54 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | Planctomycetales | Unested_Family 6 | U_Genus 220 | NA | Family | ASV990 | Unested_Family 6 | 0.19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Planctomycetes | U_Order 63 | U_Family 142 | U_Genus 221 | NA | Order | ASV2574 | U_Order 63 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Planctomycetota | Rare Planctomycetota | U_Order 68 | U_Family 139 | U_Genus 214 | NA | Class | ASV140 | Rare Planctomycetota | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Acetobacterales | Acetobacteraceae | Unested_Genus 18 | ASV314 | ASV | ASV314 | Rare Unested_Genus 18 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Acetobacterales | Acetobacteraceae | Unested_Genus 18 | Rare Unested_Genus 18 | ASV | ASV314 | Rare Unested_Genus 18 | 0.23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Caulobacterales | Caulobacteraceae | Phenylobacterium | NA | Genus | ASV1575 | Phenylobacterium | 0.03 | 0 | 0 | 0 | 0 | 0.55 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Caulobacterales | Caulobacteraceae | Rare Caulobacteraceae | NA | Genus | ASV1575 | Rare Caulobacteraceae | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Elsterales | Unested_Family 7 | U_Genus 227 | NA | Family | ASV128 | Unested_Family 7 | 0.34 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Micropepsales | Micropepsaceae | Unested_Genus 19 | NA | Family | ASV365 | Micropepsaceae | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rare Alpha-proteobacteria | U_Family 149 | U_Genus 244 | NA | Order | ASV4 | Rare Alphaproteobacteria | 0.38 | 0 | 0 | 0 | 0 | 0 | 0 | 0.52 | 0 |
| Proteobacteria | Alphaproteobacteria | Reyranellales | Reyranellaceae | Reyranella | ASV67 | ASV | ASV67 | ASV67 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Reyranellales | Reyranellaceae | Reyranella | Rare Reyranella | ASV | ASV67 | Rare Reyranella | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | A0839 | U_Genus 230 | NA | Family | ASV3518 | A0839 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | Methylobacterium-Methylorubrum | NA | Genus | ASV123 | Methylobacterium-Methylorubrum | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | Methylocapsa | NA | Genus | ASV177 | Methylocapsa | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | Microvirga | NA | Genus | ASV44 | Microvirga | 0.54 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | Rare Beijerinckiaceae | NA | Genus | ASV36 | Rare Beijerinckiaceae | 0.17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | Roseiarcus | ASV36 | ASV | ASV36 | ASV36 | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Beijerinckiaceae | U_Genus 232 | NA | Genus | ASV433 | U_Genus 232 | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Devosiaceae | Devosia | ASV378 | ASV | ASV378 | ASV378 | 0.03 | 0.47 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Devosiaceae | Rare Devosia | ASV378 | ASV | ASV378 | Rare Devosia | 0.13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Hyphomicrobiaceae | Hyphomicrobium | ASV250 | ASV | ASV250 | ASV250 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Hyphomicrobiaceae | Hyphomicrobium | Rare Hyphomicrobium | ASV | ASV250 | Rare Hyphomicrobium | 0.12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Hyphomicrobiaceae | Pedomicrobium | ASV130 | ASV | ASV130 | ASV130 | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Hyphomicrobiaceae | Pedomicrobium | Rare Pedomicrobium | ASV | ASV130 | Rare Pedomicrobium | 0.15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Labraceae | U_Genus 235 | NA | Family | ASV1980 | Labraceae | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Methyloligellaceae | U_Genus 236 | NA | Family | ASV118 | Methyloligellaceae | 0.50 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rare Rhizobiales | U_Genus 243 | NA | Family | ASV4 | Rare Rhizobiales | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae | Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium | NA | Genus | ASV184 | Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae | Mesorhizobium | ASV492 | ASV | ASV492 | ASV492 | 0.02 | 0 | 0 | 0 | 0 | 0.6 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae | Mesorhizobium | ASV498 | ASV | ASV498 | ASV498 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae | Mesorhizobium | Rare Mesorhizobium | ASV | ASV334 | Rare Mesorhizobium | 0.08 | 0 | 0 | 0 | 0 | 0.57 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae | Rare Rhizobiaceae | NA | Genus | ASV184 | Rare Rhizobiaceae | 0.14 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiaceae | U_Genus 238 | NA | Genus | ASV2688 | U_Genus 238 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiales Incertae Sedis | Bauldia | NA | Genus | ASV1802 | Bauldia | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiales Incertae Sedis | Sediis | NA | Genus | ASV1802 | Rare Rhizobiales Incertae Sedis | 0.04 | 0.56 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhizobiales Incertae Sedis | U_Genus 240 | NA | Genus | ASV1784 | U_Genus 240 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Rhodomicrobiaceae | U_Genus 241 | NA | Family | ASV22 | Rhodomicrobiaceae | 0.36 | 0 | 0 | 0.7 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | U_Family 151 | U_Genus 246 | NA | Family | ASV1035 | U_Family 151 | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Afipia | ASV4 | ASV | ASV4 | ASV4 | 1.12 | 0 | 0 | 0 | 0 | 0.49 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Afipia | Rare Afipia | ASV | ASV4 | Rare Afipia | 0.21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**File S3-1 cont.**

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA % | pH r | Texture r | SOM Quantity r | SOM Quality A r | SOM Quality B r | SOM Quality C r | SOM Quality D r | SOM Quality E r |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Bradyrhizobium | ASV131 | ASV | ASV131 | ASV131 | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Bradyrhizobium | ASV306 | ASV | ASV306 | ASV306 | 0.03 | 0 | 0 | 0 | 0 | 0.51 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Bradyrhizobium | ASV7 | ASV | ASV7 | ASV7 | 0.82 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Bradyrhizobium | ASV804 | ASV | ASV804 | ASV804 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Bradyrhizobium | ASV7 | ASV | ASV7 | Rare Bradyrhizobium | 0.16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Pseudolabrys | ASV68 | Genus | ASV68 | Pseudolabrys | 0.26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Rare Xanthobacteraceae | NA | Genus | ASV4 | Rare Xanthobacteraceae | 0.04 | 0.49 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Rhodoplanes | NA | Genus | ASV633 | Rhodoplanes | 0.14 | 0.59 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Unested_Genus 20 | ASV5 | ASV | ASV5 | Rare Unested_Genus 20 | 0.85 | 0 | 0 | 0.68 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Unested_Genus 20 | ASV5 | ASV | ASV5 | Rare Unested_Genus 20 | 3.43 | 0 | 0 | 0.56 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rhizobiales | Xanthobacteraceae | Variibacter | NA | Genus | ASV14557 | Variibacter | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Rickettsiales | U_Family 152 | NA | NA | Order | ASV7273 | Rickettsiales | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Sphingomonadales | U_Genus 248 | NA | NA | Family | ASV20 | Rare Sphingomonadales | 0.05 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae | Ellin6055 | NA | Genus | ASV273 | Ellin6055 | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae | Novosphingobium | NA | Genus | ASV1529 | Novosphingobium | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae | Sphingomonas | ASV20 | ASV | ASV20 | ASV20 | 0.25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae | Sphingomonas | ASV238 | ASV | ASV238 | ASV238 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae | Sphingomonas | ASV432 | ASV | ASV432 | ASV432 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae | Sphingomonas | Rare Sphingomonas | ASV | ASV20 | Rare Sphingomonas | 0.48 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Sphingomonadales | Sphingomonadaceae | U_Genus 249 | NA | Genus | ASV2108 | U_Genus 249 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | Tistrellales | U_Family 154 | U_Genus 250 | NA | Order | ASV2524 | Tistrellales | 0.03 | 0.53 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Alphaproteobacteria | U_Order 72 | U_Family 155 | U_Genus 251 | NA | Order | ASV370 | U_Order 72 | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | A21b | U_Genus 252 | NA | Family | ASV919 | A21b | 0.19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Burkholderiaceae | Paraburkholderia-Caballeronia-Burkholderia | NA | Genus | ASV746 | Paraburkholderia-Caballeronia-Burkholderia | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Comamonadaceae | Rare Comamonadaceae | NA | Genus | ASV368 | Rare Comamonadaceae | 0.10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Comamonadaceae | Rhizobacter | NA | Genus | ASV3051 | Rhizobacter | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Comamonadaceae | Unested_Genus 21 | NA | Genus | ASV368 | Unested_Genus 21 | 0.14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Comamonadaceae | Variovorax | NA | Genus | ASV3163 | Variovorax | 0.03 | 0.51 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Nitrosomonadaceae | Ellin6067 | NA | Genus | ASV2022 | Ellin6067 | 0.12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Nitrosomonadaceae | IS-44 | NA | Genus | ASV2220 | IS-44 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Nitrosomonadaceae | mle1-7 | NA | Genus | ASV646 | mle1-7 | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Nitrosomonadaceae | MND1 | NA | Genus | ASV4721 | MND1 | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Nitrosomonadaceae | Rare Nitrosomonadaceae | NA | Genus | ASV646 | Rare Nitrosomonadaceae | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Nitrosomonadaceae | U_Genus 256 | NA | Genus | ASV12379 | U_Genus 256 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Oxalobacteraceae | Duganella | NA | Genus | ASV5069 | Duganella | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Oxalobacteraceae | Massilia | ASV240 | ASV | ASV240 | ASV240 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Oxalobacteraceae | Massilia | ASV262 | ASV | ASV262 | ASV262 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Oxalobacteraceae | Massilia | Rare Massilia | ASV | ASV57 | Rare Massilia | 0.74 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Oxalobacteraceae | Novilimerbaspirillum | NA | Genus | ASV2112 | Novilimerbaspirillum | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Oxalobacteraceae | Rare Oxalobacteraceae | NA | Genus | ASV57 | Rare Oxalobacteraceae | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Oxalobacteraceae | Telluria | NA | Genus | ASV957 | Telluria | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Oxalobacteraceae | U_Genus 260 | NA | Genus | ASV850 | U_Genus 260 | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | Rare Burkholderiales | U_Genus 258 | NA | Family | ASV57 | Rare Burkholderiales | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0.49 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | SC-I-84 | U_Genus 261 | NA | Family | ASV2219 | SC-I-84 | 0.25 | 0 | -0.58 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | TRA3-20 | U_Genus 262 | NA | Family | ASV1445 | TRA3-20 | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Burkholderiales | U_Family 158 | U_Genus 263 | NA | Family | ASV2016 | U_Family 158 | 0.07 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | CCD24 | U_Family 159 | U_Genus 264 | NA | Order | ASV1538 | CCD24 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Coxiellales | U_Family 160 | U_Genus 265 | NA | Order | ASV2794 | Coxiellales | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**File S3-1 cont.**

| Phylum | Class | Order | Family | Genus | ASV | Level | ASVname | Name | Mean RA % | pH r | Texture r | SOM Quantity r | SOM Quality A r | SOM Quality B r | SOM Quality C r | SOM Quality D r | SOM Quality E r |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Proteobacteria | Gammaproteobacteria | Diplorickettsiales | Diplorickettsiaceae | Aquicella | NA | Genus | ASV8951 | Aquicella | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Diplorickettsiales | Diplorickettsiaceae | U_Genus 267 | NA | Genus | ASV3361 | U_Genus 267 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Incertae Sedis | Unknown Family | Acidibacter | NA | Family | ASV1417 | Unknown Family | 0.30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | JG36-TzT-191 | U_Family 163 | U_Genus 269 | NA | Order | ASV1058 | JG36-TzT-191 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Legionellales | Legionellaceae | Legionella | NA | Family | ASV5269 | Legionellaceae | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Pseudomonadales | Pseudomonadaceae | Pseudomonas | NA | Genus | ASV178 | Pseudomonas | 0.30 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Rare Gamma-proteobacteria | U_Family 157 | U_Genus 269 | NA | Order | ASV57 | Rare Gammaproteobacteria | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Steroidobacterales | Steroidobacteraceae | Steroidobacter | NA | Family | ASV641 | Steroidobacteraceae | 0.06 | 0.56 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | U_Order 74 | U_Family 169 | U_Genus 278 | NA | Order | ASV5276 | U_Order 74 | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | WD260 | U_Family 167 | U_Genus 274 | NA | Order | ASV173 | WD260 | 0.15 | 0 | 0 | 0.47 | 0 | 0 | 0 | 0 | 0.47 |
| Proteobacteria | Gammaproteobacteria | Xanthomonadales | Rhodanobacteraceae | Dokdonella | NA | Genus | ASV2037 | Dokdonella | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Xanthomonadales | Rhodanobacteraceae | Rare Rhodanobacteraceae | NA | Genus | ASV1709 | Rare Rhodanobacteraceae | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Xanthomonadales | Rhodanobacteraceae | Rhodanobacter | NA | Genus | ASV1709 | Rhodanobacter | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | Gammaproteobacteria | Xanthomonadales | Xanthomonadaceae | U_Genus 277 | NA | Family | ASV1955 | Xanthomonadaceae | 0.09 | 0.62 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Proteobacteria | U_Class 23 | U_Order 75 | U_Family 170 | U_Genus 279 | NA | Class | ASV7795 | U_Class 23 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sumerlaeota | U_Class 24 | U_Order 76 | U_Family 171 | U_Genus 280 | NA | Phylum | ASV3832 | Sumerlaeota | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| U_Phylum 1 | U_Class 27 | U_Order 82 | U_Family 184 | U_Genus 304 | NA | Phylum | ASV598 | U_Phylum 1 | 0.28 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Chlamydiae | Chlamydiales | Parachlamydiaceae | Candidatus Protochlamydia | NA | Genus | ASV4521 | Candidatus Protochlamydia | 0.03 | 0 | 0 | 0.56 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Chlamydiae | Chlamydiales | Parachlamydiaceae | Neochlamydia | NA | Genus | ASV12101 | Neochlamydia | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Chlamydiae | Chlamydiales | Parachlamydiaceae | U_Genus 283 | NA | Genus | ASV4840 | U_Genus 283 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Chlamydiae | Rare Chlamydiae | U_Family 172 | U_Genus 282 | NA | Order | ASV4521 | Rare Chlamydiae | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Chthoniobacterales | Chthoniobacteraceae | Candidatus Udaeobacter | ASV10 | ASV | ASV10 | ASV10 | 0.49 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Chthoniobacterales | Chthoniobacteraceae | Candidatus Udaeobacter | NA | ASV | ASV10 | Rare Candidatus Udaeobacter | 6.29 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Chthoniobacterales | Chthoniobacteraceae | Chthoniobacter | NA | Genus | ASV292 | Chthoniobacter | 0.59 | 0.55 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Chthoniobacterales | Chthoniobacteraceae | LD29 | NA | Genus | ASV1838 | LD29 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Chthoniobacterales | Chthoniobacteraceae | U_Genus 288 | NA | Genus | ASV4969 | U_Genus 288 | 0.06 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Chthoniobacterales | Rare Chthoniobacterales | U_Genus 285 | NA | Family | ASV10 | Rare Chthoniobacterales | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Chthoniobacterales | U_Family 176 | U_Genus 290 | NA | Family | ASV24735 | U_Family 176 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Chthoniobacterales | Xiphinematobacteraceae | U_Genus 289 | NA | Family | ASV54 | Xiphinematobacteraceae | 0.82 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Methylacidiphilales | U_Family 177 | U_Genus 291 | NA | Order | ASV8615 | Methylacidiphilales | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Opitutales | Opitutaceae | Lacunisphaera | NA | Genus | ASV8963 | Lacunisphaera | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.48 |
| Verrucomicrobiota | Verrucomicrobiae | Opitutales | Opitutaceae | Opitutus | NA | Genus | ASV3324 | Opitutus | 0.11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Opitutales | Opitutaceae | U_Genus 294 | NA | Genus | ASV6443 | U_Genus 294 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Pedosphaerales | Pedosphaeraceae | ADurb.Bin063-1 | NA | Genus | ASV1057 | ADurb.Bin063-1 | 0.16 | -0.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Pedosphaerales | Pedosphaeraceae | Pedosphaera | NA | Genus | ASV1312 | Pedosphaera | 0.09 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Pedosphaerales | Pedosphaeraceae | U_Genus 296 | NA | Genus | ASV2322 | U_Genus 296 | 0.17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Pedosphaerales | Rare Pedosphaerales | U_Genus 295 | NA | Family | ASV1057 | Rare Pedosphaerales | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | S-BQ2-57 soil group | U_Family 180 | U_Genus 297 | NA | Order | ASV4782 | S-BQ2-57 soil group | 0.03 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | U_Order 80 | U_Family 182 | U_Genus 302 | NA | Order | ASV10098 | U_Order 80 | 0.04 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Verrucomicrobiales | Rubritaleaceae | Luteolibacter | NA | Genus | ASV1364 | Luteolibacter | 0.06 | 0.54 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Verrucomicrobiales | Verrucomicrobiaceae | Roseimicrobium | NA | Genus | ASV1884 | Roseimicrobium | 0.05 | 0.64 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Verrucomicrobiota | Verrucomicrobiae | Verrucomicrobiales | Verrucomicrobiaceae | U_Genus 301 | NA | Genus | ASV6292 | U_Genus 301 | 0.02 | 0.52 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| WPS-2 | U_Class 26 | U_Order 81 | U_Family 183 | U_Genus 303 | NA | Phylum | ASV859 | WPS-2 | 0.17 | -0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Supplementary Text S3-1.** Reasons for the chosen approach of sequence processing. In recent years, studies have shown that higher taxonomic resolution is required to link environment and soil bacterial community composition (Cruz-Martínez et al., 2012; Martiny et al., 2015; Morrissey et al., 2016; Stone et al., 2023). However, the majority of large scale correlative studies known to the authors only investigated potential links at the level of taxonomic resolution of phyla to orders (Fierer et al., 2007; Lauber et al., 2009; Rousk et al., 2010; Delgado-Baquerizo et al., 2016b; Bahram et al., 2018), or after clustering of taxa into functional groups (Labouyrie et al., 2023). The reason for this is that the investigation of higher taxonomic resolution comes with several challenges. First, at a higher taxonomic resolution (or a more narrow definition of operational taxonomic units (OTUs), or a finer clustering of zero-radius OTUs), a larger fraction of soil bacterial communities becomes too rare to be reliably used in correlation analysis, because of the strongly asymmetric abundance distribution of soil bacteria (Lynch and Neufeld, 2015). In the present dataset, 75 % of the 85 692 ASVs were only present at one site, and only 0.1 % of ASVs were present in more than half of the 105 samples (Figure S3-7a). Moreover, rare taxa have high uncertainties relative to their range of observed abundances across the gradient, which means that analysis of such taxa is not sensitive for detection of real abundance patterns (shown based on the data of this study, Figure S3-7b,c). It would be problematic to just aggregate taxa at several different levels of taxonomic resolution and test at all levels. This would result in redundant testing, as the abundances of higher-resolution taxa in sum constitute the abundances of lower-resolution taxa. However, simply discarding all rare amplicon sequence variants (ASVs) or OTUs and correlating only the remaining ubiquitous ASVs/OTUs results in a large waste of information. For instance, filtering for ubiquitous ASVs that are present in at least half of the samples of this study would lead to an average loss of 73.4 % of reads. The sum of rare taxa could however still contain information if a trait or adaptation is conserved at a level of lower taxonomic resolution. The sum of rare taxa could however still contain information if a trait or adaptation is conserved at a level of lower taxonomic resolution. Due to these unresolved methodological challenges, potential links between biogeochemical variables and abundance patterns at levels of higher taxonomic resolution have remained largely unexplored at broader scales. A workflow of the approach of sequence processing is shown in Figure S3-1. In the following, we provide a brief summary of the results of the sequence processing. Across all levels of taxonomic resolution, on average across soils 68 % of relative abundance was retained in coherent taxa, and 30 % was aggregated into ΣRare-groups (Figure S3-2a). Across all samples, 44.1 % of average relative abundance was retained at the ASV level, 24.9 % at genus level, 18.8 % at family level and less than 10 % at levels of lower taxonomic resolution (Figure S3-2b). At all levels of taxonomic resolution, there were more coherent taxa than ΣRare-groups (Figure S3-2), and on

average coherent taxa contained a larger fraction of relative abundance than the ΣRare-groups, except at the ASV level (Figure S3-2b).

**Supplementary Text S3-2.** Different soil masses at DNA extraction did not bias the analysis. Adjustment of soil mass for DNA extraction could theoretically lead to an artifact in alpha diversity estimates. Greater soil mass (in the case of this study used for soils with lower TOC) and consequently greater soil volume could contain a larger number of rare ASVs. This could theoretically result in a bias towards higher richness in soils of lower TOC. Since we found a negative relationship between alpha diversity and SOM Quantity, we wanted to rule out that this observation was due to such bias. To that end, we took the exact weights of the three replicates of each site and calculated the mass difference relative to the lightest replicate (ΔDW). We further calculated the richness difference relative to the lightest replicate (ΔRichness). If there was a pronounced bias that was stronger than other spatial heterogeneity and noise, we would expect i) consistent increases of richness with mass among the triplicates, and if such bias was subject to saturation with increasing soil volume, we would expect ii) the increases (slopes of relationship ΔRichness ~ ΔDW) to decrease with increasing soil mass that was extracted. Results of this analysis show that there is no such bias (Figure S3-4). The same concern applied for relative abundance estimates: A larger number of rare ASVs could theoretically result in a bias towards lower relative abundance of other taxa in soils of lower TOC. Similar to the analysis for richness (described above), we therefore calculated the relative abundance difference relative to the lightest replicate (ΔCounts) for four different taxa which showed positive relative abundance correlations with SOM Quantity. If there was a pronounced bias that was stronger than other spatial heterogeneity and noise, we would expect i) consistent decreases of counts with increasing mass among the triplicates, and if such bias was subject to saturation with increasing soil volume, we would expect ii) the decreases (slopes of relationship ΔCounts ~ ΔDW) to get smaller with increasing soil mass that was extracted. Results of this analysis show that there is no such bias (Figure S3-6).

**Supplementary Text S3-3.** Clusters of taxonomic units with biogeochemically-related patterns of relative abundance. A detailed summary of the results of the correlation analyses can be found in Supplementary File S3-1. In the following, we discuss observed patterns in the context of previous findings, in more depth than the main text allowed. Soil pH: We found clusters of negative (in the class Acidobacteriae) and positive (in the class Blastocatellia) associations with pH within the phylum of Acidobacteriata, which is consistent with to previous studies (Rousk et al., 2010; Griffiths et al., 2011). In our study, the taxonomic groups with negative pH associations were quantitatively more dominant, which perhaps helps explain why other studies found the entire phylum to be negatively linked with pH (Lauber et al., 2009; Delgado-Baquerizo et al., 2016b). In the Chloroflexi phylum, for which opposing

pH associations have been reported (negative, Bahram et al. (2018); positive, Delgado-Baquerizo et al. (2016)), we also found a negative cluster (in the class Ktedonobacteria) and a positive cluster (in the class Chloroflexia). We could not identify any consistent patterns within the phyla Bacteroidetes and Actinobacteriota, which have been linked to pH in contrasting ways (Bacteriodetes: negative, (Bahram et al., 2018); positive, (Lauber et al., 2009; Rousk et al., 2010); Actinobacteriota: negative, (Bahram et al., 2018); neutral, (Rousk et al., 2010); positive, (Lauber et al., 2009; Delgado-Baquerizo et al., 2016b)).
SOM Quantity: We found two clusters of dominance that were linked with OM quantity in the phylum Firmicutes, namely in the families of Planococcaceae and Peptostreptococcaceae. Experimental studies have linked the phylum Firmicutes with copiotrophic traits (Cleveland et al., 2007; Pepe-Ranney et al., 2016; Stone et al., 2023), and representatives of Firmicutes possess a large number of transporters that could support fast growth (Trivedi et al., 2013). In contrast, previous regional to global studies found no conclusive patterns for Firmicutes (Fierer et al., 2007) or linked the phylum negatively with TOC (Delgado-Baquerizo et al., 2016b), and a study on dryland restoration interpreted Firmicutes to behave oligotrophic (Bastida et al., 2015). We identified another cluster in the family Xanthobacteraceae, which belongs to the order of Rhizobiales in the class of Alpha-Proteobacteria. At high taxonomic resolutions, contrasting patterns have been found for (Alpha-)Proteobacteria. Positive (Delgado-Baquerizo et al., 2016b) as well as negative (Fierer et al., 2007; Bahram et al., 2018) correlations with TOC have been observed. However, at the order-level, both an experimental study (Geyer and Barrett, 2019) as well as a study on dryland restoration (Bastida et al., 2015) found Rhizobiales to behave copiotroph. The family Xanthobacteraceae features nitrogen (N) fixers, thus their increased dominance in C-rich systems could stem from an advantage in conditions of N limitation. However, in this study, available C and N strongly correlated with the rotated component OM Quantity (Table S4). If the C:N ratio of substrate was the main driver for the dominance pattern of the Xanthobacteraceae, we would rather expect a correlation with the rotate[WD2] d compo[WD3] nent OM Quality E (C:N ratio). We could not find any patterns within subgroups of other phyla that repeatedly (but not exclusively) showed copiotroph behavior in substrate addition experiments such as Beta-Proteobacteria (Fierer et al., 2007; Eilers et al., 2010; Morrissey et al., 2016) or Gamma-Proteobacteria (Cleveland et al., 2007; Eilers et al., 2010; Morrissey et al., 2016). Similarly, we found no clusters in the phylum Actinobacteria, which has been reported to become more dominant upon substrate addition (Eilers et al., 2010; Geyer and Barrett, 2019), while also having been associated with oligotroph behavior in other studies (Bastian et al., 2009; Trivedi et al., 2013; Delgado-Baquerizo et al., 2016b).
SOM Quality: Specifically, we found a cluster of ASVs in the family *Micrococcaceae* that was dominant in soils with a higher C:N ratio (SOM quality E). Soil-dwelling members of the *Micrococcaceaea* have been repeatedly found in hydrocarbon polluted soils (Dastager et al., 2014), and have been shown to be able to decompose complex (aromatic) substrates (Sims et al., 1986; Storey et al., 2018). Further,

two groups in the order Streptosporangiales, two groups in the family *Micromonosporaceae* and two groups in the genus *Mesorhizobium* were positively correlated with SOM quality B (POM). We interpret this RC to reflect systems in which free coarse plant litter, potentially available for decomposition, accumulates. In agreement with this interpretation, members of the Streptosporangiales have been found to be involved in primary decomposition of plant material in soils (Otoguro et al., 2014). *Micromonosporaceae* have been found to be inhabit wet soils, peat, roots and plant material (such as nitrogen-fixing nodules) (Trujillo et al., 2014). While the pattern in our data contrasts with the apparent moisture preference, it matches with the observation that members of this family have been found to prefer systems where free undecomposed plant material accumulates. Lastly, the genus *Mesorhizobium* correlated with SOM quality B (POM). This genus belongs to the rhizobia, which can either leave freely in soils, or form a symbiosis as N-fixers in the nodules of legume plants (Tatsukami et al., 2013). Theoretically, the dominance of rhizobia could vary with plant species composition (i.e. the dominance of symbiotic legume plants). Further it has been shown that low temperature can physiologically hamper the formation of root nodules (discussed in (Duan et al., 2022)). However, neither does MAT load onto the variable SOM quality B (POM), nor is the relative abundance of *Mesorhizobia* high enough (maximum value 0.6 %) to suggest that legume nodules played a large role to shape bacterial communities along this gradient.

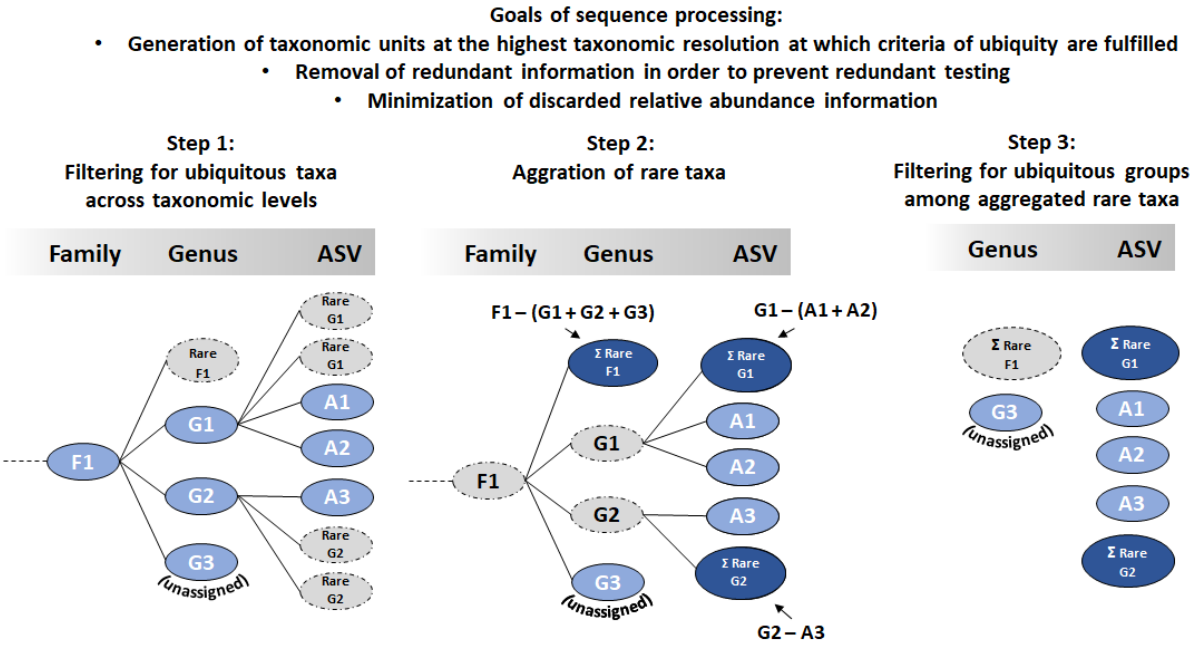



**Figure S3-1**. Summary of the workflow for sequence processing with a schematic example. In step 1, taxa were filtered for ubiquity at each level of taxonomic resolution. In step 2, rare (i.e. not ubiquitous) taxa were aggregated into ΣRare-groups, and redundant information across levels of taxonomic resolution was removed. In step 3, the newly aggregated groups of rare taxa (ΣRare-groups) were

filtered for ubiquity. Taxa and ΣRare-groups that were retained after step 3 were jointly referred to as “taxonomic units” and were used for downstream analysis. In the schematic example, taxa/groups that get removed in a step are shown in light gray with dashed outlines. Aggregated groups of rare taxa are shown in dark blue. Unchanged taxa are shown in light blue. All calculation steps were based on relative abundance data.

a)

| Taxonomic level | Taxa before | Data processing | | |
|---|---|---|---|---|
| | | Taxa | ΣRare-groups | All taxonomic units (Taxa + ΣRare-groups) |
| | (Number) | (Number) | (Number) | (Number) |
| **Phylum** | 45 | 4 | 0 | 4 |
| **Class** | 154 | 19 | 6 | 25 |
| **Order** | 383 | 40 | 11 | 51 |
| **Family** | 612 | 56 | 15 | 71 |
| **Genus** | 1277 | 134 | 25 | 159 |
| **ASV** | 85692 | 94 | 33 | 127 |
| **Total** | - | 347 | 90 | 437 |

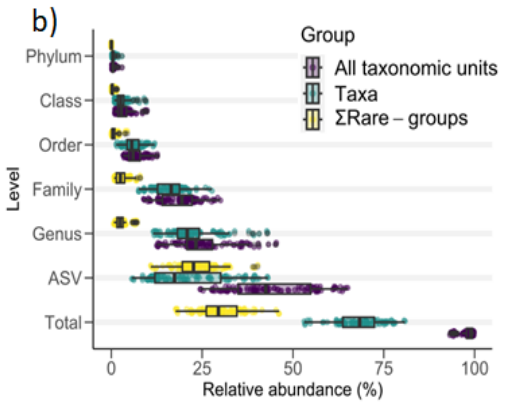



**Figure S3-2.** Summary of the results of sequence processing. Panel a) After the workflow, 436 taxonomic units were retained across all levels of taxonomic resolution. "Taxa" are coherent taxonomic units, while "ΣRare-groups" are aggregated groups of rare taxa at the given level of taxonomic resolution. Panel b) Amount of relative abundance across levels of taxonomic resolutions and groups (i.e. taxa vs. ΣRare-groups). Points show values for individual samples (n = 105), boxplots show median (center line), 25th and 75th percentiles (box limits), 1.5x interquartile range (whiskers) and outliers (points).

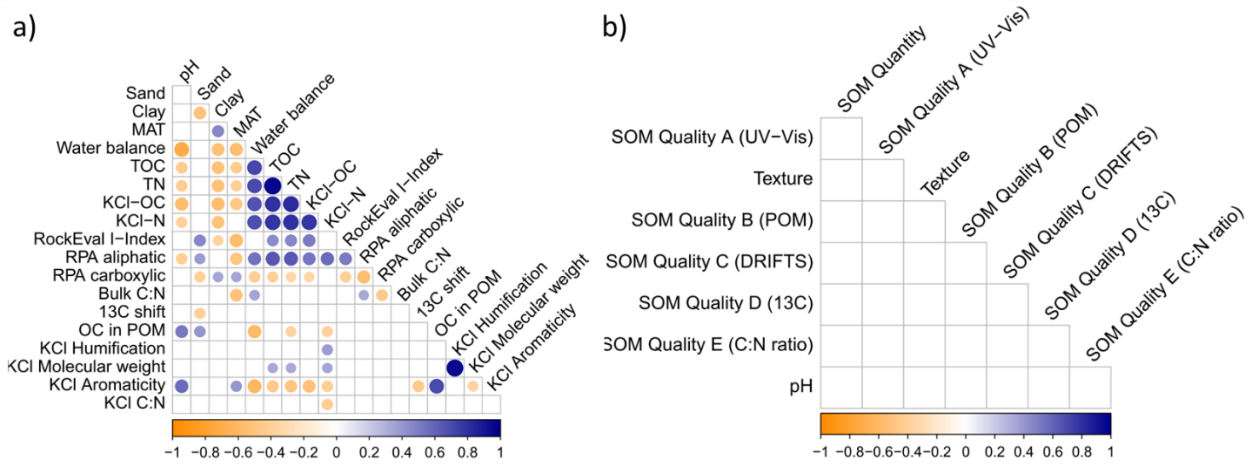



**Figure S3-3**. Panel a) Pearson correlation matrix of the measured biogeochemical variables (with significant outliers following Rosner's test replaced by mean values). Shown are significant correlations (p-value < 0.05). Circle size and color indicate the strength of the relationships (Pearson correlation coefficient). RCs are listed in the order of their Eigenvalues. Light orange to dark blue represents a gradient from negative to positive correlation coefficients. MAT = mean annual temperature; TOC = total organic carbon; TN = total nitrogen; KCl-OC = KCl-extractable organic carbon; KCl-N = KCl-extractable nitrogen; RPA = relative peak area; POM = particulate organic matter; DRIFTS = diffuse reflectance infrared spectroscopy; UV-Vis = ultraviolet-visible spectroscopy. Panel b) Pearson correlation matrix of the rotated components (RCs) obtained after rotated principal component

analysis. Note that no correlations between the RCs were found at p-value < 0.05, therefore confirming a very low degree of autocorrelation between the remaining RCs.

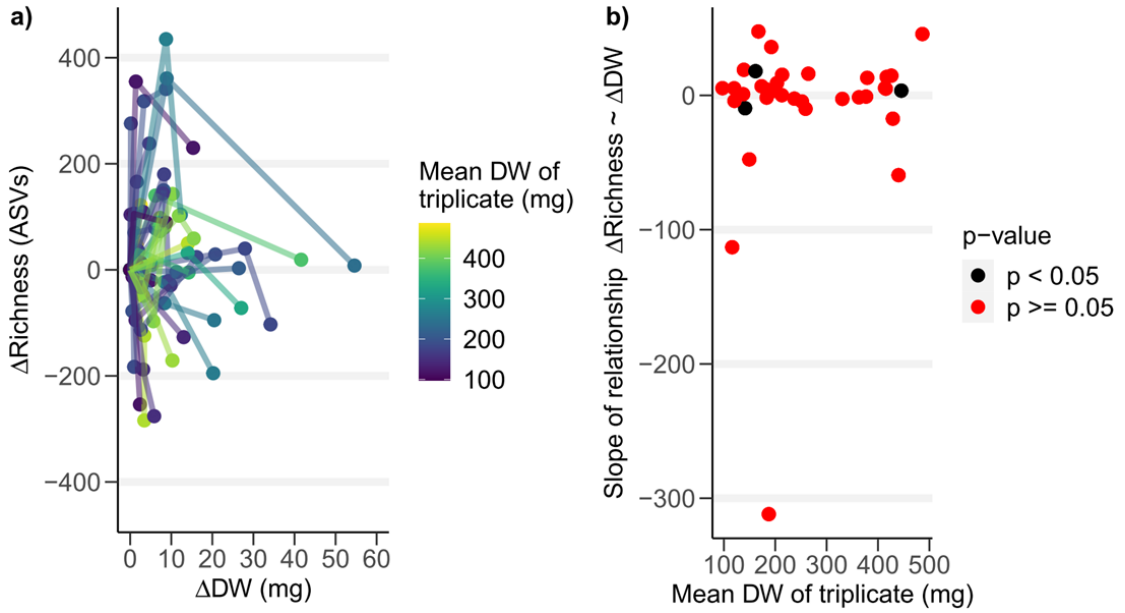



**Figure S3-4**. Variable soil mass for DNA extraction does not lead to an artifact in alpha diversity. Panel a) shows that among triplicates with variable soil mass (x-axis) there is no consistent pattern of increase of richness (y-axis). Dots are data points, lines for visual aid. Panel b) shows that rarely any within-triplicate-trends are consistent, that there are decreasing as well as increasing patterns, and that there is no shift in patterns with the mean DW of triplicates. Thus, we are confident that adjustment of soil mass to TOC at the DNA extraction step does not lead to a systematic bias in our alpha diversity analysis.

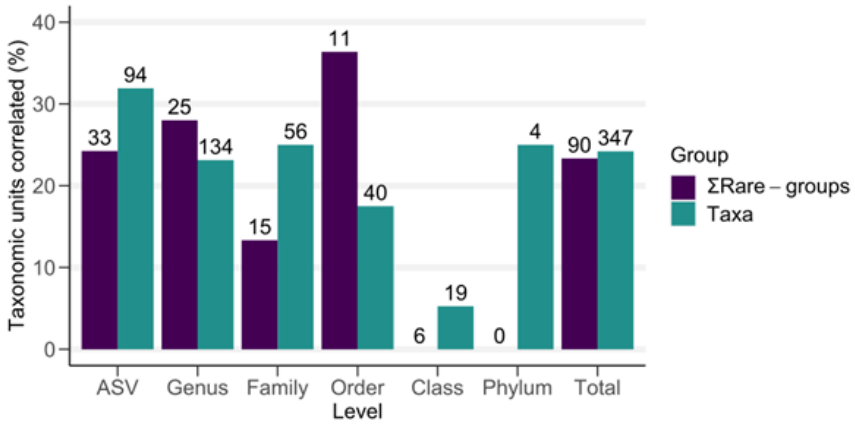



**Figure S3-5.** The percentages of taxonomic units that were significantly correlated with biogeochemical variables across all levels of taxonomic resolution. Numbers on top of the bars indicate the numbers of groups tested.

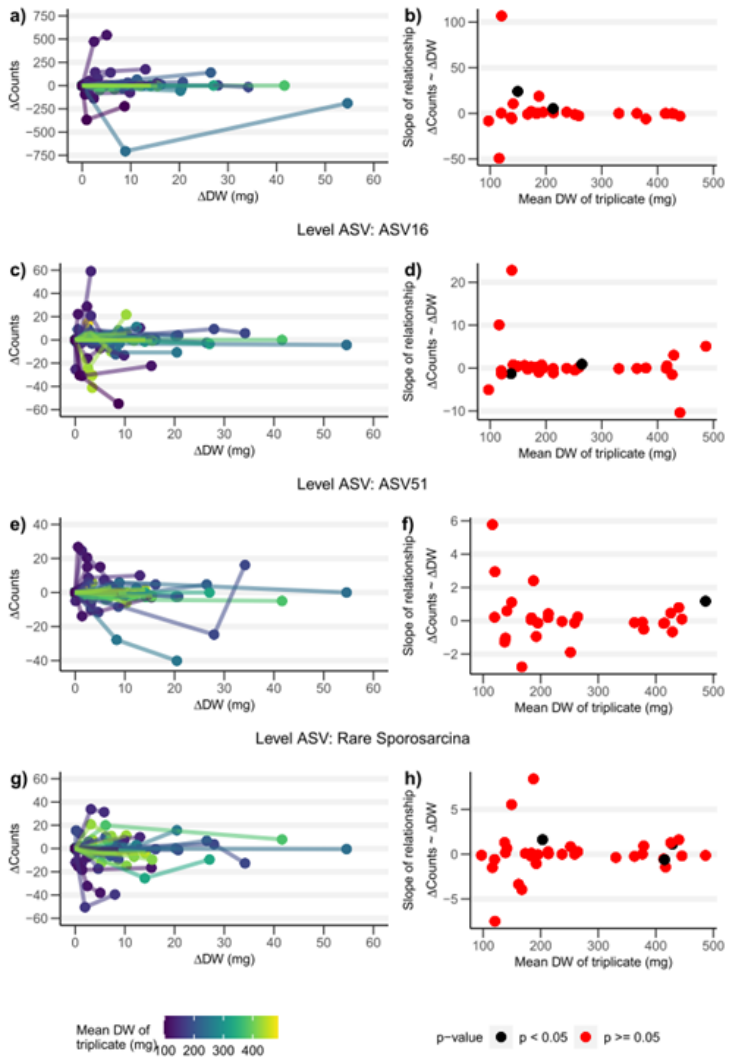



**Figure S3-6**. Variable soil mass for DNA extraction does not lead to an artifact in relative abundance. Panel a) shows that among triplicates with variable soil mass (x-axis) there is no consistent pattern of decrease of relative abundance (y-axis). Dots are data points, lines for visual aid. Panel b) shows that rarely any within-triplicate-trends are consistent, that there are decreasing as well as increasing patterns, and that there is no shift in patterns with the mean DW of triplicates. Thus, we are confident that adjustment of soil mass to TOC at the DNA extraction step does not lead to a systematic bias in our relative abundance analysis. Note that counts directly translate to relative abundance, since all samples were rarefied to 17188 reads (thus 1 count = 0.0058 % relative abundance).

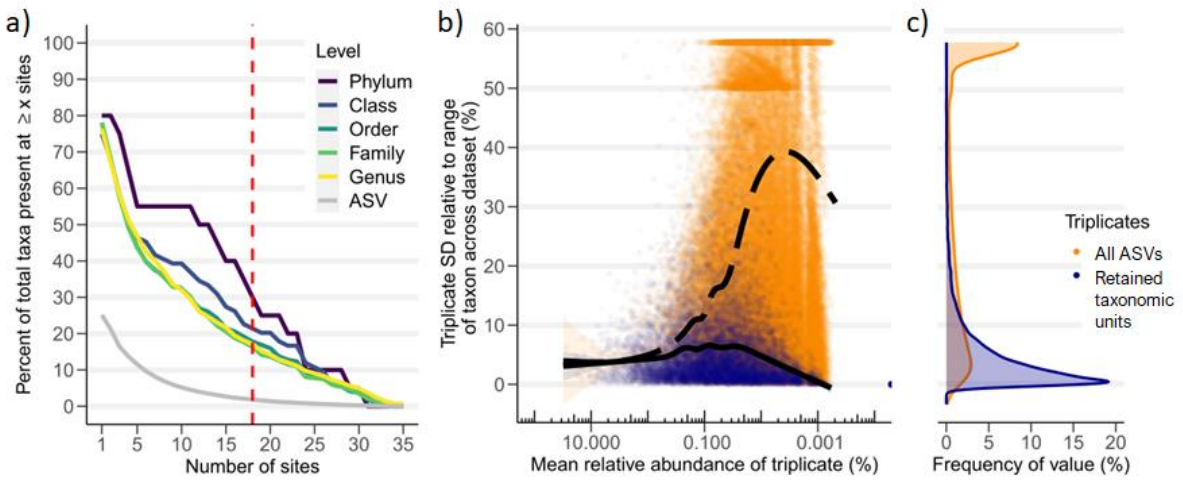



**Figure S3-7**. Panel a) Distribution of taxa sorted by the number of soils in which they were present. The dashed red line indicates presence at more than half of the sites. Different colors indicate aggregation of taxa at different levels of taxonomic resolution. For example, ~ 75 % of ASVs were only present at one site. Panel b) The uncertainty of relative abundances of reads (expressed as the standard deviation (SD) of each triplicate) relative to the observed range of the respective taxa (y-axis) at the ASV level. Light orange shows the entire (unfiltered) dataset, while dark blue shows the data points retained after filtering and processing (i.e. 436 taxonomic units). Observations with low relative abundance (x-axis) tend to have higher uncertainties. Lines are smoothed fits using a general additive models with k = 10, (solid = All ASVs, dashed = Filtered & processed taxa). $n_{\text{All ASVs}}$ = 165 019 (each ASV x site combination in the data set), $n_{\text{Retained taxonomic units}}$ = 15 260. Note that some data points with very low mean relative abundance (and SD) are not shown for the Filtered & processed data. Panel c) The density distributions of uncertainties associated with the data. The high number of values between 0.50 and 0.58 in panel a) stems from the following: If two of the triplicates have identical relative abundances, and the range across the data set is the same as in the triplicate, the result is a value of 0.58. If one triplicate has a value of zero, one triplicate has a value as large as the range across the data set, and one triplicate has a value half of that, the result is a value of 0.50.

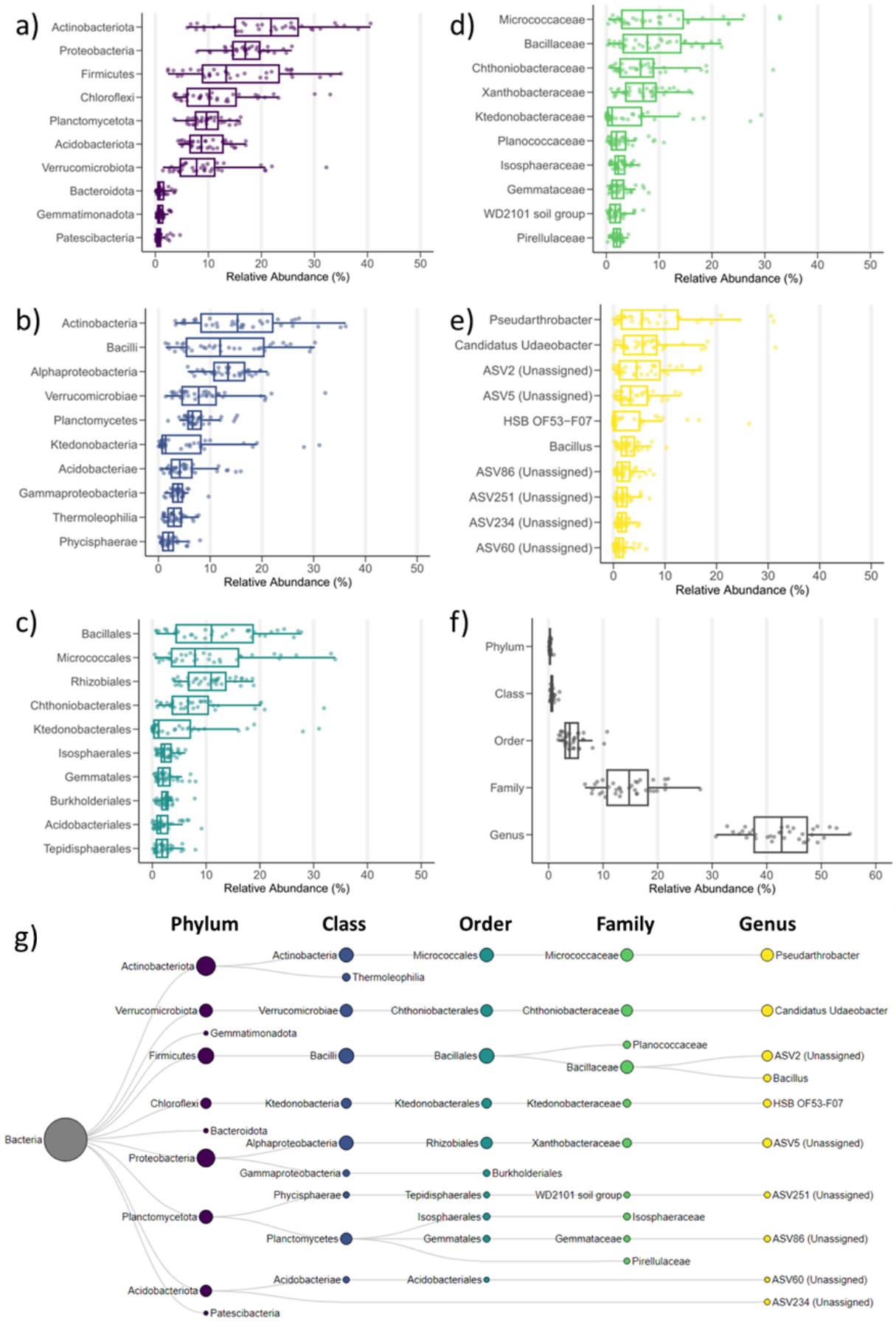

a)
Actinobacteriota
Proteobacteria
Firmicutes
Chloroflexi
Planctomycetota
Acidobacteriota
Verrucomicrobiota
Bacteroidota
Gemmatimonadota
Patescibacteria
0 10 20 30 40 50
Relative Abundance (%)
b)
Actinobacteria
Bacilli
Alphaproteobacteria
Verrucomicrobiae
Planctomycetes
Ktedonobacteria
Acidobacteriae
Gammaproteobacteria
Thermoleophilia
Phycisphaerae
0 10 20 30 40 50
Relative Abundance (%)
c)
Bacillales
Micrococcales
Rhizobiales
Chthoniobacterales
Ktedonobacterales
Isosphaerales
Gemmatales
Burkholderiales
Acidobacteriales
Tepidisphaerales
0 10 20 30 40 50
Relative Abundance (%)
d)
Micrococcaceae
Bacillaceae
Chthoniobacteraceae
Xanthobacteraceae
Ktedonobacteraceae
Planococcaceae
Isosphaeraceae
Gemmataceae
WD2101 soil group
Pirellulaceae
0 10 20 30 40 50
Relative Abundance (%)
e)
Pseudarthrobacter
Candidatus Udaeobacter
ASV2 (Unassigned)
ASV5 (Unassigned)
HSB OF53−F07
Bacillus
ASV86 (Unassigned)
ASV251 (Unassigned)
ASV234 (Unassigned)
ASV60 (Unassigned)
0 10 20 30 40 50
Relative Abundance (%)
f)
Phylum
Class
Order
Family
Genus
0 10 20 30 40 50 60
Relative Abundance (%)
g)
Phylum
Class
Order
Family
Genus
Bacteria
Actinobacteriota
Actinobacteria
Thermoleophilia
Micrococcales
Micrococcaceae
Pseudarthrobacter
Verrucomicrobiota
Verrucomicrobiae
Chthoniobacterales
Chthoniobacteraceae
Candidatus Udaeobacter
Gemmatimonadota
Firmicutes
Bacilli
Bacillales
Planococcaceae
Bacillaceae
ASV2 (Unassigned)
Bacillus
Chloroflexi
Ktedonobacteria
Ktedonobacterales
Ktedonobacteraceae
HSB OF53-F07
Bacteroidota
Proteobacteria
Alphaproteobacteria
Rhizobiales
Xanthobacteraceae
ASV5 (Unassigned)
Gammaproteobacteria
Burkholderiales
Phycisphaerae
Tepidisphaerales
WD2101 soil group
ASV251 (Unassigned)
Planctomycetota
Isosphaerales
Isosphaeraceae
Planctomycetes
Gemmatales
Gemmataceae
ASV86 (Unassigned)
Pirellulaceae
Acidobacteriae
Acidobacteriales
ASV60 (Unassigned)
Acidobacteriota
ASV234 (Unassigned)
Patescibacteria

**Figure S3-8**. Panels a) to e) show the ten taxa with highest mean relative read abundance of reads across the gradient (n = 35), resolved at five different levels of taxonomic resolution. a) Phyla, b) Classes, c) Orders, d) Families, e) Genera. Panel f) shows the relative read abundance of reads belonging to unassigned ASVs at the five levels of taxonomic resolution. Panel g) shows how the ten dominant taxa of each level are taxonomically related. Circle sizes show the mean relative read abundance of reads across the gradient. This way of illustration emphasizes that the dominance of some taxa at low levels of taxonomic resolution is due to very dominant taxonomic units at higher levels of taxonomic resolution (e.g. the dominance of the phylum Verrucomicrobiata is mostly driven by the high relative read abundance of Candidatus Udaeobacter).

**Table S3-1**. Summary of the dada2 pipeline. Listed are the numbers of reads for each processing step and each sample (n = 105), as well as the recovery of reads in percent. Colors indicate recoveries relative to all other samples (green = high recovery, red = low recovery).

| | | Reads (Number) | | | | | | Read recovery (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ID | Rep. | Input | Filtered | Denoised Forward | Denoised Reverse | Merged | Chimera Removal | Total Pipeline | Denoising Step | Merging Step | Chimera Removal Step |
| 1 | a | 55043 | 55026 | 50173 | 51058 | 35943 | 30311 | 55 | 92 | 70 | 84 |
| 1 | b | 52576 | 52564 | 48171 | 48673 | 34135 | 28778 | 55 | 92 | 70 | 84 |
| 1 | c | 50988 | 50964 | 46478 | 46809 | 32926 | 27720 | 54 | 92 | 70 | 84 |
| 2 | a | 58448 | 58429 | 53777 | 54104 | 41430 | 29544 | 51 | 92 | 77 | 71 |
| 2 | b | 55036 | 55000 | 50610 | 50119 | 38805 | 27073 | 49 | 92 | 77 | 70 |
| 2 | c | 60476 | 60461 | 55237 | 55799 | 42576 | 30570 | 51 | 92 | 76 | 72 |
| 3 | a | 52640 | 52622 | 47622 | 47947 | 35555 | 29497 | 56 | 91 | 74 | 83 |
| 3 | b | 54484 | 54462 | 49163 | 49099 | 36484 | 30038 | 55 | 90 | 74 | 82 |
| 3 | c | 66361 | 66340 | 60859 | 60980 | 45895 | 37825 | 57 | 92 | 75 | 82 |
| 4 | a | 48059 | 48038 | 43737 | 43826 | 32165 | 24456 | 51 | 91 | 73 | 76 |
| 4 | b | 45177 | 45169 | 41719 | 41844 | 31249 | 24381 | 54 | 93 | 75 | 78 |
| 4 | c | 50018 | 50002 | 45759 | 46240 | 33947 | 25711 | 51 | 92 | 73 | 76 |
| 5 | a | 58352 | 58326 | 53166 | 53989 | 39443 | 32501 | 56 | 92 | 73 | 82 |
| 5 | b | 62017 | 62005 | 56438 | 57458 | 41865 | 34017 | 55 | 92 | 73 | 81 |
| 5 | c | 74219 | 74197 | 68475 | 68959 | 50764 | 41137 | 55 | 93 | 74 | 81 |
| 6 | a | 68179 | 68162 | 63471 | 63886 | 50181 | 36913 | 54 | 93 | 79 | 74 |
| 6 | b | 61061 | 61049 | 56540 | 56513 | 44107 | 32925 | 54 | 93 | 78 | 75 |
| 6 | c | 63192 | 63169 | 58747 | 59078 | 46491 | 35071 | 55 | 93 | 79 | 75 |
| 7 | a | 51544 | 51532 | 48057 | 48353 | 36640 | 27415 | 53 | 94 | 76 | 75 |
| 7 | b | 56270 | 56249 | 51801 | 52076 | 38843 | 28410 | 50 | 92 | 75 | 73 |
| 7 | c | 60164 | 60135 | 55518 | 55849 | 41600 | 30615 | 51 | 93 | 74 | 74 |
| 8 | a | 46264 | 46247 | 42305 | 42650 | 31006 | 24974 | 54 | 92 | 73 | 81 |
| 8 | b | 54245 | 54223 | 49463 | 49746 | 36187 | 29564 | 55 | 91 | 73 | 82 |
| 8 | c | 69953 | 69934 | 64764 | 65354 | 49052 | 39497 | 56 | 93 | 75 | 81 |
| 9 | a | 51685 | 51660 | 47003 | 46813 | 34004 | 27358 | 53 | 91 | 73 | 80 |
| 9 | b | 50777 | 50768 | 46392 | 46707 | 33718 | 26813 | 53 | 92 | 72 | 80 |
| 9 | c | 46878 | 46868 | 45615 | 46120 | 33934 | 27035 | 54 | 92 | 74 | 80 |
| 10 | a | 38917 | 38870 | 33836 | 34188 | 28431 | 23582 | 61 | 88 | 83 | 83 |
| 10 | b | 37349 | 37318 | 31888 | 32018 | 25269 | 20392 | 55 | 86 | 79 | 81 |
| 10 | c | 43778 | 43752 | 38521 | 38394 | 31416 | 25626 | 59 | 88 | 82 | 82 |
| 11 | a | 51812 | 51792 | 47109 | 46958 | 34054 | 26597 | 51 | 91 | 73 | 78 |
| 11 | b | 59103 | 59091 | 54331 | 54123 | 39737 | 30752 | 52 | 92 | 73 | 77 |
| 11 | c | 48359 | 48339 | 44148 | 44073 | 32723 | 25237 | 52 | 91 | 74 | 77 |
| 12 | a | 66037 | 66013 | 61514 | 62222 | 47581 | 35174 | 53 | 94 | 76 | 74 |
| 12 | b | 58565 | 58553 | 54421 | 54977 | 41581 | 30643 | 52 | 93 | 76 | 74 |
| 12 | c | 52444 | 52431 | 48850 | 49226 | 37472 | 27607 | 53 | 94 | 76 | 74 |
| 13 | a | 46854 | 46843 | 42708 | 42853 | 31696 | 24779 | 53 | 91 | 74 | 78 |
| 13 | b | 76014 | 75993 | 70623 | 70536 | 54140 | 41650 | 55 | 93 | 77 | 77 |
| 13 | c | 76815 | 79783 | 73888 | 73716 | 55956 | 42572 | 53 | 93 | 76 | 76 |
| 14 | a | 55829 | 55803 | 51619 | 51399 | 40989 | 29019 | 52 | 92 | 80 | 71 |
| 14 | b | 49643 | 49631 | 46207 | 46011 | 36618 | 25947 | 52 | 93 | 80 | 71 |
| 14 | c | 54363 | 54353 | 50385 | 50333 | 39969 | 28461 | 52 | 93 | 79 | 71 |
| 15 | a | 52878 | 52862 | 48350 | 48448 | 35668 | 28329 | 54 | 92 | 74 | 79 |
| 15 | b | 51272 | 51257 | 47054 | 47192 | 35017 | 27602 | 54 | 92 | 74 | 79 |
| 15 | c | 50022 | 50008 | 45585 | 45607 | 33670 | 26627 | 53 | 91 | 74 | 79 |
| 16 | a | 53134 | 53119 | 48895 | 49380 | 35554 | 30586 | 58 | 93 | 72 | 86 |
| 16 | b | 60505 | 60489 | 55852 | 55975 | 39908 | 33664 | 56 | 92 | 71 | 84 |
| 16 | c | 53599 | 53581 | 49126 | 49278 | 35543 | 30033 | 56 | 92 | 72 | 84 |
| 17 | a | 53702 | 53691 | 48093 | 48554 | 33381 | 27958 | 52 | 90 | 69 | 84 |
| 17 | b | 72247 | 72218 | 65953 | 66769 | 46576 | 38628 | 53 | 92 | 70 | 83 |
| 17 | c | 69773 | 69745 | 63018 | 63597 | 44201 | 36771 | 53 | 91 | 70 | 83 |
| 18 | a | 56643 | 56623 | 52812 | 52833 | 42201 | 33120 | 58 | 93 | 80 | 78 |
| 18 | b | 65302 | 65256 | 61059 | 61092 | 48916 | 37925 | 58 | 94 | 80 | 78 |
| 18 | c | 58700 | 58684 | 54947 | 55263 | 44198 | 33268 | 57 | 94 | 80 | 75 |
| 19 | a | 54132 | 54099 | 49582 | 49642 | 37020 | 30084 | 56 | 92 | 75 | 81 |
| 19 | b | 54628 | 54607 | 50208 | 50382 | 37592 | 30368 | 56 | 92 | 75 | 81 |
| 19 | c | 57906 | 57888 | 53101 | 53404 | 39555 | 32007 | 55 | 92 | 74 | 81 |
| 20 | a | 47233 | 47220 | 43493 | 43723 | 33560 | 25228 | 53 | 92 | 77 | 75 |
| 20 | b | 52815 | 52798 | 48606 | 48976 | 37383 | 28311 | 54 | 92 | 76 | 76 |
| 20 | c | 49429 | 49422 | 45419 | 45788 | 35051 | 26022 | 53 | 92 | 77 | 74 |
| 21 | a | 79537 | 79503 | 73871 | 74084 | 58961 | 43179 | 54 | 93 | 80 | 73 |
| 21 | b | 69321 | 69306 | 64695 | 65219 | 52129 | 38148 | 55 | 94 | 80 | 73 |
| 21 | c | 57313 | 57305 | 53540 | 53726 | 42776 | 31784 | 55 | 94 | 80 | 74 |
| 22 | a | 63004 | 62985 | 58754 | 58664 | 46195 | 32001 | 51 | 93 | 79 | 69 |
| 22 | b | 56037 | 56024 | 52376 | 52586 | 41580 | 28853 | 51 | 94 | 79 | 69 |
| 22 | c | 44713 | 44698 | 41458 | 41531 | 33402 | 24435 | 55 | 93 | 80 | 73 |
| 23 | a | 49597 | 49574 | 46098 | 45918 | 35736 | 24351 | 49 | 93 | 78 | 68 |
| 23 | b | 50359 | 50542 | 47063 | 47293 | 36873 | 26090 | 52 | 93 | 78 | 71 |
| 23 | c | 43379 | 43358 | 40156 | 40062 | 31204 | 22347 | 52 | 93 | 78 | 72 |
| 24 | a | 47310 | 47294 | 42354 | 42434 | 31606 | 26673 | 56 | 90 | 74 | 84 |
| 24 | b | 37447 | 37443 | 33655 | 33599 | 25070 | 21015 | 56 | 90 | 75 | 84 |
| 24 | c | 45518 | 45508 | 40950 | 41340 | 30578 | 25704 | 56 | 90 | 74 | 84 |
| 25 | a | 51948 | 51931 | 47506 | 48086 | 34579 | 27599 | 53 | 92 | 72 | 80 |
| 25 | b | 56743 | 56725 | 52506 | 52961 | 38487 | 30599 | 54 | 93 | 73 | 80 |
| 25 | c | 54592 | 54583 | 50263 | 50853 | 36440 | 29542 | 54 | 93 | 72 | 81 |
| 26 | a | 41943 | 41932 | 38642 | 38995 | 31021 | 23671 | 56 | 93 | 80 | 76 |
| 26 | b | 40915 | 40902 | 37560 | 37903 | 30386 | 23522 | 57 | 92 | 80 | 77 |
| 26 | c | 29792 | 29781 | 27240 | 27318 | 21741 | 17188 | 58 | 92 | 80 | 79 |
| 27 | a | 45197 | 45182 | 40753 | 41148 | 30768 | 23819 | 53 | 91 | 75 | 77 |
| 27 | b | 48150 | 48127 | 43563 | 43657 | 32453 | 25623 | 53 | 91 | 74 | 79 |
| 27 | c | 47797 | 47772 | 42817 | 42904 | 32017 | 25496 | 53 | 90 | 75 | 80 |
| 28 | a | 57359 | 57349 | 52686 | 53280 | 40067 | 33535 | 58 | 92 | 75 | 84 |
| 28 | b | 53102 | 53086 | 48821 | 49128 | 36958 | 31224 | 59 | 92 | 75 | 84 |
| 28 | c | 58344 | 58321 | 53273 | 53341 | 40294 | 33585 | 58 | 91 | 76 | 83 |
| 29 | a | 41270 | 41251 | 37404 | 37525 | 27722 | 22812 | 55 | 91 | 74 | 82 |
| 29 | b | 37633 | 37627 | 34306 | 34592 | 26050 | 22142 | 59 | 92 | 75 | 85 |
| 29 | c | 57958 | 57948 | 53599 | 54132 | 39044 | 31666 | 55 | 93 | 72 | 81 |
| 30 | a | 48703 | 48676 | 44518 | 44344 | 34966 | 29151 | 60 | 91 | 79 | 83 |
| 30 | b | 42232 | 42215 | 38251 | 38291 | 30351 | 25328 | 60 | 91 | 79 | 83 |
| 30 | c | 52068 | 52049 | 47718 | 48005 | 38270 | 31930 | 61 | 92 | 80 | 83 |
| 31 | a | 41136 | 41119 | 37038 | 37332 | 28015 | 23273 | 57 | 90 | 75 | 83 |
| 31 | b | 54070 | 54060 | 49283 | 49682 | 37178 | 30913 | 57 | 92 | 75 | 83 |
| 31 | c | 44480 | 44475 | 40226 | 40328 | 30420 | 25165 | 57 | 91 | 75 | 83 |
| 32 | a | 46684 | 46667 | 41395 | 41544 | 30297 | 26826 | 57 | 89 | 73 | 89 |
| 32 | b | 48504 | 48492 | 43270 | 43692 | 32565 | 29061 | 60 | 90 | 75 | 89 |
| 32 | c | 44499 | 44491 | 39606 | 40005 | 29386 | 25971 | 58 | 89 | 73 | 88 |
| 33 | a | 38141 | 38130 | 34836 | 34643 | 26606 | 22299 | 58 | 91 | 77 | 84 |
| 33 | b | 44310 | 44297 | 40662 | 41063 | 31527 | 26364 | 59 | 92 | 77 | 84 |
| 33 | c | 58659 | 58645 | 54029 | 54396 | 40178 | 31886 | 54 | 92 | 74 | 79 |
| 34 | a | 40647 | 40636 | 37535 | 37701 | 29403 | 23792 | 59 | 93 | 78 | 81 |
| 34 | b | 44686 | 44676 | 41117 | 41323 | 32067 | 25823 | 58 | 92 | 78 | 81 |
| 34 | c | 54040 | 54025 | 49930 | 49854 | 39119 | 31350 | 58 | 92 | 78 | 80 |
| 35 | a | 43498 | 43482 | 39715 | 40101 | 30851 | 24731 | 57 | 92 | 77 | 80 |
| 35 | b | 52980 | 52960 | 48658 | 48745 | 37161 | 29799 | 56 | 92 | 76 | 80 |
| 35 | c | 45120 | 45107 | 41279 | 41743 | 32259 | 26918 | 60 | 92 | 77 | 83 |
| Average | | 53056 | 53038 | 48640 | 48877 | 36918 | 29066 | 55 | 92 | 76 | 79 |

**Table S3-2**. Ranges of measured biogeochemical variables across all sites (outlier corrected). TOC = total organic carbon; TN = total nitrogen; RPA = relative peak area.

| Variable | Unit | Minimum | Mean | Median | Maximum |
|---|---|---|---|---|---|
| **pH** | - | 4.1 | 5.0 | 5.0 | 6.7 |
| **Sand** | % | 15.6 | 59.1 | 62.6 | 93.4 |
| **Clay** | % | 0.8 | 3.0 | 2.5 | 8.7 |
| **MAT** | °C | 3.0 | 10.9 | 10.7 | 17.1 |
| **Water balance (MAP-PET)** | mm | -1382 | 207 | 81 | 1704 |
| **TOC** | $g\ kg^{-1}$ | 5.9 | 68.1 | 57.4 | 187.4 |
| **TN** | $g\ kg^{-1}$ | 0.5 | 5.4 | 4.1 | 15.0 |
| **KCl-OC** | $mg\ kg^{-1}$ | 40.7 | 244.9 | 214.2 | 632.8 |
| **KCl-N** | $mg\ kg^{-1}$ | 3.6 | 122.2 | 107.9 | 402.9 |
| **RockEval I-Index** | - | -0.09 | 0.18 | 0.15 | 0.41 |
| **RPA aliphatic** | % | 41.9 | 75.1 | 77.7 | 90.5 |
| **RPA carboxylic** | % | 5.4 | 11.1 | 10.8 | 23.2 |
| **Bulk C:N** | - | 12.1 | 15.4 | 15.2 | 20.7 |
| **13C shift** | - | 0.96 | 1.02 | 1.02 | 1.07 |
| **OC in POM fraction** | % | 3.5 | 12.6 | 8.2 | 38.8 |
| **KCl Humification** | - | 1.2 | 2.2 | 2.0 | 3.9 |
| **KCl Molecular weight** | - | 1.3 | 2.6 | 2.6 | 4.3 |
| **KCl Aromaticity** | $L\ mol^{-1}\ cm^{-1}$ | 38 | 206 | 158 | 471 |
| **KCl C:N** | - | 2.8 | 6.1 | 5.3 | 13.4 |

**Table S3-3**. Results of the rPCA for dimension reduction. The 8 retained rotated components (RCs) with cumulative variance of 80 % that resulted from this rPCA. The top shows Eigenvalues and individual as well as cumulative variability explained by the RCs. Loadings are shown for all 19 input variables. Interpretation and naming of the RCs was based on variables with loadings ≥ 0.5 (highlighted), and the assigned names for the RCs are given at the top. POM = particulate organic matter; DRIFTS = diffuse reflectance infrared Fourier transform spectroscopy; TOC = total organic carbon; TN = total nitrogen; RPA = relative peak area.

| | RC1 | RC2 | RC3 | RC4 | RC5 | RC6 | RC7 | RC8 |
|---|---|---|---|---|---|---|---|---|
| | SOM quantity | SOM quality A (UV-Vis) | Texture | SOM quality B (POM) | SOM quality C (DRIFTS) | SOM quality D ($^{13}$C) | SOM quality E (C:N ratio) | pH |
| **Eigenvalue** | **7.41** | **3.09** | **2.09** | **1.32** | **1.1** | **1.06** | **0.76** | **0.61** |
| **Proportion var.** | **0.23** | **0.11** | **0.09** | **0.09** | **0.07** | **0.07** | **0.07** | **0.06** |
| **Cumulative var.** | **0.23** | **0.34** | **0.44** | **0.52** | **0.60** | **0.67** | **0.73** | **0.80** |
| pH | -0.28 | -0.07 | 0.05 | 0.33 | -0.16 | 0.08 | -0.16 | 0.86 |
| Sand | 0.10 | 0.10 | 0.81 | 0.28 | 0.21 | 0.31 | 0.08 | 0.04 |
| Clay | -0.39 | -0.08 | -0.88 | 0.09 | -0.12 | 0.17 | -0.11 | 0.01 |
| MAT | -0.26 | 0.09 | -0.22 | 0.06 | -0.14 | 0.24 | -0.34 | 0.08 |
| Water balance (MAP-PET) | 0.56 | 0.00 | 0.20 | -0.49 | 0.21 | -0.09 | 0.23 | -0.36 |
| TOC | 0.90 | 0.15 | 0.13 | -0.13 | 0.15 | -0.01 | 0.21 | -0.07 |
| TN | 0.93 | 0.13 | 0.14 | -0.15 | 0.12 | -0.02 | 0.02 | -0.08 |
| KCl-OC | 0.81 | 0.15 | 0.17 | -0.07 | 0.08 | -0.04 | -0.01 | -0.33 |
| KCl-N | 0.87 | 0.18 | 0.14 | -0.19 | 0.10 | 0.05 | -0.03 | -0.07 |
| RockEval I-Index | 0.29 | 0.12 | 0.21 | 0.06 | 0.19 | 0.15 | 0.06 | -0.09 |
| RPA aliphatic | 0.49 | 0.19 | 0.04 | 0.05 | 0.61 | -0.03 | 0.13 | -0.16 |
| RPA carboxylic | -0.23 | 0.10 | -0.28 | -0.05 | -0.85 | -0.17 | -0.23 | 0.12 |
| Bulk C:N | 0.10 | -0.10 | 0.11 | 0.11 | 0.19 | 0.05 | 0.94 | -0.12 |
| 13C shift | 0.00 | -0.02 | -0.03 | -0.12 | -0.09 | -0.98 | -0.04 | -0.06 |
| OC in POM fraction | -0.26 | -0.13 | 0.13 | 0.90 | 0.06 | 0.11 | 0.15 | 0.21 |
| KCl Humification | 0.19 | 0.96 | 0.05 | -0.06 | -0.02 | 0.01 | -0.06 | 0.03 |
| KCl Molecular weight | 0.16 | 0.96 | 0.07 | -0.09 | 0.00 | 0.01 | -0.05 | -0.11 |
| KCl Aromaticity | -0.35 | -0.22 | -0.03 | 0.54 | 0.18 | 0.34 | 0.02 | 0.31 |
| KCl C:N | -0.19 | -0.13 | -0.13 | -0.03 | 0.03 | 0.04 | -0.08 | 0.06 |

## 7.4 Appendix Chapter 4

**Supplementary Text S4-1.**

Stepwise regression is an approach to build linear regression models based on iterative automated variable selection. Due to the resulting constrained complexity, the models are simple to interpret. However, because of multiple comparisons during model fitting, p-values can be biased too low. We performed stepwise regression with Monte Carlo cross validation to select the most important predictor variables for prediction of the 11 microbial traits and functions. For this, independent variables were scaled, and the data was split 100 times into a 75 % training set and a 25 % validation set. Model performance was assessed using the RMSE relative to the validation set, and a maximum of five predictor variables (out of 8, 11 and 19 potential predictors for ENV, MIC and ENV+MIC, respectively) were retained in order to constrain model complexity and avoid overfitting. For every model, homoscedasticity (Breusch-Pagan test, R-package "lmtest", Zeileis & Hothorn (2002)) and normal distribution of residuals (Shapiro-Wilk test) were tested, and in case of violation of either assumption, the target variables were log-transformed. Scaled variable importance (sVI) was quantified for each significant predictor (p-value ≤ 0.05).

Least absolute shrinkage and selection operator (lasso) regression, also known as L1 regularization regression, uses penalties to constrain model complexity. In this regression approach, coefficients of independent variables can be shrunk to zero if their explanatory value is outweighed by their addition to model complexity. The strength of this penalty is controlled by the model parameter λ. We determined the optimum value for λ for each model with Monte Carlo cross validation, as described in above. Tested λ values ranged from 0 to 5 at intervals of 0.1, from 5 to 10 at intervals of 0.5, and from 10 to 100 at intervals of 1. The best model was chosen based on RMSE relative to the validation set, and in case of violation of normal distribution and homoscedasticity of residuals, the target variables were log-transformed. For each retained predictor, sVI was calculated as described above. The standard errors of the predictors, which are required for sVI calculation, are not saved by the caret-implementation of the glmnet package. We therefore estimated them based on 100-fold bootstrapping of lasso models using the optimum λ value. In cases where all predictor variables were shrunk to zero and only an intercept was retained, we did not consider the resulting models for further analysis.

In addition to the two linear regression approaches, we also conducted random forest regression in order to allow for non-linear relationships between microbial traits and functions and predictor variables. As a trade-off, random forest models can be difficult to interpret, as they retain all predictor variables, and do not give coefficients. We conducted Monte Carlo Cross validation as described above, in order to tune the models for the number of trees (100 or 1000), node size (2, 7, 15) and the number

of variables to randomly sample as candidates at each split (mtry, from 1 to the number of predictors, in steps of 1). Model evaluation was based on RMSE. For each variable, models were built for untransformed and log-transformed target variables, and the model with better goodness of fit ($R^2$) was retained. The sVI values for the random forest models were obtained with the function varImp() from the caret package.

Adjusted $R^2$ vs. unadjusted $R^2$: While in simple linear regression approaches such as stepwise regression, $R^2$ is generally adjusted to account for the number of predictors, this is not readily possible for lasso regression or random forest regression. We therefore used the unadjusted $R^2$ to compare the regreesion approaches, while cautioning the reader that in the case of stepwise regression, a larger number of retained predictors artificially inflates unadjusted $R^2$ values.

Coefficient signs of model predictors: Notably, a large number of (less important) predictors only featured in the random forest models due to the gradient of model complexity from stepwise regression to random forest. In contrast to the non-linear random forest models, stepwise models and lasso models feature coefficient signs (i.e. + and -). For predictor variables that featured in corresponding stepwise models as well as lasso models, we verified that coefficient signs were consistent for predictor variables. Subsequently, for predictor variables that featured in one or both of the linear models, we assigned the signs to the respective mean sVI values. Mean sVI values of predictor variables that only featured in random forest models could not be assigned with signs.

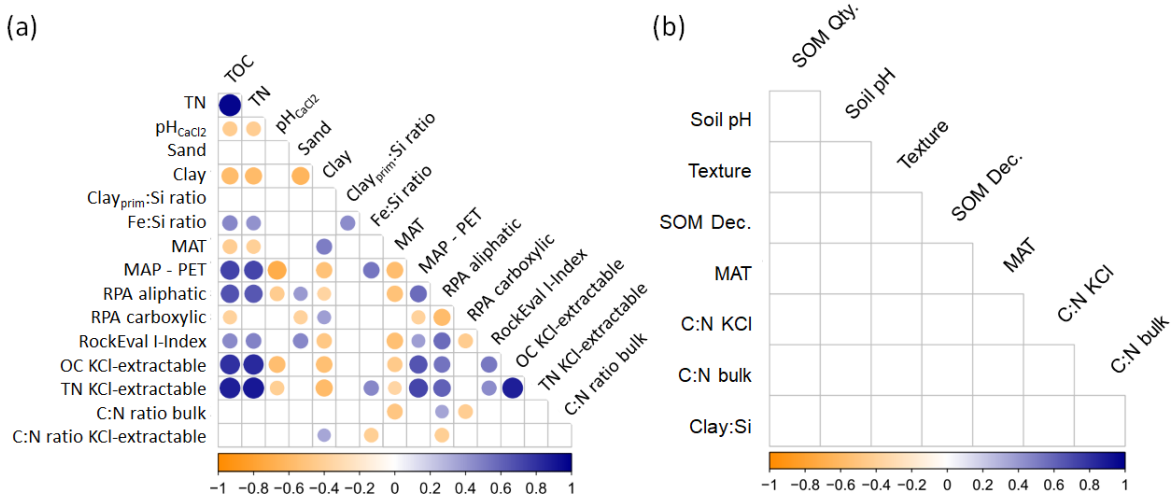



**Figure S4-1**. Correlation matrix showing significant Pearson correlations (p-value < 0.05) between the (a) raw environmental variables and (b) the eight retained rotated components of the ENV dataset. Circle size and color indicate the strength of the relationships (Pearson correlation coefficient). Shown

are significant correlations (Pearson correlation, $p < 0.05$). SOM Qty. = SOM quantity, SOM Dec. = SOM Decomposition.

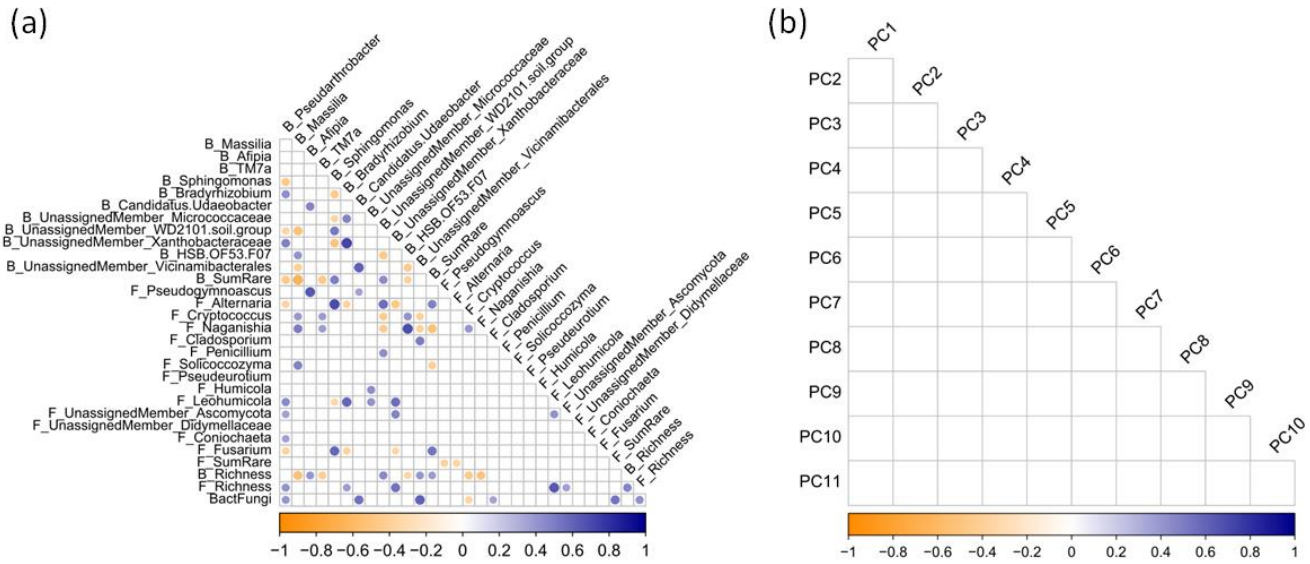



**Figure S4-2**. Correlation matrix showing significant Pearson correlations (p-value < 0.05) between the (a) raw microbial community variables and (b) the 11 retained principal components of the MIC dataset. Circle size and color indicate the strength of the relationships (Pearson correlation coefficient). Shown are significant correlations (Pearson correlation, $p < 0.05$). B = Bacteria, F = Fungi, BactFungi = Bacteria:Fungi ratio, PC = principal component.

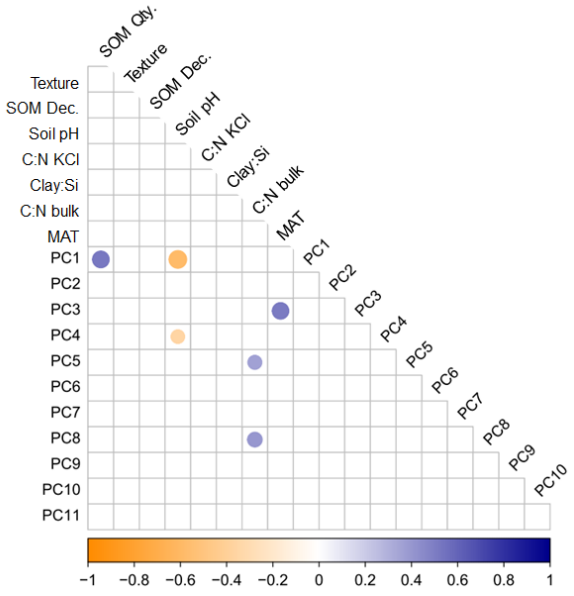



**Figure S4-3.** Correlation matrix showing significant Pearson correlations (p-value < 0.05) between the eight rotated components of the ENV dataset and the 11 principal components of the MIC dataset. Circle size and color indicate the strength of the relationships (Pearson correlation coefficient). Shown

are significant correlations (Pearson correlation, $p < 0.05$). SOM Qty. = SOM quantity, SOM Dec. = SOM Decomposition, PC = Principal component.

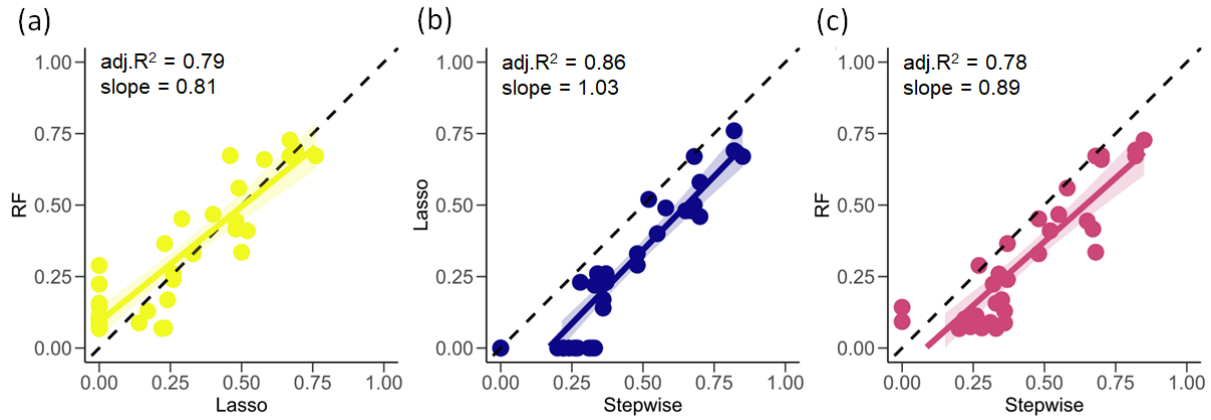



**Figure S4-4.** Pairwise combinations of $R^2$ values of the models. The linear relationships with slopes close to 1 show that all models result in comparable patterns of goodness of fit. The dashed black indicates the 1:1 line, points show data points, solid lines show linear regression fits with 95 % confidence intervals. RF = Random Forest. n = 33.

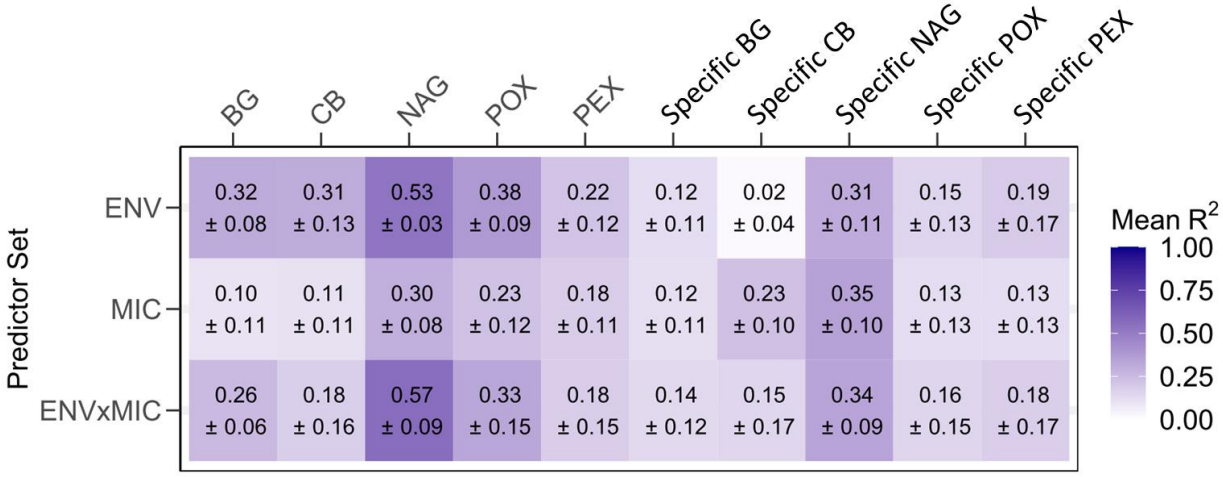



**Figure S4-5.** Proportion of the variation of 10 potential extracellular enzyme activities (x-axis) that can be explained with three different sets of predictors (y-axis). Shown are mean ± S.D. of $R^2$ across three different regression approaches (stepwise, lasso, and random forest). ENV = Environmental data, MIC = Microbial community data, BG = ß-Glucosidase, CB = cellobiosidase, NAG = N-acetyl-β-glucosaminidase, POX = phenoloxidase, PEX = peroxidase, Spec. = specific.

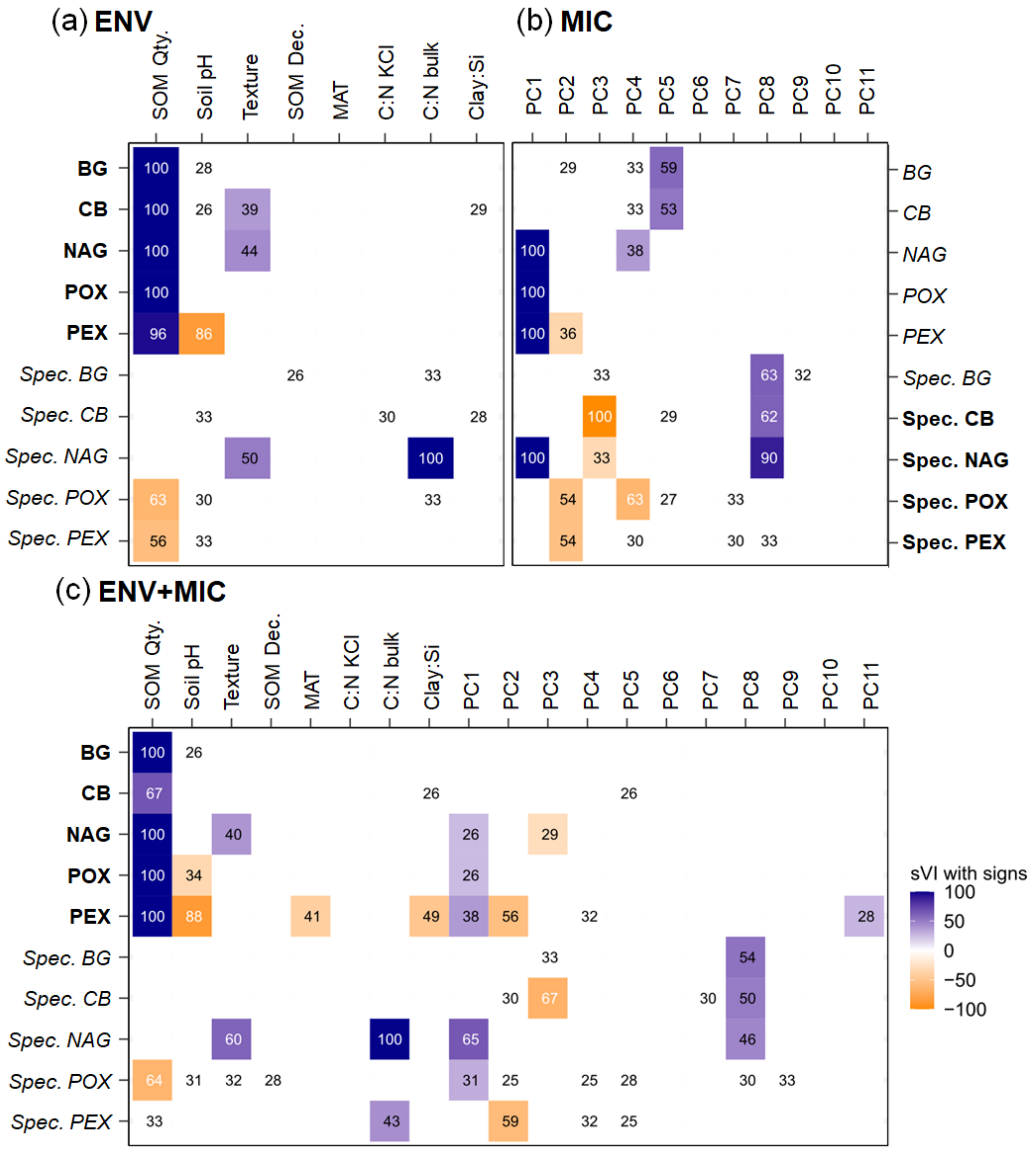



**Figure S4-6.** Scaled variable importance (sVI) ≥ 25 in the prediction of 10 potential extracellular enzyme activities (y-axis). (a) Prediction with environmental data (ENV); (b) prediction with microbial community data (MIC); (c) Prediction with both datasets (ENV+MIC). Shown are mean sVI values across three different regression approaches. Colors indicate the signs of the coefficients (negative: orange; positive: blue). Note that variables which only featured in random forest models do not have signs. The best models for each microbial trait or function are highlighted by bold font. SOM Qty. = SOM quantity, SOM Dec. = SOM Decomposition, PC = Principal component, BG = ß-Glucosidase, CB = cellobiosidase, NAG = N-acetyl-β-glucosaminidase, POX = phenoloxidase, PEX = peroxidase, Spec. = specific.

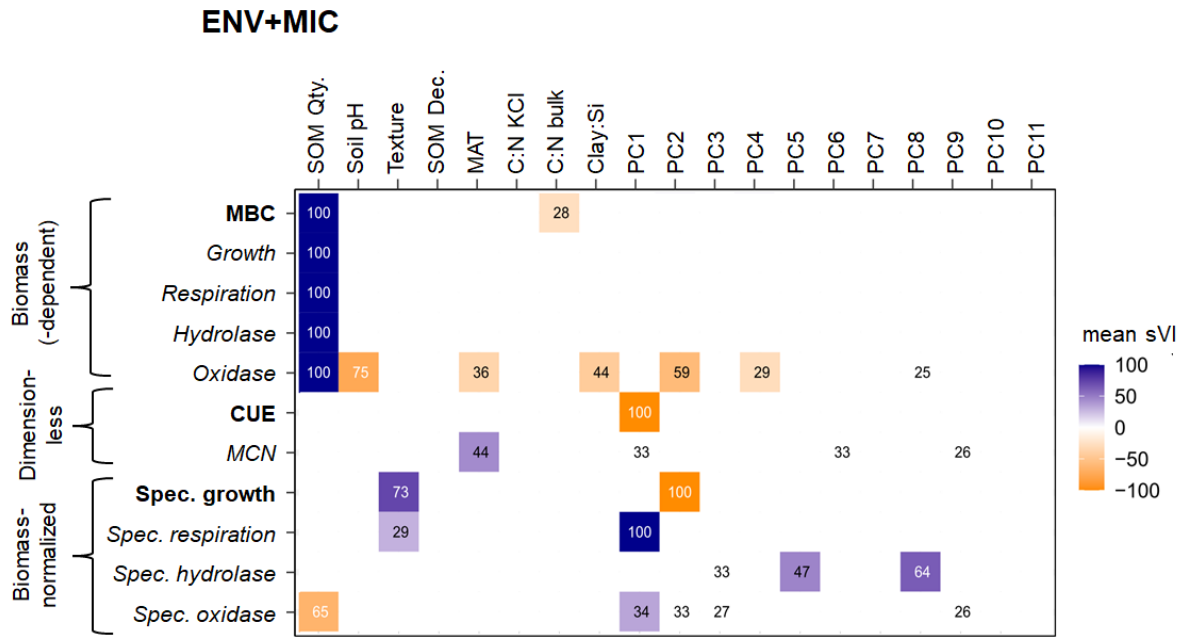



**Figure S4-7.** Scaled variable importance (sVI) ≥ 25 in the prediction of 11 microbial traits and functions (y-axis). Prediction with environmental data and microbial community data (ENV+MIC). Shown are mean sVI values across three different regression approaches (corresponding S.D. values are shown in Figure S4-8). Colors indicate the signs of the coefficients (negative: orange; positive: blue). Note that variables which only featured in random forest models do not have signs. The best models for each microbial trait or function are highlighted by bold font. SOM Qty. = SOM quantity, SOM Dec. = SOM Decomposition, MBC = Microbial biomass carbon, CUE = Carbon use efficiency, MCN = Microbial C:N, Spec. = specific, KCl = extractable with KCl, PC = principal component.

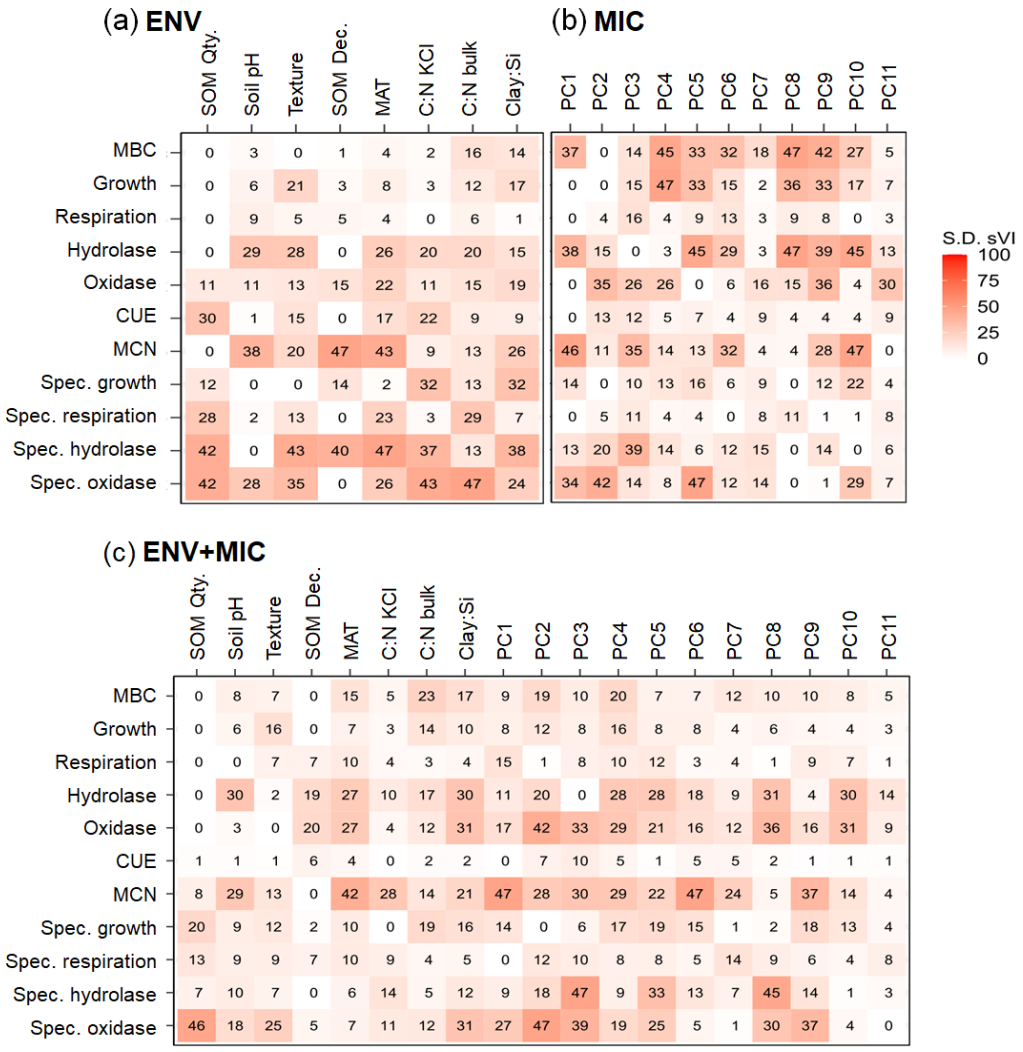



**Figure S4-8.** Standard deviations (S.D.) of the scaled variable importances (sVI) across three different regression approaches (stepwise, lasso, and random forest) (y-axis). (a) Prediction with environmental data (ENV); (b) prediction with microbial community data (MIC); (c) Prediction with both datasets (ENV+MIC). Shown are mean sVI values across three different regression approaches. Lighter red indicates a lower S.D., and thus a higher consensus among the models regarding the respective sVI. SOM Qty. = SOM quantity, SOM Dec. = SOM Decomposition, MBC = Microbial biomass carbon, CUE = Carbon use efficiency, MCN = Microbial C:N, Spec. = specific, KCl = extractable with KCl, PC = principal component.

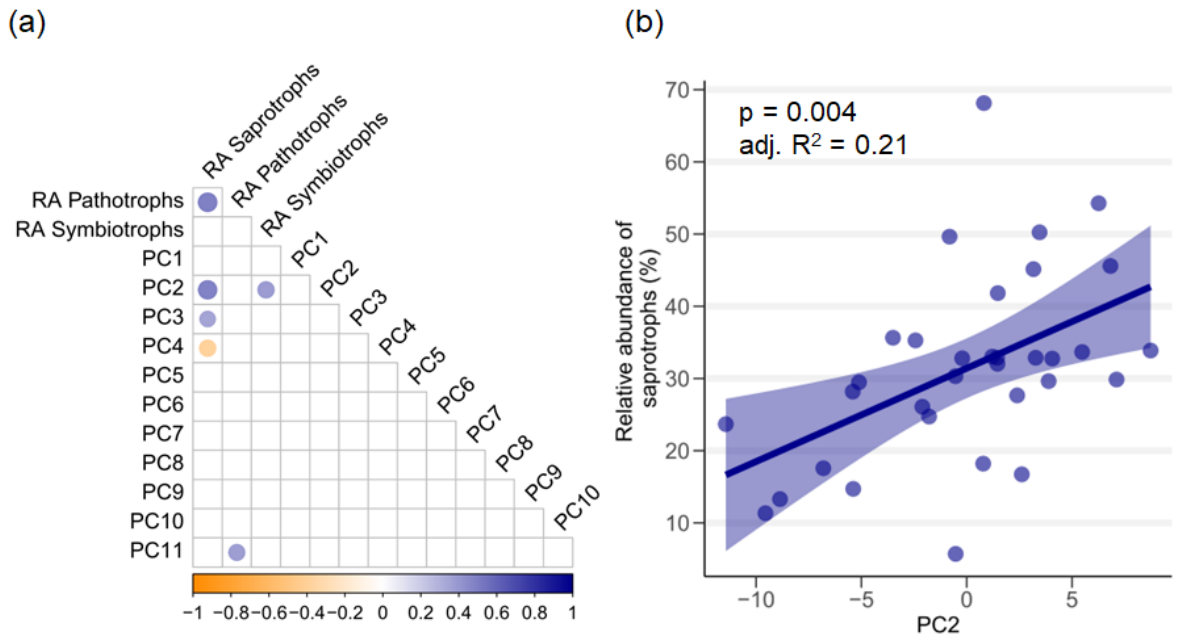



**Figure S4-9.** Correlation between the relative abundance of FUNGuild assigned fungal trophic modes and principal components of the MCI dataset. Points show data points, solid lines show linear regression fits with 95 % confidence intervals, RA = relative abundance, PC = principal component.

**Table S4-1.** Soil physicochemical and climatic variation along the gradient. Ranges of all soil physicochemical and climatic variables that were used for the rPCA for dimension reduction of the ENV data set.

| Variable | Unit | Min. | Mean | Median | Max. |
|---|---|---|---|---|---|
| SOC | g $kg^{-1}$ | 5.9 | 68.3 | 57.4 | 187.4 |
| TN | g $kg^{-1}$ | 0.5 | 5.4 | 3.9 | 15.0 |
| OC KCl-extractable | mg $kg^{-1}$ | 40.7 | 249.1 | 214.2 | 632.8 |
| TN KCl-extractable | mg $kg^{-1}$ | 3.6 | 131.6 | 109.9 | 402.9 |
| $pH_{CaCl2}$ | - | 4.1 | 5.0 | 4.9 | 6.7 |
| Sand | % | 15.6 | 59.3 | 62.6 | 93.4 |
| Clay | % | 0.8 | 3.0 | 2.4 | 8.7 |
| $Clay_{prim}$:Si ratio | - | 0.01 | 0.04 | 0.04 | 0.08 |
| Fe:Si ratio | - | 0.06 | 0.22 | 0.23 | 0.45 |
| MAT | °C | 3.0 | 10.9 | 10.7 | 17.1 |
| MAP - PET | mm | -1207 | 261 | 133 | 1704 |
| RPA aliphatic | % | 41.9 | 75.0 | 77.8 | 90.5 |
| RPA carboxylic | % | 5.4 | 11.0 | 10.6 | 23.2 |
| RockEval I-Index | - | -0.09 | 0.17 | 0.15 | 0.40 |
| C:N ratio bulk | - | 12.1 | 15.4 | 15.2 | 20.7 |
| C:N ratio KCl-extractable | - | 2.8 | 6.1 | 5.2 | 13.4 |

**Table S4-2.** Results of the rPCA for dimension reduction of the ENV dataset. The top shows Eigenvalues and individual as well as cumulative variability explained by the RCs. Loadings are shown for all input variables. Interpretation and naming of the RCs was based on variables with loadings ≥ 0.5 (highlighted), and the assigned names for the RCs are given at the top. SOM Decomp. = SOM Decomposition.

| | RC1 | RC2 | RC3 | RC4 | RC5 | RC6 | RC7 | RC8 |
|---|---|---|---|---|---|---|---|---|
| | **SOM Quantity** | **Soil pH** | **Texture** | **SOM Decomp.** | **MAT** | **C:N KCl** | **C:N bulk** | **Clay:Si** |
| **Eigenvalue** | 7.05 | 2.16 | 1.49 | 1.30 | 0.81 | 0.75 | 0.72 | 0.48 |
| **Proportion var.** | 0.28 | 0.08 | 0.08 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 |
| **Cumulative var.** | 0.28 | 0.36 | 0.44 | 0.51 | 0.58 | 0.65 | 0.72 | 0.79 |
| **SOC** | 0.90 | -0.11 | 0.10 | -0.14 | -0.08 | -0.09 | 0.18 | -0.18 |
| **TN** | 0.94 | -0.12 | 0.09 | -0.10 | -0.10 | -0.06 | 0.01 | -0.08 |
| **OC KCl-extractable** | 0.83 | -0.29 | 0.05 | -0.09 | -0.19 | -0.04 | -0.03 | -0.03 |
| **TN KCl-extractable** | 0.91 | -0.13 | 0.09 | -0.08 | -0.07 | -0.14 | -0.04 | 0.06 |
| **$pH_{CaCl2}$** | -0.29 | 0.92 | 0.10 | 0.08 | 0.08 | 0.09 | -0.11 | 0.07 |
| **Sand** | 0.13 | 0.10 | 0.91 | -0.16 | -0.07 | -0.17 | 0.11 | 0.11 |
| **Clay** | -0.46 | 0.01 | -0.47 | 0.12 | 0.24 | 0.14 | -0.09 | -0.15 |
| **$Clay_{prim}$:Si ratio** | 0.10 | -0.06 | -0.11 | 0.09 | -0.05 | -0.06 | 0.12 | -0.95 |
| **Fe:Si ratio** | 0.33 | -0.08 | -0.09 | 0.03 | -0.08 | -0.20 | 0.04 | -0.23 |
| **MAT** | -0.23 | 0.13 | -0.10 | 0.10 | 0.87 | 0.12 | -0.24 | 0.06 |
| **MAP - PET** | 0.58 | -0.50 | 0.00 | -0.16 | -0.27 | -0.08 | 0.15 | 0.00 |
| **RPA aliphatic** | 0.48 | -0.17 | 0.16 | -0.36 | -0.20 | -0.26 | 0.13 | -0.09 |
| **RPA carboxylic** | -0.21 | 0.10 | -0.17 | 0.90 | 0.09 | -0.07 | -0.21 | -0.11 |
| **RockEval I-Index** | 0.34 | -0.13 | 0.24 | -0.16 | -0.26 | 0.05 | 0.04 | 0.06 |
| **C:N ratio bulk** | 0.04 | -0.11 | 0.11 | -0.20 | -0.20 | -0.09 | 0.93 | -0.13 |
| **C:N ratio KCl-extractable** | -0.15 | 0.09 | -0.16 | -0.05 | 0.10 | 0.94 | -0.09 | 0.06 |

**Table S4-3.** Variation of microbial community composition along the gradient. Ranges of all microbial community composition variables that were used for the PCA for dimension reduction of the MIC data set. Variables with "B_" are relative abundances (expressed as %) of bacterial genera, variables with "F_" are relative abundances (expressed as %) of fungal genera. "B_Richness" and "F_Richness" are numbers of different bacterial and fungal amplicon sequencing variants, respectively. "BactFungi" is the ratio of 16S rRNA gene counts over ITS2 region counts.

| Variable | Min. | Mean | Median | Max. |
|---|---|---|---|---|
| B_Pseudarthrobacter | 0.27 | 9.38 | 5.97 | 28.06 |
| B_Massilia | 0.26 | 8.31 | 4.12 | 65.21 |
| B_Afipia | 0.17 | 1.43 | 1.21 | 4.26 |
| B_TM7a | 0.01 | 2.37 | 0.38 | 42.85 |
| B_Sphingomonas | 0.27 | 3.67 | 1.23 | 22.80 |
| B_Bradyrhizobium | 0.00 | 1.47 | 1.11 | 4.30 |
| B_Candidatus.Udaeobacter | 0.06 | 8.78 | 6.08 | 33.51 |
| B_UnassignedMember_Micrococcaceae | 0.03 | 1.87 | 0.73 | 8.97 |
| B_UnassignedMember_WD2101.soil.group | 0.07 | 6.69 | 5.96 | 23.59 |
| B_UnassignedMember_Xanthobacteraceae | 0.11 | 1.78 | 1.61 | 4.50 |
| B_HSB.OF53.F07 | 0.00 | 3.20 | 0.17 | 28.46 |
| B_UnassignedMember_Vicinamibacterales | 0.00 | 1.51 | 1.19 | 4.85 |
| B_SumRare | 13.88 | 49.94 | 49.45 | 78.80 |
| F_Pseudogymnoascus | 0.00 | 5.70 | 0.02 | 87.07 |
| F_Alternaria | 0.00 | 5.93 | 0.08 | 44.98 |
| F_Cryptococcus | 0.00 | 5.60 | 0.78 | 59.65 |
| F_Naganishia | 0.00 | 3.99 | 0.27 | 46.17 |
| F_Cladosporium | 0.00 | 2.53 | 0.52 | 22.11 |
| F_Penicillium | 0.00 | 4.27 | 2.69 | 23.98 |
| F_Solicoccozyma | 0.21 | 3.10 | 1.86 | 11.89 |
| F_Pseudeurotium | 0.00 | 1.92 | 0.16 | 29.33 |
| F_Humicola | 0.00 | 2.94 | 0.28 | 34.78 |
| F_Leohumicola | 0.00 | 3.08 | 0.95 | 20.09 |
| F_UnassignedMember_Ascomycota | 0.00 | 4.06 | 0.77 | 38.42 |
| F_UnassignedMember_Didymellaceae | 0.00 | 2.15 | 0.15 | 35.45 |
| F_Coniochaeta | 0.00 | 2.35 | 0.61 | 22.98 |
| F_Fusarium | 0.00 | 3.45 | 0.76 | 25.03 |
| F_SumRare | 8.73 | 49.27 | 49.02 | 87.98 |
| B_Richness | 287 | 778 | 811 | 1342 |
| F_Richness | 67 | 130 | 117 | 206 |
| BactFungi | 1.4 | 88.6 | 53.4 | 310.3 |

**Table S4-4.** Results of the rPCA for dimension reduction of the ENV dataset. The top shows Eigenvalues and individual as well as cumulative variability explained by the RCs. Loadings are shown for all input variables.

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | PC10 | PC11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Eigenvalue** | 6.14 | 4.96 | 2.73 | 2.33 | 1.94 | 1.48 | 1.39 | 1.28 | 1.13 | 1.04 | 0.96 |
| **Proportion var.** | 0.20 | 0.16 | 0.09 | 0.08 | 0.06 | 0.05 | 0.04 | 0.04 | 0.04 | 0.03 | 0.03 |
| **Cumulative var.** | 0.20 | 0.36 | 0.45 | 0.52 | 0.58 | 0.63 | 0.68 | 0.72 | 0.75 | 0.79 | 0.82 |
| B_Pseudarthrobacter | 0.58 | 0.40 | 0.24 | -0.29 | 0.19 | -0.28 | 0.06 | 0.26 | -0.12 | 0.08 | 0.19 |
| B_Massilia | 0.56 | -0.51 | -0.23 | -0.24 | 0.11 | 0.18 | 0.15 | -0.15 | -0.03 | 0.01 | -0.04 |
| B_Afipia | -0.06 | 0.36 | -0.38 | 0.72 | 0.14 | -0.02 | 0.13 | 0.12 | 0.13 | 0.01 | 0.06 |
| B_TM7a | 0.25 | -0.41 | -0.32 | -0.17 | -0.03 | -0.33 | 0.00 | 0.47 | -0.32 | 0.17 | -0.17 |
| B_Sphingomonas | -0.74 | -0.30 | 0.29 | -0.09 | 0.21 | 0.09 | 0.07 | -0.03 | -0.07 | 0.08 | -0.04 |
| B_Bradyrhizobium | 0.59 | 0.37 | 0.34 | 0.34 | -0.06 | 0.13 | -0.19 | 0.19 | 0.11 | 0.01 | 0.05 |
| B_Candidatus.Udaeobacter | -0.16 | 0.57 | -0.63 | 0.08 | -0.07 | 0.03 | -0.08 | -0.15 | -0.05 | -0.14 | 0.17 |
| B_UnassignedMember_Micrococcaceae | 0.45 | 0.05 | 0.42 | 0.43 | -0.36 | -0.25 | 0.26 | -0.10 | 0.01 | -0.17 | 0.17 |
| B_UnassignedMember_WD2101.soil.group | -0.73 | 0.17 | 0.11 | 0.17 | 0.24 | 0.09 | -0.22 | 0.05 | -0.10 | 0.00 | 0.22 |
| B_UnassignedMember_Xanthobacteraceae | 0.53 | 0.61 | 0.17 | 0.14 | 0.06 | 0.20 | -0.22 | -0.08 | -0.05 | 0.19 | 0.00 |
| B_HSB.OF53.F07 | 0.60 | -0.35 | 0.11 | 0.14 | 0.04 | 0.38 | -0.04 | -0.12 | 0.36 | -0.10 | 0.05 |
| B_UnassignedMember_Vicinamibacterales | -0.34 | 0.65 | -0.35 | -0.24 | -0.23 | 0.02 | -0.08 | 0.10 | 0.15 | 0.12 | -0.09 |
| B_SumRare | -0.74 | 0.12 | 0.38 | 0.19 | -0.25 | -0.05 | -0.07 | -0.13 | 0.17 | -0.04 | -0.20 |
| F_Pseudogymnoascus | -0.12 | 0.10 | -0.55 | 0.57 | 0.28 | -0.33 | -0.10 | -0.01 | -0.03 | 0.07 | -0.11 |
| F_Alternaria | -0.67 | -0.29 | 0.35 | -0.03 | 0.27 | -0.01 | 0.09 | 0.12 | -0.05 | -0.25 | -0.09 |
| F_Cryptococcus | 0.40 | -0.55 | -0.03 | 0.19 | -0.22 | 0.08 | -0.10 | 0.30 | 0.06 | -0.03 | -0.13 |
| F_Naganishia | 0.50 | -0.55 | -0.17 | 0.02 | -0.02 | 0.28 | -0.01 | 0.17 | 0.22 | 0.00 | 0.05 |
| F_Cladosporium | -0.33 | 0.23 | -0.10 | -0.17 | -0.48 | -0.03 | 0.44 | 0.27 | 0.30 | -0.01 | -0.21 |
| F_Penicillium | -0.45 | 0.05 | 0.06 | 0.02 | 0.02 | 0.39 | 0.16 | 0.38 | 0.01 | -0.12 | 0.51 |
| F_Solicoccozyma | 0.27 | -0.16 | -0.42 | -0.10 | 0.43 | 0.09 | 0.27 | -0.47 | 0.14 | 0.07 | -0.05 |
| F_Pseudeurotium | 0.21 | 0.16 | 0.39 | -0.04 | 0.06 | -0.36 | -0.39 | -0.07 | 0.24 | 0.38 | 0.07 |
| F_Humicola | -0.11 | 0.00 | 0.09 | 0.28 | -0.70 | -0.05 | 0.24 | -0.40 | -0.24 | 0.08 | 0.05 |
| F_Leohumicola | 0.52 | 0.42 | 0.40 | 0.20 | 0.10 | 0.14 | 0.17 | 0.11 | -0.09 | 0.01 | -0.26 |
| F_UnassignedMember_Ascomycota | 0.33 | 0.37 | 0.16 | -0.03 | 0.09 | 0.29 | 0.08 | -0.11 | -0.65 | -0.11 | -0.10 |
| F_UnassignedMember_Didymellaceae | -0.32 | -0.08 | 0.03 | -0.03 | -0.05 | 0.23 | 0.32 | -0.02 | -0.04 | 0.74 | 0.26 |
| F_Coniochaeta | 0.08 | -0.01 | 0.36 | -0.16 | 0.40 | -0.49 | 0.46 | -0.11 | 0.18 | -0.10 | 0.18 |
| F_Fusarium | -0.62 | -0.13 | 0.20 | -0.21 | -0.02 | 0.17 | -0.22 | -0.09 | 0.04 | 0.08 | -0.23 |
| F_SumRare | 0.21 | 0.43 | 0.00 | -0.60 | -0.16 | -0.04 | -0.30 | -0.19 | 0.10 | -0.19 | 0.16 |
| B_Richness | -0.41 | 0.63 | 0.05 | 0.19 | 0.37 | 0.10 | 0.12 | 0.06 | 0.09 | -0.07 | -0.14 |
| F_Richness | 0.28 | 0.68 | 0.17 | -0.15 | 0.19 | 0.22 | 0.25 | 0.05 | 0.13 | 0.13 | -0.31 |
| BactFungi | 0.10 | 0.72 | -0.30 | -0.39 | -0.15 | -0.01 | 0.15 | 0.10 | 0.04 | -0.12 | 0.06 |

**Table S4-5.** Measured microbial traits and functions along the gradient, sorted by latitude. For visual aid, all variables are colored light (low values) to dark (high values). MBC = microbial biomass carbon, CUE = carbon use efficiency, MCN = microbial C:N ratio, Spec. = specific, resp. = respiration, MUF = 4-Methylumbelliferyl, DOPA = L-3,4-dihydroxyphenylalanine.

| Site ID | Site name | Coordinates WGS84 | | WRB soil group | MBC | Growth | Respiration | Hydrolases | Oxidases |
|---|---|---|---|---|---|---|---|---|---|
| | | (°S) | (°W) | | µg C $g^{-1}$ soil | µg C $g^{-1}$ soil $h^{-1}$ | µg C $g^{-1}$ soil $h^{-1}$ | nmol MUF $g^{-1}$ soil $h^{-1}$ | nmol DOPA $g^{-1}$ soil $h^{-1}$ |
| 1 | Los Vilos | 31.8128 | 71.5001 | Arenosol | 60.0 | 0.5 | 0.6 | 718.6 | 1267.5 |
| 2 | Catapilco | 32.5265 | 71.2608 | Planosol | 65.4 | 0.5 | 0.5 | 1566.8 | 3865.6 |
| 4 | Bochinche | 33.3390 | 71.6124 | Kastanozem | 97.6 | 1.0 | 0.6 | 1272.0 | 2767.8 |
| 5 | Clarillo | 33.6804 | 70.5857 | Kastanozem | 59.3 | 0.3 | 0.5 | 1294.9 | 1971.9 |
| 6 | Matanzas | 33.9686 | 71.8756 | Chernozem / Kastanozem / Phaeozem | 70.6 | 0.2 | 0.5 | 1936.6 | 4062.5 |
| 7 | Pudahuel | 33.9991 | 71.1405 | Chernozem / Kastanozem | 99.2 | 0.6 | 1.0 | 1594.7 | 2228.3 |
| 8 | Pimpinela | 34.3239 | 70.7297 | Luvisol | 107.3 | 0.6 | 0.6 | 1187.5 | 2445.1 |
| 9 | Teno | 34.8247 | 71.0918 | Chernozem | 63.8 | 0.1 | 0.4 | 322.6 | 2082.8 |
| 10 | Bramaderos | 35.6133 | 71.3146 | Andosol | 45.4 | 0.1 | 0.2 | 1057.4 | 1113.5 |
| 11 | Arauco | 37.2534 | 73.2669 | Cambisol | 92.4 | 0.4 | 0.8 | 566.4 | 2271.9 |
| 12 | Collipulli | 37.9206 | 72.4872 | Luvisol | 41.7 | 0.1 | 0.2 | 361.6 | 3458.3 |
| 13 | Puerto Saavedra | 38.7741 | 73.3897 | Gleysol | 250.4 | 0.7 | 0.6 | 2330.8 | 5731.6 |
| 14 | Melipeuco | 38.8744 | 71.8555 | Andosol | 143.9 | 0.4 | 1.7 | 3072.3 | 2651.1 |
| 15 | Choshuenco | 39.8594 | 72.1112 | Cambisol | 167.9 | 0.6 | 2.8 | 2659.3 | 4763.3 |
| 16 | Hueicoya | 39.9325 | 73.4128 | Acrisol | 124.6 | 0.6 | 2.3 | 948.5 | 2248.7 |
| 17 | Bahia Mansa | 40.6089 | 73.7489 | Cambisol | 60.1 | 0.2 | 1.4 | 1005.0 | 4031.5 |
| 18 | Corte Alto | 40.9034 | 73.1540 | Andosol | 158.5 | 0.5 | 1.7 | 3097.2 | 5261.0 |
| 19 | Chapo | 41.4271 | 72.6478 | Andosol | 161.6 | 0.7 | 3.5 | 1527.0 | 1792.2 |
| 20 | Chiloe Norte | 42.0530 | 73.7991 | Andosol | 89.4 | 0.4 | 1.8 | 1726.1 | 5675.6 |
| 21 | Pachabrán | 42.4219 | 73.8243 | Andosol | 180.9 | 0.8 | 2.4 | 2737.7 | 5077.3 |
| 22 | Aitui | 43.0579 | 73.6171 | Andosol | 302.4 | 1.7 | 6.8 | 3276.7 | 4900.4 |
| 23 | La Junta | 43.9639 | 72.3965 | Andosol | 246.6 | 1.5 | 6.3 | 3636.5 | 5286.0 |
| 24 | La Tapera | 44.6585 | 71.7844 | Andosol | 162.2 | 0.5 | 1.5 | 3881.3 | 2140.7 |
| 25 | Puerto Cisnes | 44.7569 | 72.6146 | Andosol | 415.2 | 2.4 | 7.5 | 2483.0 | 10332.9 |
| 26 | Puerto Aysen | 44.9020 | 72.6246 | Andosol | 348.4 | 2.2 | 6.1 | 2497.5 | 4545.5 |
| 27 | Simpson | 45.7864 | 72.9076 | Cambisol | 185.7 | 0.5 | 1.2 | 4267.0 | 3103.7 |
| 28 | Cerro Castillo | 46.1076 | 72.0736 | Andosol | 47.5 | 0.1 | 0.4 | 372.9 | 2988.2 |
| 29 | Bahia Exploradores | 46.4954 | 73.1277 | Andosol | 69.4 | 0.3 | 1.9 | 1796.1 | 5918.2 |
| 30 | Puerto Sánchez | 46.5682 | 72.6128 | Andosol | 159.1 | 0.8 | 1.9 | 2682.9 | 1910.8 |
| 31 | Cochrane | 47.3235 | 72.6455 | Cambisol | 152.8 | 0.5 | 2.2 | 2264.3 | 2582.0 |
| 32 | Caleta Tortel | 47.7915 | 73.5179 | Gleysol | 6.7 | 0.1 | 0.5 | 341.6 | 4249.6 |
| 33 | Ultima Esperanza | 51.8070 | 72.1649 | Leptosol | 139.8 | 0.6 | 1.9 | 2955.2 | 5375.7 |
| 34 | Porvenir | 53.3148 | 70.3611 | Leptosol | 463.6 | 1.8 | 4.0 | 9514.8 | 4022.7 |

**Table S4-5 cont.**

| Site ID | Site name | CUE | MCN | Spec. growth | Spec. resp. | Spec. hydrolases | Spec. oxidases | Residuals | Residuals | Residuals |
|---|---|---|---|---|---|---|---|---|---|---|
| | | - | - | % of MBC $d^{-1}$ | % of MBC $d^{-1}$ | nmol MUF $\mu g^{-1}$ MBC $h^{-1}$ | nmol DOPA $\mu g^{-1}$ MBC $h^{-1}$ | Spec. resp. ~ RC3 Texture % of MBC d-1 | Spec. growth ~ RC3 Texture % of MBC d-1 | Spec. growth ~ Spec. resp. % of MBC d-1 |
| 1 | Los Vilos | 0.60 | 11.1 | 8.8 | 5.9 | 12.0 | 21.1 | -12.5 | 2.2 | 4.5 |
| 2 | Catapilco | 0.52 | 26.3 | 8.2 | 7.6 | 24.0 | 59.1 | -8.7 | 2.8 | 3.8 |
| 4 | Bochinche | 0.59 | 7.6 | 10.9 | 7.5 | 13.0 | 28.4 | -8.7 | 5.5 | 6.6 |
| 5 | Clarillo | 0.38 | 7.3 | 6.1 | 10.0 | 21.8 | 33.3 | -6.9 | 0.4 | 1.6 |
| 6 | Matanzas | 0.31 | 14.5 | 3.6 | 8.1 | 27.4 | 57.5 | -5.7 | -0.5 | -0.8 |
| 7 | Pudahuel | 0.43 | 21.0 | 6.1 | 8.0 | 16.1 | 22.5 | -8.1 | 0.7 | 1.7 |
| 8 | Pimpinela | 0.43 | 13.2 | 6.4 | 8.4 | 13.8 | 22.8 | -5.7 | 2.2 | 2.0 |
| 9 | Teno | 0.32 | 6.9 | 2.4 | 5.2 | 5.1 | 32.6 | -7.7 | -1.2 | -1.8 |
| 10 | Bramaderos | 0.35 | 10.1 | 2.3 | 4.3 | 11.3 | 24.5 | -6.7 | -0.2 | -1.8 |
| 11 | Arauco | 0.28 | 5.3 | 5.2 | 13.3 | 6.1 | 24.6 | -3.3 | -0.4 | 0.4 |
| 12 | Colipulli | 0.16 | 39.4 | 1.4 | 7.5 | 8.7 | 82.9 | -4.6 | -1.8 | -3.0 |
| 13 | Puerto Saavedra | 0.53 | 15.6 | 2.9 | 2.5 | 17.7 | 22.9 | -10.1 | -0.6 | -1.1 |
| 14 | Melipeuco | 0.21 | 11.4 | 3.0 | 11.7 | 21.4 | 18.4 | -5.0 | -2.6 | -1.6 |
| 15 | Choshuenco | 0.20 | 8.9 | 4.0 | 15.6 | 15.8 | 28.4 | 1.5 | -0.3 | -1.0 |
| 16 | Hueicoya | 0.21 | 9.7 | 5.3 | 19.6 | 7.6 | 18.0 | 3.8 | 0.2 | 0.1 |
| 17 | Bahia Mansa | 0.12 | 35.0 | 3.1 | 21.7 | 16.7 | 67.1 | 6.9 | -1.5 | -2.3 |
| 18 | Corte Alto | 0.21 | 12.4 | 3.4 | 12.5 | 19.5 | 33.2 | -4.8 | -2.6 | -1.3 |
| 19 | Chapo | 0.20 | 15.2 | 4.9 | 19.4 | 9.4 | 11.1 | 3.8 | -0.1 | -0.3 |
| 20 | Chiloe Norte | 0.17 | 17.7 | 4.5 | 22.4 | 19.3 | 63.5 | 6.4 | -0.7 | -1.0 |
| 21 | Pachabrán | 0.24 | 13.2 | 4.6 | 14.4 | 15.1 | 28.1 | -1.5 | -0.6 | -0.3 |
| 22 | Altui | 0.21 | 12.2 | 5.9 | 22.5 | 10.8 | 16.2 | 7.4 | 1.1 | 0.4 |
| 23 | La Junta | 0.18 | 11.9 | 6.4 | 29.9 | 14.7 | 21.4 | 13.4 | 0.9 | 0.4 |
| 24 | La Tapera | 0.18 | 9.2 | 3.1 | 14.3 | 23.9 | 13.2 | -2.1 | -2.4 | -1.8 |
| 25 | Puerto Cisnes | 0.25 | 10.7 | 6.2 | 18.8 | 6.0 | 24.9 | 2.7 | 0.9 | 1.0 |
| 26 | Puerto Aysen | 0.26 | 13.6 | 6.7 | 19.3 | 7.2 | 13.0 | 5.4 | 2.5 | 1.5 |
| 27 | Simpson | 0.29 | 6.3 | 2.7 | 6.7 | 23.0 | 16.7 | -8.4 | -2.1 | -1.6 |
| 28 | Cerro Castillo | 0.17 | 4.8 | 2.6 | 12.5 | 7.9 | 62.9 | -3.7 | -2.8 | -2.1 |
| 29 | Bahia Exploradores | 0.10 | 18.2 | 4.6 | 39.8 | 16.6 | 85.3 | 23.3 | -0.8 | -2.1 |
| 30 | Puerto Sánchez | 0.24 | 6.1 | 5.3 | 16.5 | 16.9 | 12.0 | -1.8 | -1.2 | 0.3 |
| 31 | Cochrane | 0.21 | 6.0 | 3.4 | 13.2 | 14.8 | 16.9 | -2.1 | -1.5 | -1.4 |
| 32 | Caleta Tortel | 0.14 | 7.6 | 10.6 | 68.1 | 51.0 | 634.3 | 53.4 | 6.1 | 1.8 |
| 33 | Ultima Esperanza | 0.29 | 6.8 | 4.4 | 10.6 | 21.1 | 38.5 | -4.5 | -0.4 | -0.2 |
| 34 | Porvenir | 0.28 | 8.7 | 4.3 | 11.0 | 20.5 | 8.7 | -5.6 | -1.3 | -0.3 |

**Table S4-6.** Detailed model description of all 99 models, sorted by dependent variable and predictor sets. MBC = microbial biomass carbon, CUE = carbon use efficiency, MCN = microbial C:N ratio, Spec. = specific, ENV = environmental dataset, MIC = microbial community dataset, ENVxMIC = environmental and microbial community datasets combined, SW = stepwise regression, LA = lasso regression, RF = random forest, Transf. = log-transformed, D.f. = degrees of freedom, SW = p-value of Shapiro-Wilk test, BP = p-value of Breusch-Pagan test.

| | | | All | | | Stepwise & Lasso | | | | | Lasso | Random Forest | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Predictor Set | Model | Transf. | $R^2$ | RMSE | adj. $R^2$ | D.f. | p-value | SW | BP | lambda | mtry | ntree | nodesize |
| MBC | ENV | SW | yes | 0.52 | 91.92 | 0.49 | 30 | 0.00 | 0.32 | 0.09 | | | | |
| MBC | ENV | LA | yes | 0.52 | 0.67 | 0.51 | 30 | 0.00 | 0.01 | 0.22 | 0.3 | | | |
| MBC | ENV | RF | no | 0.41 | 85.1 | | | | | | | 8 | 100 | 15 |
| MBC | MIC | SW | yes | 0.24 | 106.59 | 0.19 | 30 | 0.02 | 0.37 | 0.05 | | | | |
| MBC | MIC | LA | | | | | 30 | 0.00 | | | 100 | | | |
| MBC | MIC | RF | no | 0.07 | 105.85 | | | | | | | 1 | 100 | 15 |
| MBC | ENVxMIC | SW | yes | 0.68 | 82.53 | 0.63 | 28 | 0.00 | 0.34 | 0.91 | | | | |
| MBC | ENVxMIC | LA | yes | 0.5 | 0.67 | 0.49 | 30 | 0.00 | 0.01 | 0.22 | 0.3 | | | |
| MBC | ENVxMIC | RF | no | 0.34 | 90.74 | | | | | | | 19 | 1000 | 7 |
| Growth | ENV | SW | yes | 0.58 | 0.42 | 0.55 | 30 | 0.00 | 0.97 | 0.01 | | | | |
| Growth | ENV | LA | yes | 0.49 | 0.57 | 0.39 | 30 | 0.00 | 0.72 | 0.05 | 0.1 | | | |
| Growth | ENV | RF | no | 0.56 | 0.41 | | | | | | | 8 | 100 | 15 |
| Growth | MIC | SW | yes | 0.2 | 0.59 | 0.15 | 30 | 0.03 | 0.78 | 0.64 | | | | |
| Growth | MIC | LA | | | | | 30 | 0.00 | | | 100 | | | |
| Growth | MIC | RF | yes | 0.07 | 0.92 | | | | | | | 1 | 100 | 2 |
| Growth | ENVxMIC | SW | no | 0.65 | 0.34 | 0.63 | 30 | 0.00 | 0.20 | 0.10 | | | | |
| Growth | ENVxMIC | LA | yes | 0.48 | 0.52 | 0.36 | 30 | 0.00 | 0.55 | 0.09 | 0.1 | | | |
| Growth | ENVxMIC | RF | no | 0.44 | 0.47 | | | | | | | 19 | 1000 | 2 |
| Respiration | ENV | SW | no | 0.85 | 0.77 | 0.83 | 28 | 0.00 | 0.35 | 0.19 | | | | |
| Respiration | ENV | LA | yes | 0.67 | 0.49 | 0.61 | 30 | 0.00 | 0.27 | 0.25 | 0.1 | | | |
| Respiration | ENV | RF | no | 0.73 | 1.01 | | | | | | | 8 | 1000 | 2 |
| Respiration | MIC | SW | yes | 0.67 | 1.6 | 0.62 | 28 | 0.00 | 0.62 | 0.49 | | | | |
| Respiration | MIC | LA | yes | 0.48 | 0.6 | 0.4 | 30 | 0.00 | 0.91 | 0.49 | 0.1 | | | |
| Respiration | MIC | RF | yes | 0.42 | 0.8 | | | | | | | 9 | 100 | 15 |
| Respiration | ENVxMIC | SW | no | 0.82 | 0.83 | 0.81 | 30 | 0.00 | 0.45 | 0.68 | | | | |
| Respiration | ENVxMIC | LA | no | 0.76 | 0.78 | 0.67 | 30 | 0.00 | 0.91 | 0.10 | 0.2 | | | |
| Respiration | ENVxMIC | RF | no | 0.67 | 1.17 | | | | | | | 18 | 100 | 2 |
| Hydrolases | ENV | SW | yes | 0.34 | 1712.27 | 0.3 | 30 | 0.00 | 0.83 | 0.24 | | | | |
| Hydrolases | ENV | LA | yes | 0.26 | 0.7 | 0.18 | 30 | 0.00 | 0.39 | 0.14 | 0.2 | | | |
| Hydrolases | ENV | RF | yes | 0.26 | 0.69 | | | | | | | 2 | 100 | 7 |
| Hydrolases | MIC | SW | yes | 0.22 | 1774.61 | 0.17 | 30 | 0.02 | 0.97 | 0.87 | | | | |
| Hydrolases | MIC | LA | | | | | 30 | 0.06 | | | 100 | | | |
| Hydrolases | MIC | RF | yes | 0.1 | 0.78 | | | | | | | 2 | 100 | 2 |
| Hydrolases | ENVxMIC | SW | yes | 0.35 | 1716.49 | 0.31 | 30 | 0.00 | 0.81 | 0.49 | | | | |
| Hydrolases | ENVxMIC | LA | yes | 0.24 | 0.74 | 0.22 | 30 | 0.00 | 0.16 | 0.28 | 0.3 | | | |
| Hydrolases | ENVxMIC | RF | yes | 0.17 | 0.72 | | | | | | | 3 | 100 | 15 |
| Oxidases | ENV | SW | no | 0.36 | 1455.76 | 0.32 | 30 | 0.00 | 0.46 | 0.08 | | | | |
| Oxidases | ENV | LA | no | 0.17 | 1403.5 | -0.02 | 30 | 0.02 | 0.53 | 0.18 | 100 | | | |
| Oxidases | ENV | RF | no | 0.13 | 1882.72 | | | | | | | 1 | 100 | 15 |
| Oxidases | MIC | SW | no | 0.28 | 1547.52 | 0.23 | 30 | 0.01 | 0.09 | 0.16 | | | | |
| Oxidases | MIC | LA | yes | 0.23 | 0.44 | 0.18 | 30 | 0.01 | 0.88 | 0.91 | 0.1 | | | |
| Oxidases | MIC | RF | no | 0.07 | 1969.32 | | | | | | | 1 | 100 | 7 |
| Oxidases | ENVxMIC | SW | no | 0.36 | 1455.76 | 0.32 | 30 | 0.00 | 0.46 | 0.08 | | | | |
| Oxidases | ENVxMIC | LA | no | 0.14 | 1301.24 | -0.46 | 30 | 0.13 | 0.56 | 0.66 | 100 | | | |
| Oxidases | ENVxMIC | RF | no | 0.09 | 1938.73 | | | | | | | 2 | 100 | 2 |
| CUE | ENV | SW | yes | 0.48 | 0.11 | 0.43 | 29 | 0.00 | 0.61 | 0.67 | | | | |
| CUE | ENV | LA | yes | 0.33 | 0.36 | 0.26 | 30 | 0.00 | 0.27 | 0.67 | 0.1 | | | |
| CUE | ENV | RF | yes | 0.33 | 0.4 | | | | | | | 1 | 1000 | 2 |
| CUE | MIC | SW | no | 0.7 | 0.07 | 0.68 | 30 | 0.00 | 0.32 | 0.12 | | | | |
| CUE | MIC | LA | no | 0.46 | 0.06 | 0.18 | 30 | 0.00 | 0.68 | 0.58 | 0 | | | |
| CUE | MIC | RF | no | 0.67 | 0.08 | | | | | | | 10 | 1000 | 15 |
| CUE | ENVxMIC | SW | no | 0.7 | 0.07 | 0.68 | 30 | 0.00 | 0.32 | 0.12 | | | | |
| CUE | ENVxMIC | LA | yes | 0.58 | 0.28 | 0.55 | 30 | 0.00 | 0.38 | 0.82 | 0.1 | | | |
| CUE | ENVxMIC | RF | no | 0.66 | 0.08 | | | | | | | 19 | 100 | 15 |
| MCN | ENV | SW | yes | 0.31 | 7.05 | 0.24 | 29 | 0.01 | 0.96 | 0.01 | | | | |
| MCN | ENV | LA | | | | | 30 | 0.00 | | | 100 | | | |
| MCN | ENV | RF | no | 0.09 | 7.8 | | | | | | | 1 | 100 | 2 |
| MCN | MIC | SW | yes | 0.16 | 7.92 | 0.11 | 30 | 0.07 | 0.13 | 0.84 | | | | |
| MCN | MIC | LA | | | | | 30 | 0.00 | | | 100 | | | |
| MCN | MIC | RF | yes | 0.09 | 0.49 | | | | | | | 1 | 100 | 15 |
| MCN | ENVxMIC | SW | yes | 0.22 | 7.75 | 0.17 | 30 | 0.02 | 0.20 | 0.05 | | | | |
| MCN | ENVxMIC | LA | | | | | 30 | 0.00 | | | 100 | | | |
| MCN | ENVxMIC | RF | yes | 0.08 | 0.49 | | | | | | | 1 | 100 | 7 |
| Spec. growth | ENV | SW | yes | 0.37 | 2.05 | 0.33 | 30 | 0.00 | 0.07 | 0.50 | | | | |
| Spec. growth | ENV | LA | yes | 0.26 | 0.38 | 0.18 | 30 | 0.00 | 0.95 | 0.30 | 0.1 | | | |
| Spec. growth | ENV | RF | yes | 0.24 | 0.41 | | | | | | | 1 | 100 | 7 |
| Spec. growth | MIC | SW | no | 0.48 | 1.62 | 0.44 | 30 | 0.00 | 0.71 | 0.48 | | | | |
| Spec. growth | MIC | LA | yes | 0.29 | 0.37 | 0.16 | 30 | 0.00 | 0.83 | 0.04 | 0.1 | | | |
| Spec. growth | MIC | RF | yes | 0.45 | 0.35 | | | | | | | 9 | 100 | 7 |
| Spec. growth | ENVxMIC | SW | yes | 0.55 | 1.77 | 0.52 | 30 | 0.00 | 0.41 | 0.62 | | | | |
| Spec. growth | ENVxMIC | LA | yes | 0.4 | 0.32 | 0.29 | 30 | 0.00 | 0.74 | 0.33 | 0.1 | | | |
| Spec. growth | ENVxMIC | RF | yes | 0.47 | 0.35 | | | | | | | 16 | 100 | 2 |
| Spec. respiration | ENV | SW | yes | 0.37 | 12.23 | 0.27 | 28 | 0.01 | 0.03 | 0.07 | | | | |
| Spec. respiration | ENV | LA | yes | 0.23 | 0.51 | 0.12 | 30 | 0.00 | 0.33 | 0.16 | 0.1 | | | |
| Spec. respiration | ENV | RF | yes | 0.37 | 0.57 | | | | | | | 4 | 1000 | 7 |
| Spec. respiration | MIC | SW | yes | 0.82 | 5.51 | 0.78 | 27 | 0.00 | 0.05 | 0.28 | | | | |
| Spec. respiration | MIC | LA | yes | 0.69 | 0.36 | 0.63 | 30 | 0.00 | 0.00 | 0.28 | 0.1 | | | |
| Spec. respiration | MIC | RF | yes | 0.69 | 0.41 | | | | | | | 9 | 1000 | 2 |
| Spec. respiration | ENVxMIC | SW | yes | 0.68 | 9.69 | 0.66 | 30 | 0.00 | 0.08 | 0.39 | | | | |
| Spec. respiration | ENVxMIC | LA | yes | 0.67 | 0.34 | 0.57 | 30 | 0.00 | 0.01 | 0.63 | 0.1 | | | |
| Spec. respiration | ENVxMIC | RF | yes | 0.67 | 0.43 | | | | | | | 18 | 100 | 2 |
| Spec. hydrolases | ENV | SW | yes | 0.14 | 8.17 | 0.09 | 30 | 0.10 | 0.23 | 0.18 | | | | |
| Spec. hydrolases | ENV | LA | | | | | 30 | 0.00 | | | 100 | | | |
| Spec. hydrolases | ENV | RF | no | 0.14 | 9.23 | | | | | | | 1 | 1000 | 15 |
| Spec. hydrolases | MIC | SW | yes | 0.33 | 7.43 | 0.29 | 30 | 0.00 | 0.02 | 0.08 | | | | |
| Spec. hydrolases | MIC | LA | yes | 0.22 | 0.42 | 0.13 | 30 | 0.00 | 0.22 | 0.30 | 0.1 | | | |
| Spec. hydrolases | MIC | RF | no | 0.07 | 8.89 | | | | | | | 1 | 1000 | 15 |
| Spec. hydrolases | ENVxMIC | SW | yes | 0.33 | 7.43 | 0.29 | 30 | 0.00 | 0.02 | 0.08 | | | | |
| Spec. hydrolases | ENVxMIC | LA | | | | | 30 | 0.00 | | | 100 | | | |
| Spec. hydrolases | ENVxMIC | RF | no | 0.16 | 8.99 | | | | | | | 1 | 1000 | 15 |
| Spec. oxidases | ENV | SW | yes | 0.27 | 104.01 | 0.22 | 30 | 0.01 | 0.12 | 0.01 | | | | |
| Spec. oxidases | ENV | LA | | | | | 30 | 0.00 | | | 100 | | | |
| Spec. oxidases | ENV | RF | no | 0.29 | 90.49 | | | | | | | 1 | 1000 | 2 |
| Spec. oxidases | MIC | SW | yes | 0.26 | 101.17 | 0.21 | 30 | 0.01 | 0.02 | 0.26 | | | | |
| Spec. oxidases | MIC | LA | | | | | 30 | 0.00 | | | 100 | | | |
| Spec. oxidases | MIC | RF | yes | 0.11 | 0.85 | | | | | | | 1 | 100 | 2 |
| Spec. oxidases | ENVxMIC | SW | yes | 0.32 | 99.37 | 0.27 | 30 | 0.00 | 0.62 | 0.03 | | | | |
| Spec. oxidases | ENVxMIC | LA | | | | | 30 | 0.00 | | | 100 | | | |
| Spec. oxidases | ENVxMIC | RF | no | 0.22 | 90.19 | | | | | | | 1 | 100 | 2 |

## 7.5 Appendix Chapter 5

**Credits for Figure 5-3.** Several icons in Figure 5-3 were taken from thenounproject.com. The following artists have created these icons: grass: Hero Arts; forest: ida ratnaningrum; plowing: Symbolon; snowflake: Jordy Madueño; ripples: cindy clegane; heat: Adrien Coquet; hourglass: yusuf kara.

# 8. List of Figures

# 9. List of Tables