



## Review

## Soil metabolomics: Deciphering underground metabolic webs in terrestrial ecosystems



Yang Song <sup>a,b,\*</sup>, Shi Yao <sup>a,b</sup>, Xiaona Li <sup>c</sup>, Tao Wang <sup>d</sup>, Xin Jiang <sup>a,b</sup>, Nanthi Bolan <sup>e,f,g</sup>, Charles R. Warren <sup>h</sup>, Trent R. Northen <sup>i,j</sup>, Scott X. Chang <sup>k,\*</sup>

<sup>a</sup> State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China

<sup>b</sup> University of Chinese Academy of Sciences, Beijing 100049, China

<sup>c</sup> School of Environment and Ecology, Jiangnan University, Wuxi 225127, China

<sup>d</sup> Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610299, China

<sup>e</sup> School of Agriculture and Environment, The University of Western Australia, Nedland, WA-6009, Australia

<sup>f</sup> The UWA Institute of Agriculture, The University of Western Australia, Nedland, WA-6009, Australia

<sup>g</sup> Healthy Environments and Lives (HEAL) National Research Network, Australia

<sup>h</sup> School of Life and Environmental Sciences, University of Sydney, Heydon-Laurence Building A08, NSW 2006, Australia

<sup>i</sup> Environmental Genomics and System Biology Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, CA 94720, USA

<sup>j</sup> Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

<sup>k</sup> Department of Renewable Resources, University of Alberta, Edmonton, Alberta T6G 2E3, Canada

## ARTICLE INFO

## ABSTRACT

## Keywords:

Dissolved organic matter  
Carbon cycling  
Metabolomes  
Extraction method  
Soil microbiome  
Rhizosphere ecology

Soil metabolomics is an emerging approach for profiling diverse small molecule metabolites, i.e., metabolomes, in the soil. Soil metabolites, including fatty acids, amino acids, lipids, organic acids, sugars, and volatile organic compounds, often contain essential nutrients such as nitrogen, phosphorus, and sulfur and are directly linked to soil biogeochemical cycles driven by soil microorganisms. This paper presents an overview of methods for analyzing soil metabolites and the state-of-the-art of soil metabolomics in relation to soil nutrient cycling. We describe important applications of metabolomics in studying soil carbon cycling and sequestration, and the response of soil organic pools to changing environmental conditions. This includes using metabolomics to provide new insights into the close relationships between soil microbiome and metabolome, as well as responses of soil metabolome to plant and environmental stresses such as soil contamination. We also highlight the advantage of using soil metabolomics to study the biogeochemical cycles of elements and suggest that future research needs to better understand factors driving soil function and health.

## 1. Introduction

The soil is the largest carbon (C) pool in terrestrial ecosystems on Earth [1]. Soil organic matter (SOM), where the soil organic C is stored, is the essential substrate that influences various microbially mediated biogeochemical processes and maintains soil function and health [1,2]. The composition of SOM represents a continuum of progressively decomposing organic compounds and a wide range of organic compounds produced through chemical reactions, microbial synthesis, and root exudation [1,3]. Chemical characterization and fingerprinting of the molecular composition of SOM have always been a research interest for soil and environmental scientists, as SOM makes up the most complex and important substrates in the environment [4]. The SOM is composed

of a large number of small molecules derived from plants and microbes, with their abundance and composition being some of the most important mediators of soil health and plant productivity [2,5,6]. Dissolved organic matter (DOM), the water-extractable organic matter fraction in the soil solution that passes through a 0.45 µm filter, constitutes less than 2% of the total SOM [7–9]. However, DOM is the most actively cycling organic matter fraction and influences most of the critical soil biogeochemical cycling processes of major elements, including C, nitrogen (N), phosphorus (P), and sulfur (S), and the conversion and mineralization of SOM [7]. Hence, the metabolism of DOM by soil microbes is a key determinant of soil organic C residence time, and approaches that decrease microbial mineralization could be used for soil C sequestration. In addition, understanding the composition of organic compounds in DOM can improve

\* Corresponding authors.

E-mail addresses: [ysong@issas.ac.cn](mailto:ysong@issas.ac.cn) (Y. Song), [scott.chang@ualberta.ca](mailto:scott.chang@ualberta.ca) (S.X. Chang).

our knowledge of soil ecosystem services, e.g., C stabilization and transformation, plant growth and food productivity, contamination remediation, and climate change regulation (Fig. 1).

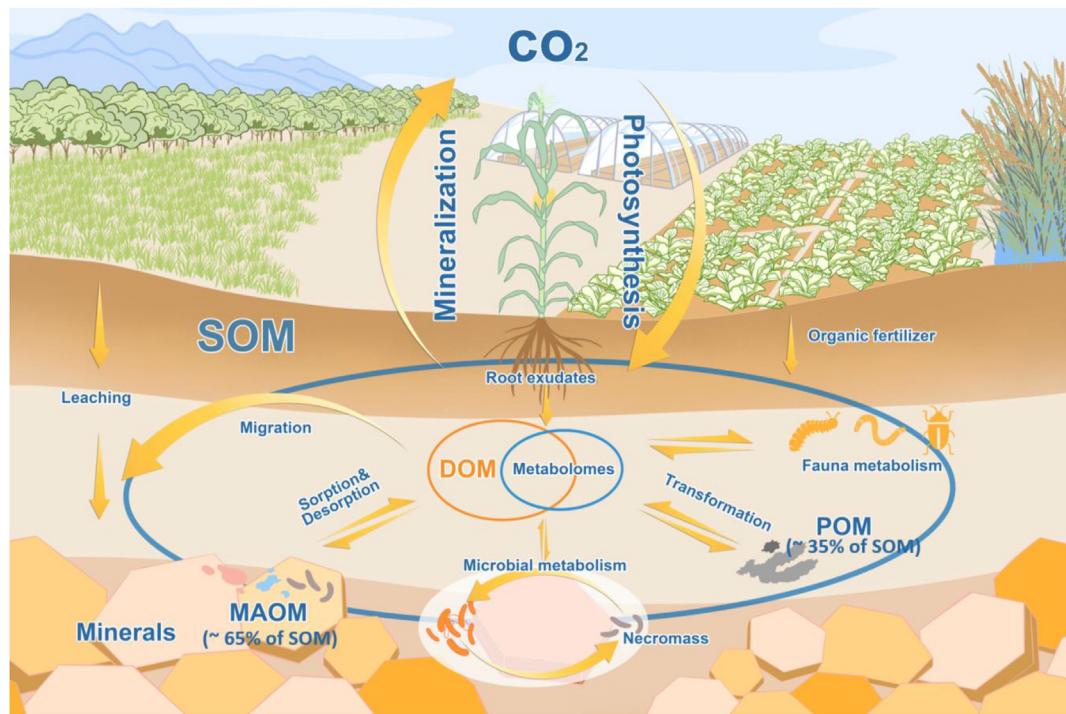
Metabolites are an important component of DOM and can be a direct measure of the biochemical activity of the soil biota. Much biochemical activity happens at the level of individual metabolites, so those compounds are the best measure of biochemical activities [11]. Similarly, the responses of metabolites (at the molecular level) to environmental stress are more sensitive than with traditional indicators such as microbial biomass C (MBC) [12]. Metabolomics, the global profiling of small molecular mass compounds in a sample [13–16], can often detect thousands of metabolite features within a single run (though the number identified is typically a small fraction of this). Metabolomics has found a wide range of applications in the field of environmental science. For example, it has been applied in aquatic ecosystems, including marine and freshwater environments, to elucidate relationships between DOM and microbial communities [17,18]. Soil metabolomics, which focuses on the study of metabolites in the soil, is a relatively new subdiscipline in soil science; however, it is gaining popularity given its ability to provide molecular-level information on C and nutrient cycling and its relationship with soil communities [19,20]. Soil metabolomics has been used in various studies for assessing soil quality and function, including but not limited to studying the biogeochemical cycles of soil C and nutrients, the relationship between metabolites and soil microbes and plants, as well as the responses of soil microbes to environmental stresses and climate change.

Soils contain various compounds produced by plants, microorganisms, and fauna. The small molecule metabolites (<1,000 Da), whether as a direct input or degradation product, are typically rapidly cycled by soil microbial communities, placing metabolites at the heart of C cycling in many ways [21–23]. It is, therefore, important to first understand the range of metabolites found in soils to explore the application of soil metabolomics in soil biogeochemical dynamics. Common classes of soil metabolites include fatty acids, amino acids, lipids, sterols, alcohols, sugars, amino-sugars, sugar

alcohols, sugar acids, organic phosphates, aromatics, purines, and other organic acids, etc., as well as many yet-to-be-identified compounds [6,20,24–26]. Ascertaining the source of those compounds presents a big challenge.

The approaches used in soil metabolomics have evolved over the years, with shifts in applications signposted by varying definitions. Early soil metabolomics studies were performed to characterize DOM, e.g., dissolved organic C (DOC) and dissolved organic N (DON), in the soil [20,27,28]. The term “soil meta-metabolome” was proposed by Warren to describe the pool of metabolites directly related to soil processes, such as N and C cycling [27]. Moreover, the metabolic profiles of entire communities, including microbes and invertebrates, living in the soil could be analyzed by “community metabolomics” [29]. Swenson et al. proposed “untargeted soil metabolomics” for the first time and established the working procedure for the measurement of soil metabolites [24]. Then, “exometabolomics” was developed to study the microbial extracellular metabolites in soil science [30,31]. Recently, “environmental metabolic footprinting” has been used to analyze soil metabolomes’ response to stresses caused by contaminants [32], and “untargeted rhizosphere metabolomics” has been proposed to analyze entire soil metabolomes [33].

The studies mentioned above constitute the main milestones in the development of soil metabolomics, including the definition, methodology, and interpretation of soil metabolomics. The timelines of these pioneer studies indicate that soil metabolomics developed quickly, from soil metabolite extraction methods and data analysis to the application of metabolomics in various research areas. To the best of our knowledge, there is no comprehensive review focusing on the research progress of soil metabolomics and its applications. Therefore, this paper aims to address the following three aspects: 1) methods for effectively extracting and identifying soil metabolites; 2) new insights into soil biogeochemical cycling gained from soil metabolomics; and 3) emerging research directions in soil metabolomics. At the end of the review, we provide some perspectives on opportunities to greatly improve the performance and impact of soil metabolomics.



**Fig. 1.** A conceptual diagram showing how soil metabolomics can enhance our understanding of terrestrial carbon cycling. The percentages of MAOM and POM to SOM can be referred to [10]. SOM, soil organic matter; DOM, dissolved organic matter; MAOM, mineral-associated organic matter; POM, particulate organic matter.

## 2. Methodologies in soil metabolomics

Generally, soil metabolomic analysis includes the extraction of metabolites from the soil matrix, metabolite characterization using various analytical instruments, and data analysis and interpretation.

### 2.1. Extraction of soil metabolites

Soils are a highly complex mixture of minerals, biopolymers, biomass, and small molecule substances, often in aggregated structures [34]. Extensive literature exists on the extraction of SOM and DOM/DOC. Soil organic matter extraction methods and the subsequent measurement of extractable SOM can be benchmarked against the total organic C determined through combustion. Common extraction methods include treating soils with potassium sulfate, sodium pyrophosphate, sodium borate, hydroxylamine, hydrochloric acid, and so on, to desorb and dissolve soil organics. Even extreme methods, such as treating the soil with hydrofluoric acid to dissolve minerals, often do not release all of the organic C associated with minerals [35–37]. Critically, all of these salts and acids can interfere with the characterization of soil metabolites when using mass spectrometry. Moreover, some of these procedures (e.g., extraction with sulfuric acid) will change the composition of metabolites, for example, by hydrolyzing biopolymers to form small molecule metabolites [38].

Given these considerations, most metabolomic methods have used chemically “gentle” approaches similar to those used in biomedical and biotechnological applications, especially when using inorganic and/or organic solvents. Aqueous extraction methods have the advantage of extracting the fraction of DOM/DOC that is most microbially accessible. Some of the common extraction methods used in soil metabolomics are summarized in Table 1.

Soil metabolites include both extracellular and intracellular organic products. Initially, the protocol used for determining total microbial biomass was adopted for extracting metabolites [60]; in this method, intracellular metabolites are extracted by fumigating with ethanol-free chloroform to lyse microbial cells [5]. Accordingly, fumigation of the soil with chloroform significantly increases the abundance of some metabolites. In one study, 2'-deoxycytidine, proline, and thymine were detected only in the fumigated compared to unfumigated soil [24]. By lysing the microbial cells, the fumigation method, in principle, creates a “whole soil extract” that contains both the extracellular and intracellular metabolites [5]. Subsequent studies suggested that chloroform fumigation can create artifacts and result in altered concentrations of some metabolites due to their oxidation via exoenzymes that remain active [61]. Direct extraction with an organic solvent or a water-solvent emulsion avoids artefacts associated with prolonged chloroform fumigation [61,62]. A promising alternative is using solvent-water systems that inhibit microbial activity. For example, low molecular weight organics were several times more abundant in extracts from soil extracted with 0.5 M K<sub>2</sub>SO<sub>4</sub> that contained 1.5% CHCl<sub>3</sub> than those extracted with either 0.5 M K<sub>2</sub>SO<sub>4</sub> or H<sub>2</sub>O [39]. This is presumably due to both improved solubility of target metabolites and cell lysis, resulting in extracts that contain both the extracellular and intracellular metabolites [24]. Common organic solvents include methanol, isopropanol, ethyl acetate, and chloroform, to name a few. In another study, methanol-chloroform-water (first methanol:chloroform = 2:1, and then chloroform:water = 1:1, i.e., the Bligh-Dyer 2-phase extraction procedure) was used to extract polar and nonpolar metabolites, with the advantage of efficiently extracting and separating lipid and water-soluble metabolites [29]. However, it should be noted that some metabolites are likely partitioned in both phases, which can introduce an additional source of experimental error.

As reported in one comparison study, although the abundance (i.e., intensity) of some extracted compounds varied among different extractants, including water, dilute salts (10 mM K<sub>2</sub>SO<sub>4</sub> or 10 mM NH<sub>4</sub>HCO<sub>3</sub>), as well as organic solvent mixtures such as isopropanol/methanol/water (3:3:2 v/v/v) and 10%, 25%, 50%, 100% methanol in water, the total

number of compounds (i.e., diversity) detected in each extraction was similar [24]. Buyer et al. refined the Bligh-Dyer 2-phase method through the inclusion of a volatile salt, i.e., ammonium bicarbonate, to improve the desorption of metabolites from soil minerals and to avoid the interference of the phosphate or citrate used in the traditional Bligh-Dyer 2-phase extraction [25]. As discussed above, salts can be included as part of the extractant to help desorb organics from mineral surfaces; however, high salt concentrations can retain water, interfering with derivatization chemistries used with gas chromatography-mass spectrometry (GC-MS) and resulting in ion suppression during electrospray ionization, depending on their concentrations [19,24].

Generally, water is a good choice for extracting many extracellular metabolites derived from plants or microbes [40], providing information on the composition of the DOC pool with minimal cell lysis, assuming the soil has sufficient salt to avoid cell rupture from osmotic stress. However, organic solvents are preferred for extracting metabolites from microbial biomass and/or analyzing non-polar metabolites such as fatty acids and sterols. The acetonitrile:isopropanol:water solution at a 3:3:2 ratio was used for extracting soil metabolites [6]. A 50% methanol solution extracted more metabolite species from the rhizosphere soil than the water and 95% methanol solution [44]. Moreover, methanol in water (1:1) for extracting polar metabolites, and a 2:1 acetone:0.2 M HEPES pH 7.7 buffer for extracting nonpolar metabolites have also been proven effective for extracting metabolites from the soil [45]. The diverse extractants used in early studies (Table 1) make it difficult for early career researchers to determine which extractant to use when studying soil metabolomics. More researchers used the methanol/water mixture than other extractants (Table 1). More recently, a protocol was proposed to extract polar and nonpolar soil metabolites with water and a mixture of ethyl acetate:water (1:1), respectively [23]. This will provide guidance to those who are not familiar with the methodologies in soil metabolomics. A standard soil metabolite extraction protocol is the basis for comparing metabolomes from different soils and will be helpful for future global soil metabolite data analysis.

There is growing interest in profiling volatile metabolites in soils, and “the soil volatilome”, which includes compounds such as alcohols, ketones, organic acids, and so on, has been proposed [59,63]. Volatiles are typically captured by passing soil gas through a porous polymer resin based on 2,6-diphenylene oxide (Tenax TA) [45], or extracted by solid-phase micro extraction [59].

In addition to the different extractants used to extract metabolites, there is a diversity of extraction procedures used in soil metabolomics studies, including the single or combined use of sonication [20], shaking [39], vortexing [64], grinding [65], and crushing [66]. Generally, grinding in a partially stabilized zirconia ring and puck bowl, combined with extraction with sonication, breaks up the aggregates and cells, leading to the extraction of both extracellular and intracellular soil metabolites [20]. Compared with sonication and shaking following chloroform fumigation, shaking alone was recommended for untargeted DOM analysis with the smallest loss of soil enzymatic activity indicators [64]. Other less commonly used methods include centrifugation without extractants [27] and microdialysis [67].

Given the diversity of soil types, it is unlikely that there will ever be a single extraction method that has high performance for all soils. Broadly acceptable and cost-effective soil metabolome extraction methods that are suitable for the complete characterization of various components of soil metabolites need to be developed. For a particular soil, sequential chemical extraction methods may be developed to separate extracellular and intracellular metabolites or polar and nonpolar metabolites. For different types of soils, different extraction methods could be developed, similar to the different P extraction methods designed for acidic vs. alkaline soils.

### 2.2. Identification of soil metabolites

The most widely used techniques for profiling soil metabolites are mass spectrometry, such as GC-MS, liquid chromatography-mass

**Table 1**

The extractants and identification instruments used for soil metabolome extraction.

	Extractant	Method of identification	References	Traits
Inorganic solvent extraction	0.01 M and 0.5 M K <sub>2</sub> SO <sub>4</sub>	GC-MS CE-MS	[24,39]	<ul style="list-style-type: none"> <li>Suitable for polar metabolites</li> <li>Simple and reflective of actual microbe-accessible extracellular metabolites</li> <li>The salt content of the resultant samples complicates metabolite analysis because of salt crystal formation during dry-down and ion suppression</li> <li>The efficiency of salt-based extractants compared with water is dependent on the ionic strength of the solution and soil properties</li> </ul>
	0.01M NH <sub>4</sub> HCO <sub>3</sub>	GC-MS	[24]	
	Water	1H NMR GC-MS CE-MS LC-Q-TOF-MS	[24,28,39,40]	
Mixed organic-inorganic solvent extraction	Chloroform:0.01M K <sub>2</sub> SO <sub>4</sub> = 1:4	GC-MS	[39,41]	<ul style="list-style-type: none"> <li>Suitable for both polar and non-polar metabolites</li> <li>Extracting both intracellular and extracellular metabolites</li> <li>The polarity of the extractant determines the polarity of the analyzable metabolites</li> <li>Organic solvents efficiently extract fatty acids and sterols</li> </ul>
	Chloroform:0.5M K <sub>2</sub> SO <sub>4</sub> = 1.5:100	CE-MS		
	Methanol:chloroform:water = 2.5:1:1, 2:2:1, 2:1:1	1H NMR	[29]	
	Isopropanol:methanol:water = 3:3:2	GC-MS	[42]	
	Methanol:chloroform:phosphate buffer = 2:1:0.8	GC-MS	[43]	
	Methanol:0.05% formic acid = 1:1, 19:1	UPLC-Q-TOF-MS	[44]	
	Methanol:chloroform:0.25M NH <sub>4</sub> HCO <sub>3</sub> = 4:10:5	GC-MS	[25]	
	Acetone:HEPES buffer = 2:1	LC-Q-TOF-MS	[45]	
	Ethyl acetate:water = 1:1	LC-MS GC-MS	[23]	
	Acetonitrile:isopropanol:water = 3:3:2	GC-TOF-MS	[46]	
	Methanol:water = 1:1, 2:1, 3:1, 4:1, 9:1	1H NMR GC-MS HPLC-MS UPLC-Q-TOF-MS	[20,24,33,47, 48,49,50,51, 52,53]	
Organic solvent extraction	Ethyl acetate	LC-Q-MS	[32]	<ul style="list-style-type: none"> <li>Suitable for non-polar metabolites</li> <li>High frequencies of lipids, as well as amino acids and organic nitrogen compounds, can be detected</li> </ul>
	Methanol	GC-MS	[54,55]	
	Methanol:ethyl acetate = 3:1	UPLC-Orbitrap-MS		
	Methanol:dichloromethane:ethyl acetate:acetonitrile = 1:1:1:1	GC-TOF-MS	[56]	
	Methanol:acetonitrile = 1:1	LC-Q-Orbitrap-MS	[57]	
	Tenax: porous polymer resin based on 2,6-diphenylene oxide	GC-MS	[58]	
Solid extraction	Solid-phase microextraction	GC-MS	[45]	<ul style="list-style-type: none"> <li>Suitable for volatile metabolites</li> </ul>
		GC-MS	[59]	

GC-MS, gas chromatography-mass spectrometry; CE-MS, capillary electrophoresis-mass spectrometry; 1H NMR, H nuclear magnetic resonance spectra; LC-Q-TOF-MS, liquid chromatograph-quadrupole-time-of-flight-mass spectrometry; UPLC-Q-TOF-MS, ultra-performance-liquid chromatograph-quadrupole-time-of-flight-mass spectrometry; LC-MS, liquid chromatograph-mass spectrometry; GC-TOF-MS, gas chromatography-time-of-flight-mass spectrometry; UPLC-Orbitrap-MS, ultra-performance liquid chromatography-orbitrap-mass spectrometry, LC-Q-MS, liquid chromatograph-quadrupole-mass spectrometry; LC-Q-Orbitrap-MS, liquid chromatograph-quadrupole-orbitrap-mass spectrometry.

spectrometry (LC-MS), and nuclear magnetic resonance (NMR) (Table 1) [68]. The LC-MS and GC-MS techniques have an advantage over NMR methods in characterizing complex and relatively dilute mixtures in profiling soil metabolites [19,69]. However, NMR techniques can determine structures *de novo* and detect metabolites that do not ionize well by LC-MS or GC-MS [70].

The GC-MS technique is well suited for analyzing a wide range of metabolites, including volatiles, typically using electron ionization [59]. Generally, GC-MS has outstanding chromatographic resolution and produces extensive fragmentation information, which is easily compared with reference databases for metabolite identification [5]. However, many metabolites require derivatization prior to GC-MS analysis [43], the yield of which can be affected by water retained in salts [19]. The profile of soil metabolites can be highly correlated with the derivatization technique used, with derivatizations containing methoxyamination solution showing less variability [43]. Nevertheless, many molecules cannot be derivatized (e.g., due to steric hindrance), or are degraded during passage through the hot injection port and column, or have too low a vapor pressure [71]. GC-MS has excellent performance for metabolites that do not require derivatization or are readily derivatized, but the selectivity afforded by derivatization renders GC-MS unable to detect many other molecules.

The LC-MS technique is typically performed either using normal or reversed-phase chromatography [72]. Electrospray Ionization (ESI) is

used to generate gas phase ions for manipulation using electromagnetic fields, often achieving resolutions  $> 100,000$  full-width half max and mass accuracy  $< 5$  parts per million. Tandem mass spectrometry (MS/MS) is now commonly used to provide additional orthogonal information (in addition to accurate mass and retention time) on the chemical structures of the metabolites [73]. These fragmentation spectra are compared to standard databases such as the National Institute of Standards and Technology, Metabolite Link, and Global Natural Products Social Molecular Networking platform, to provide identifications including spectral relationships to known metabolites or other metabolites in the experiment [74].

Several other techniques are used for soil metabolite analysis. Capillary electrophoresis-mass spectrometry is another technique that has been used to study soil metabolites [27]. It affords exquisite separations of polar, charged metabolites but is not widely used [13]. Recently, proton transfer reaction time of flight mass spectrometry (PTR-TOF-MS) was used to study soil microbial volatile metabolites in response to snowmelt [75], showing a potential non-destructive approach to study the soil biogeochemical cycles. Finally, NMR is becoming increasingly sensitive and able to characterize more complex mixtures [70]. It has, for example, been used to study the response of soil metabolomes to contamination and amendments, such as biochar [20,28,29].

Several studies used multiple soil metabolomic methods and compared their advantages and weaknesses. Van Dam and Bouwmeester found that

LC-MS or NMR could be used to analyze water-soluble secondary metabolites [16]. Swenson and Northen [23] and Tang et al. [76] found that LC-MS detected more compounds than GC-MS. However, Jenkins et al. noticed that 11 soil metabolites that were identified by GC-MS were not detected by LC-MS [5]. Sugars, such as hexoses, dihexoses and the sugar alcohol mannitol, would be best identified chromatographically by GC-MS [5]. The LC-MS and NMR were integrated for soil metabolome analysis to analyze sugars and organic acids precisely [77].

The “targeted” data analysis and absolute metabolite quantification of soil metabolites are challenging because they both require the use of authentic standards, which are often not available for many soil metabolites. In “targeted” data analysis, knowledge of the mass spectral properties of standards is used to identify those compounds in soil metabolomics data. Similarly, standards are used to construct calibration curves for absolute quantification. Absolute quantification of metabolite concentrations is challenging because a typical metabolomics experiment might resolve 100 s of metabolites, with concentrations spanning  $>3$  orders of magnitude. Purchasing and analyzing purified standards for 100 s of metabolites is often prohibitively expensive, highlighting the challenge posed by the limited availability of purified standards for many metabolite compounds. It is reported that some of the metabolites with high abundance mass ion signals, including amino acids, sugars, nucleobases, and nucleosides, could be quantified by LC triple quadrupole MS. However, only 25 individual metabolites could be quantified among 96 metabolites [5]. Owing to these challenges, e.g., lack of authentic standards, it is common to use the “untargeted” metabolomics methods for relative quantitation or semi-quantitative analysis using chemically related standards. Thus, limited quantitative work has been performed to identify the fraction of the DOC pool that can be identified using metabolomics. It should be noted that combining untargeted and targeted soil metabolomics would be a promising strategy in soil science. The potentially different metabolites between treatments could be screened by untargeted metabolomics, and the changes of certain metabolites could be verified by targeted metabolomics if standards are available. For example, both untargeted and targeted soil metabolomics were combined to successfully elucidate the eco-corona formation of soil metabolites on the surface of microplastics [78].

### 3. Research progress in soil metabolomics

#### 3.1. Metabolomics for characterizing SOM and for understanding the behavior of DOM

There is a notable overlap between DOM and metabolome in soil (Fig. 1). Soil metabolomics focuses on the measurement of the composition of compounds with low molecular weight (e.g., sugars, fatty acids, and amino acids) in DOM, which helps elucidate the roles of metabolites in SOM stabilization and mineralization, and the transformation of SOM-associated nutrients. Therefore, the early stage application of soil metabolomics focuses on the characterization of the molecular composition of DOM in soil [24,27,28,39,79,58]. For example, it was discovered that the pool of N-containing small organic compounds was more diverse than generally recognized, including abundant quaternary ammonium compounds in addition to the well-known amino acids [27].

Metabolomics can provide insights into changes in DOM composition even when there is no change in overall DOM concentration (total C or N). For example, adding engineered nanomaterials to the soil did not change total DOM concentration but altered the composition and redistribution of soil metabolites in DOM [47]. The amendment of biochar to the soil indeed decreased carbohydrate and peptide concentrations in DOM, even though the total concentration of DOM increased [28]. Thus, soil metabolomics can provide a higher resolution of the changes of specific compounds in DOM in response to environmental conditions than that conducted with the measurement of overall DOM concentration. Hassanpour et al. also stated that the metabolomics-based analysis could help with high-resolution evaluation of the spatial divergence of the DOM in water

[80]. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS), an ultrahigh-resolution electrospray, has been used to characterize DOM [68,81,82]. Unlike metabolomics analysis that typically uses LC-MS/MS to identify metabolite features  $<1,500$  Da—a fraction of which is definitively identified—FT-ICR MS typically does not use chromatography but instead takes advantage of the ultra-high resolution to directly assign chemical formulas for all of DOM with a broad range of molecular sizes [68]. A few limitations of FT-ICR MS are that this technique typically stops at the level of identifying chemical formulas, which is less available, hard to operate, and more costly than LC-MS or GC-MS.

Sorption of DOM by minerals is an important abiotic factor affecting the environmental behavior of DOM in the soil. Soil metabolomic analysis provides the fingerprinting of the redistribution of various metabolites of DOM in soil solution (i.e., pore water) and solid (i.e., adsorbed on soil minerals) phases. Bacterial lysates, such as phosphate-containing and dicarboxylate metabolites, could be strongly sorbed by clay minerals, including ferrihydrite, thereby influencing the mobilization of phosphate [21]. The sorption and desorption of soil metabolites on Fe oxides are quite different from that on silicate minerals such as kaolinite [79]. Using  $^{13}\text{C}$ -labeled lysates from a soil bacterium, Swenson et al. found that cationic and anionic metabolites had the lowest recoveries, while non-ionized metabolites exhibited high recovery after the lysates interacted with sterilized soil [24]. This indicates that the sorption of metabolites to soil may cause the under estimation of DOM fractions in the soil. As mentioned above, the desorption of metabolites from minerals can be achieved with an extractant of high ionic strength. More generally, in the last 50+ years, soil chemistry has developed means to extract different soil pools preferentially by altering extractant chemistry. These range from complex sequential extraction procedures yielding three or more operationally defined pools (e.g., sequential extraction scheme for soil P [83]) through simple contrasts such as extraction with water versus 2 M KCl to quantify free and free + adsorbed pools, respectively. Soil metabolomics should capitalize on this wealth of knowledge; however, classic soil chemistry approaches are not readily compatible with mass spectrometry because they involve inorganic salts, sometimes in high concentrations (e.g., 0.5 M  $\text{K}_2\text{SO}_4$ , or 2 M KCl) and sometimes in complex mixtures (e.g., the popular Mehlich-3 extract, which uses a combination of acetic acid, ammonium nitrate, ammonium fluoride, nitric acid, and ethylenediaminetetraacetic acid). Therefore, future research needs to focus on developing mass spectrometry-compatible extraction techniques to analyze metabolite distribution in both the solution and solid phases of the soil.

#### 3.2. Soil microbiome and metabolome

There is a significant relationship between soil metabolome and soil microbiome, the latter being defined as a characteristic microbial community occupying habitats in the soil and referring to microorganisms and their theater of activity [84]. Thus, this relationship includes but is not limited to i) the contribution of microbial metabolites to soil metabolome, ii) the networks and correlation between soil metabolites and microbial community composition, and iii) the relationship between soil metabolites and microbial community function, activity, etc. Soil microbes play a major role in the biogeochemical cycling of elements, including plant nutrients such as C, N, P, and S, and the cycles of soil organic contaminants and heavy metals; in addition, microbial metabolites make up a significant part of soil metabolomes (Fig. 1). Soil metabolites in DOM with molecular weight lower than 1,000 Da are the most accessible growth substrates for soil microbes [5]. It is difficult to assess the total contribution of soil organics that are derived from microbial metabolites. However, a current view is that much of soil organics originate from microbial necromass [2,85]. It has been estimated that soil microbial metabolites can account for more than 15% of the total MBC [41]. Since microbes use small molecules directly as substrates and different groups of microbes target different subsets of metabolites, the abundance, diversity, and availability of soil metabolites will likely affect

the structure, activity, and function of soil microbial communities [31]. Researchers are increasingly exploring the relationship between soil microbiomes and metabolomes to better understand soil biogeochemical cycling processes, especially the C cycle in relation to soil C stabilization and sequestration [19], nutrient cycles in relation to bioavailability and nutrient flux in soil [40,44], and contaminant cycles in relation to soil remediation [86].

Metabolites occupy a central position in C stabilization in soil. There continues to be active debate about the mechanisms of soil C stabilization/preservation. One common view is that simple organic compounds that are converted by microbes into microbial necromass (which often includes waste products) contribute to the stable C content in the soil [87]. Necromass and metabolites can be stabilized via physicochemical associations with clay minerals (such as Fe and Al oxides, and silicates), resulting in what is referred to as mineral-associated organic matter (MAOM, accounting for approximately 65% of SOM [8,9,10], Fig. 1). In the case of MAOM, the interaction of microbial metabolism with soil mineralogy determines the proportion of plant inputs and microbially derived C that are incorporated into the MAOM and get stabilized in soil (i.e., C sequestration).

Sequencing technologies have enabled extensive investigation of soil microbial community structure [88,89]. The use of multi-omics, including metagenomics and metabolomics, has also been expanded to study soil microbial community networks [90]. Metabolic network analysis is one of the complementary tools that critically aid in exploring cross-species metabolite exchanges and the activities of non-culturable microbial populations in the soil [91]. The relationships between soil microbes and soil metabolites can be examined using techniques such as Procrustes tests combined with Mantel tests [48], or other network correlation analyses. For example, it was shown that during the wetting of soil biocrust, microbes of *Microcoleus* sp. released metabolites such as adenosine and adenine. At the same time, those in *Bacillus* sp. consumed metabolites such as glutamate and myristate [54]. More recently, the application of machine learning to these data sourced from desert soil biocrust wetting [54] helps identify additional associations between soil microbes and metabolites [92]. Significant positive or negative co-occurrences based on Pearson correlation exist between soil metabolites and certain microbes [93]. Similar approaches have been used to show that sucrose had the greatest correlation with bacterial members in the rhizosphere soil of green pepper [56] and that cerium dioxide ( $\text{CeO}_2$ ) nanoparticles altered the co-occurrence network between soil microbial communities and metabolites [42]. Metabolic cross-feeding between microbes, i.e., microbes using metabolites of other members to maintain their own growth in the environment [94], is an area of intense interest and is suggested by positive or negative correlations between soil metabolites and certain microbes [45,49]. It is important to note that there may be uncertainties in the correlation between the soil metabolome and microbiomes. These uncertainties could arise from the lack of pure strain or standard metabolite. Thus, stable isotope-based technologies are needed to test these predicted relationships under field conditions.

Understanding the relationships between soil metabolites and microbial activity and function can help design interventions [28]. For example, anaerobic soil disinfestation, which creates anaerobic soil conditions through the incorporation of easily decomposable soil amendments, application of plastic mulch, or irrigation to form saturated soil conditions, reduces soil metabolites, resulting in changes in microbial metabolic pathways, including pyruvate metabolism, glycolysis, butanoate metabolism encompassing tricarboxylic acid cycle [40]. Environmentally relevant concentrations of testosterone can down-regulate the microbes' amino acid metabolism processes and decrease the contents of soil metabolites such as isozeaxanthin, hydroxyatrazine, and L-isoleucine [48]. The carbohydrate metabolism of soil microbial communities was found to be promoted by sulfadiazine antibiotics, which may reduce the content of carbohydrates, thus reducing the protection for SOM [95]. Recently, Yang et al. observed that plant growth-promoting microbes such as *Arthrobacter ureafaciens* DnL1-1 and

*Trichoderma harzianum* LTR-2 induced a specific wheat rhizosphere metabolite pool, including upregulated lipids, benzoic acids, and amino acids [50], suggesting that microbial metabolism can be a sensitive indicator of environmental changes [96,97]. Lu et al. recently found that metabolomes reach equilibrium faster than soil microbial community structures in response to the cultivation of woody species [96].

Tracing the fate of metabolites during microbial metabolism in the soil is difficult because of their transient flux, complex cycling with multiple sources and sinks for many metabolites, and various compounds in the soil that may interfere with the analysis. The fate of some metabolites can be deduced via isotope label approaches. For example, where labeled isotopologues are available, one can determine bidirectional fluxes of metabolites across plant roots [98], while the flux of metabolites through soil solution can be determined via an isotope pool dilution experiment [99]. Inputs of metabolites to the soil via plant root exudates can be deduced from pulse labeling of  $^{13}\text{CO}_2$  incorporated during photosynthesis appearing in the rhizosphere [100]; however, most studies have examined total rhizosphere  $^{13}\text{C}$  rather than incorporation into individual metabolites [101,102]. Soil exometabolomics, also known as "metabolic footprinting" [103], is a powerful approach for studying the metabolism of a set of small molecule compounds by the microbial community or microbial strains under laboratory conditions [30]. This method uses metabolomics to compare inoculated vs. uninoculated media to directly explore the causal relationship between microbiomes and metabolites under relatively simple conditions that avoid external environmental interference, e.g., a metabolite is produced or consumed by a microbe. It was observed that soil bacteria have different substrate preferences, suggesting that exometabolite pools contribute to microbial diversity and community structure [31]. Some efforts have been made to create soil-relevant defined media that enable complete tracking of the various organic metabolites, including a soil-defined medium [5] and, more recently, a new medium that takes elemental stoichiometry into account and has been shown to support the growth of diverse soil microbes [104]. Using soil exometabolomic footprinting methods, Cyle et al. found that there was no clear relationship between the compounds' nominal oxidation state of C and the order of substrate depletion by a pure strain, *Paraburkholderia* sp. 1N, which belongs to copiotroph [77]. They further confirmed the complexities of the metabolomic strategies of microbial communities in metabolite cycling in soils [105]. The accumulation of the short-chain carboxylic acid in the soil anoxic incubation condition was found to be greater than that in the oxic condition by measuring the metabolism of  $^{13}\text{C}$ -glucose through the exometabolomics approach [106]. We anticipate that large-scale exometabolite profiling of soil microbes in combination with metagenomic sequencing will significantly help predict metabolic niches and interactions [107].

### 3.3. Rhizosphere soil metabolomics

The rhizosphere is the most dynamic interface between the soil and the plant, with a large amount of substrates and much energy flowing through the soil, forming a hotspot for soil nutrient transformation and biological activity [108]. About 10%–40% of plant photosynthetic products are released into the rhizosphere soil as root exudates [109]. However, the diversity of compounds released in the soil rhizosphere remains poorly understood. The rhizosphere chemicals contain root exudates and products from their breakdown, microbial metabolites, and compounds derived from biogeochemical cycles of various products such as those released through the decomposition of SOM (Fig. 1). Hence, studying metabolites of the rhizosphere can be challenging because many metabolites can have multiple sources and sinks. These compounds can be studied by untargeted rhizosphere metabolomics [33], while isotope labeling techniques can be employed to study fluxes [98,99]. Rhizosphere soil metabolomics has been used to study the distribution and speciation of root exudates, root-microbial interaction, and the response of roots to both abiotic and biotic stresses, including exposure to toxic

contaminants [16]. Zhalmina et al. used exometabolomic analysis of root exudates and demonstrated that bacteria enriched in the rhizosphere preferred to use aromatic organic acids [109].

Metabolomics analysis of plant root exudates helps better understand metabolites' role as signals between plant roots and nematodes, microbes, and competing plants. Metabolomics analysis showed that flavonoids, lipids, and alkaloids were enriched in *Arabidopsis* soil, and compounds that were not annotated were more abundant in non-rhizosphere than in rhizosphere soil [44]. By analyzing the variation of metabolites from different genotypes of sorghum (*Sorghum bicolor*) growing in soil, clay, or sand, it was shown that belowground substrates influenced profiles of root exudates and that sucrose was greatly enhanced in the rhizosphere environment [110]. Seven soil metabolites were significantly higher in the soil with alpine meadow plants growing on *Floccularia luteovirens* fairy rings compared to those outside of the fairy rings [111]. Moreover, the rhizosphere metabolites with maize (*Zea mays*) grown in soil differed from those grown in water incubation experiments [33].

Rhizosphere exudates are closely and dynamically connected with soil microbes that reside within the rhizosphere [109]. For example, a negative correlation has been observed between the carbohydrates and the soil bacteria of *Chloroflexia* and *Actinobacteria* in the rhizosphere of *Camellia oleifera* [96]. It has also been found that iron-oxidizing bacteria could change the distribution of soil metabolites in rhizosphere soil [66], while some of the key metabolites were consistent with those in non-rhizosphere soil. Song et al. measured metabolites from the pepper rhizosphere and bulk soils under plastic greenhouse vegetable cultivation and found that starch and sucrose metabolism pathways varied the most between rhizosphere and bulk soils, with down-regulation of the functional genes participating in this pathway in the rhizosphere [56]. Concentrations of trehalose and betaine, as well as choline-like and carnitine-like compounds detected by NMR, and amino acids such as aspartic acid and glutamine detected by LC-MS, were higher in rhizosphere soils of *Burkea* tree than those in non-rhizosphere soils [112].

A major challenge in interpreting rhizospheric metabolomic data lies in differentiating between metabolites produced by the plant host vs. rhizosphere bacteria [69]. The limited availability of data on the quantity of metabolites further complicates the understanding of the dynamic of the rhizosphere metabolites. It has been suggested that metabolomics may not be able to resolve all signaling relations in rhizosphere soil [16]. Thus, using a metabolomic approach, the term “signalomics” has been proposed to describe the chemical communications between plants and microbes and among microbes themselves in the rhizosphere [108].

#### 3.4. Soil metabolomes and soil contamination

The response of soil metabolome to contamination has been identified as a key ecological indicator and molecular marker in soil remediation. Soil contamination by persistent organic pollutants, potentially toxic elements, and emerging contaminants such as nanomaterials and microplastics has become a serious environmental issue. The response of individual biota or entire communities to contaminants could be predicted or assessed by sensitive compounds, i.e., biomarkers [113]. Recently, soil metabolites have been investigated as biomarkers for soil contamination. The least variation in the soil's entire community metabolites occurs in soils with similar pollution profiles based on principal component analysis [29]. Patil et al. proposed environmental metabolic footprinting (EMF) to analyze soil metabolomes in response to the stress of contaminants, such as herbicides [32]. They found that the impact period of leptospermone, the natural  $\beta$ -triketone herbicide, on the soil microbiomes was shorter than that of systemic post-emergence herbicides, e.g., sulcotriione [32]. EMF can be used to generate both endogenous and exogenous metabolite profiles that can serve as biomarkers for contaminants in soils. For example, EMF reflected more information, e.g., the by-product formation for pesticides and microbial metabolism modification, than the measurement of the half-life of pesticides in the environment [114]. In a case study, phytosphingosine was

detected as a specific metabolite in soils contaminated with phenol and hydrogen fluoride [55]. The responses of soil metabolomes to the contamination of polycyclic aromatic hydrocarbons (PAHs) [93], chloropyrifos [115], epoxiconazole [116], and CeO<sub>2</sub> nanoparticles in soil [117] were also investigated to elucidate the soil microbial response to the stress of contaminants. Starch and sucrose metabolism, N metabolism, S metabolism, propanoate metabolism, fatty acid metabolism, and the urea cycle were fluctuated in soil exposed to CeO<sub>2</sub> nanoparticles, resulting in enhanced concentrations of fatty acids, amino acids, and amines [42].

Additionally, the rhizosphere acts as a hotspot for the remediation of soil contaminants. The C metabolism has been found to affect the degradation of PAHs in the rhizosphere [86,118]. The rhizosphere metabolites, including levoglucosan, linolenic acid, 4-hydroxycinnamic acid, and allo-inositol, were significantly increased in response to the addition of engineered nanomaterials such as silicon dioxide (SiO<sub>2</sub>), titanium dioxide (TiO<sub>2</sub>), or ferroferric oxide (Fe<sub>3</sub>O<sub>4</sub>), in soils planted with maize [47]. Similarly, metabolites (including sugars and sugar alcohols, fatty acids, and small organic acids) increased, while amino acids decreased in the *Brassica rapa* subsp. *chinensis* (Pakchoi) rhizosphere soil exposed to SiO<sub>2</sub> nanoparticles [51]. The effect of Ag nanoparticles on the soil metabolite profiles in soil planted with cucumber was also studied, showing that the negative impact of Ag nanoparticles on microbes, such as decreasing the abundances of fatty acids, was not altered by the presence of the plant [52].

Microplastic contamination in terrestrial environments is an emerging concern. Polybutylene succinate and polylactide microplastics were reported to alter the lipid metabolism in soil [119]. Polyethylene microplastics significantly changed the soil metabolite spectrum in a short period, including organic N compounds, lipids and lipid-like molecules, and organic acids and derivatives [120]. The coexistence of oxytetracycline and polyethylene microplastic in soil significantly increased the relative concentrations of arachidic acid and indoleacetic acid while decreasing dodecanoic and heptadecanoic acids [121]. It remains to be clarified whether there are dose-effect relationships between certain soil metabolites and contaminants. Further research is essential to test whether soil metabolomes could act as molecular-level biomarkers sensitive to soil contamination. The impact of the contaminants of emerging concern on soil metabolomes needs to be further studied for potential risk assessment in the future.

#### 3.5. Metabolomics for the analysis of soil health and climatic stresses

Soil metabolomes can be a sensitive indicator of soil quality. Soil metabolome has the potential to provide more sensitive soil quality assessment than the traditional soil quality indicators, including pH, cation exchange capacity, SOM, MBC, enzyme activity (e.g., dehydrogenase), and so on [6]. For example, it was reported that the soil metabolites tend to cluster better to land use than the location of sampling, contrasting with findings based on soil chemical elemental composition [20]. Similarly, Nguyen et al. measured soil metabolomes from 188 backyard soils across 14 U.S. States, demonstrating that soil metabolomes could reflect the effects of local factors such as temperature, light level, and human activities on the soil [57].

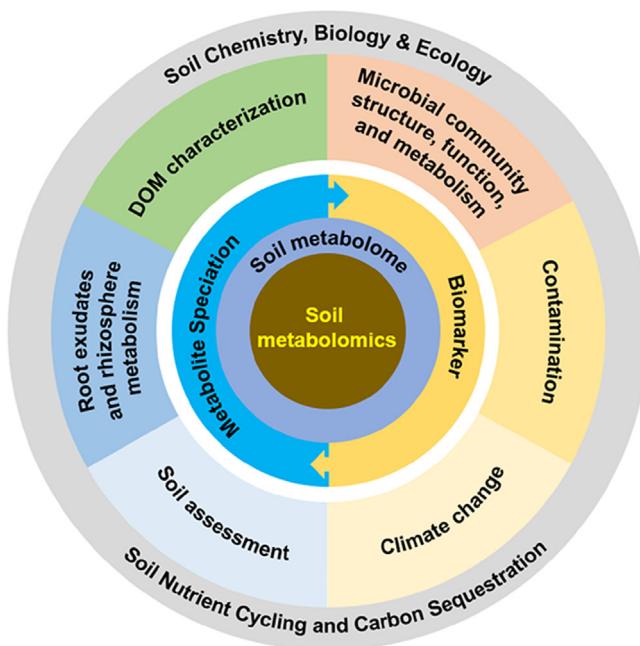
The responses of soil microbes to climatic stress and soil management can also be characterized by soil metabolomics. The soil metabolism process was strongly impacted by climate change-induced extreme weather events, including drought [46]. Kakumanu et al. found that the soil metabolites such as sugars and alcohols did not increase consistently with soil drying [41]. The drought-stressed soil contained >10-fold more known microbial osmolytes than those in the control [39]. Soil microbes could enhance the production of osmolytes, such as accumulating sugars and polyols, to alleviate environmental stresses [122]. The impacts of global warming [53], drying and rewetting [99], and P fertilization [123] on the soil metabolomes have also been studied. Therefore, characterizing the changes in soil microbiomes will further elucidate the responses of soil microorganisms to climate changes in the future.

#### 4. Conclusions and future perspectives

Soil metabolomics has developed rapidly, attracting growing interest for its application in researching soil biogeochemical cycles. Studies on the characterization of DOM, microbe–plant interaction in the rhizosphere, and microbial response to environmental stress and climate changes have been conducted via soil metabolomics, using nontargeted or targeted techniques. Soil metabolomics is a powerful approach connecting soil chemistry, soil biology, and soil ecology (Fig. 2), poised for further development along with new techniques and instruments. In the following section, we highlight a number of areas of standardization and research that we believe will greatly improve the performance of soil metabolomics.

##### 1) Standardization of methods and reporting

The metabolomics field, including its subfields such as soil metabolomics, has a well-acknowledged need for standardization [124–127]. For example, an urgent need for standardization can be observed in the area of soil extraction methods, which are highly variable between labs. A plethora of methods have been used in different labs to extract metabolites from the soil (Table 1). Since the metabolites detected depend on the exact extraction method, standardization is needed to enable cross-study comparisons, for example, to enable global data comparison in soil science and the development of consistent bioindicators [19]. With this said, we also acknowledge that there is a need for improved extraction methods, especially for soils with high clay contents as well as soil components such as MAOM. Given the diversity of global soil types, we envision the development of standard extracting methods tailored to different soil types with different SOM, clay mineralogy, pH, etc.



**Fig. 2.** The research progress and application of soil metabolomics. Soil metabolomics studies the abundance, diversity, composition, and dynamics of soil metabolome in response to temporal and spatial variations of soil conditions. Soil metabolites can act as biomarkers for stresses such as climate change and environmental contamination. In the future, the relationship between soil metabolome and microbial community and the inheritance of soil metabolome and nutrient cycling and carbon sequestration will be broadly considered, and soil health assessment based on soil metabolome will be developed.

##### 2) New metabolites need to be identified

Most compounds detected in metabolomics experiments are not assigned (estimated to be <14% [128]), particularly in less studied systems such as soils. This greatly hampers the understanding of the entire soil metabolome and their roles and relationships with soil microbes, plants, and biogeochemical cycling of elements. We emphasize the need to expand databases and cheminformatic tools for assigning the chemical structure of metabolites, a step critical to improving the performance of soil metabolomics. We further anticipate that the application of artificial intelligence will facilitate the analysis of mass spectrometric data for the identification of metabolites. For example, a recently report method employed machine learning to determine chemical formulas based on fragmentation data, enabling conventional instruments to provide information comparable to those obtained using more specialized FT-ICR instruments [129].

##### 3) Gaining new insights into mechanisms of soil C stabilization

We anticipate that soil metabolomics will become increasingly important, given the significant interest in using soils to sequester carbon. This is because plants and microbes produce and act on specific chemical structures (e.g., glucose vs. galactose), and therefore, in many cases, it will be important to characterize SOM and DOM at the molecular level. It will be important to develop targeted methods (likely using isotopic standards) to determine the absolute quantities of metabolites in the soil, for example, by using stable isotope additions [130]. In the future, the high-resolution analysis of the DOM, coupled with targeted soil metabolomics, will greatly deepen our understanding of the components and dynamics of DOM in the soil. Moreover, since, in general, MAOM contributes the largest fraction of SOM, further research on the molecular composition of MAOM should be conducted with soil metabolomics to understand the contribution of small molecular compounds to the formation of MAOM. As described above, this will require new methodologies for soil MAOM extraction that are compatible with the next generation of molecular identification techniques.

##### 4) Isotope-based metabolomics

Using isotope-labeled metabolites to study their transformations will help expand research capabilities, especially for those metabolites that are released in trace amounts and are rapidly degraded in the soil [131]. For example, <sup>13</sup>C [100], <sup>15</sup>N [132], and D<sub>2</sub>O labeling [133] have been used to study the fluxes of root exudates, gross fluxes of substrates, and *de novo* synthesis of metabolites in soil. Isotope-based metabolomics will be a powerful technique to elucidate soil metabolites' environmental fate and behavior in future research [106]. These can also provide important insights into the flux of metabolites within soils [99,130].

##### 5) Investigation of spatiotemporal dynamics of soil metabolome

The soil metabolite profile is very dynamic. However, most reported studies are based on soil metabolite profiles at selected sampling points. The spatial and temporal changes in the concentration and composition of metabolomes should be linked with the soil's biochemical properties and environmental conditions. The relationship between soil metabolomes and microbiomes, especially in response to climate change, needs further research. The dynamics of soil metabolomes will provide new insights into interactions among soil biota, such as microbe–microbe, microbe–environment, and microbe–plant interactions through chemical dialogue or food web [94].

##### 6) Exometabolic profiling of soil microbes using big-data

With the development of molecular biology and machine learning techniques, exometabolomics may be extended from culturable to unculturable microbial strains or communities, allowing us to identify their

substrate preferences and secreted products to help interpret the metabolic web of microbes in soils. The integration of exometabolomics and spatial information through techniques such as mass spectrometry imaging will be another powerful method for elucidating microbial interactions in the soil.

## 7) Multitechnology approaches in soil metabolomics

Both NMR and MS data could be combined and used to analyze soil metabolites. Combining multiple analytical methods for metabolomics, such as NMR and GC-MS, has been suggested [70]. Employing multi-omics tools to study the transformation and stabilization of SOM is an emerging aspect. Metabolomics could verify genomics results to indicate what has happened in the soil driven by microbes. Therefore, integrating soil metabolomics with metagenomics, metatranscriptomics, and metaproteomics, coupled with machine learning technology, should be used broadly to deepen our understanding of soil biogeochemistry and their responses to management and climate change in the future.

## CRediT authorship contribution statement

Y.S. and S.X.C. developed the idea for this manuscript. Y.S. organized the manuscript. X.J. and S.X.C. supervised this work. Y.S., S.Y. and X.L. designed the figures. T.W., N.B., C.W., T.N edited the manuscript and made substantial contributions. All authors discussed and reviewed the manuscript.

## Declaration of competing interests

The authors declare no competing financial interest.

## Acknowledgments

This study was financially supported by the Youth Innovation Promotion Association, CAS (2021309), the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2018-05700), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA28010501), and the National Natural Science Foundation of China (42277303, 42107244). TRN is supported by m-CAFEs Microbial Community Analysis & Functional Evaluation in Soils, an SFA led by Lawrence Berkeley National Laboratory and supported by the U.S. Department of Energy, Office of Science, Office of Biological & Environmental Research under contract number DE-AC02-05CH11231. The work conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. This work was also partially supported by a Discovery grant from the Australian Research Council (DP220101620).

## References

- [1] J. Lehmann, M. Kleber, The contentious nature of soil organic matter, *Nature* 528 (2015) 60–68.
- [2] M.W.I. Schmidt, M.S. Torn, S. Abiven, T. Dittmar, G. Guggenberger, I.A. Janssens, et al., Persistence of soil organic matter as an ecosystem property, *Nature* 478 (2011) 49–56.
- [3] C.M. Kallenbach, S.D. Frey, A.S. Grandy, Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls, *Nat. Commun.* 7 (2016) 3929.
- [4] I. Koegel-Knabner, C. Rumpel, Advances in molecular approaches for understanding soil organic matter composition, origin, and turnover: a historical overview, *Adv. Agron.* 149 (2018) 1–48.
- [5] S. Jenkins, T.L. Swenson, R. Lau, A.M. Rocha, A. Aaring, T.C. Hazen, et al., Construction of viable soil defined media using quantitative metabolomics analysis of soil metabolites, *Front. Microbiol.* 8 (2017) 2618.
- [6] E. Withers, P.W. Hill, D.R. Chadwick, D.L. Jones, Use of untargeted metabolomics for assessing soil quality and microbial function, *Soil Biol. Biochem.* 143 (2020) 107758.
- [7] N.S. Bolan, D.C. Adriano, A. Kunhikrishnan, T. James, R. McDowell, N. Senesi, Dissolved organic matter: biogeochemistry, dynamics, and environmental significance in soils, *Adv. Agron.* 110 (2011) 1–75.
- [8] W. Yu, W. Huang, S.R. Weintraub-Leff, S.J. Hall, Where and why do particulate organic matter (POM) and mineral-associated organic matter (MAOM) differ among diverse soils? *Soil Biol. Biochem.* 172 (2022) 108765.
- [9] J.M. Lavallee, J.L. Soong, M.F. Cotrufo, Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century, *Global Change Biol.* 26 (2020) 261–273.
- [10] N.W. Sokol, E.D. Whalen, A. Jilling, C. Kallenbach, J. Pett-Ridge, K. Georgiou, Global distribution, formation and fate of mineral-associated soil organic matter under a changing climate: a trait-based perspective, *Funct. Ecol.* 36 (2022) 1411–1429.
- [11] K. Peters, A. Worrich, A. Weinhold, O. Alka, G. Balcke, C. Birkemeyer, et al., Current challenges in plant eco-metabolomics, *Int. J. Mol. Sci.* 19 (2018) 1385.
- [12] L. Pagano, A.D. Serving, R. De La Torre-Roche, A. Mukherjee, S. Majumdar, J. Hawthorne, et al., Molecular response of crop plants to engineered nanomaterials, *Environ. Sci. Technol.* 50 (2016) 7198–7207.
- [13] M. de Raad, C.R. Fischer, T.R. Northen, High-throughput platforms for metabolomics, *Curr. Opin. Chem. Biol.* 30 (2016) 7–13.
- [14] J.L. Wolfender, G. Marti, A. Thomas, S. Bertrand, Current approaches and challenges for the metabolite profiling of complex natural extracts, *J. Chromatogr. A* 1382 (2015) 136–164.
- [15] O.A.H. Jones, M.L. Maguire, J.L. Griffin, D.A. Dias, D.J. Spurgeon, C. Svendsen, Metabolomics and its use in ecology, *Austral Ecol.* 38 (2013) 713–720.
- [16] N.M. van Dam, H.J. Bouwmeester, Metabolomics in the Rhizosphere: tapping into belowground chemical communication, *Trends Plant Sci.* 21 (2016) 256–265.
- [17] N. Jiao, G.J. Herndl, D.A. Hansell, R. Benner, G. Kattner, S.W. Wilhelm, et al., Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean, *Nat. Rev. Microbiol.* 8 (2010) 593–599.
- [18] E.C. Minor, M.M. Swenson, B.M. Mattson, A.R. Oyler, Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis, *Environ. Sci. Process Impacts* 16 (2014) 2064–2079.
- [19] D.P. Overy, M.A. Bell, J. Habtewold, B.L. Helgason, E.G. Gregorich, “Omics” technologies for the study of soil carbon stabilization: a review, *Front. Environ. Sci.* 9 (2021) 617952.
- [20] S. Rochfort, V. Ezernieks, P. Mele, M. Kitching, NMR metabolomics for soil analysis provide complementary, orthogonal data to MIR and traditional soil chemistry approaches - a land use study, *Magn. Reson. Chem.* 53 (2015) 719–725.
- [21] T.L. Swenson, B.P. Bowen, P.S. Nico, T.R. Northen, Competitive sorption of microbial metabolites on an iron oxide mineral, *Soil Biol. Biochem.* 90 (2015) 34–41.
- [22] P.A.W. van Hees, D.L. Jones, R. Finlay, D.L. Godbold, U.S. Lundstrom, The carbon we do not see - the impact of low molecular weight compounds on carbon dynamics and respiration in forest soils: a review, *Soil Biol. Biochem.* 37 (2005) 1–13.
- [23] T.L. Swenson, T.R. Northen, Untargeted soil metabolomics using liquid chromatography-mass spectrometry and gas chromatography-mass spectrometry, *Methods Mole. Biol.* 1859 (2019) 97–109.
- [24] T.L. Swenson, S. Jenkins, B.P. Bowen, T.R. Northen, Untargeted soil metabolomics methods for analysis of extractable organic matter, *Soil Biol. Biochem.* 80 (2015) 189–198.
- [25] J.S. Buyer, B. Vinyard, J. Maul, K. Selmer, R. Lupitskyy, C. Rice, et al., Combined extraction method for metabolomic and PLFA analysis of soil, *Appl. Soil Ecol.* 135 (2019) 129–136.
- [26] T. Kind, G. Wohlgemuth, D.Y. Lee, Y. Lu, M. Palazoglu, S. Shahbaz, et al., FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry, *Anal. Chem.* 81 (2009) 10038–10048.
- [27] C.R. Warren, High diversity of small organic N observed in soil water, *Soil Biol. Biochem.* 57 (2013) 444–450.
- [28] P.J. Mitchell, A.J. Simpson, R. Soong, M.J. Simpson, Shifts in microbial community and water-extractable organic matter composition with biochar amendment in a temperate forest soil, *Soil Biol. Biochem.* 81 (2015) 244–254.
- [29] O.A.H. Jones, S. Sdepanian, S. Lofts, C. Svendsen, D.J. Spurgeon, M.L. Maguire, et al., Metabolomic analysis of soil communities can be used for pollution assessment, *Environ. Toxicol. Chem.* 33 (2014) 61–64.
- [30] L.P. Silva, T.R. Northen, Exometabolomics and msI: deconstructing how cells interact to transform their small molecule environment, *Curr. Opin. Biotechnol.* 34 (2015) 209–216.
- [31] R. Baran, E.L. Brodie, J. Mayberry-Lewis, E. Hummel, U.N. Da Rocha, R. Chakraborty, et al., Exometabolite niche partitioning among sympatric soil bacteria, *Nat. Commun.* 6 (2015) 8289.
- [32] C. Patil, C. Calvayrac, Y. Zhou, S. Romdhane, M.-V. Salvia, J.-F. Cooper, et al., Environmental metabolic footprinting: a novel application to study the impact of a natural and a synthetic beta-triketone herbicide in soil, *Sci. Total Environ.* 566 (2016) 552–558.
- [33] N. Cheng, Y. Peng, Y. Kong, J. Li, C. Sun, Combined effects of biochar addition and nitrogen fertilizer reduction on the rhizosphere metabolomics of maize (*Zea mays* L.) seedlings, *Plant Soil* 433 (2018) 19–35.
- [34] J. Lv, Z. Huang, L. Luo, S. Zhang, Y. Wang, Advances in molecular and microscale characterization of soil organic matter: current limitations and future prospects, *Environ. Sci. Technol.* 56 (2022) 12793–12810.
- [35] L. Lopez-Sangil, P. Rovira, Sequential chemical extractions of the mineral-associated soil organic matter: an integrated approach for the fractionation of organo-mineral complexes, *Soil Biol. Biochem.* 62 (2013) 57–67.

- [36] M. Kleber, J. Lehmann, Humic substances extracted by alkali are invalid proxies for the dynamics and functions of organic matter in terrestrial and aquatic ecosystems, *J. Environ. Qual.* 48 (2019) 207–216.
- [37] A. Nebbiuso, A. Piccolo, Basis of a humomics science: chemical fractionation and molecular characterization of humic biosuprastructures, *Biomacromolecules* 12 (2011) 1187–1199.
- [38] M. Roman, W.T. Winter, Effect of sulfate groups from sulfuric acid hydrolysis on the thermal degradation behavior of bacterial cellulose, *Biomacromolecules* 5 (2004) 1671–1677.
- [39] C.R. Warren, Response of osmolytes in soil to drying and rewetting, *Soil Biol. Biochem.* 70 (2014) 22–32.
- [40] C.W. Johns, A.B. Lee, T.I. Springer, E.N. Rosskopf, J.C. Hong, W. Turechek, et al., Using NMR-based metabolomics to monitor the biochemical composition of agricultural soils: a pilot study, *Eur. J. Soil Biol.* 83 (2017) 98–105.
- [41] M.L. Kakumanu, C.L. Cantrell, M.A. Williams, Microbial community response to varying magnitudes of desiccation in soil: a test of the osmolyte accumulation hypothesis, *Soil Biol. Biochem.* 57 (2013) 644–653.
- [42] C. Pan, Y. Bao, A. Guo, J. Ma, Environmentally relevant-level CeO<sub>2</sub> NP with ferrous amendment alters soil bacterial community compositions and metabolite profiles in rice-planted soils, *J. Agric. Food Chem.* 68 (2020) 8172–8184.
- [43] C. Willers, P.J.J. van Rensburg, S. Claassens, Can a metabolomics-based approach be used as alternative to analyse fatty acid methyl esters from soil microbial communities? *Soil Biol. Biochem.* 103 (2016) 417–428.
- [44] P. Petriacq, A. Williams, A. Cotton, A.E. McFarlane, S.A. Rolfe, J. Ton, Metabolite profiling of non-sterile rhizosphere soil, *Plant J.* 92 (2017) 147–162.
- [45] S.S. Hewavitharana, E. Klarer, A.J. Reed, R. Leisso, B. Poirier, L. Honnaas, et al., Temporal dynamics of the soil metabolome and microbiome during simulated anaerobic soil disinfection, *Front. Microbiol.* 10 (2019) 2365.
- [46] R.W. Brown, D.R. Chadwick, H. Zang, D.L. Jones, Use of metabolomics to quantify changes in soil microbial function in response to fertiliser nitrogen supply and extreme drought, *Soil Biol. Biochem.* 160 (2021) 108351.
- [47] L. Zhao, H. Zhang, J.C. White, X. Chen, H. Li, X. Qu, et al., Metabolomics reveals that engineered nanomaterial exposure in soil alters both soil rhizosphere metabolite profiles and maize metabolic pathways, *Environ. Sci.: Nano* 6 (2019) 1716–1727.
- [48] J. Ding, H. Sun, A. Liang, J. Liu, L. Song, M. Lv, et al., Testosterone amendment alters metabolite profiles of the soil microbial community, *Environ. Pollut.* 272 (2021) 115928.
- [49] S.S. Hewavitharana, E. Klarer, J. Muramoto, C. Shennan, M. Mazzola, Analysis of environmental variables and carbon input on soil microbiome, metabolome and disease control efficacy in strawberry attributable to anaerobic soil disinfection, *Microorganisms* 9 (2021) 1638.
- [50] K. Yang, H. Li, L. Li, J. Hu, Y. Wei, H. Yang, et al., Soil metabolomics reveal complex interactions between *Arthrobacter ureafaciens* and *Trichoderma harzianum* when co-inoculated on wheat, *Pedobiologia* 85–86 (2021) 150723.
- [51] L. Tian, J. Shen, G. Sun, B. Wang, R. Ji, L. Zhao, Foliar application of SiO<sub>2</sub> nanoparticles alters soil metabolite profiles and microbial community composition in the Pakchoi (*Brassica chinensis* L.) rhizosphere grown in contaminated mine soil, *Environ. Sci. Technol.* 54 (2020) 13137–13146.
- [52] H. Zhang, M. Huang, W. Zhang, J.L. Gardea-Torresdey, J.C. White, R. Ji, et al., Silver nanoparticles alter soil microbial community compositions and metabolite profiles in unplanted and cucumber-planted soil, *Environ. Sci. Technol.* 54 (2020) 3334–3342.
- [53] A. Gargallo-Garriga, J. Sardans, M. Ayala-Roque, B.D. Sigurdsson, N.I.W. Leblans, M. Oravec, et al., Warming affects soil metabolome: the case study of Icelandic grasslands, *Eur. J. Soil Biol.* 105 (2021) 103317.
- [54] T.L. Swenson, U. Karaoz, J.M. Swenson, B.P. Bowen, T.R. Northen, Linking soil biology and chemistry in biological soil crust using isolate exometabolomics, *Nat. Commun.* 9 (2018) 19.
- [55] N. Kim, Y. Ahn, J. Jo, H. Pyo, J. Lee, J. Choi, Soil assessment after chemical accidents using metabolic profiling and microbial community evaluation, *Chemosphere* 268 (2021) 129362.
- [56] Y. Song, X. Li, S. Yao, X. Yang, X. Jiang, Correlations between soil metabolomics and bacterial community structures in the pepper rhizosphere under plastic greenhouse cultivation, *Sci. Total Environ.* 728 (2020), 138439–138439.
- [57] T.D. Nguyen, M. Lesani, I. Forrest, Y. Lan, D.A. Dean, Q.M.R. Gibaut, et al., Local phenomena shape backyard soil metabolite composition, *Metabolites* 10 (2020) 86.
- [58] N. Ueberschaar, K. Lehmann, S. Meyer, C. Zerfass, B. Michalzik, K.U. Totsche, et al., Soil solution analysis with untargeted GC-MS-a case study with different lysimeter types, *Front. Earth Sci.* 8 (2021) 563379.
- [59] D. Schenkel, A. Deveau, J. Niimi, P. Mariotte, A. Vitra, M. Meisser, et al., Linking soil's volatileome to microbes and plant roots highlights the importance of microbes as emitters of belowground volatile signals, *Environ. Microbiol.* 21 (2019) 3313–3327.
- [60] P.C. Brookes, A. Landman, G. Pruden, D.S. Jenkinson, Chloroform fumigation and the release of soil nitrogen: a rapid direct extraction method to measure microbial biomass nitrogen in soil, *Soil Biol. Biochem.* 17 (1985) 837–842.
- [61] C.R. Warren, Comparison of methods for extraction of organic N monomers from soil microbial biomass, *Soil Biol. Biochem.* 81 (2015) 67–76.
- [62] B.R. Whelan, N.J. Barrow, A study of a method for displacing soil solution by centrifuging with an immiscible liquid, *J. Environ. Qual.* 9 (1980) 315–319.
- [63] D. Schenkel, M.C. Lemfack, B. Piechulla, R. Splivallo, A meta-analysis approach for assessing the diversity and specificity of belowground root and microbial volatiles, *Front. Plant Sci.* 6 (2015) 707.
- [64] M.A. Bell, U. McKim, A. Sproule, R. Tobalt, E. Gregorich, D.P. Overy, Extraction methods for untargeted metabolomics influence enzymatic activity in diverse soils, *Sci. Total Environ.* 828 (2022) 154433.
- [65] C. Wu, Y. Ma, D. Wang, Y. Shan, X. Song, H. Hu, et al., Microbiology combined with metabolomics revealing the response of soil microorganisms and their metabolic functions exposed to phthalic acid esters, *Ecotoxicol. Environ. Saf.* 233 (2022) 113338.
- [66] Z. Qian, C. Wu, W. Pan, X. Xiong, L. Xia, W. Li, Arsenic transformation in soil-rice system affected by Iron-Oxidizing strain (*Ochrobactrum sp.*) and related soil metabolomics analysis, *Front. Microbiol.* 13 (2022) 794950.
- [67] S. Buckley, R. Brackin, S. Jamtgard, T. Nasholm, S. Schmidt, Microdialysis in soil environments: current practice and future perspectives, *Soil Biol. Biochem.* 143 (2020) 107743.
- [68] W. Bahureksa, M.M. Tfaily, R.M. Boiteau, R.B. Young, M.N. Logan, A.M. McKenna, et al., Soil organic matter characterization by fourier transform ion cyclotron resonance mass spectrometry (FTICR MS): a critical review of sample preparation, analysis, and data interpretation, *Environ. Sci. Technol.* 55 (2021) 9637–9656.
- [69] R.A. White III, A. Rivas-Ubach, M.I. Borkum, M. Köberl, A. Bilbao, S.M. Colby, et al., The state of rhizospheric science in the era of multi-omics: a practical guide to omics technologies, *Rhizosphere* 3 (2017) 212–221.
- [70] F. Bhinderwala, N. Wase, C. DiRussro, R. Powers, Combining mass spectrometry and NMR improves metabolite detection and annotation, *J. Proteome Res.* 17 (2018) 4017–4022.
- [71] H. Kaspar, K. Dettmer, Q. Chan, S. Daniels, S. Nimkar, M.L. Davighi, et al., Urinary amino acid analysis: a comparison of iTRAQ (R)-LC-MS/MS, GC-MS, and amino acid analyzer, *J. Chromatogr. B* 877 (2009) 1838–1846.
- [72] R. Jiang, Y. Jiao, F. Xu, Chemical derivatization-based LC-MS for metabolomics: advantages and challenges, *Bioanalysis* 8 (2016) 1881–1883.
- [73] J.F. Xiao, B. Zhou, H.W. Ressom, Metabolite identification and quantitation in LC-MS/MS-based metabolomics, *Trac-Trends Anal. Chem.* 32 (2012) 1–14.
- [74] M. Wang, J.J. Carver, V.V. Phelan, L.M. Sanchez, N. Garg, Y. Peng, et al., Sharing and community curation of mass spectrometry data with global natural products social molecular networking, *Nat. Biotechnol.* 34 (2016) 828–837.
- [75] J. Kim, A.H. Goldstein, R. Chakraborty, K. Jardine, R. Weber, P.O. Sorensen, et al., Measurement of volatile compounds for real-time analysis of soil microbial metabolic response to simulated snowmelt, *Front. Microbiol.* 12 (2021) 679671.
- [76] S. Tang, Q. Ma, J. Zhou, W. Pan, D.R. Chadwick, A.S. Gregory, et al., Use of untargeted metabolomics to analyse changes in extractable soil organic matter in response to long term fertilization, *Biol. Fertil. Soils* 59 (2023) 301–316.
- [77] K.T. Cyle, A.R. Klein, L. Aristilde, C.E. Martinez, Ecophysiological study of *Paraburkholderia* sp. strain 1N under soil solution conditions: dynamic substrate preferences and characterization of carbon use efficiency, *Appl. Environ. Microbiol.* 86 (2020) e01851–20.
- [78] S. Yao, X. Li, T. Wang, X. Jiang, Y. Song, H.P.H. Arp, Soil metabolome impacts the formation of the eco-corona and adsorption processes on microplastic surfaces, *Environ. Sci. Technol.* 57 (2023) 8139–8148.
- [79] M.R. Gmach, M.R. Cherubin, K. Kaiser, C.E. Pellegrino Cerri, Processes that influence dissolved organic matter in the soil: a review, *Sci. Agric.* 77 (2020) e20180164.
- [80] B. Hassanpour, N. Blair, L. Aristilde, Metabolomics analysis of unresolved molecular variability in stoichiometry dynamics of a stream dissolved organic matter, *Water Res.* 223 (2022) 118923.
- [81] A. Hu, M. Choi, A.J. Tanentzap, J. Liu, K.-S. Jang, J.T. Lennon, et al., Ecological networks of dissolved organic matter and microorganisms under global change, *Nat. Commun.* 13 (2022) 3600.
- [82] A. Hu, F. Meng, A.J. Tanentzap, K.-S. Jang, J. Wang, Dark matter enhances interactions within both microbes and dissolved organic matter under global change, *Environ. Sci. Technol.* 57 (2023) 761–769.
- [83] M.J. Hedley, J.W.B. Stewart, B.S. Chauhan, Changes in inorganic and organic soil phosphorus fractions induced by cultivation practices and by laboratory incubations, *Soil Sci. Soc. Am. J.* 46 (1982) 970–976.
- [84] G. Berg, D. Rybakova, D. Fischer, T. Cernava, M.-C.C. Verges, T. Charles, et al., Microbiome definition re-visited: old concepts and new challenges, *Microbiome* 8 (2020) 103.
- [85] R.A. Neurath, J. Pett-Ridge, I. Chu-Jacoby, D. Herman, T. Whitman, P.S. Nico, et al., Root carbon interaction with soil minerals is dynamic, leaving a legacy of microbially derived residues, *Environ. Sci. Technol.* 55 (2021) 13345–13355.
- [86] X. Li, Y. Song, Y. Bian, C. Gu, X. Yang, F. Wang, et al., Insights into the mechanisms underlying efficient Rhizodegradation of PAHs in biochar-amended soil: from microbial communities to soil metabolomics, *Environ. Int.* 144 (2020) 105995.
- [87] N.W. Sokol, M.A. Bradford, Microbial formation of stable soil carbon is more efficient from belowground than aboveground input, *Nat. Geosci.* 12 (2019) 46–53.
- [88] S.G. Tringe, E.M. Rubin, Metagenomics: DNA sequencing of environmental samples, *Nat. Rev. Genet.* 6 (2005) 805–814.
- [89] J. Handelsman, Metagenomics: application of genomics to uncultured microorganisms, *Microbiol. Mol. Biol. Rev.* 69 (2005), 195–195.
- [90] S.R. Lindemann, H.C. Bernstein, H.-S. Song, J.K. Fredrickson, M.W. Fields, W. Shou, et al., Engineering microbial consortia for controllable outputs, *ISME J.* 10 (2016) 2077–2084.

- [91] C. Cardona, P. Weisenhorn, C. Henry, J.A. Gilbert, Network-based metabolic analysis and microbial community modeling, *Curr. Opin. Microbiol.* 31 (2016) 124–131.
- [92] J.T. Morton, A.A. Aksenen, L.F. Nothias, J.R. Foulds, R.A. Quinn, M.H. Badri, et al., Learning representations of microbe-metabolite interactions, *Nat. Methods* 16 (2019) 1306–1314.
- [93] X. Li, C. Qu, Y. Bian, C. Gu, X. Jiang, Y. Song, New insights into the responses of soil microorganisms to polycyclic aromatic hydrocarbon stress by combining enzyme activity and sequencing analysis with metabolomics, *Environ. Pollut.* 255 (2019) 113312.
- [94] S. Pontrelli, R. Szabo, S. Pollak, J. Schwartzman, D. Ledezma-Tejeida, O.X. Cordero, et al., Metabolic cross-feeding structures the assembly of polysaccharide degrading communities, *Sci. Adv.* 8 (2022) eabk3076.
- [95] L. Qiu, T.J. Daniell, S.A. Banwart, M. Nafees, J. Wu, W. Du, et al., Insights into the mechanism of the interference of sulfadiazine on soil microbial community and function, *J. Hazard Mater.* 419 (2021) 126388.
- [96] S. Lu, Y. He, Y. Chen, L. Chen, Z. Wang, J. Yuan, et al., Co-analysis of rhizosphere metabolomics and bacterial community structures to unfold soil ecosystem health in Camellia oleifera land under long-term cultivation, *Appl. Soil Ecol.* 171 (2022) 104336.
- [97] A. Tejero Rioseras, K.D. Singh, N. Nowak, M.T. Gaugg, T. Bruderer, R. Zenobi, et al., Real-time monitoring of tricarboxylic acid metabolites in exhaled breath, *Anal. Chem.* 90 (2018) 6453–6460.
- [98] C.R. Warren, Simultaneous efflux and uptake of metabolites by roots of wheat, *Plant Soil* 406 (2016) 359–374.
- [99] C.R. Warren, Pools and fluxes of osmolytes in moist soil and dry soil that has been re-wet, *Soil Biol. Biochem.* 150 (2020) 108012.
- [100] R.I. Griffiths, M. Manefield, N. Ostle, N. McNamara, A.G. O'Donnell, M.J. Bailey, et al., ( $\text{CO}_2$ )-C-13 pulse labelling of plants in tandem with stable isotope probing: methodological considerations for examining microbial function in the rhizosphere, *J. Microbiol. Methods* 58 (2004) 119–129.
- [101] S. Chowdhury, M. Farrell, N. Bolan, Photoassimilated carbon allocation in a wheat plant-soil system as affected by soil fertility and land-use history, *Plant Soil* 383 (2014) 173–189.
- [102] S. Chowdhury, M. Farrell, N. Bolan, Priming of soil organic carbon by malic acid addition is differentially affected by nutrient availability, *Soil Biol. Biochem.* 77 (2014) 158–169.
- [103] D.B. Kell, M. Brown, H.M. Davey, W.B. Dunn, I. Spasic, S.G. Oliver, Metabolic footprinting and systems biology: the medium is the message, *Nat. Rev. Microbiol.* 3 (2005) 557–565.
- [104] M. de Raad, Y.V. Li, J.V. Kuehl, P.F. Andeer, S.M. Kosina, A. Hendrickson, et al., A defined medium for cultivation and exometabolite profiling of soil bacteria, *Front. Microbiol.* 13 (2022) 855331.
- [105] K.T. Cyle, A.R. Klein, L. Aristilde, C.E. Martinez, Dynamic utilization of low-molecular-weight organic substrates across a microbial growth rate gradient, *J. Appl. Microbiol.* 133 (2022) 1479–1495.
- [106] B. Hassaniour, L. Aristilde, Redox-related metabolic dynamics imprinted on short-chain carboxylic acids in soil water extracts: a C-13-Exometabolomics analysis, *Environ. Sci. Technol. Lett.* 8 (2021) 183–191.
- [107] S.M. Kosina, A.M. Greiner, R.K. Lau, S. Jenkins, R. Baran, B.P. Bowen, et al., Web of microbes (WoM): a curated microbial exometabolomics database for linking chemistry and microbes, *BMC Microbiol.* 18 (2018) 115.
- [108] M.I. Mhlongo, L.A. Piater, N.E. Madala, N. Labuschagne, I.A. Dubery, The chemistry of plant-microbe interactions in the rhizosphere and the potential for metabolomics to reveal signaling related to defense priming and induced systemic resistance, *Front. Plant Sci.* 9 (2018) 112.
- [109] K. Zhalmina, K.B. Louie, Z. Hao, N. Mansoori, U.N. da Rocha, S. Shi, et al., Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly, *Nat. Microbiol.* 3 (2018) 470–480.
- [110] S.B. Miller, A.L. Heuberger, C.D. Broeckling, C.E. Jahn, Non-targeted metabolomics reveals sorghum rhizosphere-associated exudates are influenced by the belowground interaction of substrate and sorghum genotype, *Int. J. Mol. Sci.* 20 (2019) 431.
- [111] M. Cao, F. Liu, L. Sun, Y. Wang, J. Wan, R. Wang, et al., Floccularia luteovirens modulates the growth of alpine meadow plants and affects soil metabolite accumulation on the Qinghai-Tibet Plateau, *Plant Soil* 459 (2021) 125–136.
- [112] L.E. Nemadzzi, J. Vervoort, G. Prinsloo, NMR-based metabolomic analysis and microbial composition of soil supporting Burkea africana growth, *Metabolites* 10 (2020) 402.
- [113] J.E. Kammenga, R. Dallinger, M.H. Donker, H.R. Kohler, V. Simonsen, R. Triebeskorn, et al., Biomarkers in terrestrial invertebrates for ecotoxicological soil risk assessment, *Rev. Environ. Contam. Toxicol.* 164 (2000) 93–147.
- [114] M.-V. Salvia, A. Ben Jrad, D. Raviglione, Y. Zhou, C. Bertrand, Environmental Metabolic Footprinting (EMF) vs. half-life: a new and integrative proxy for the discrimination between control and pesticides exposed sediments in order to further characterise pesticides' environmental impact, *Environ. Sci. Pollut. Res.* 25 (2018) 29841–29847.
- [115] A. Guo, C. Pan, J. Ma, Y. Bao, Linkage of antibiotic resistance genes, associated bacteria communities and metabolites in the wheat rhizosphere from chlorpyrifos-contaminated soil, *Sci. Total Environ.* 741 (2020) 140457.
- [116] P. Xue, X. Liu, L. Zhao, J. Zhang, Z. He, Integrating high-throughput sequencing and metabolomics to investigate the stereoselective responses of soil microorganisms to chiral fungicide cis-epoxiconazole, *Chemosphere* 300 (2022) 134198.
- [117] W. Li, P. Zhang, H. Qiu, C.A.M. Van Gestel, W.J.G.M. Peijnenburg, X. Cao, et al., Commonwealth of soil health: how do earthworms modify the soil microbial responses to  $\text{CeO}_2$  nanoparticles? *Environ. Sci. Technol.* 56 (2021) 1138–1148.
- [118] X. Li, S. Yao, Y. Bian, X. Jiang, Y. Song, The combination of biochar and plant roots improves soil bacterial adaptation to PAH stress: insights from soil enzymes, microbiome, and metabolome, *J. Hazard Mater.* 400 (2020) 123227.
- [119] Y. Sun, C. Duan, N. Cao, C. Ding, Y. Huang, J. Wang, Biodegradable and conventional microplastics exhibit distinct microbiome, functionality, and metabolome changes in soil, *J. Hazard Mater.* 424 (2022) 127282.
- [120] C. Wu, Y. Ma, D. Wang, Y. Shan, X. Song, H. Hu, et al., Integrated microbiology and metabolomics analysis reveal plastic mulch film residue affects soil microorganisms and their metabolic functions, *J. Hazard Mater.* 423 (2022) 127258.
- [121] A. Guo, C. Pan, X. Su, X. Zhou, Y. Bao, Combined effects of oxytetracycline and microplastic on wheat seedling growth and associated rhizosphere bacterial communities and soil metabolite profiles, *Environ. Pollut.* 302 (2022) 119046.
- [122] M. Miura, P.W. Hill, D.L. Jones, Impact of a single freeze-thaw and dry-wet event on soil solutes and microbial metabolites, *Appl. Soil Ecol.* 153 (2020) 103636.
- [123] H. Cheng, M. Yuan, L. Tang, Y. Shen, Q. Yu, S. Li, Integrated microbiology and metabolomics analysis reveal responses of soil microorganisms and metabolic functions to phosphorus fertilizer on semiarid farm, *Sci. Total Environ.* 817 (2022) 152878.
- [124] R.M. Salek, C. Steinbeck, M.R. Viant, R. Goodacre, W.B. Dunn, The role of reporting standards for metabolite annotation and identification in metabolomic studies, *GigaScience* 2 (2013) 13.
- [125] S. Alseehk, A. Aharoni, Y. Brotman, K. Contrepois, J. D'Auria, J. Ewald, et al., Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices, *Nat. Methods* 18 (2021) 747–756.
- [126] J.A. Kirwan, H. Gika, R.D. Beger, D. Bearden, W.B. Dunn, R. Goodacre, et al., Quality assurance and quality control reporting in untargeted metabolic phenotyping: mQACC recommendations for analytical quality management, *Metabolomics* 18 (2022) 70.
- [127] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, et al., Proposed minimum reporting standards for chemical analysis, *Metabolomics* 3 (2007) 211–221.
- [128] W. Bitremieux, M. Wang, P.C. Dorrestein, The critical role that spectral libraries play in capturing the metabolomics community knowledge, *Metabolomics* 18 (2022) 94.
- [129] S. Xing, S. Shen, B. Xu, X. Li, T. Huan, BUDDY: molecular formula discovery via bottom-up MS/MS interrogation, *Nat. Methods* 20 (2023) 881–890.
- [130] Y. Hu, Q. Zheng, W. Wanek, Flux analysis of free amino sugars and amino acids in soils by isotope tracing with a novel liquid chromatography/high resolution mass spectrometry platform, *Anal. Chem.* 89 (2017) 9192–9200.
- [131] P. Baldrian, The known and the unknown in soil microbial ecology, *FEMS Microbiol. Ecol.* 95 (2019) fiz005.
- [132] W. Wanek, M. Mooshammer, A. Bloechl, A. Hanreich, K. Keiblinger, S. Zechmeister-Boltenstern, et al., Determination of gross rates of amino acid production and immobilization in decomposing leaf litter by a novel N-15 isotope pool dilution technique, *Soil Biol. Biochem.* 43 (2011) 221.
- [133] C.R. Warren,  $\text{D}_2\text{O}$  labelling reveals synthesis of small, water-soluble metabolites in soil, *Soil Biol. Biochem.* 165 (2022) 108543.