Cover plant functional types alter the abundance and composition of hydrophobic compounds: the relationship with soil water repellency on the Chinese Loess Plateau

Xiaohong Chai (cxhnwafu@163.com)
Northwest A&F University https://orcid.org/0009-0008-6592-9765
Weiwei Wang
Xiuzi Ren
Junfeng Wang
Qi Zhang
Gaohui Duan
Yuanyuan Qu
Xuexuan Xu
https://orcid.org/0000-0002-1889-9646
Feng Du

Research Article

Keywords: Plant functional types, Soil water repellency, Soil organic carbon, Hydrophobic compounds, Biomarkers

Posted Date: June 7th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2954393/v1

Abstract

Background and aims

It is widely accepted that soil water repellency (SWR) is mainly caused by plant-derived hydrophobic compounds in soils. The relation between these hydrophobic compounds, which are defined as SWR biomarkers, and SWR has been rarely known and the knowledge of their sources remains controversial. We aimed to select key SWR biomarkers predicting SWR and to trace their origin.

Methods

Topsoils under/around five dominant plant species (DPS) belonging to various plant functional types (PFTs) on the Chinese Loess Plateau were sampled, together with plant samples, i.e., plant leaves and roots. A sequential extraction procedure and hydrolysis approach was applied to obtain three organic fractions: dichloromethane (DCM)/MeOH soluble fraction (D), DCM/MeOH soluble fraction of isopropanol/ammonia solution (IPA/NH_3) extract (AS), and DCM/MeOH insoluble fraction of IPA/NH_3 extract (AI), which were analyzed by gas chromatography-mass spectrometry.

Results

The two-way hierarchical clustering analysis revealed a distinct division of soil organic carbon composition under different DPS, and the leaves of evergreen trees offered more cutin than those of other PFTs. In addition, structural equation modeling showed that AS cutin (path coefficient = 0.30) and AI cutin (path coefficient = 0.47) had direct and positive effects on SWR. Moreover, there was a strong link between SWR and the ratio of the two separate compound groups when AS cutin or AI cutin were taken as the numerators.

Conclusion

After considering the SWR behavior during extraction and the chemical composition of each fraction, we concluded that leaf-derived cutin appears to have the greatest effect.

Introduction

Soil water repellency (SWR) is a widespread phenomenon that occurs in forest and agricultural soils under various plant cover and climatic conditions (Doerr et al. 2000; Jordán et al. 2013). It refers to the reduced ability of soil to absorb water, leading to the formation of water-repellent zones (Doerr et al. 2000; González-Pérez et al. 2008). The presence of SWR has significant negative impacts on soil health and ecosystem functioning. One of the most significant impacts is soil waterlogging in lowland areas, as water is not absorbed by the soil but instead carries soil particles with it, increasing surface runoff and soil erosion in highlands (Coelho et al. 2005; Urbanek et al. 2015; DeBano et al. 2000a; Doerr and Thomas 2000; Smettem et al. 2021; Zheng et al. 2017). Moreover, it hinders plant growth and lowers crop yields by reducing the water availability to plants (Doerr and Thomas 2000; Li et al. 2019; Müller et al. 2016). SWR also affects the cycling of nutrients, the storage of carbon, and microbial activity & function in soils, all of which have a cascading effect on ecosystem services (Doerr et al. 2000; Schonsky et al. 2014). For example, microbial activity can be inhibited in hydrophobic soils, leading to decreased nutrient cycling and organic matter decomposition (Muñoz-Rojas et al. 2012; Kraemer et al. 2019). This may in turn affects plant growth and soil health. In addition, the increased runoff and erosion leveled up by SWR contributes to the loss of nutrients locally and the contamination elsewhere (Kodešová et al. 2008). This also affects the quality of surface and groundwater, as contaminants are carried away by runoff. In conclusion, SWR has negative impacts on soil and plant health, as well as on the hydrology and ecology of ecosystems. Understanding the mechanism and management of SWR is an important area of research with broad implications for agriculture, forestry, environmental science, and urban development. By developing effective management strategies to mitigate the negative impacts of SWR, we can promote sustainable land use and protect the health and productivity of our ecosystems.

Although SWR is common in soils and has many detrimental effects on ecosystems, its causes are still not completely understood. Some studies have focused on investigating the factors that contribute to SWR, including soil texture, moisture content, temperature, microbial activity, and soil organic carbon (SOC) content (DeBano 2000b; Arcenegui et al. 2007; Mataix-Solera et al. 2007; Chau et al. 2014; Karunarathna et al. 2010; White et al. 2017; Hermansen et al. 2019; Balshaw et al. 2020). For example, Dekker and Ritsema (1996) showed that SWR may be influenced by the soil texture and structure, as well as the concentration and chemical composition of hydrophobic compounds. González-Pérez et al. (2004) presented a new view of the chemical structure of soil organic matter (SOM) and its relationship with SWR. Jiménez-Morillo et al. (2022) developed a new test procedure to evaluate the influence of particle-size distribution on SWR. Specifically, researches have shown that the presence of certain bacterial species can contribute to SWR (Chai et al. 2022; Seaton et al. 2019; Song et al. 2019). Therefore, the exact mechanism causing SWR remains a subject of debate, but an increasing number of studies consider hydrophobic organic compounds (i.e., SWR biomarkers), which are produced by plant exudates, microbial activity, or fire-induced changes in soil organic matter (Doerr et al. 2000; Jiménez-Morillo et al. 2022), as a critical factor (Jiménez-Pinilla et al. 2016; González-Peñaloza et al. 2013; Doerr et al. 2005). For instance, de Blas et al. (2013) argued that the concentration of plant-derived extractable free lipids in soils under pine and eucalyptus was positively liked to SWR, while those of bound lipids were not linked to SWR. Mao et al. (2014) speculated that high molecular weight plant-derived suberin was more capable of predicting SWR than low molecular weight lipids. However, the mechanisms of how hydrophobic substances interact with soil particles and alter SWR are still uncertain. Several theories have been proposed to explain this phenomenon, but the most widely accepted one is the “coating” or “film” theory. This theory contends that soil particles are covered in hydrophobic materials, which form aggregates that repel water (DeBano 2000a; Doerr et al. 2000). According to this theory, hydrophobic substances bind to soil particles through van der Waals forces, hydrogen bonding, or chemical reactions, forming a hydrophobic film that repels water (DeBano 2000a; Doerr et al. 2000; Doerr and Thomas 2000; Bayry et al. 2012; Mao et al. 2019; Smettem et al. 2021; Leighton-Boyce et al. 2007). Overall, the understanding of the mechanisms underlying SWR formation is still evolving, and further research is needed to fully elucidate this complex phenomenon.

In natural ecosystems, various dominant plant species (DPS) that belong to distinct plant functional types (PFTs) may differentially alter the degree of SWR. For instance, Rodríguez-Alleres and Benito (2011) noted that soils under *Pinus pinaster* and *Eucalyptus globulus* primary forests always showed more severe SWR than those under other plant species during the summer dry conditions in NW Spain. Lozano et al. (2013) found that higher occurrence of SWR under natural *Pinus* and *Quercus*, while all *Cistus* and bare soil samples were wettable. Alanís et al. (2016) observed that the intensity of SWR under shrubs could be wettable to slight WR, but soils under native trees (fir, fir-pine-oak, and pine-oak forest) were very wettable to extreme WR. Similarly, Jordán et al. (2009) and Jordán et al. (2013) reported relatively high values of SWR in soils from Mexican volcanic highlands, particularly beneath the canopies of fir and pines, while inexistent or slight WR was observed in soils at the bare ash-covered areas. Based on the previous works (Butzen et al. 2015; Doerr et al. 2005; de Blas et al. 2010, 2013; Mao et al. 2014; Zavala et al. 2009; Lozano et al. 2013; Alanís et al. 2016; Walden et al. 2015), the proportion of water-repellent soil around the distinct PFTs is as follows: bare soil < shrubs and herbs < evergreen trees. Most studies investigating SWR have focused on indigenous species-dominated plants, such as grassland in the United Kingdom (Seaton et al. 2019), permanent pastures in the Netherlands (Morley et al. 2005; Mao et al. 2014; Dekker et al. 2018), and forest in Spain (Rodríguez-Alleres and Benito 2011, 2012; Lozano et al. 2013, 2014). The distribution of SWR in artificial forest ecosystems with various plant cover is still unknown, nevertheless.

The Chinese Loess Plateau ($6.4 \times 105\ km^2$) is a semi-arid region and is considered one of the most eroded areas (mean soil loss rate: $2860\ t\ km^{-2}\ year^{-1}$); it is characterized by an extremely complex soil-eroding catena (Wang et al. 2017). Loess is the most widely distributed soil in the Chinese Loess Plateau. However, little is known about the behavior of SWR in loess soils, which are the initial soils in erosive environments with very low SOC content (Chai et al. 2022). In recent decades, with the accumulation of SOC after the world’s largest vegetation restoration project “Grain for Green”, the SWR of loess has emerged as an important topic (Liu and Zhan 2019). Improved studying of the effects of cover PFTs in loess on SWR and clarifying the possible relationship between SWR and SWR biomarkers could have significant implications for developing effective strategies to manage soil water availability, particularly in arid and semi-arid regions where water is scarce.

To investigate the main reasons for SWR differences under different DPS (i.e., *Pinus tabulaeformis* Carr., *Robinia pseudoacacia* L., *Hippophae rhamnoides* L., *Coronilla varia* L., *Agropyron cristatum* (L.) Gaertn.) belonging to various PFTs (i.e., evergreen trees, deciduous trees, shrubs, legumes, grasses). In this paper, we studied the SOM accrual and chemical composition of topsoils under/around these five DPS on the Chinese Loess Plateau. More importantly, we used a sequential extraction to divide SWR

biomarkers into three individual extraction fractions (EFs) and quantified the major plant-derived (cutin, suberin, and long-chain fatty acids) and microbial-derived biomarkers (short-chain fatty acids) in these different solvent EFs to assess the relationship between the input and sequestration of source-specific SWR biomarkers with SWR, thus improving our mechanistic understanding of SWR occurrence at the molecular level. Based on the SWR performance during extraction, the different chemical composition of each EF, and the distinct PFTs among the five DPS, we hypothesized that (1) evergreen trees would be more likely to cause more severe SWR than other PFTs; (2) all soils under/around different DPS would have a greater proportion of plant-derived C than microbial-derived C, and plant-derived C would be the main cause of SWR; and (3) the abundance and composition of SWR biomarkers vary depending on PFTs; evergreen trees would provide more SWR-inducing cutin and less long-chain fatty acids (LFA) compared with deciduous trees, shrubs, legumes, and grasses.

Materials and methods

Sampling

This study was conducted in the Wangdonggou watershed of Changwu Country, Shaanxi Province, China (35°13'N, 107°40'E; 1220 m a.s.l.). The study watershed has an area of 6.3 km^2, and the total length of the main gully was 4.97 km (Wang et al. 2021). Wangdonggou is located in the typical loess hilly and gully region of the Chinese Loess Plateau, and the topography is highly fragmented. The region features a semi-arid continental climate, with an annual mean temperature of 9.1°C. The annual mean precipitation is 568 mm, and predominantly occurs from July to August as short heavy storms. The frost-free period is 194 days, and potential evapotranspiration is 967 mm (Huang et al. 2003). There are four vegetation types in the watershed, including two typical artificial forests and two herbaceous vegetation. One artificial forest is dominated by *Pinus tabulaeformis* Carr. (PT), and the other one is *Robinia pseudoacacia* L. - *Hippophae rhamnoides* L. (RP-HR) forest; the two herbaceous vegetation are dominated by C*oronilla varia* L. (CV), and *Agropyron cristatum* (L.) Gaertn. (AC), respectively. The plant functional types (PFTs) of these five dominant species are as follows: evergreen trees, deciduous trees, shrubs, legumes, and grasses. All soils were classified as typical silt loam (USDA soil texture classes), developed from aeolian loess parent materials. More details on soil texture are given in Chai et al. (2022). Expect for the topsoil under PT and RP, most topsoil showed a light brownish or yellow-brown tint, of which the litter layer was scarcely existent and grass leaves decompose quickly, but debris of plant leaves and many fine roots were observed. The topsoil beneath PT and RP, however, had a dark grey or brownish hue, and a 1–3 cm litter layer was discovered, although only a few roots were visible.

Locations of the four vegetation types were chosen for the collection of plant and soil samples in the watershed. The selected sampling locations shouldn't have any human activity interfere with them, and have comparable soil type, vegetation restoration years, elevation, slope gradient, and previous tillage operations, to eliminate possible effects of non-experimental factors on SWR. The study region and sampling locations are summarized in Fig. 1 and Table 1. For all of the five dominant species, six stems for replicates were selected to take plant samples, and soil samples under/around them, as well. Although high variability of SWR was observed in our previous work (Chai et al. 2022), to avoid as far as possible distance-dependent sampling of SWR and other measurements, the six selected stems should not be in the neighborhood. Under/around each stem of each species, three micro-sampling positions with more than 50 cm spacing separation were placed to eliminate sampling errors. For trees and shrubs, the soil micro-sampling positions were located at an intermediate distance between the trunk and the edge of the canopy, where the microhabitat should be most favorable for the degradation of plant litter by soil microorganisms (Moro et al. 1997). However, micro-sampling positions of herbaceous plants were as close to the stem as possible to ensure that the withered litter in the collected soil samples were formed by the research objects. Additionally, in the RP-HR forest, it was nevertheless possible to sample at micro-sampling places under each stem of each species without interfering with the others because of the relatively low density of vegetation. In May 2021, when the topsoil was almost air-drying, soil samples were taken from 0–3 cm thicknesses of topsoil at micro-sampling positions under/around each of the five dominating plant species (n = 18 per species). The living plant leaf and root samples were taken separately from each plant species (n = 6 per species). Prior to further analysis, all collected soils were air-dried to constant weight at room temperature (20–25 °C) and sieved (2 mm) to remove coarse soil particles. All the plant samples were freeze-dried and stored in a dry place.

Table 1
The basic information for each sampling location

Sampling locations	Vegetation restoration years (a)	Longitude	Latitude	Altitude (m)	Slope (°)	Coverage (%)	Accompanying species
PT	30	107°38'22"E	35°10'9"N	1126	10	85.4	*Bothriochloa ischaemum* (L.) Keng
RP-HR	30	107°41'55"E	35°12'32"N	1062	31	88.2	*Portulaca oleracea* L.; *Rubus corchorifolius* L. f.; *Agropyron cristatum* (L.) Gaertn.; *Arundo donax* L.
CV	30	107°42'7"E	35°15'3"N	1100	5	83.4	*Agropyron cristatum* (L.) Gaertn.
AC	30	107°41'21"E	35°12'54"N	1098	9	70.7	*Arundo donax* L.; *Artemisia gmelinii* Weber ex Stechm.

Soil characteristics and water repellency measurements

The soil organic matter (SOM) and total nitrogen (TN) were measured by the K_2CrO_7-H_2SO_4 oxidation method and the Kjeldahl method, respectively (Fu et al. 2011; Nelson and Sommers 1982; Bremner and Mulvaney 1982). The SOM value was converted to a total organic carbon (TOC) value using the established 1.724:1 SOM: TOC ratio.

For measuring SWR, approximately 15 g of soil was placed on a 50-mm diameter aluminum box and exposed to a controlled laboratory (20°C, ~ 50% relative humidity) for one week to eliminate the potential impacts of preceding atmospheric humidity on SWR. The SWR was performed using the water drop penetration time (WDPT) test. This simple test measures the average time taken for three 100 µl water droplets to penetrate into the soil. The method is widely known and accepted to assess the persistence of SWR (Van't Woudt 1959; Wessel 1988; Dekker and Ritsema 1994). The water-repellent classes were divided into five grades with different intervals according to Bisdom et al. (1993), with WDPT ≤ 5 s representing wettable and WDPT > 5 s water-repellent conditions. According to Lozano et al. (2013), to linearly rate the SWR, the logarithm of the WDPT value was used to determine a sample was water-repellent or not. Specially, if the value of $\log_{10}$ WDPT is > 0.7, it was considered as water-repellent (Table 2). After the determination of the WDPT value, a series of samples under/around the same DPS with different $\log_{10}$ WDPT intervals were obtained for subsequent SWR biomarkers extraction experiments. Finally, a total of 18, 12, 9, 9, and 6 soil samples were obtained from PT, RP, HR, CV, and AC. To compare the changes after the extraction processes, the WDPT measurements of air-dried soils were also conducted after each extraction.

Table 2
WDPT classes and class increments used in the present study.

Repellency rating	Wettable	Water repellency[a]											
		Slight				Strong				Severe			Extreme
WDPT classes	≤ 5	10	20	30	60	120	180	300	600	900	1800	3600	> 3600
WDPT interval (s)	≤ 5	6–10	11–20	21–30	31–60	61–120	121–180	181–300	301–600	601–900	901–1800	1801–3600	> 3600
$\log_{10}$ WDPT interval[b]	≤ 0.7	0.7–1.0	1.0–1.3	1.3–1.5	1.5–1.8	1.8–2.1	2.1–2.3	2.3–2.5	2.5–2.8	2.8–3.0	3.0–3.3	3.3–3.6	> 3.6

[a] based on Bisdom et al. (1993); [b] based on Lozano et al. (2013).

Soil and plant extraction procedures

To investigate SWR biomarkers, sequential extraction methods have been applied to all the soils and the plant samples. lipids were extracted by subsequently dichloromethane/methanol (DCM/MeOH) (9:1 *v: v*) and isopropanol/ammonia solution (IPA/NH_3) (7:3 *v:*

v, 25–28% ammonia solution).

For extracting free lipids from samples, according to the method of Bull et al. (2000), each dried soil, leaf, and root sample was accurately weighed and extracted using a Soxhlet apparatus (AI-ZFCDY-6Z, Na ai Co., Ltd., Shanghai, China) containing DCM/MeOH (9:1 *v: v*) at 70°C for 24 h to obtain the D fraction. The solvent was removed with a rotary evaporator (BUCHI Lab. AG, Flawil, Switzerland R-215). After redissolving the lipids in the solvent, extracts were passed through an SPE column filled with anhydrous Na_2SO_4 (2000 mg, 6 ml) to remove any water and were dried using a gentle stream of nitrogen, and then stored at -20°C.

The residual soils were air-dried and extracted by IPA/NH_3 (7:3 *v: v*) using the Soxhlet apparatus at 95°C for 48 h. After extraction, the solvent also needs to be concentrated using the rotary evaporator. The soluble lipids (AS fraction) were separated from the dried IPA/NH_3 extracts by DCM/MeOH (9:1 *v: v*), and the residues resulted in AI fractions, which involved ester bonds. All solutions were combined and both this combined solution and the insoluble residue were dried under nitrogen, and then stored at -20°C.

Prior to analysis, all the D and AS fractions of the soils and DCM/MeOH extracts of the plants were methylated using 100 µl of (trimethylsilyl)diazomethane (TMS-CH_2N_2) at room temperature. The AI fractions and the lipid-free dried leaves and roots were depolymerized through trans-methylation using BF_3–MeOH at 70°C for 24 h (Riederer et al., 1993). Then, all the extracts were eluted over a small silica gel (100–200 mesh) column with ethyl acetate and were silylated using N, O-bis (trimethylsilyl)trifluoroacetamide (BSTFA) at 70°C for 30 min.

Derivatized extracts were analyzed using a triple quadrupole gas chromatography-mass spectrometry (GC-MS) GCMS-TQ8050NX instrument (Shimadzu Production Co., Kyoto, Japan) with a mass range of m/z 50–800. 1 µl of the derivatized extracts was injected onto an SH-Rxi-5Sil MS capillary column (Shimadzu 30 m × 0.25 mm inner diameter × 0.25 µm film thickness), using helium at a 1.0 ml min^{-1} constant flow rate as the carrier gas. The oven heating program started with an initial temperature of 70°C, increased to 130°C at 20°C min^{-1}, heated at 4°C min^{-1} from 130°C to 320°C, and finally held at 320°C for 20 min. Retention times combined with mass spectra were employed to identify compounds. A known amount of squalane was added to the extracts as an internal standard. Identification of the compounds was carried out by their mass spectra using NIST libraries, interpreted spectra, retention times, or comparison to literature data, and quantified by GC-MS chromatographic peak area integration and according to the following formula to correct for possible co-eluting compounds.

$$m_i = A_i / A_{is} \times m_{is} \times f_{i'} \times \alpha$$

1

Where m_i = the quality of the object to be measured; A_i = the peak area of the object to be measured; A_{is} = the peak area of the internal standard; m_{is} = the quality of the internal standard; $f_{i'}$ = the relative correction factor; α = the conversion coefficient.

Statistical data analysis

All data were tested for normality and homogeneity of variance using SPSS Statistics 26 (IBM crop., USA), and log-transformed when necessary, and One-way ANOVA combined with LSD test was used to evaluate the significance of differences in soil properties and SWR biomarkers among different DPS. To better visualize the differences in specific SWR biomarkers across soil samples with different log_{10} WDPT (s) under/around different DPS in each EF, the two-way hierarchical clustering analysis (HCA) presented in heatmaps was performed using the “pheatmap” package in R version 4.1.2 (R Core Team 2021). HCA was conducted using Euclidean distance, with one individual/sample corresponding to one column of the data matrix and each row of the data matrix representing the composition variables, such as the concentrations of metabolites, proteins, or biomarkers (Meunier et al. 2007; Suseela et al. 2015). Simple linear regression was performed to quantify the relationship between soil characteristics, the concentrations of single SWR biomarkers, and SWR (i.e., the measured WDPT value) on a logarithmic scale (log_{10} (s)). Here regression analysis was carried out on the absolute amount (µg g^{-1} soil) and the relative amount (µg g^{-1} TOC) of SWR biomarkers to evaluate both the quantitative and qualitative effects. Correlation analysis between each compound group and SWR was presented in heatmaps using the “corrplot” package in R version 4.1.2. Meanwhile, a structural equation modeling (SEM) was constructed to further explore the possible associations between SWR and SWR biomarkers at a grouped level using the Amos 21.0.0 software package (IBM SPSS Amos, Chicago, IL, USA). This approach involves proposing a causal model, considering both direct and indirect paths, then fitting the data and critically evaluating the proposed causal model (Seaton et al. 2019). All the data were

normalized before modeling. In this study, the quantity of each compound group was defined as an absolute amount (μg g^{-1} soil), and the quality as the ratio of the concentrations of the two different compound groups. In addition, functional compound groups were distinguished according to EFs (D, AI, and AS) and their compound classes, i.e., fatty acids, alkanols, alkanes, *ω*-hydroxyalkanoic acids, or *α, ω*-dicarboxylic acids.

Results

Soil chemical properties and SWR

The TOC content, TN content, C/N ratio, and WDPT value of all samples are shown in Table S1. Only soils under PT have strong to extreme WR, while most soils under/around other DPS were classified as slight WR (Table S1). The samples had TOC contents of 16.22 to 69.52 g kg^{-1} of soil, TN contents of 1.44 to 4.92 g kg^{-1} of soil, and C/N ratios of 7.46 to 20.41, respectively. PT6 had the highest TOC content and SWR, and AC1 had the lowest TOC content and SWR. After the WDPT test, most of the measured samples showed water repellency (Table S1). On the other hand, the occurrence of SWR was different among species and was higher under PT. Soil TOC content under/around the different DPS is shown in Fig. 2a, being higher in samples of PT and RP, with mean values of 48.01 g kg^{-1} soil and 52.92 g kg^{-1} soil respectively. Soil TN content and C/N ratio under/around the different DPS are shown in Fig. 2b–c. Due to the nitrogen fixation function of leguminous plants, RP and CV had a higher TN content. However, samples under CV had the lowest C/N ratio, with an average value of 8.05%

Before extraction, $\log_{10}$ TOC showed a strong positive linear correlation with SWR ($r^2 = 0.66$, $p < 0.001$) (Fig. 2d). Moreover, Pearson's correlation was positive but weak between SWR and $\log_{10}$ TN ($r^2 = 0.32$, $p < 0.05$) (Fig. 2e) However, there was no relationship between C/N ratio and SWR (Fig. 2f). After DCM/MeOH extraction, SWR of most soils increased ($p > 0.05$) (Fig. 3), while the average TOC content, TN content and C/N ratio of all soils decreased ($p > 0.05$). Significant positive correlations were found between $\log_{10}$ TOC and SWR for the soils ($r^2 = 0.67$, $p < 0.001$) (Fig. 2d), and $\log_{10}$ TN also showed a weak positive linear correlation with SWR ($r^2 = 0.40$, $p < 0.01$) (Fig. 2e). Most soils became wettable after IPA/NH_3 extraction (Fig. 3), again the average TOC content, TN content and C/N ratio of all soils decreased ($p > 0.05$), and the mean SWR of all soils significantly reduced ($p < 0.001$) (Fig. 3). Meanwhile, there was no discernible relationship between TOC, TN, or C/N ratio and SWR ($p > 0.05$) (Fig. 2d–f).

Single biomarkers and SWR

Single biomarkers from soils

The SWR biomarkers identified by sequential extraction methods from all soils are presented in Table S2. Alkanoic (alkenoic) acids (C_{16}–C_{32}), alkanols (C_{18}–C_{28}), and alkanes (C_{17}–C_{31}) were the three main types of SWR biomarkers identified in the D fractions. As can be seen clearly in the distribution of single SWR biomarkers in the D fractions (Fig. 4a), identified alkanoic/alkenoic acids and alkanols from all soils showed a strong even-over-odd preference, whereas the alkanes discovered performed an odd-over-even predominance. Typical root biomarkers such as C_{22} *ω*-hydroxyalkanoic acids and C_{22} *α, ω*-dicarboxylic acids were also found in D fractions from all soils. Most compounds in the AS/AI fractions also appeared in the D fractions, e.g., alkanoic acids and alkanols comprised the two largest groups of compounds in three fractions for all soils. Alkanoic/alkenoic acids (C_{16}–C_{32}), alkanols (C_{16}–C_{32}), *ω*-hydroxyalkanoic acids (C_{16}–C_{25}), and *α, ω*-dicarboxylic acids (C_{16}–C_{30}) were observed in AS (Fig. 4b) and AI fractions (Fig. 4c), whereas no alkanes were found in these fractions. *ω*-Hydroxyalkanoic acids with dominantly even-carbon-numbered chains were found to be the second most abundant SWR biomarkers in AS and AI fractions, and the *α, ω*- dicarboxylic acids were also dominated by even-numbered carbon chains. Generally, D alkanoic/alkenoic acids, AS alkanoic/alkenoic acids and D alkanols were the three main compound groups in all soils (Fig. 4d), in which the relative amounts of D alkanoic acids were on average 35.67% and 63.74% higher than those of AS alkanoic/alkenoic acids and D alkanols, respectively.

The two-way hierarchical clustering of the concentrations of specific biomarkers further revealed different grouping patterns across different PFTs with different $\log_{10}$ WDPT (s) treatments based on different compound classes (Fig. 5 and Table S2). Apart from PT1 and RP2, the absolute amount (μg g^{-1} soil) and relative amount (μg g^{-1} TOC) of biomarkers both showed a clear separation across soils under/around the different DPS, suggesting that differences in biomarkers after DCM/MeOH extraction were dependent on PFTs rather than treatments with different $\log_{10}$ WDPT (s) (Fig. 5a and d), e.g., HR showed a clear separation from PT, RP, CV, and AC,

which had higher concentrations of alkanes and alkanoic acids (< C_{22}) and alkanol, while PT and RP with higher WDPT value contained more even-carbon-numbered chains alkanoic acids (> C_{22}) and alkanols, especially C_{22} *α, ω*-dicarboxylic acid (a suberin-derived biomarker) was abundant in PT (Fig. 5a). In AS fractions (Fig. 5b and e) and AI fractions (Fig. 5c and f), the soil biomarkers belonging to the same DPS clustered more clearly, showing a predominant division of SOC composition across DPS after IPA/NH_3 extraction. Specifically, PT with higher concentrations of even-carbon-numbered chains alkanoic acids (> C_{22}), alkanols, and *ω*-hydroxyalkanoic acids and *α, ω*-dicarboxylic acids clustered separately from RP, HR, CV, and AC in AS fractions. In contrast, RP, HR, and CV had higher concentrations of even-carbon-numbered chains alkanoic acids (> C_{22}), alkanols, and *ω*-hydroxyalkanoic acids, and *α, ω*-dicarboxylic acids in AI fractions.

Single biomarkers from plants

Alkanoic acids, alkanols, and alkanes were identified from DCM/MeOH extracts of plant leaves/needles and roots (Fig. S1), which were also the three main groups mentioned in soils. For the identified alkanoic acids and alkanols in all leaves/needles and roots, a strong even-over-odd preference was found, whereas the alkanes discovered performed an odd-over-even predominance. The number of different alkanols found in roots was larger than in the leaves, which is similar to found for the alkanoic acids and alkanes, but the abundance of the alkanols in the leaves was much higher, e.g., C_{22} alkanols found in HR leaves had a mean of 165.40 μg g^{-1} higher than that in roots. In addition, variation in abundance of single biomarkers with DPS observed in leaves/needles and roots, e.g., C_{24} alkanols were the most dominating in PT and RP leaves, but not in other species. In contrast, C_{26} alcohols were predominant in HR, CV, and AC roots, but not in PT and RP. Interestingly, the odd-over-even-numbered alkanes (C_{21}–C_{31}) were observed in the leaves and roots, except for PT needles in which no alkanes were identified.

Alkanoic acids, alkanols, *ω*-hydroxyalkanoic acids were released from the ester-bound lipids (cutin and suberin) upon BF_3–MeOH hydrolysis of all leaves/needles and roots (Fig. S2a–g), and only *α, ω*-dicarboxylic acids were observed in the roots. Short-chain (C_{16-18}) alkanoic acids were the predominant ester-bound alkanoic acids in all leaves at relatively high concentrations. Compared with leaves, a larger number of ester-bound alkanols in greater abundance were found in the roots, e.g., only C_{22} alkanols were found in PT needles, but even-numbered alkanols (C_{20}–C_{28}) were all distributed in the roots. Specifically, PT needles had the highest abundance of *ω*-hydroxyalkanoic acid (C_{12-18}), which was 425.62 μg g^{-1} higher than that of other plant leaves. In addition, even-numbered *ω*-hydroxyalkanoic acid (C_{20-28}) and *α, ω*-dicarboxylic acids (C_{16-30}) as typical suberin-derived biomarkers were only found in the plant roots. No *α, ω*-dicarboxylic acids were found in CV while in the roots of the other DPS the dominating *α, ω*-dicarboxylic acid differs: PT (C_{18}), RP (C_{16}), HR (C_{18}), and AC (C_{20}).

Relations between single soil–plant linked biomarkers and SWR

For all soils, the majority of compounds had negative but no significant correlations between their relative concentrations (μg g^{-1} TOC) and SWR (Fig. S3). In Table 3 and Table S3, only the significant correlations between absolute concentrations and relative concentrations of individual biomarkers and SWR are given. For the D fractions, SWR positively related to the absolute concentrations (μg g^{-1} soil) of even-carbon-numbered chains alkanoic acids (C_{16-24} and C_{32}) and alkanols (C_{18}, C_{26}, and C_{28}) (Table S3), while the relative concentrations (μg g^{-1} TOC) of more than half of the compounds were negatively linked to SWR. In particular, comparing the D fractions with extractable lipids of plants, C_{16}, C_{17}, and C_{18} alkanoic acids in the D fractions of soils were negatively related to SWR for all soils, which were most abundant in CV and AC roots. The RP, PT, and CV leaves/needles contained the highest concentration of C_{24} and C_{28} alkanols respectively, which were the only two compounds in the D fractions of soils that were positively correlated with SWR (Table 3).

Compared with extractable lipids, ester-bound lipids (AS and AI fractions) contained a larger number of different biomarkers, except for the alkanes (Table S2). There was a positive relationship between SWR and the absolute concentrations of cutin-derived (e.g., C_{16} *ω*-hydroxyalkanoic acids) and suberin-derived (e.g., C_{16}, C_{30} *α, ω*-dicarboxylic acids, and C_{25} *ω*-hydroxyalkanoic acids) biomarkers of soils in AS fractions (Table S3). Comparing the ester-bound lipids of plants, the relative concentrations of cutin-derived (e.g., C_{16} *ω*-hydroxyalkanoic acids) and suberin-derived (e.g., C_{18} *α, ω*-dicarboxylic acids) biomarkers were observed to be positively linked to SWR (Table 3), which were most abundant in PT needles and RP roots.

Table 3

The relative concentrations (log_{10} (μg g^{-1} TOC)) of all identified biomarkers in D fractions, AS fractions, and AI fractions significantly related to SWR (log_{10} WDPT (s)), $n = 54$

SWR biomarkers[a]	Coef.[b]	Sig.[c]	SWR biomarkers	Coef.[b]	Sig.[c]
D Heptadecanoic acid	-0.289	0.034	AS Pentacosanoic acid	-0.399	0.003
D Hexadecenoic acid	-0.694	0.000	AS Octadecosanoic acid	-0.483	0.000
D Octadecenoic acid	-0.299	0.028	AS Nonacosanoic acid	0.595	0.000
D Eicosanoic acid	-0.306	0.024	AS Dotriacontanoic acid	-0.289	0.034
D Heneicosanoic acid	-0.673	0.000	AS Docosanol	-0.300	0.028
D Tetracosanoic acid	-0.392	0.003	AS Tetracosanol	0.511	0.000
D Hexacosanoic acid	-0.645	0.000	AS Pentacosanol	0.365	0.007
D Octadecosanoic acid	-0.751	0.000	AS Octacosanol	-0.387	0.004
D Dotriacontanoic acid	0.457	0.001	AS *ω*-Hydroxyhexadecanoic acid	0.358	0.008
D Eicosanol	-0.305	0.025	AS *ω*-Hydroxydocosanoic acid	-0.342	0.011
D Heneicosanol	-0.394	0.003	AS *ω*-Hydroxypentacosanoic acid	0.374	0.005
D Docosanol	0.332	0.014	AS *ω*-Hydroxyhexacosanoic acid	-0.489	0.000
D Octacosanol	0.367	0.006	AS *α, ω*-Hexadecanedioic acid	0.636	0.000
D Nonadecane	-0.701	0.000	AS *α, ω*-Triacontanedioic acid	0.465	0.000
D Eicosane	-0.311	0.022	AI Docosanoic acid	-0.294	0.031
D Heneicosane	-0.573	0.000	AI Tricosanoic acid	-0.491	0.000
D Tetracosane	-0.462	0.000	AI Pentacosanoic acid	-0.357	0.008
D Nonacosane	-0.734	0.000	AI Hexacosanoic acid	-0.538	0.000
D *ω*-Hydroxydocosanoic acid	-0.536	0.000	AI Octadecosanoic acid	-0.329	0.015
D *α, ω*-Docosanedioic acid	0.311	0.022	AI Triacontanoic acid	-0.402	0.003
AS Hexadecanoic acid	-0.401	0.003	AI Dotriacontanoic acid	-0.298	0.029
AS Heptadecanoic acid	-0.316	0.020	AI Octadecanol	0.335	0.013
AS Hexadecenoic acid	-0.530	0.000	AI Heptacosanol	-0.427	0.001
AS Nonadecanoic acid	-0.350	0.009	AI *ω*-Hydroxyhexadecanoic acid	0.460	0.000
AS Eicosanoic acid	0.353	0.009	AI *ω*-Hydroxytricosanoic acid	0.278	0.042
AS Heneicosanoic acid	-0.342	0.011	AI *α, ω*-Octadecanedioic acid	0.433	0.001

[a] D, AS, and AI refer to DCM/MeOH soluble fraction, DCM/MeOH soluble fraction of IPA/NH_3 extract, and DCM/MeOH insoluble fraction of IPA/NH_3 extract, respectively. [b] Linear correlation coefficient. [c] Significant *P*-values

Biomarker groups and SWR

Soil biomarker groups

To get a more general view of the relationship between certain compounds and SWR, we have analyzed compound groups (i.e., the sum of all compounds of the same class and the same EF) (Table S2). The absolute concentrations (μg g^{-1} soil) of LFA were higher

in D fractions than in the AS and AI fractions (Fig. 6a). Compared with PT, RP, CV and AC, the soils under HR had higher LFA concentrations in D fractions ($p < 0.05$), whereas the LFA concentrations in the AS and AI fractions were lowest (Fig. 6a). Cutin, primarily derived from leaf tissues, was only observed in the AS and AI fractions, in which the soils under HR had 88.68% lower cutin concentration than that under PT in the AS fractions ($p < 0.05$) (Fig. 6b), while in the AI fractions, the absolute amounts of cutin were similar between the soils beneath different DPS. Similarly, the absolute amounts of root-derived suberin in AI fractions were higher than in the D and AS fractions ($p < 0.01$) (Fig. 6c), In addition, the absolute concentrations of total plant-derived biomarkers (sum of LFA, cutin, suberin) (Table S3) were also higher in the AI fractions than in the D and AS fractions ($p < 0.01$) (Fig. 6d), and the legumes and grasses soils (i.e., around CV and AC) had higher concentrations of total plant-derived biomarkers in the AS and AI fractions than the soils under/around other PFTs ($p < 0.05$) (Fig. 6d). To trace the microbial origin of TOC, we calculated the absolute concentrations of SFA. In the AS fractions, the absolute amounts of SFA were highest across all the DPS, which were on average 57.54% higher than that of other EFs with no significant difference among soils under/around the different DPS ($p > 0.05$) (Fig. 6e). Overall, the absolute total amounts of LFA and suberin were significantly higher than those of cutin and SFA when the EFs were not considered. Furthermore, differences in absolute total concentrations of compound groups were observed only in suberin, where the absolute total concentrations of the soils around CV and AC ranged from 124.46 to 163.90 μg g^{-1} soil, which was higher than those under PT, RP, and HR (Fig. 6f).

The variation trend of LFA relative abundance was similar to that of LFA absolute abundance, and the soils under HR had a mean of 2.08 mg g^{-1} TOC higher than other DPS in the D fractions ($p < 0.01$) (Fig. 7a). The relative abundance of LFA in the AS and AI fractions were on average 39.77% higher in the soils under RP relative to the soils under HR ($p < 0.01$) (Fig. 7a). The cutin relative abundance showed a trend contrary to that of cutin absolute abundance, while in the AS fractions, the soils under PT also had 71.16% and 92.76% higher cutin concentrations than that under HR and AC ($p < 0.01$) (Fig. 7b). In addition, the variation trend of the suberin relative amounts were as follows: AS fractions > AI fractions > D fractions ($p < 0.01$) (Fig. 7c), and the legumes and grasses soils (i.e., around CV and AC) with lower SWR had higher suberin relative amounts than evergreen trees, deciduous trees, and shrubs soils (i.e., under PT, RP, and HR) among the three EFs. As opposed to the absolute concentrations, the relative concentrations of total extractable plant-derived biomarkers were higher in D fractions than in AS and AI fractions across all soils under different PFTs ($p < 0.01$) (Fig. 7d), while the variation trend in different DPS showed a similar pattern to the relative concentrations across each EF. The relative amounts of SFA in D and AS fractions were higher than in the AI fractions, and the shrubs soils (i.e., under HR) in the AS fractions had a mean of 435.82 μg g^{-1} TOC higher than other PFTs (Fig. 7e). Overall, the relative concentration of total LFA under shrubs (HR), total cutin under evergreen trees (PT) and total suberin under deciduous trees, legumes and grasses (AC and CV) were highest, respectively ($p < 0.05$) (Fig. 7f).

Plant biomarker groups

For group abundances of extractable lipids of leaves (Fig. S4a–c), the LFA concentration of RP was on average 67.71% higher than other DPS, and the cutin concentration of PT was the highest, followed by RP, and then HR, CV, and AC. The roots of legumes and grasses (i.e., CV and AC) in LFA were highest, and the roots of CV were richer in suberin and SFA than the other roots (Fig. S4d–f). Similar to the extractable lipids of leaves, the ester-bound lipids of PT leaves/needles in cutin were the highest, followed by RP, and then HR, CV, and AC. Moreover, the LFA concentrations of HR were on average 53.84% and 71.54% higher than other DPS in the leaves/needles and roots, and the SFA concentrations of legumes and grasses (i.e., CV and AC) in the roots were highest, whereas the suberin concentrations of HR roots were also higher than other DPS (Fig. 8a–f).

Relations between groups of soil–plant linked biomarkers and SWR

In Fig. 9a–b, only the significant correlations between absolute concentrations and relative concentrations of the compound groups in all soils and SWR are given. Interestingly, although the composition within each compound group is different, SWR almost showed a positive correlation with the absolute concentration of compound groups (Fig. 9a) and a negative correlation with the relative concentration (Fig. 9b). The absolute amounts of D LFA, D SFA, AS cutin, AS suberin, AI cutin, AI SFA, total LFA, total SFA, and total cutin were positively related to SWR, and the relative amounts of AS cutin, AI cutin, and total cutin were positively related to SWR, but the relative amounts of D LFA, D SFA, AS SFA, AI LFA, total LFA, and total SFA were negatively related to SWR (Fig. 9a–b). Furthermore, SEM analysis was performed to simulate the direct and indirect effects of compound groups on SWR (Fig. 9c). The model showed that SWR biomarker groups could explain 50% of the variance in SWR. For all soils, AS cutin (path coefficient = 0.30; $p < 0.01$) and AI cutin (path coefficient = 0.47; $p < 0.001$) had direct and positive effects on SWR, which were consistent with the

correlation analysis results. However, other compound groups had no direct effect on SWR ($p > 0.05$). In addition, Al LFA had positive effects on Al suberin, but negative effects on AS cutin and Al cutin, and AS LFA had negative effects on Al cutin (Fig. 9c). We analyzed the standardized total effects of different individual compound groups to further assess the comprehensive regulatory effect of the driving factors on SWR (Fig. 9d). Our data confirmed that Al cutin had the greatest positive and integrated effect on SWR, followed by AS cutin, and then Al suberin.

Quality relationship of two compound groups to SWR

In order to understand whether the quality parameters of biomarkers can describe the SWR, the correlation between the ratio of the two separate compound groups and SWR was analyzed. For all soils, AS/Al cutin was essential for a significant combination, when AS cutin or Al cutin were acted as the numerator, and the correlation between the ratio of two separate groups and SWR was positive, otherwise, no significant relationship was observed. Furthermore, AS cutin / D LFA and Al cutin / D LFA were the strongest positive linked to SWR ($r^2 = 1$, $p < 0.001$) (Figure S5).

Discussion

Soil characteristics and SWR

By supplying additional OC inputs to the soil, the Grain for Green Program has long been acknowledged to play a significant role in preserving or boosting soil C sequestration (Wang et al., 2016b). The higher cumulative C inputs in all sampling locations may be primarily attributed to the Grain for Green Program's favorable effect on SOC accumulation. Due to the introduction of hydrophobic materials into the soil, one of the causes of the high WR in the examined soils may be the high SOC content (Zavala et al., 2009). According to multiple studies (Scott et al., 2000; Harper et al., 2000; Mataix-Solera and Doerr, 2004; Mataix-Solera et al., 2007; Doerr et al., 2005; Lozano et al., 2013; Hermansen et al., 2019; Mao et al., 2019; Wang et al., 2016a; Seaton et al., 2019), there is a positive relationship between SOC and SWR, particularly when samples were taken from the same soil type and under or near the same DPS. In the present study, we found that some soil characteristics were highly correlated with SWR, in particular, TOC, TN, and C/N ratios. However, the very strong relation between TOC and SWR determined by WPDT found in our study is not supported by other studies (Horne and McIntosh, 2000; de Blas et al., 2010; Dekker and Ritsema, 1994; Doerr et al., 2005). This discrepancy may be primarily due to the soils used in others' work were from different soil types and geographical locations, having different effects on SWR, while the loess we used were all from the same area. The positive relation between TOC and SWR demonstrated that SOC greatly contributes to SWR. Consequently, TOC will be considered an important soil parameter to predict the WR levels of loess after returning farmland to forest (grassland).

TN was positively but weakly correlated with SWR, as earlier found by Lachacz et al. (2009), which is probably caused by the ability of legumes to fix atmospheric N, high-C and high-N litter inputs may shift the stoichiometry of substrates, accelerating the mineralization rate of microorganisms and increasing the yield of polysaccharides, thereby improving SWR. However, the C/N ratio was not correlated with SWR, which is contrasting to previous results (Mao et al., 2014). This divergent result may be our study was different to the studies which on the specific coarse-textured soils, such as the sandy soils of the Netherlands (Mao et al., 2014), New Zealand (Wallis et al., 1990), southeastern Australia (Atanassova and Doerr, 2010) and southeastern Spain (Lozano et al., 2013). Due to its low soil-specific surface, the characteristics of these soils are also classically related to WR, which is prone to be coated by hydrophobic substances (Doerr et al., 2000).

Single SWR biomarkers and SWR

In the present work, sequential extraction methods have been used to identify SWR biomarkers, including n-alkanoic acids, n-alkanols, n-alkanes, alkenoic acids, sterols, and some monomers derived from cutin and suberin, which are all hydrophobic compounds (Franco et al., 2000; Horne and McIntosh, 2000; Mao et al., 2014; Mao et al., 2015; Bull et al., 2000; Lozano et al., 2013); however, more than half of the biomarkers in relative quantity were still negatively linked to SWR (Table 3). Similar results have also been described for other sites. Mao et al. (2015) report for sandy Netherlands soil a more negative linear relationship between short-chain compounds and SWR than long-chain ones. Atanassova and Doerr (2010) showed in southeastern Australia the greater abundance of polar compounds in the less repellent sandy loam soil compared with the high-repellent and medium-repellent sandy soils. Therefore, the severity of SWR is not related to all the constituents of SOC (Doerr et al., 2005), or the effectiveness of individual organic compounds on SWR decreases with increasing concentration. The above can be attributed to the low molecular weight polar

compounds that are supposed to be more mobile and diffuse more quickly through soil water, inducing a relatively lower SWR (Mainwaring et al., 2004).

However, Lozano et al. (2013) found strong correlations between extractable lipid content and WR for Quercus and Pinus samples establishing a clear relationship. de Blas et al. (2010) reported different concentrations in extractable lipid content could explain differences in the WR persistence of samples from the same plant species. In agreement with these previous works, our results showed the positive relations between the absolute concentrations (μg g^{-1} soil) of even-carbon-numbered chains alkanoic acids (C_{16-24} and C_{32}) and alkanols (C_{18}, C_{26}, and C_{28}) with SWR (Table S3). In addition, the two-way hierarchical clustering of the absolute and relative concentrations further revealed the SWR biomarkers belonging to the same DPS clustered more clearly, showing a predominant division of SOC composition between different DPS (Fig. 5). Specifically, in AS fractions, PT with higher WDPT value contained higher concentrations of even-carbon-numbered chains alkanoic acids (> C_{22}), alkanols, and ω-hydroxyalkanoic acids and α, ω-dicarboxylic acids clustered separately from RP, HR, CV and AC. According to previous studies (Doerr et al., 2000; Zavala et al., 2009; Kaiser et al., 2015; Mao et al., 2019), the main reasons for this phenomenon may be as follows. First, the emergence of SWR is primarily due to the interaction between water molecules and polar molecules, where the polar molecules are composed of a hydrophilic group (head) and a hydrophobic chain (tail). When the cohesion between water molecules is greater than the force between water and the soil surface, the soil surface shows water repellence. When the hydrophobic coating encounters water droplets, the force between them changes, reorienting and organizing the amphiphilic molecules. Once the attraction between water molecules and the soil surface is greater than the cohesion among the water molecules, the hydrophilic heads of polar amphiphilic molecules face outward, making the soil wettable. Second, higher hydrophobic compounds cause longer interaction time between water molecules and soil surface, resulting in greater SWR measured by WDPT. Although the content and composition of SWR biomarkers vary across soils under/around the different PFTs, our study has pointed out that the effects of single biomarkers on SWR cannot be accurately quantified, and therefore single biomarkers are not good predictors of SWR levels.

SWR biomarker groups and SWR

Since an individual SWR biomarker may not be able to predict SWR well, we analyzed the possible relations between compound groups and SWR. In our work, the compounds in the D fraction were mainly the extractable free lipids, while the AS fraction and AI fraction is a combination of free lipids and ester lipids hydrolyzed by microbes and BF_3–MeOH, which agreed with some previous studies (Van Bergen et al., 1997; DeBano, 2000b; Franco et al., 2000; Mao et al., 2014). In terms of the extent to which SWR was represented during the extraction sequence, the SWR of most soils increased after DCM/MeOH extractions, while the SWR significantly reduced after IPA/NH_3 extraction (Fig. 3). Additionally, in all soils, the LFA concentrations in the D fractions were higher than those in the AS and AI fractions, whereas cutin and suberin concentrations of them showed a trend contrary to LFA (Fig. 6–7). These results indicated that it is ester-bound lipids, not free lipids, that play a major role in causing SWR. Although no clear evidence for such molecular behavior has been given, this paradox can be explained by a certain degree of mobility of free lipids. Due to the longer chain and less mobility of ester-bound lipids, they may be more hydrophobic than free, low-molecular-weight lipids. DCM/MeOH solvent, by extracting more lipid material, leaves more hydrophobic adsorption sites on the surface. Once the free lipids were removed, the adsorbed extent of ester-bound lipids onto high-affinity hydrophobic sites increased and as a consequence of that SWR increased as well. Because ester-bound lipids exist as polyfunctional macromolecules, they may change their structural conformations due to interactions involving the carboxyl and alcoholic hydroxyl functional groups. These interactions are disturbed by the extraction of polar compounds by IPA/NH_3 solvent, causing conformational disruption, hence desorption of polar compounds and exposure of surface polar adsorption sites, resulting in a wettable soil (Atanassova and Doerr, 2010; Riederer et al., 1993; Mao et al., 2019). Thus, SWR increased in all the soils after DCM/MeOH solvent extraction and was eliminated by IPA/NH_3. Although the IPA/NH_3 solvent extracted less free lipids and more ester-bound lipids than the DCM/MeOH solvent, it was more suitable for extracting compounds of certain associated with SWR elimination.

In this work, we separately analyzed the quantity and quality of compound groups in these different extract fractions. For all soils, the absolute concentrations of most compound groups associated with SWR (Fig. 9a), which is supported by Mao et al. (2015) who also noted that the positive relationships between the absolute amounts of all the compound groups and SWR follow the positive correlations between TOC and SWR. Obviously, the relationships between absolute amounts of compound groups and SWR are determined by the TOC content, we cannot identify which compound group affects SWR by analyzing the absolute concentration of the compound group.

Regarding the relative amounts of SWR biomarker groups, AS cutin, Al cutin, and total cutin were the groups to show positive relations with SWR, while the groups of D LFA, D SFA, AS SFA, Al LFA, total LFA, and total SFA showed negatively relations with SWR for all soils. Similar results have been reported in many previous studies (Mao et al., 2014; Atanassova and Doerr, 2010; Hansel et al., 2008), which showed that the differences in SWR studied cannot be attributed solely to a single factor or group, but also the interaction of multiple compound groups. Therefore, we simulated the direct and indirect effects of compound groups on SWR through SEM (Fig. 9c and Table S4). In agreement with the above research, our results showed the direct and positive effects of AS cutin and Al cutin on SWR. In addition, Al LFA had negative effects on AS cutin and Al cutin, and AS LFA had negative effects on Al cutin further revealing the indirect effects of Al LFA and AS LFA on SWR level. However, SWR is a complex property caused by numerous interconnected soil parameters. The variability (50%) could account for the SWR identified by structural equation modeling. It is intuitive to speculate that other biotic and abiotic factors may be able to further improve the prediction of SWR, such as soil water content and microorganisms. In our study, we have tried to explain which SWR biomarkers are the most relevant in the development of SWR in loess. Our correlation data of the ratio of the two separate compound groups and SWR revealed that when AS cutin and Al cutin were acted as the numerators, the ratio of the two separate compound groups was always positively correlated with SWR, and positive correlations between AS cutin / D LFA and Al cutin / D LFA with SWR were strongest ($r^2 = 1$, $p < 0.001$) (Figure S5). This finding might be mainly due to the increased adsorption of cutin onto high-affinity hydrophobic sites on the soil surface after more D fraction compounds were extracted, resulting in a larger SWR. According to the above results, SWR occurring on the loess surface seems to be the most influenced by cutin-derived C, and cutin / LFA is considered an important soil quality parameter for predicting SWR levels on the loess surface (Fig. 10).

Link between plants, microbes and SWR

Cover plant types may largely alter the content and composition of SOC by providing additional OC residues (Van Bergen et al., 1997; Nierop, 2001; Kögel-Knabner, 2002; Zhang et al., 2022). In the present work, the main groups of the extractable and ester-bound lipids present in the leaves and roots were all identified in D, AS, and Al fractions of the soils under/around the given plant species, which might be due to the cover plant are the main sources of the SWR biomarkers (Nierop et al., 2003; Naafs et al., 2004). In line with previous studies (Naafs et al., 2004; Nierop et al., 2006; Zhang et al., 2022), the predominance of even-over-odd numbered D LFA, AS LFA, and Al LFA (> C_{24} alkanes, >C_{22} n-alkanoic acids and alkanols) indicated that the primary sources are plants (Jansen et al., 2006; Zhang et al., 2022). By contrast, D SFA, AS SFA, and Al SFA (C_{16}–C_{18} n-alkanoic and n-alkenoic acid), which were identified in all of our topsoil and plant samples, maybe both from plants and microbes (Jaffé et al., 1996; Jansen et al., 2006; Zhang et al., 2022). AS cutin and AS suberin may be from microbially hydrolyzed ester lipids, while Al cutin and Al suberin are polymers hydrolyzed by BF_3–MeOH from leaves/needles and roots of plants, respectively (Riederer et al., 1993; Mao et al., 2014). It can be inferred that plant-derived C contributes more to surface SWR generation, but the role of microorganisms in the hydrolysis of AS cutin and AS suberin in the natural environment cannot be underestimated, as AS cutin and AS suberin was significantly higher than Al cutin and Al suberin, respectively (Fig. 7b–c). In addition, the relationship between other microbial-derived C, such as ergosterol and glomalin-related soil protein, and SWR was not considered in our work. These carbon sources in soil may be produced by filamentous fungi (Feeney et al., 2004; Wessels, 1996, 2000; White et al., 2000; Rillig et al., 2010; Bayry et al., 2012; Lozano et al., 2013), which directly affect SWR. Thus, our research emphasizes the role of microorganisms in the hydrolysis of SWR biomarkers, and we will focus on the distribution of more microbial-derived SWR biomarkers in loess and their relationship with SWR in subsequent studies.

From the above analysis, the ester-bound lipid biomarkers after IPA/NH_3 extraction represent the cutin and suberin-derived compounds in the plant leaves/needles and roots, respectively. For the Al fractions, differences in absolute concentrations of compound groups were only observed in suberin, in which legumes, grasses soils, and shrubs soils were richer than evergreen trees and deciduous trees (Fig. 6c), which is contrary to the performance of SWR. This may be primarily due to plants with shallow and horizontal extended roots easily accumulating suberin than deep-rooted plants in topsoil. In our work, the small amounts of α, ω-dicarboxylic acids under deep-rooted plants may derive from grasses providing suberin to the topsoil. Another source could be bark, which also contains suberin although they contribute less to the soil than roots (Preston et al., 1994). Compared with deciduous trees, legumes, grasses and shrubs, soils under evergreen trees contained more cutin (Fig. 7f), and had the highest SWR (Table S1). These results are in line with Rodríguez-Alleres et al. (2011) and Badía et al. (2013), who both manifested that the SWR persistence of grassland soil was much less than that under forest soil. Seaton et al. (2019) and Smettem et al. (2021) have successively reached the same conclusion that some deep-rooted plants can take advantage of SWR under drought stress, leading to the emergence of co-evolutionary behavior in natural ecosystems. This observation is also found in previous studies, which showed

SWR not only allows rainwater to form a preferential flow, penetrate into deep soil, and store in large quantities (De Boeck and Verbeeck, 2011; Zeppenfeld et al., 2017; Alanís et al., 2016; Lozano et al., 2013), but also reduces soil water evapotranspiration loss through various mechanisms, making deep-rooted plants more resistant to drought than shallow-rooted plants (Doerr et al., 2006; Rye and Smettem, 2018; Shahidzadeh-Bonn et al., 2007; Gupta et al., 2015). This can be inferred that deep-rooted evergreen plants tend to accumulate cutin on the loess surface, forming a hydrophobic layer and hence they were more likely to cause SWR than other PFTs.

Conclusions

This study presents the first molecular investigation to explore the relationship between SOM composition and SWR output in the Chinese Loess Plateau, the extremely complex soil-eroding region. The sequential extraction procedure and hydrolysis approach applied in this work was DCM/MeOH, IPA/NH_3, and BF_3–MeOH transmethylation to clearly identify the classes and origin of the different SWR biomarkers. The free lipids in the D fractions are mainly consisted of plants and a small part from microbes or plants, while the ester-bound lipids in the AS fractions are mainly originated from microbially hydrolyzed leaf/needles-derived cutin and root-derived suberin, and the ester-bound lipids in the AI fractions were polymers hydrolyzed by BF_3–MeOH from leaves and roots. Thus, we speculated the cover plant is the main source of the SWR biomarkers, but emphasized the significant role of soil microorganisms in the hydrolysis of SWR biomarkers. In addition, compared to deciduous trees, shrubs, legumes, and grasses, the soils under evergreen trees offered more cutin, although being smaller in concentration, which was stronger than the SWR induced by free lipids. To understand whether the quality parameters of two different compound groups can describe the SWR, the SEM analysis was performed and the quality parameters, i.e., the ratio of the two separate compound groups was proposed. For our topsoil, the combination cutin / LFA is considered an important soil parameter for predicting SWR levels on the loess surface. These results shed new light shed new light that the differentiation of cover PFTs alter the abundance and composition of hydrophobic compounds in soils, thereby affecting SWR levels in a semi-arid region of the Chinese Loess Plateau. Furthermore, this study highlights the important role of plant-soil interactions in mediating the high variability of SWR in typical loess soils, which could have significant implications for developing effective strategies to manage soil water availability at a later stage, particularly in arid and semi-arid regions where water is scarce.

Declarations

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant number 41977426). We would like to thank the staff of Changwu Agro-Ecological Experiment Station for their assistance in field investigations and sample collections.

Data Availability

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

References

1. Alanís N, Hernández-Madrigal VM, Cerdà A, Muñoz-Rojas M, Zavala LM, Jordán A (2016) Spatial gradients of intensity and persistence of soil water repellency under different forest types in central Mexico. Land Degrad Dev 28(1):317–327. https://doi.org/10.1002/ldr.2544
2. Arcenegui V, Mataix-Solera J, Guerrero C, Zornoza R, Mayoral AM, Morales J (2007) Factors controlling the water repellency induced by fire in calcareous mediterranean forest soils. Eur J Soil Sci 58:1254–1259. https://doi.org/10.1111/j.1365-2389.2007.00917.x

3. Atanassova I, Doerr S (2010) Organic compounds of different extractability in total solvent extracts from soils of contrasting water repellency. Eur J Soil Sci 61:298–313. https://doi.org/10.1111/j.1365-2389.2009.01224.x
4. Badía D, Aguirre JA, Martí C, Márquez MA (2013) Sieving effect on the intensity and persistence of water repellency at different soil depths and soil types from NE-Spain. Catena 108:44–49. https://doi.org/10.1016/j.catena.2012.02.003
5. Balshaw HM, Douglas P, Davies ML, Doerr SH (2020) Pyrene and nile red fluorescence probes for in-situ study of polarity and viscosity of soil organic coatings implicated in soil water repellency. Eur J Soil Sci 71:868–879. https://doi.org/10.1111/ejss.12925
6. Bayry J, Aimanianda V, Guijarro JI, Sunde M, Latgé J-P (2012) Hydrophobins–unique fungal proteins. PloS Pathog 8(5):e1002700. https://doi.org/10.1371/journal.ppat.1002700
7. Bisdom EBA, Dekker LW, Schoute JFTh (1993) Water repellency of sieve fractions from sandy soils and relationships with organic material and soil structure. Geoderma 56:105–118. https://doi.org/10.1016/0016-7061(93)90103-R
8. Bremner J, Mulvaney C (1982) Nitrogen total content. In: Methods of soil analysis, 2nd edn. USA, pp 595–624
9. Bull ID, Bergen PF, Nott CJ, Poulton PR, Evershed RP (2000) Organic geochemical studies of soils from the Rothamsted classical experiments—V. The fate of lipids in different long-term experiments. Org Geochem 31(5):389–408. https://doi.org/10.1016/s0146-6380(00)00008-5
10. Butzen V, Seeger M, Marruedo A, Jonge LD, Wengel R, Ries JB, Casper MC (2015) Water repellency under coniferous and deciduous forest—experimental assessment and impact on overland flow. Catena 133:255–265. https://doi.org/10.1016/j.catena.2015.05.022
11. Chai X, Xu X, Li L, Wang W, Li S, Geming P, Qu Y, Zhang Q, Ren X, Xu Y, Li M (2022) Physicochemical and biological factors determining the patchy distribution of soil water repellency among species of dominant vegetation in loess hilly region of China. Front Plant Sci. https://doi.org/10.3389/fpls.2022.908035
12. Chau W H, Biswas A, Vujanovic V, Si BC (2014) Relationship between the severity, persistence of soil water repellency and the critical soil water content in water repellent soils. Geoderma 221–222:113–120. https://doi.org/10.1016/j.geoderma.2013.12.025
13. Coelho COA., Laouina A, Regaya K, Ferreira AJD, Carvalho TMM, Chaker M, Naafa R, Naciri R, Boulet AK, Keizer JJ (2005) The impact of soil water repellency on soil hydrological and erosional processes under Eucalyptus and evergreen Quercus forests in the Western Mediterranean. Aust J Soil Res 43:309–318. https://doi.org/10.1071/SR04083
14. DeBano LF (2000a) Water repellency in soils: a historical overview. J Hydrol 231–232:4–32. https://doi.org/10.1016/s0022-1694(00)00180-3
15. DeBano LF (2000b) The role of fire and soil heating on water repellency in wildland environments: a review. J Hydrol 231–232:195–206. https://doi.org/10.1016/S0022-1694(00)00194-3
16. de Blas E, Rodríguez-Alleres M, Almendros G (2010) Speciation of lipid and humic fractions in soils under pine and eucalyptus forest in northwest Spain and its effect on water repellency. Geoderma 155:242–248. https://doi.org/10.1016/j.geoderma.2009.12.007
17. de Blas E, Almendros G, Sanz J (2013) Molecular characterization of lipid fractions from extremely water-repellent pine and eucalyptus forest soils. Geoderma 206:75–84. https://doi.org/10.1016/j.geoderma.2013.04.027
18. De Boeck HJ, Verbeeck H (2011) Drought-associated changes in climate and their relevance for ecosystem experiments and models. Biogeosciences 8(5):1121–1130. https://doi.org/10.5194/bg-8-1121-2011
19. Dekker LW, Ritsema CJ (1994) How water moves in a water repellent sandy soil: 1. Potential and actual water repellency. Water Resour Res 30(9):2507–2517. https://doi.org/10.1029/94WR00749
20. Dekker LW, Ritsema CJ (1996) Variation in water content and wetting patterns in Dutch water repellent peaty clay and clayey peat soils. Catena 28:89–105. https://doi.org/10.1016/s0341-8162(96)00047-1
21. Doerr SH, Shakesby RA, Walsh RPD (2000) Soil water repellency: its causes, characteristics and hydro-geomorphological significance. Earth-Sci Rev 51:33–65. https://doi.org/10.1016/S0012-8252(00)00011-8
22. Doerr SH, Thomas A (2000) The role of soil moisture in controlling water repellency: new evidence from forest soils in Portugal. J Hydrol 231–232:134–147. https://doi.org/10.1016/s0022-1694(00)00190-6

23. Doerr SH, Llewellyn CT, Douglas P, Morley CP, Mainwaring KA, Haskins C, Johnsey L, Ritsema CJ, Stagnitti F, Allinson G, Ferreira AJD, Keizer JJ, Ziogas AK, Diamantis J (2005) Extraction of compounds associated with water repellency in sandy soils of different origin. Aust J Soil Res 43:225–237. https://doi.org/10.1071/SR04091
24. Doerr SH, Shakesby RA, Dekker LW, Ritsema CJ (2006) Occurrence, prediction and hydrological effects of water repellency amongst major soil and land-use types in a humid temperate climate. Eur J Soil Sci 57(5):741–754. https://doi.org/10.1111/j.1365-2389.2006.00818.x
25. Feeney DS, Paul D, Hallett TD, Karl Ritz JI, Young IM (2004) Does the presence of glomalin relate to reduced water infiltration through hydrophobicity? Can J Soil Sci 84(4):365–372. https://doi.org/10.4141/s03-095
26. Franco CM, Clarke P, Tate M, Oades J (2000) Hydrophobic properties and chemical characterisation of natural water repellent materials in Australian sands. J Hydrol 231–232:47–58. https://doi.org/10.1016/s0022-1694(00)00182-7
27. Fu B, Yu L, Lue Y, He C, Yuan Z, Wu B (2011) Assessing the soil erosion control service of ecosystems change in the loess plateau of China. Ecol Complex 8(4):284–293. https://doi.org/10.1016/j.ecocom.2011.07.003
28. Giovannini G, Lucchesi S (1983). Effect of fire on hydrophobic and cementing substances of soil aggregates. Soil Sci 136:231–236. https://doi.org/10.1097/00010694-198310000-00006
29. González-Pérez JA, González-Vila FJ, Almendros G, Knicker H (2004) The effect of fire on soil organic matter—a review. Environ Int 30:855–870. https://doi.org/10.1016/j.envint.2004.02.003
30. González-Pérez JA, de Andrés JR, Clemente L, Martín J A, González-Vila F J (2008) Organic carbon and environmental quality of riverine and off-shore sediments from the gulf of cádiz, spain. Environ Chem Lett 6(1):41–46. https://doi.org/10.1007/s10311-007-0107-0
31. González-Peñaloza FA, Cerdà A, Zaval LM, Jordán A, Giménez-Morera A, Arcenegui V (2013) Do conservative agriculture practices increase soil water repellency? A case study in citrus-cropped soils. soil Till Res 124(2012):233–239. https://doi.org/10.1016/j.still.2012.12.001
32. Gupta B, Shah DO, Mishra B, Joshi PA, Gandhi VG, Fougat RS (2015) Effect of top soil wettability on water evaporation and plant growth. J Colloid Interf Sci 449:506–513. https://doi.org/10.1016/j.jcis.2015.02.018
33. Hansel FA, Aoki CT, Maia CMBF, Cunha JA, Dedecek RA (2008) Comparison of two alkaline treatments in the extraction of organic compounds associated with water repellency in soil under Pinus taeda. Geoderma 148(2):167–172. https://doi.org/10.1016/j.geoderma.2008.10.002
34. Harper RJ, McKissock I, Gilkes RJ, Carter DJ, Blackwell PS (2000) A multivariate framework for interpreting the effects of soil properties, soil management and landuse on water repellency. J Hydrol 231–232:371–383. https://doi.org/10.1016/s0022-1694(00)00209-2
35. Hermansen C, Moldrup P, Müller K, Jensen PW, Carlo VDD, Jeyakumar P, Jonge LWD (2019) Organic carbon content controls the severity of water repellency and the critical moisture level across New Zealand pasture soils. Geoderma 338:281–290. https://doi.org/10.1016/j.geoderma.2018.12.007
36. Hooper D, Coughlan J, Mullen M (2008) Structural equation modelling: guidelines for determining model fit. Electron J Bus Res Methods 6:53–60.
37. Horne D, McIntosh J (2000) Hydrophobic compounds in sands in New Zealand–extraction, characterisation and proposed mechanisms for repellency expression. J Hydrol 231–232:35–46. https://doi.org/10.1016/s0022-1694(00)00181-5
38. Huang MB, Shao MG, Zhang L, Li YS (2003) Water use efficiency and sustainability of different long-term crop rotation systems in the Loess Plateau of China. Soil Till Res 72:95–104. https://doi.org/10.1016/S0167-1987(03)00065-5
39. Jaffé R, Elismé T, Cabrera AC (1996). Organic geochemistry of seasonally flooded rain forest soils: molecular composition and early diagenesis of lipid components. Org Geochem 25:9–17. https://doi.org/10.1016/S0146-6380(96)00103-9
40. Jansen B, Nierop KGJ, Hageman JA, Cleef AM, Verstraten JM (2006) The straight-chain lipid biomarker composition of plant species responsible for the dominant biomass production along two altitudinal transects in the Ecuadorian Andes. Org Geochem 37:1514–1536. https://doi.org/10.1016/j.orggeochem.2006.06.018
41. Jiménez-Morillo NT, Almendros G, Miller AZ, Hatcher PG, González-Pérez JA (2022) Hydrophobicity of soils affected by fires: an assessment using molecular markers from ultra-high resolution mass spectrometry. Sci Total Environ 817:152957. https://doi.org/10.1016/j.scitotenv.2022.152957

42. Jiménez-Pinilla P, Doerr SH, Ahn S, Lozano E, Mataix-Solera J, Jordán A, Zavala LM, Arcenegui V (2016) Effects of relative humidity on the water repellency of fire-affected soils. Catena 138:68–76. https://doi.org/10.1016/j.catena.2015.11.012
43. Jordán A, Zavala LM, Nava AL, Alanís N (2009) Occurrence and hydrological effects of water repellency in different soil and land use types in Mexican volcanic highlands. Catena 79:60–71. https://doi.org/10.1016/j.catena.2009.05.013
44. Jordán A, Zavala LM, Mataix-Solera J, Doerr SH (2013) Soil water repellency: origin, assessment and geomorphological consequences. Catena 108:1–5. https://doi.org/10.1016/10.1016/s0167-1987(03) 00065-5
45. Kaiser M, Kleber M, Berhe AA (2015) How air-drying and rewetting modify soil organic matter characteristics: An assessment to improve data interpretation and inference. Soil Biol Biochem 80:324–340. https://doi.org/10.1016/j.soilbio.2014.10.018
46. Karunarathna A K, Kawamoto K, Moldrup P, Jonge LWD, Komatsu T (2010) A simple beta-function model for soil-water repellency as a function of water and organic carbon contents. Soil Sci 175:461–468. https://doi.org/10.1097/SS.0b013e3181f55ab6
47. Kodešová R, Kočárek M, Kodeš V, Šimůnek J, Kozák J (2008) Impact of soil micromorphological features on water flow and herbicide transport in soils. Vadose Zone J 7:798–809. https://doi.org/10.2136/vzj2007.0079
48. Kögel-Knabner I (2002) The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. Soil Biol Biochem 34:139–162. https://doi.org/10.1016/s0038-0717(01)00158-4
49. Kraemer FB, Hallett PD, Morrás H, Garibaldi L, Cosentino D, Duval M, Galantini J (2019) Soil stabilisation by water repellency under no-till management for soils with contrasting mineralogy and carbon quality. Geoderma 355:113902. https://doi.org/10.1016/j.geoderma.2019.113902
50. Lachacz A, Nitkiewicz M, Kalisz B (2009) Water repellency of post-boggy soils with a various content of organic matter. Biologia 64:634–638. https://doi.org/10.2478/s11756-009-0096-5
51. Leighton-Boyce G, Doerr SH, Shakesby RA, Walsh RPD (2007) Quantifying the impact of soil water repellency on overland flow generation and erosion: a new approach using rainfall simulation and wetting agent onin situ soil. Hydrol Process 21:2337–2345. https://doi.org/10.1002/hyp.6744
52. Li Y, Yao N, Tang DX, Chau HW, Feng H (2019) Soil water repellency decreases summer maize growth. Agr Forest Meteorol 266–267:1–11. https://doi.org/10.1016/j.agrformet.2018.12.001
53. Liu F, Zhan Y (2019) Soil water repellency in China and Israel: synthesis of observations and experiments. Appl Ecol Environ Res 17(4):8599–8614. https://doi.org/10.15666/aeer/1704_85998614
54. Lozano E, Jiménez-Pinilla P, Mataix-Solera J, Arcenegui V, Bárcenas GM, González-Pérez JA, García-Orenes F, Torres MP, Mataix-Beneyto J (2013) Biological and chemical factors controlling the patchy distribution of soil water repellency among plant species in a Mediterranean semiarid forest. Geoderma 207–208:212–220. https://doi.org/10.1016/j.geoderma.2013.05.021
55. Lozano E, García-Orenes F, Bárcenas-Moreno G, Jiménez-Pinilla P, Mataix-Solera J, Arcenegui V, Morugán-Coronado A, Mataix-Beneyto J (2014) Relationships between soil water repellency and microbial community composition under different plant species in a mediterranean semiarid forest. J Hydrol Hydromech 62(2):101–107. https://doi.org/10.2478/johh-2014-0017
56. Mainwaring KA, Morley CP, Doerr SH, Douglas P, Llewellyn CT, Llewellyn G, Matthew I, Stein BK (2004) Role of heavy polar organic compounds for water repellency of sandy soils, Environ Chem Lett 2:35–39. https://doi.org/10.1007/s10311-004-0064-9
57. Mao J, Nierop KGJ, Sinninghe Damsté JS, Dekker SC (2014) Roots induce stronger soil water repellency than leaf waxes. Geoderma 232–234:328–340. https://doi.org/10.1016/j.geoderma2014.05.024
58. Mao J, Nierop KGJ, Rietkerk M, Dekker SC (2015) Predicting soil water repellency using hydrophobic organic compounds and their vegetation origin. Soil 1(1):411–425. https://doi.org/10.5194/soil-1-411-2015
59. Mao J, Nierop KGJ, Dekker SC, Dekker LW, Chen B (2019) Understanding the mechanisms of soil water repellency from nanoscale to ecosystem scale: a review. J Soil Sediment 19:171–185. https://doi.org/10.1007/s11368-018-2195-9
60. Mataix-Solera J, Doerr S (2004) Hydrophobicity and aggregate stability in calcareous topsoils from fire-affected pine forests in southeastern Spain. Geoderma 118(1–2):77–88. https://doi.org/10.1016/s0016-7061(03)00185-x
61. Mataix-Solera J, Arcenegui V, Guerrero C, Mayoral AM, Morales J, González J, Garcia-Orenes F, Gómez I (2007) Water repellency under different plant species in a calcareous forest soil in a semiarid Mediterranean environment. Hydrol Process 21:2300–2309. https://doi.org/10.1002/hyp.6750

62. Meunier B, Dumas E, Piec I, Béchet D, Hébraud M, Hocquette J (2007) Assessment of hierarchical clustering methodologies for proteomic data mining. J Proteome Res 6:358–366. https://doi.org/10.1021/pr060343h
63. Morley CP, Mainwaring KA, Doerr SH, Douglas P, Llewellyn CT, Dekker LW (2005) Organic compounds at different depths in a sandy soil and their role in water repellency. Aust J Soil Res 43:239–249. https://doi.org/10.1071/SR04094
64. Moro MJ, Pugnaire FI, Haase P, Puigdefabregas J (1997) Effect of the canopy of Retama sphaerocarpa on its understorey in a semiarid environment. Funct Ecol 11:425–431. https://doi.org/10.1046/j.1365-2435.1997.00106.x
65. Müller K, Carrick S, Meenken E, Clemens G, Hunter D, Rhodes P, Thomas S (2016) Is subcritical water repellency an issue for efficient irrigation in arable soils? Soil Till Res 161:53–62. https://doi.org/10.1016/j.still.2016.03.010
66. Muñoz-Rojas M, Jordán A, Zavala LM, Rosa DDL, Abd-Elmabod SK, Anaya-Romero M (2012) Organic carbon stocks in Mediterranean soil types under different land uses (southern spain). Solid Earth Discuss 4:1095–1128. https://doi.org/10.5194/sed-4-1095-2012
67. Naafs DFW, Van Bergen PF, Boogert SJ, De Leeuw JW (2004). Solvent-extractable lipids in an acid andic forest soil; variations with depth and season. Soil Biol Biochem 36:297–308. https://doi.org/10.1016/j.soilbio.2003.10.005
68. Nelson DW, Sommers LE (1982) Total carbon, organic carbon, and organic matter. In: Methods of Soil Analysis, 3rd edn. USA, pp 961–1010
69. Nierop KGJ (2001) Temporal and vertical organic matter differentiation along a vegetation succession as revealed by pyrolysis and thermally assisted hydrolysis and methylation. J Anal Appl Pyrol 61:111–132, https://doi.org/10.1016/S0165-2370(01)00132-2
70. Nierop KGJ, Naafs DFW, Verstraten JM (2003) Occurrence and distribution of ester-bound lipids in Dutch coastal dune soils along a pH gradient. Org Geochem 34:719–729. https://doi.org/10.1016/S0146-6380(03)00042-1
71. Nierop KGJ, Jansen B, Hageman JA, Verstraten JM (2006) The complementarity of extractable and ester-bound lipids in a soil profile under pine. Plant Soil 286:269–285. https://doi.org/10.1007/s11104-006-9043-1
72. Preston CM, Hempfling R, Schulten HR, Schnitzer M, Trofymow JA, Axelson DE (1994) Characterization of organic matter in a forest soil of coastal British Columbia by NMR and pyrolysis-field ionization mass spectrometry. Plant Soil 158:69–82. https://doi.org/10.1007/BF00007919
73. R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
74. Riederer M, Matzke K, Ziegler F, Kögel-Knabner I (1993) Occurrence, distribution and fate of the lipid plant biopolymers cutin and suberin in temperate forest soils. Org Geochem 20:1063–1076. https://doi.org/10.1016/0146-6380(93)90114-Q
75. Rillig MC, Mardatin NF, Leifheit EF, Antunes PM (2010) Mycelium of arbuscular mycorrhizal fungi increases soil water repellency and is sufficient to maintain water-stable soil aggregates. Soil Biol Biochem 42(7):1189–1191. https://doi.org/10.1016/j.soilbio.2010.03.027
76. Rodríguez-Alleres M, Benito E (2011) Spatial and temporal variability of surface water repellency in sandy loam soils of NW Spain under *Pinus pinaster* and *Eucalyptus globulus* plantations. Hydrol Process 25:3649–3658. https://doi.org/10.1002/hyp.8091
77. Rodríguez-Alleres M, Benito E (2012) Temporal fluctuations of water repellency in forest soils of Galicia, NW Spain. Do soil samples dried at laboratory reflect the potential soil water repellency? Hydrol Process 26:1179–1187. https://doi:10.1002/hyp.8209
78. Rye CF, Smettem KRJ (2018) Seasonal variation of subsurface flow pathway spread under a water repellent surface layer. Geoderma 327:1–12. https://doi.org/10.1016/j.geoderma.2018.04.008
79. Scott DF (2000) Soil wettability in forested catchments in South Africa as measured by different methods and as affected by vegetation cover and soil characteristics. J Hydrol 231–232:87–104. https://doi.org/10.1016/S0022-1694(00)00186-4
80. Schonsky H, Peters A, Wessolek G (2014) Effect of soil water repellency on energy partitioning between soil and atmosphere: A conceptual approach. Pedosphere 24(4):498–507. https://doi.org/10.1016/S1002-0160(14)60036-9
81. Seaton FM, Jones DL, Creer S, George PBL, Smart SM, Lebron I, Barrett G, Emmett BA, Robinson DA (2019) Plant and soil communities are associated with the response of soil water repellency to environmental stress. Sci Total Environ 687:929–938. https://doi.org/10.1016/j.scitotenv.2019.06.052

82. Shahidzadeh-Bonn N, Azouni A, Coussot P (2007) Effect of wetting properties on the kinetics of drying of porous media. J Phys-Condens Mat 19(11):112101. https://doi.org/10.1088/0953-8984/19/11/112101
83. Smettem K, Rye C, Henry DJ, Sochacki SJ, Harper RJ (2021) Soil water repellency and the five spheres of influence: a review of mechanisms, measurement and ecological implications. Sci Total Environ 787:147429. https://doi.org/10.1016/j.scitotenv.2021.147429
84. Song E, Pan X, Kremer RJ, Goyne KW, Anderson SH, Xiong X (2019) Influence of repeated application of wetting agents on soil water repellency and microbial community. Sustainability 11:4505. https://doi.org/10.3390/su11164505
85. Suseela V, Tharayil N, Xing B, Dukes JS (2015) Warming and drought differentially influence the production and resorption of elemental and metabolic nitrogen pools in Quercus rubra. Global Change Biol 21(11):4177–4195. https://doi.org/10.1111/gcb.13033
86. Urbanek E, Walsh RPD, Shakesby RA (2015) Patterns of soil water repellency change with wetting and drying: the influence of cracks, roots and drainage conditions. Hydrol Process 29(12):2799–2813. https://doi.org/10.1002/hyp.10404
87. Van Bergen PF, Bull ID, Poulton PR, Evershed RP (1997) Organic geochemical studies of soils from the Rothamsted Classical Experiments—I. Total lipid extracts, solvent insoluble residues and humic acids from Broadbalk Wilderness. Org Geochem 26:117–135, https://doi.org/10.1016/S0146-6380(96)00134-9
88. Van't Woudt BD (1959) Particle coatings affecting the wettability of soils. J Geophys Res 64:263–267. https://doi.org/10.1029/JZ064i002p00263
89. Walden LL, Harper RJ, Mendham DS, Henry DJ, Fontaine JB (2015) Eucalyptus reforestation induces soil water repellency. Soil Res 53(2):168–177. https://doi.org/10.1071/sr13339
90. Wallis MG, Horne DJ, McAuliffe KW (1990) A study of water repellency and its amelioration in a yellow-brown sand: 1. Severity of water repellency and the effects of wetting and abrasion. NZ J Agric Res 33:139–144. https://doi.org/10.1080/00288233.1990.10430671
91. Wang J, Zheng Y, Hu H, Li J, Zhang L, Chen B, Chen W, He J (2016a) Coupling of soil prokaryotic diversity and plant diversity across latitudinal forest ecosystems. Sci Rep 6:19561. https://doi.org/10.1038/srep19561
92. Wang S, Fu BJ, Piao SR, Lü YH, Ciais P, Feng XM, Wang YF (2016b) Reduced sediment transport in the Yellow River due to anthropogenic changes. Nat Geosci 9(1):38–41. https://doi.org/10.1038/ngeo2602
93. Wang Z, Hu Y, Wang R, Guo S, Du L, Zhao M, Yao Z (2017) Soil organic carbon on the fragmented Chinese Loess Plateau: Combining effects of vegetation types and topographic positions. Soil Till Res 174:1–5. https://doi.org/10.1016/j.still.2017.05.005
94. Wang R, Hu Y, Khan A, Du L, Wang Y, Hou F, Guo S (2021) Soil prokaryotic community structure and co-occurrence patterns on the fragmented Chinese Loess Plateau: Effects of topographic units of a soil eroding catena. Catena 198:105035. https://doi.org/10.1016/j.catena.2020.105035
95. Wessel AT (1988) On using the effective contact angle and the water drop penetration time for classification for water repellency in dune soils. Earth Surf Proc Land 13(6):555–561. https://doi.org/10.1002/esp.3290130609
96. Wessels JGH (1996) Hydrophobins: Proteins that change the nature of the fungal surface. Adv Microb Physiol 38:1–45. https://doi.org/10.1016/s0065-2911(08)60154-x
97. Wessels JGH (2000) Hydrophobins, unique fungal proteins. Mycologist 14(4):153–159. https://doi.org/10.1016/s0269-915x(00)80030-0
98. White AM, Lockington DA, Gibbes B (2017) Does fire alter soil water repellency in subtropical coastal sandy environments? Hydrol Process 31:341–348. https://doi.org/10.1002/hyp.11000
99. Zavala LM, González FA, Jordán A (2009) Intensity and persistence of water repellency in relation to vegetation types and soil parameters in Mediterranean SW Spain. Geoderma 152:361–374. https://doi.org/10.1016/j.geoderma.2009.07.011
100. Zeppenfeld T, Balkenhol N, Kóvacs K, Carminati A (2017) Rhizosphere hydrophobicity: A positive trait in the competition for water. PloS One 12(7):e0182188. https://doi.org/10.1371/journal.pone.0182188
101. Zhang Z, Kaye JP, Bradley BA, Amsili JP, Suseela V (2022) Cover crop functional types differentially alter the content and composition of soil organic carbon in particulate and mineral-associated fractions. Global Change Biol 28:5831–5848. https://doi.org/10.1111/gcb.16296

102. Zheng S, Lourenço SDN, Cleall PJ, Chui TFM, Ng AKY, Millis SW (2017) Hydrologic behavior of model slopes with synthetic water repellent soils. J Hydrol 554:582–599. https://doi.org/10.1016/j.jhydrol.2017.09.013

Figures

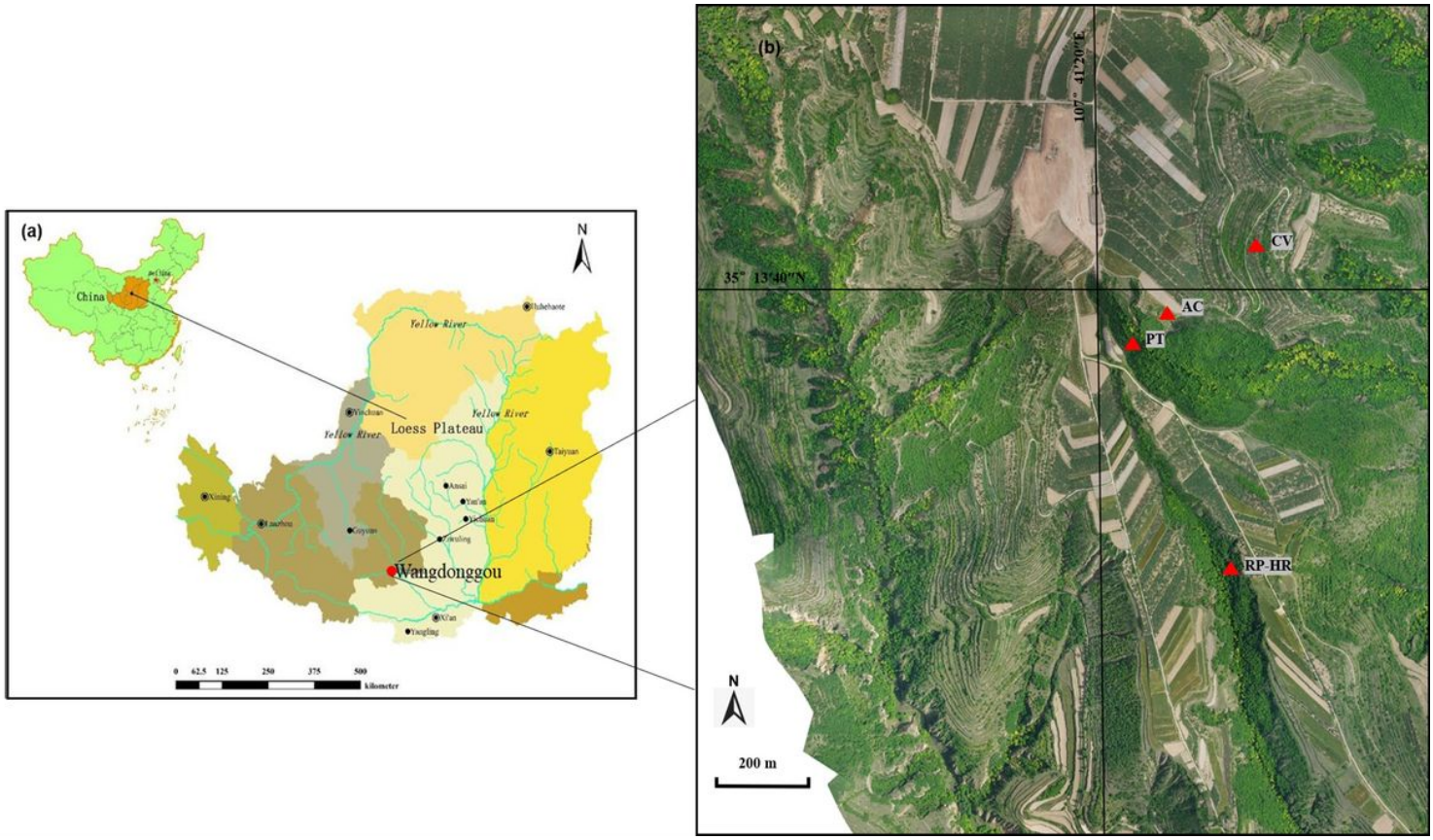

Figure 1

Information for the study area (cited from Wang et al. 2016a) (**a**) and sampling locations (**b**). PT, *Pinus tabulaeformis* Carr.; RP-HR, *Robinia pseudoacacia* L. - *Hippophae rhamnoides* L.; CV, *Coronilla varia* L.; AC, *Agropyron cristatum* (L.) Gaertn.

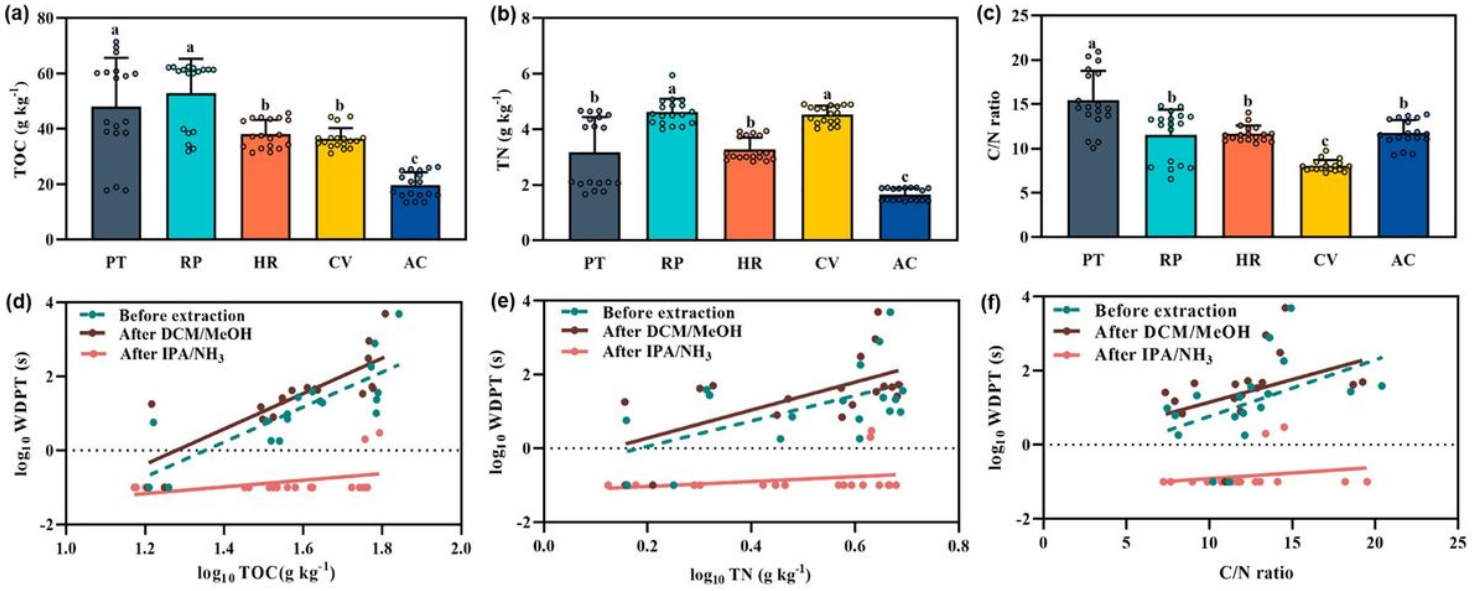

Figure 2

The soil TOC content, TN content, and C/N ratio under/around the different DPS (**a–c**). SWR (log $_{10}$ WDPT(s)) as a function of log_{10} TOC, log_{10} TN, and C/N ratio before and after DCM/MeOH and IPA/NH_3 extractions (**d–f**). Error bars in (a–c) represent standard deviations of the mean values (n = 18) with lowercase letters above bars indicating significant differences among plant species treatments in specific fractions at $p < 0.05$. The solid dots in (d–f) are the mean rates included in the correlation analysis

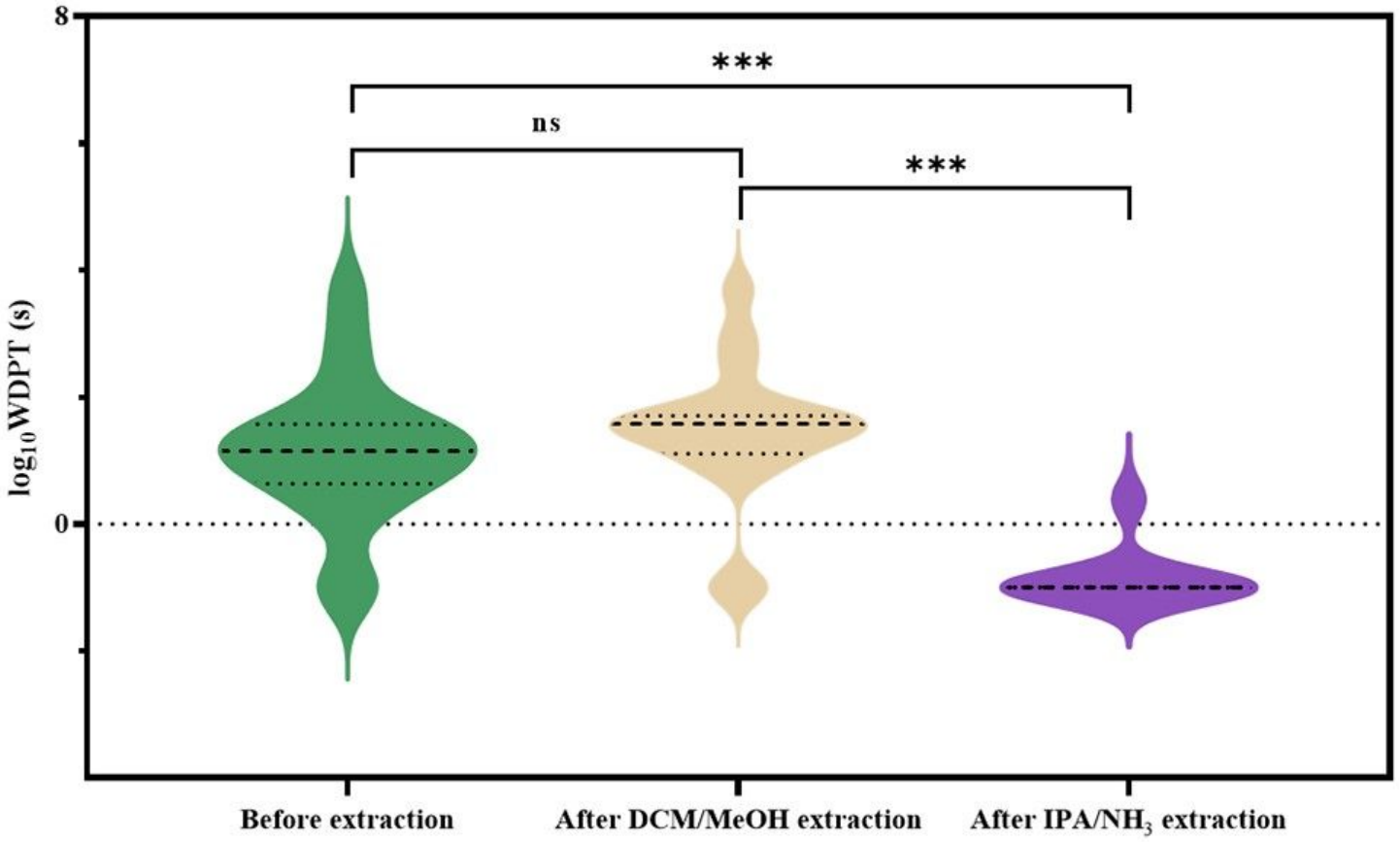

Figure 3

Mean values (log_{10} WDPT(s), n = 54) of the differences of SWR before extraction, after DCM/MeOH extraction, and after IPA/NH_3 extraction. The long-dotted lines are the medians, and the short-dotted lines are the quartiles. ns represents no significant difference ($p > 0.05$), *** represents the significant difference $p < 0.001$

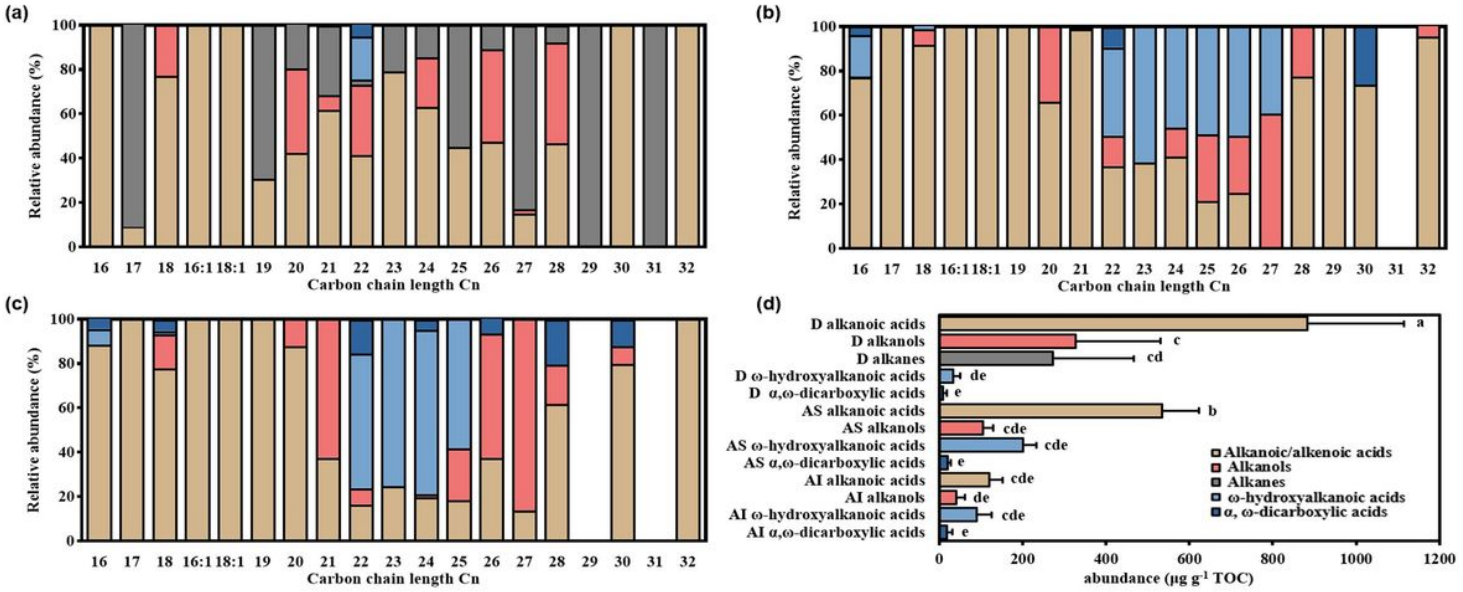

Figure 4

The relative average concentrations (μg g^{-1} TOC) of SWR biomarkers (n-alkanoic/alkenoic acids, n-alkanols, n-alkanes, *ω*-hydroxyalkanoic acids, and *α, ω*-dicarboxylic acids) in D fractions (**a**), AS fractions (**b**) and AI fractions (**c**) from all soils; and the relative abundance of SWR biomarker groups in different EFs (**d**). The different letters in (d) indicate significant differences in different EFs ($p < 0.05$). Error bars in (d) represent standard deviations of the mean values ($n = 54$). D, AS, and AI refer to DCM/MeOH soluble fraction, DCM/MeOH soluble fraction of IPA/NH_3 extract, and DCM/MeOH insoluble fraction of IPA/NH_3 extract, respectively

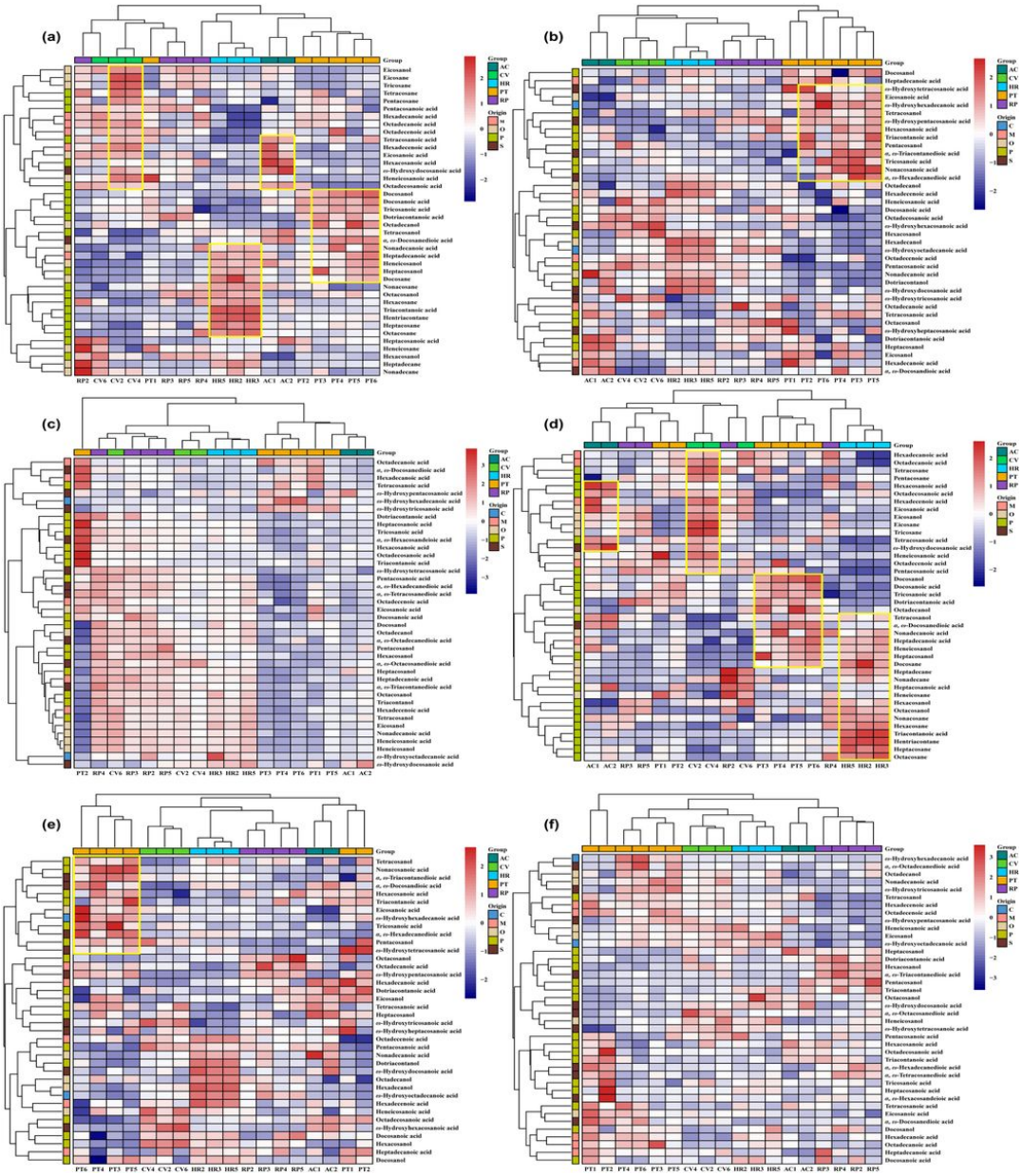

Figure 5

Heatmap and two-way hierarchical clustering of the absolute (log_{10} (μg g^{-1} soil)) and relative concentrations (log_{10} (μg g^{-1} TOC)) of all identified biomarkers in D fractions (**a, d**), AS fractions (**b, e**) and AI fractions (**c, f**) under/around the different DPS with different log_{10} WDPT (s) treatments. Each column represents a specific treatment under/around the different DPS, and each row represents a positively identified biomarker based on mass spectrum information. The red and blue colors in different cells indicate relatively higher and lower relative average concentration (μg g^{-1} soil) of specific biomarkers in each treatment, respectively. Biomarkers belonging to different clustering are highlighted in (a) and (d), showing a predominant division of soil organic carbon composition through DCM/MeOH extraction between different DPS. The soil biomarkers belonging to different clustering are also highlighted in

(b) and (e), and some biomarkers belonging to higher relative average concentrations (μg g^{-1} soil) are separately marked, showing an obvious difference between treatments based on cutin/suberin-derived biomarkers. C, cutin. M, short-chain fatty acids (SFA). O, others. P, long-chain fatty acids (LFA). S, suberin

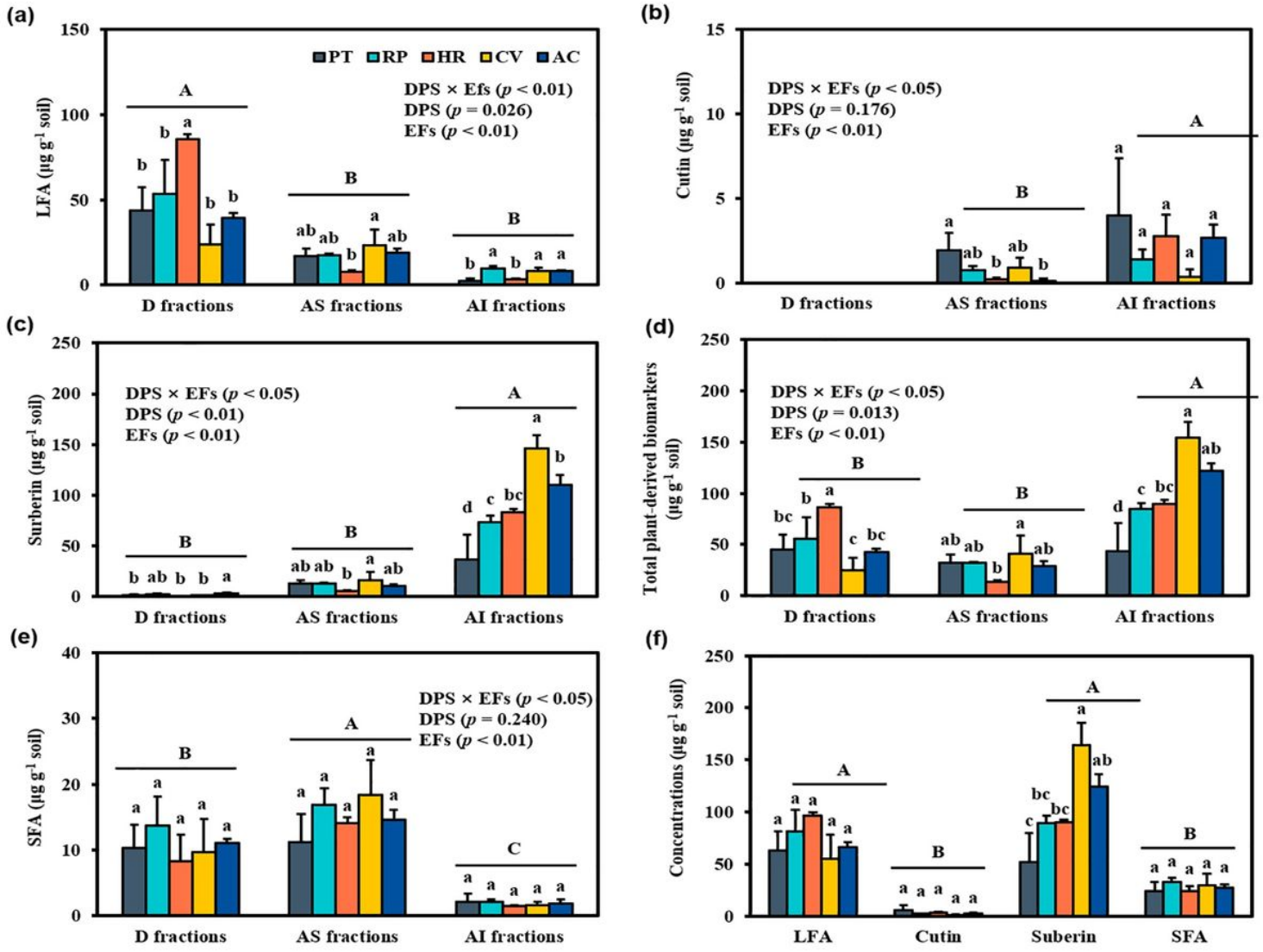

Figure 6

The absolute concentrations of biomarker groups (μg g^{-1} soil) in the D, AS, and AI fractions (**a–e**), and the total absolute concentrations (sum of D, AS, and AI fractions) of LFA, cutin, suberin, and SFA under/around the different DPS (**f**). **a** Long-chain fatty acids (>C_{24} alkanes, >C_{22} n-alkanoic acids and alkanols). **b** Cutin (C_{14}–C_{18} *ω*-hydroxyalkanoic acids). **c** Suberin (*α, ω*-dicarboxylic acids [C_{16}–C_{24}; saturated and substituted] and *ω*-hydroxyalkanoic acids [C_{20}–C_{30}; saturated and substituted]). **d** Total plant-derived biomarkers (sum of LFA, cutin, suberin). **e** Short-chain fatty acids (C_{16}–C_{18} n-alkanoic and n-alkenoic acid). DPS × EFs, the interaction effect of dominant plant species and extraction fractions. Error bars are ± SD (n = 18, 12, 9, 9, 6 correspond to the number of soil samples from PT, RP, HR, CV, and AC) of the mean with lowercase letters above bars indicating significant differences between DPS in specific fraction at $p < 0.05$. The different uppercase letter above bars indicates significant differences among EFs across all soils at $p < 0.05$

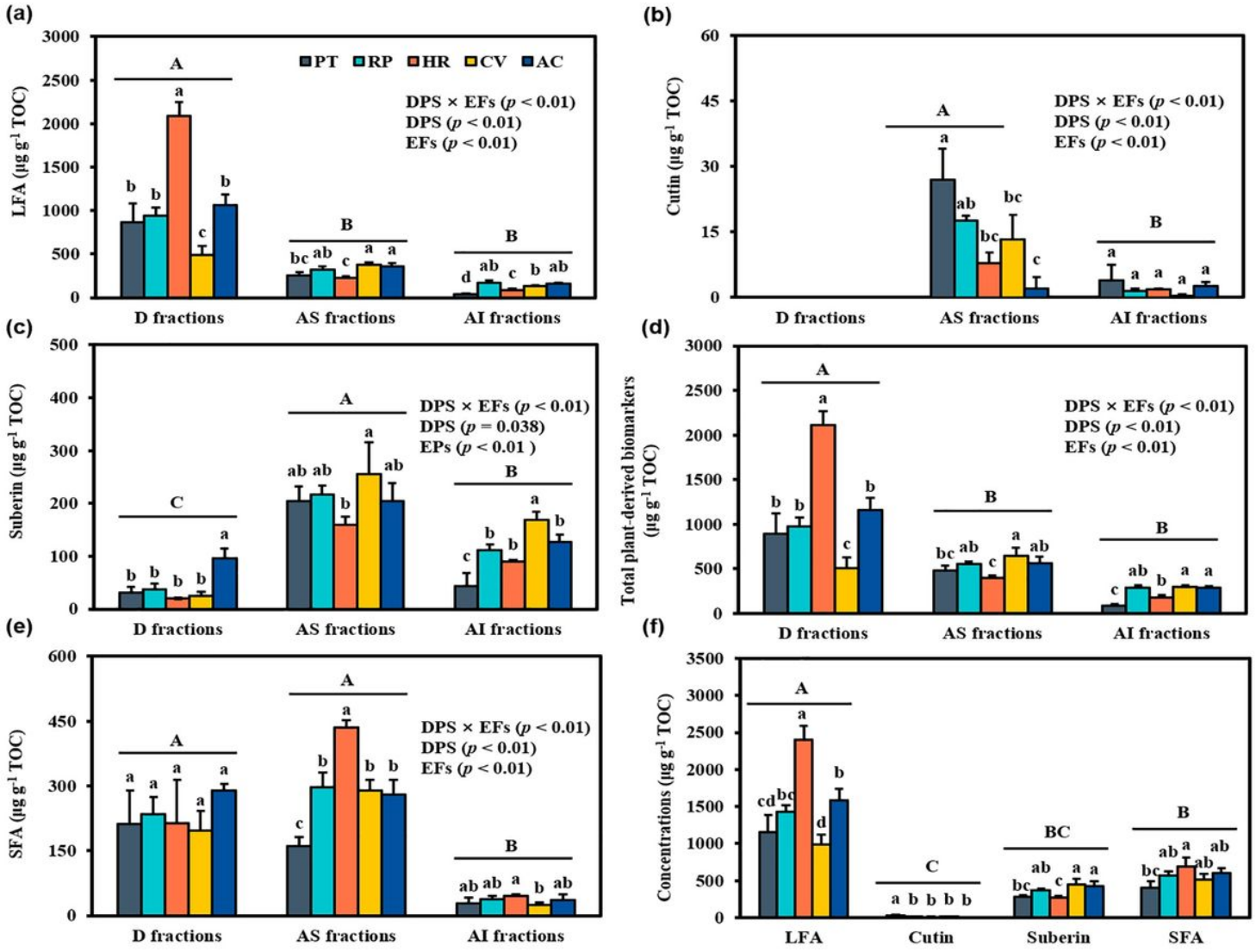

Figure 7

The relative concentrations of biomarker groups (μg g^{-1} TOC) in D, AS, and AI fractions (**a–e**), and the total relative concentrations (sum of D, AS, and AI fractions) of LFA, cutin, suberin, and SFA under/around the different DPS (**f**). **a** Long-chain fatty acids (>C_{24} alkanes, >C_{22} n-alkanoic acids, and alkanols). **b** Cutin (C_{14}–C_{18} *ω*-hydroxyalkanoic acids). **c** Suberin (*a, ω*-dicarboxylic acids [C_{16}–C_{24}; saturated and substituted] and *ω*-hydroxyalkanoic acids [C_{20}–C_{30}; saturated and substituted]). **d** Total plant-derived biomarkers (sum of LFA, cutin, suberin). **e** Short-chain fatty acids (C_{16}–C_{18} n-alkanoic and n-alkenoic acid). DPS × EFs, the interaction effect of dominant plant species and extraction fractions. Error bars are ± SD (n = 18, 12, 9, 9, 6 correspond to the number of soil samples from PT, RP, HR, CV, and AC) of the mean with lowercase letters above bars indicating significant differences between DPS in specific fraction at $p < 0.05$. The different uppercase letter above bars indicates significant differences among EFs across all soils at $p < 0.05$

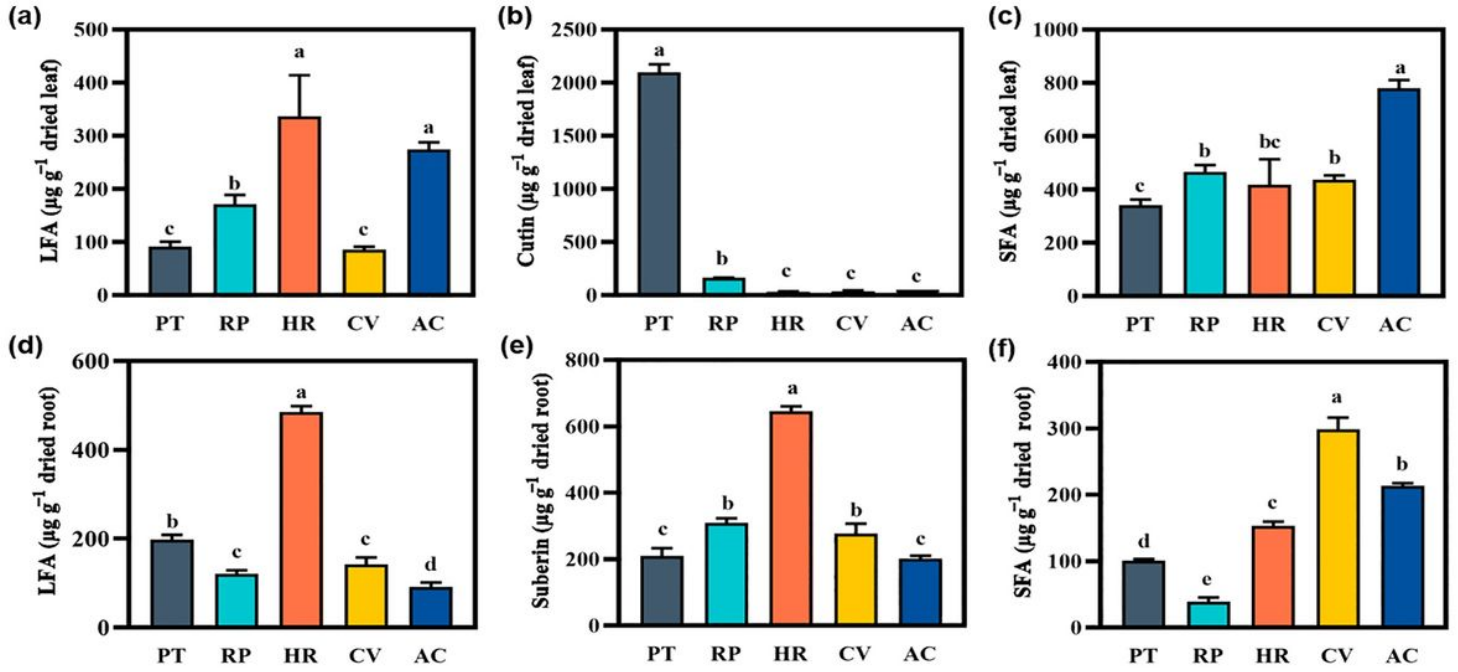

Figure 8

The group abundances of ester-bound lipids upon BF_3–MeOH hydrolysis of leaves (**a–c**) and roots(**d–e**). Error bars represent standard deviations of the mean values (n = 6) with lowercase letters above bars indicating significant differences among plant species treatments in specific fractions at $p < 0.05$

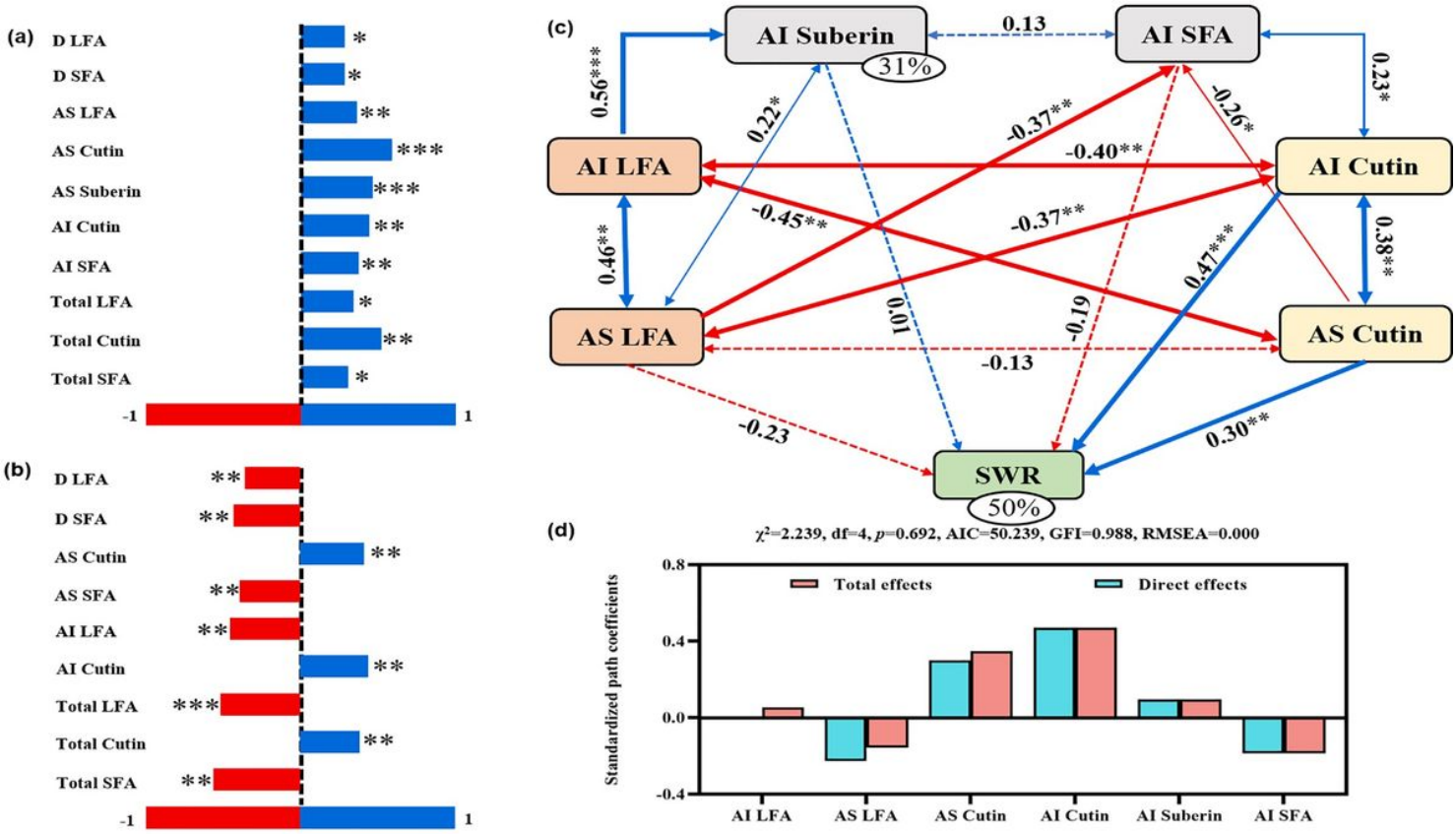

Figure 9

Correlations between selected absolute (log_{10} (μg g^{-1} soil)) and relative concentrations (log_{10} (μg g^{-1} TOC)) of compound groups and SWR in all soils (**a–b**), structural equation modeling (SEM) reveals the direct and indirect effects of compound groups on SWR (**c**), and the total and direct effect of each predictor on SWR as estimated from the model parameters (**d**). The width of the arrow joining two boxes in (c) is proportional to the strength of the relationship, numbers on arrows are normalized path coefficients, positive relationships are represented by a blue arrow, negative by red. Solid arrows in (c) indicate significant standardized path coefficients, dotted arrows indicate non-significant standardized path coefficients ($p > 0.05$), and the percentage near the circle indicates the variance R^2 explained by the model. The criteria for the evaluation of structural equation modeling fits, such as the *p*-values, model chi-square (χ^2), goodness-of-fit index (GFI), and the root mean square error of approximation (RMSEA), were adopted according to Hooper et al. (2008), and the full output from the model is in Table S4. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

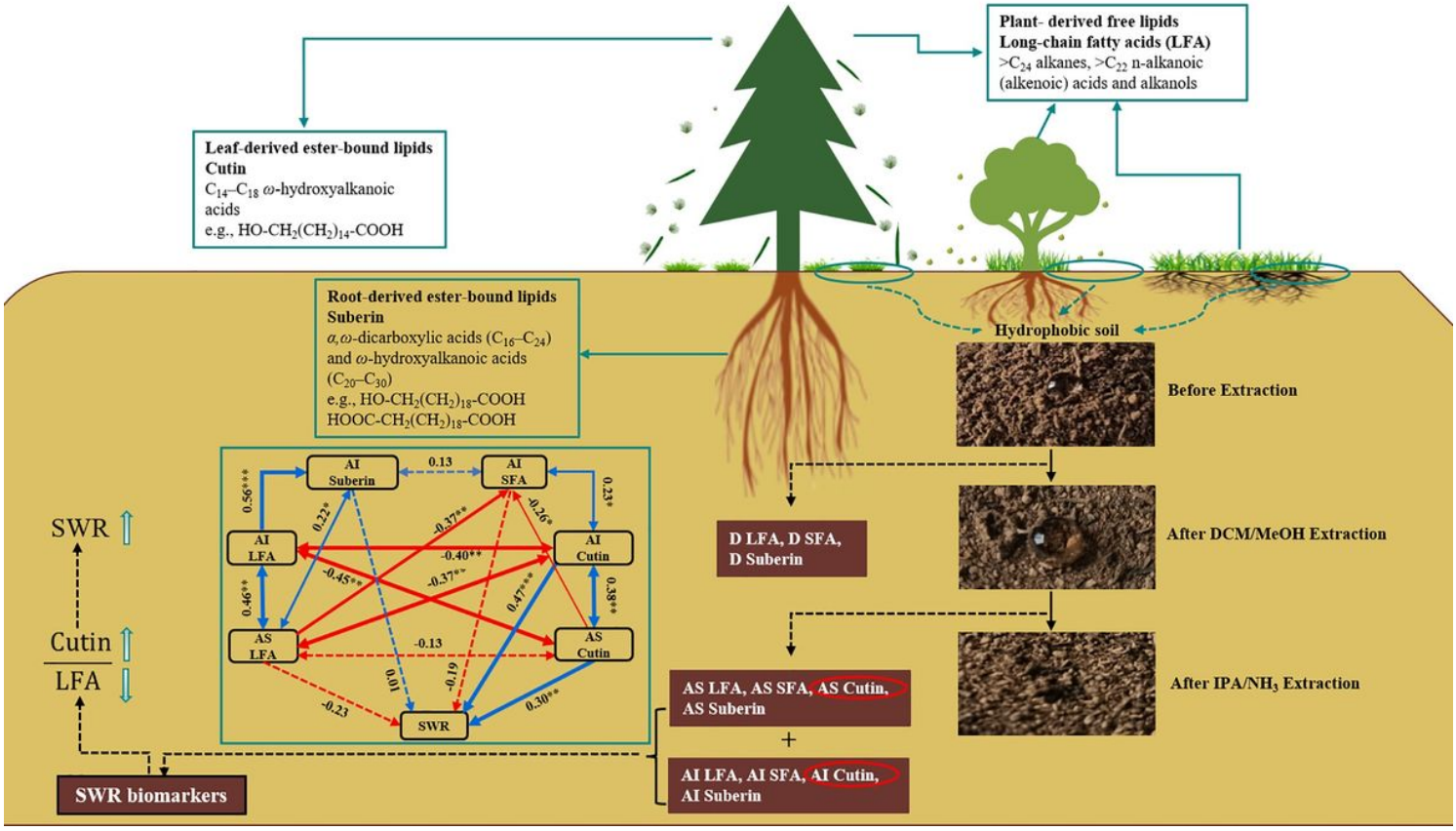

Figure 10

Conceptual framework of the effect of biomarker groups on SWR

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformationSI.docx