Lipid biogeochemistry of paddy soils

Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität

zu Kiel

vorgelegt von

Cornelia Müller-Niggemann

Kiel, 2015

Erster Gutachter:	Prof. Dr. Lorenz Schwark
Zweiter Gutachter:	Prof. Dr. Volker Thiel
Tag der mündlichen Prüfung:	26.01.2016
Zum Druck genehmigt:	26.01.2016

gez. Prof. Dr. Wolfgang J. Duschl, Dekan

Always remember…
You're Braver than you believe,
Stronger than you seem
and Smarter than you think.

(Christopher Robin)

Contents

Index of figures

Index of tables

Index of supplementary figures

Index of supplementary tables

Abstract

Rice constitutes one of the most important staple foods for more than half of the World's population. Presently 157 million ha are under rice cultivation with a demand expected to increase continuously. Typical wet rice field management includes ploughing and puddling of soils under submergence as well as alternate flooding and draining of fields, processes that initiate the development of so-called paddy soils. On global scale paddy soils may evolve on most different soil substrates. Soil redox-conditions associated with paddy management are considered to diminish decomposition rates of soil organic matter (SOM), which in favours its accumulation as well as the emission of greenhouse gases (N_2O, CO_2 and CH_4). Hence, it is of paramount importance to study the effect of wetland rice cultivation on biogeochemical processes. Due to the large spatial extent of and high carbon turnover in rice paddies their role in global carbon cycling as carbon (CO_2) sinks or sources needs to be investigated. Up to now, biogeochemical cycling in such agroecosystems has been investigated intensively, whereas molecular lipid biomarker studies remained scarce. Investigating biogeochemical processes in paddy soil based on lipid and xenobiotic constitution may provide further insights into the impact of rice paddy management onto SOM composition and preservation. In soils, lipids are either decomposition products of standing biomass (mostly plants and microorganism) or represent viable microbial biomass itself.

This dissertation aimed at a comprehensive study of management-induced effects on lipid influx into paddy soils and subsequent lipid evolution under the environmental conditions prevailing. A crucial finding was that management and climatic conditions dramatically influenced the lipid composition in paddy compared to upland (non-paddy) soils. In this context, recurring anaerobic conditions due to periodical submergence of paddy soils had a profound important influence on SOM preservation potential and microbial community structures. Climatic conditions affected biogeochemical lipid cycling to a larger extent than the substrate on which the soil had developed. Accelerated cycling in tropical versus subtropical locations was amply reflected in preservation and composition of plant lipids and microbial membrane lipids.

One part of this dissertation addressed changes in lipid composition due to the development of a 2000 yr old paddy soil and a 700 yr old upland soil chronosequence. Both chronosequences evolved on identical substrate, tidal wetland sediment in land reclamation areas of the Chinese Zhejiang Province. Agricultural utilization of former tidal sediments affected topsoil lipid composition, with plant wax derived lipid constitution adopting rapidly within only 50 yr. Topsoil *n*-alkane patterns corresponded to those of the crops planted, as confirmed by a parallel chemotaxonomic study of wax alkane composition in wetland rice and various upland crop plants (rice, maize, sorghum, rape, mustard, beans, and cotton). Elevated proportions of $n\text{-}C_{25}$ and $n\text{-}C_{33}$ alkanes in soils under paddy management indicated a predominant input and preservation of rice root derived lipids. Exclusively in paddies an increased accumulation of plant lipids was linked to long-term utilization identifying soil management as an important control factor in agroecosystem biogeochemistry.

Paddy management exerts a profound control on microbial lipids via humidity, soil pH, temperature, and redox conditions, all of which regulate organic matter influx into soil. Analogous to plant wax lipids, the membrane lipid composition of archaea and bacteria revealed a rapid change in microbial consortia by adaption to soil management practises during evolution of the paddy soil chronosequence. Relative distributions of microbial lipids, present as isoprenoid and branched glycerol dialkyl glycerol tetraethers (GDGTs) discriminated all soils according to management type. Elevated abundance of bacterial membrane lipids and concomitant increase of methanogen marker molecules characterized paddy soils. The generally assumption that anaerobic and acid-tolerant bacteria (supposed to be Acidobacteria) synthesize branched tetraether lipids was supported by exceptionally high abundances of branched GDGTs (brGDGTs) in paddy soil. In tropical paddy soil higher air temperature and humidity primarily stimulated accelerated biogeochemical cycling but enhanced management intensity seemed to be of complementary importance. The isoprenoidal GDGT crenarchaeol, exclusively associated to ammonia-oxidizing *Thaumarchaeota,* occurred in higher proportions in dry upland soils compared to submerged paddy soils. A comparative study of lipids and gene abundances of the functional gene *amoA* AOA, which encodes the enzyme for ammonia oxidation in archaea, exhibited an inconsistent

relationship in paddy soils. This observation may either have indicated a higher content of fossil crenarchaeol or suggested the presence of *Thaumarchaeota* lacking these genes. It was previously anticipated that cyclization and methylation of brGDGTs in soil predominantly relates to mean annual air temperature (MAT) and/or soil pH. However, the results presented here identified soil moisture as a further important environmental variable to affect brGDGTs distributions. The brGDGT-based temperature (T_{MC}) of non-flooded upland soils in general was found to be higher than in adjacently located paddy soils. This was attributed to differences in soil moisture that may also have had an effect on the soil temperature.

In conclusion, data presented in this dissertation demonstrated crop plant as well as microbial derived lipid patterns of paddy soils and non-paddy soils to differ distinctively. This confirms especially organic matter input and specific microbial consortia to have adapted to the conditions prevailing. The results received here broadened our understanding of agroecosystems soil development in general, microbial processes occurring in rice paddies, and the chemotaxonomy of crop plants.

XX

Zusammenfassung

Reis ist eines der wichtigsten Grundnahrungsmittel für mehr als die Hälfte der stetig wachsenden Weltbevölkerung. Derzeit wird weltweit eine Fläche von 157 Millionen ha für den Anbau von Reis genutzt, welche sich aufgrund der steigenden Nachfrage weiter ausdehnen wird. Die typische Bodenbewirtschaftung eines Reisfeldes umfasst das Nasspflügen, einem Verfahren aus Pflügen (ploughing) unter Wasser und dem anschließenden Abdichten (puddling) der Bodenoberfläche, sowie das abwechselnde Fluten und Entwässern der Felder, welche zusammen zur Herausbildung des sogenannten „paddy soil“ führen. Diese paddy Böden können global gesehen auf den unterschiedlichsten Substraten entstehen. Die mit der paddy Bewirtschaftung in Zusammenhang stehenden Redoxbedingungen im Boden, werden generell mit einer geringeren Zersetzungsrate und der bevorzugten Anreicherung der organischen Bodensubstanz sowie mit der Freisetzung von Treibhausgasen (N_2O, CO_2 und CH_4) in Verbindung gebracht. Deshalb ist es von großer Bedeutung, die generellen Auswirkungen des Nassreisanbaues auf die biochemischen Konsequenzen hin zu untersuchen. Aufgrund der großen Flächenausdehnung und des hohen Kohlenstoffumsatzes von Reisböden ist es besonders für die Erforschung des globalen Kohlenstoffkreislaufes wichtig, zu wissen, ob diese paddy Böden eher als Kohlenstoff-Senke (CO_2) oder -Quelle fungieren. Derzeit existieren zahlreiche Untersuchungen, die sich mit den biochemischen Stoffkreisläufen in solchen Agrarökosystemen befasst haben, aber molekulare Biomarker-Studien sind bisher äußerst selten. Die Untersuchung von biogeochemischen Prozessen in Reisböden basierend auf der Analyse von Lipiden und Xenobiotika kann einen viel detaillierteren Einblick in die Auswirkungen der paddy Bodenbewirtschaftung auf die Zusammensetzung und Erhaltung der organischen Bodensubstanz geben. In Böden sind Lipide entweder Zersetzungsprodukte abgestorbener Organismen und deren abgelagerte Biomasse (hauptsächlich Pflanzen und Mikroorganismen) oder sie entstammen der noch lebenden mikrobiellen Biomasse.

Das Ziel dieser Dissertation war es, eine umfassende Studie über den Einfluss der paddy Bodenbewirtschaftung auf den Eintrag und die Entwicklung der Lipidzusammensetzung unter verschiedenen gegebenen Umweltbedingungen zu erhalten. Zu den wichtigsten Ergebnissen dieser Arbeit zählen, dass sowohl die Bewirtschaftung als auch die klimatischen

Gegebenheiten die Lipidzusammensetzung im Boden beeinflussen, wie es z.B. in den signifikant unterschiedlichen Ergebnissen von paddy Böden und upland Böden (non-paddy Böden sind Agrarböden ohne Nassreisbewirtschaftung) angezeigt wurde. Hierbei spielt die periodische Überflutung der paddy Böden, bei welchen zyklisch anaerobe Bedingungen auftreten, eine zentrale Rolle für die Erhaltung der organischen Bodensubstanz sowie für die Ausbildung der mikrobiellen Vergesellschaftung. Zudem beeinflussten die klimatischen Bedingungen die biogeochemischen Lipidkreisläufe intensiver als das entsprechende Substrat, auf dem sich die Böden entwickelten. Hierbei schreiten diese Kreisläufe in den Tropen deutlich beschleunigter voran als in den Subtropen, was sich auch in der Erhaltung und Zusammensetzung von Pflanzenlipiden und der mikrobiellen Membranlipide widerspiegelte.

Ein Teil der Arbeit befasste sich mit der Entwicklung der Lipidzusammensetzung innerhalb einer 2000 Jahre alten paddy Chronosequenz und einer 700 Jahre alten upland Chronosequenz, die sich beide parallel auf den ehemaligen Gezeitensedimenten (tidal wetland Sedimente) einer Landgewinnungsfläche in der chinesischen Provinz Zhejiang entwickelt haben. Hierbei wurde festgestellt, dass die landwirtschaftliche Nutzung von ehemaligen Gezeitensedimenten einen Einfluss auf die Lipidzusammensetzung des Oberbodens hat. Rasche Veränderung, bereits innerhalb der ersten 50 Jahre, der Zusammensetzung der im Oberboden vorkommenden Pflanzenwachse wurde beobachtet. Die *n*-Alkan Verteilungsmuster der Böden entsprachen hierbei jeweils der Wax-Signatur der hauptsächlich angebauten Nutzpflanzen, was durch die chemotaxonomische Studie von Reis- und verschiedensten upland-Kulturpflanzen (Mais, Sorghum, Raps, Senf, Bohnen und Baumwolle) belegt wurde. So zeigte insbesondere der erhöhte Anteil an *n*-C_{25} und *n*-C_{33} Alkanen einen überwiegenden Eintrag von Reiswurzeln in paddy Böden an. Zudem konnte nur in den paddy Böden eine deutliche Anreicherung dieser Lipide über die Nutzungszeit beobachtet werden, was den deutlichen Einfluss der Bewirtschaftung auf die Biogeochemie von Agroökosystmen dokumentiert.

Die Zusammensetzung mikrobieller Bodenlipide ist ebenfalls stark von der Bewirtschaftungsform beeinflusst, die z.B. die Feuchtigkeit, den Boden pH, die Bodentemperatur, die Redoxbedingungen und den Eintrag organischer Substrate steuern kann.

Analog zur Entwicklung bodenbürtiger Pflanzenwachse, zeigten die Membranlipide der Archaeen und Bakterien eine schnelle Anpassung der mikrobiellen Vergesellschaftung an die veränderte Nutzungsart während der Bodenentwicklung innerhalb der paddy Chronosequenz an. Hierzu wurde die relative Verteilung von mikrobiellen Membranlipiden wie isoprenoidalen und verzweigten Glycerol Dialkyl Glycerol Tetraethern (GDGTs) herangezogen, welche es erlaubten, Böden unterschiedlicher Bewirtschaftungsform voneinander zu diskriminieren. Dabei waren insbesondere das vermehrte Vorkommen von bakteriellen Membranlipiden und der Anstieg von methanogenen Markern typisch für paddy Böden. Die deutlich höheren Konzentrationen von verzweigten GDGTs (brGDGT) unterstützen außerdem die Annahme, dass diese Membranlipide von anaeroben und säuretoleranten Bakterien synthetisiert werden. Diese sind derzeit nicht identifiziert, werden jedoch den Acidobakterien zugeschrieben. Die beschleunigten biogeochemischen Kreisläufe in tropischen paddy Böden wurden primär durch die allgemein höhere Temperatur und Luftfeuchtigkeit initiiert, aber auch die erhöhte Intensität der Bewirtschaftung (bis zu drei Ernten per annum) scheint hierbei eine Rolle zu spielen. Eine erhöhte Präsenz von Crenarchaeol, einem isoprenoidalen GDGT welches nur mit Ammonium-oxidierenden *Thaumarchaeota* assoziiert ist, tritt interessanterweise in höheren Konzentrationen in den trockenen upland Böden gegenüber den gefluteten paddy Böden auf. Die vergleichende Studie von Lipiden und den Genabundanzen des funktionellen Genes *amoA* AOA, welches das Enzym für die Ammoniumoxidation in Archaeen kodiert, zeigen eine voneinander unabhängige Verteilung im paddy Boden. Diese Beobachtung kann einerseits auf einen deutlich höheren Anteil an fossilem Crenarchaeol hindeuten oder suggeriert vielmehr die Anwesenheit von *Thaumarchaeota,* die dieses funktionelle Gen nicht beinhalten. Bisher wurde angenommen, dass die Cyclisierung und die Methylierung von brGDGTs in Böden hauptsächlich von der Jahresdurchschnittstemperatur (MAT) und/oder des Boden pHs gesteuert werden. Die Ergebnisse dieser Arbeit zeigen hingegen, dass die Bodenfeuchte ebenfalls eine Umweltvariable ist, die einen großen Einfluss auf die Verteilung der brGDGTs hat. Die mittels brGDGTs berechnete Temperatur (T_{MC}) war generell höher in den nicht gefluteten upland Böden im Vergleich zu den benachbarten paddy Böden, was auf die unterschiedliche Bodenfeuchte, welche die Bodentemperatur mitreguliert, zurückzuführen ist.

Zusammenfassend kann gesagt werden, dass die in dieser Arbeit vorgestellten Daten es ermöglichten, den Pflanzeneintrag sowie die mikrobielle Vergesellschaftung in unterschiedlich bewirtschafteten Agrarböden (paddy Böden und upland Böden) mittels spezifischer Lipidmuster zu unterscheiden. Die hier erhaltenen Ergebnisse erweitern unser Verständnis über die (landwirtschaftliche) Bodenentwicklung, über mikrobiologische Prozesse in paddy Böden und der Chemotaxonomie von Kulturpflanzen.

1. Introduction

1.1 Soils

Soil is the unconsolidated biologically active part of the upper Earth's crust consisting of weathered minerals, liquids, gases, organic materials in various stages of decay and living organisms (Soil Survey Staff, 1999; Schaetzl and Thompson, 2015). It is a natural body that may develop on various parent materials (Schaetzl and Thompson, 2015) as e.g. igneous, metamorphic or sedimentary rocks as well as unconsolidated aeolian, alluvial or lacustrine sediments. Pedogenesis is the process of soil formation, which is induced by the interaction of topographical conditions, climate and living matter upon initial material over time (Brady and Weil, 2002). The upper boundary is defined as the interface to air or the cover of vegetation, while the lower boundary separates the soil layers from the parent material. Soils composed of different distinctive mineral layers that were formed through the environmental conditions prevailing and soil processes (e.g. illuviation of clay, decalcification, gleization). Depending on its stage of development, the natural mineral soil-body can be vertically divided into an A-horizon (topsoil) at the top, a subjacent B-horizon (subsoil) and a C-horizon (slightly weathered parent substrate) at the bottom, all of which can be further subdivided into various subcategories (Press and Siever, 1995, Blume et al., 2010). Biological as well as human activities affect topsoils most commonly. Thereby, characteristic features may be influenced, including the composition and accumulation of organic matter (litter, humins, lipids), various transformation processes (humification, weathering of minerals), and mobilization or relocation processes (leaching of clay minerals, eluviation, podsolization). Main processes occurring in deeper subsoils are *in situ* weathering and illuviation of topsoil dissolved minerals (Schaetzl and Thompson, 2015). Soil processes are affected by soil properties (e.g. pH, ion exchange capacity, porosity, temperature, moisture and biota), which are not static but rather dynamic with fluctuation during the seasons. A soil can be studied on different time-scales e.g. in a short-term (days, weeks, months) or in long-term (years, decades, centuries) perspective.

Human activities are also important soil-forming factors, which contribute to the development of so called Anthrosols. These soil types include human-made and human-influenced soils

with anthropogenic processes as e.g. mechanical operations, long and intensive agricultural usage, fertilization (organic/inorganic) and wet cultivation (IUSS Working Group WR, 2007). Especially the long-term utilization of arable soils, comprising e.g. cultivation of crops, irrigation, artificial drainage and application of fertilizers, have a strong effect on most soil properties and processes (Schaetzl and Thompson, 2015). The history of anthropogenic soil modification starts amongst others with the cultivation and domestication of plants (cereal grasses as e.g. einkorn wheat, emmer wheat, barley) already during the Neolithic, around 10500 calibrated years before present (cal BP) in south-west Asia (Fuller et al., 2010; Zohary et al., 2012). In 2012 the FAO statistical division (2015) reported that approximately 12% of the World´s surface area was used as arable land.

1.2 Paddy soils

One of the most important cereals worldwide was and is rice. The recent Asian rice (*Oryza sativa*) belongs to the Poaceae and is an annual grass, which generally grows on flooded fields (water from irrigation, rainfed or floodplain systems) under various climatic conditions. Numerous rice cultivars have been developed via centuries of rice domestication, including approximately 100000 different rice varieties (http://www.knowledgebank.irri.org). Currently more than fifty percent of the world population feeds on rice. Approximately an area of 157 million ha is covered by soils under rice paddy management (Fig. 1.1), which contributes more than 18% to the total area of the ten major crops worldwide (FAO, 2003).

To date, the earliest evidence of rice was found in the Middle Yangtze region, as e.g. in the Yuchanyan Cave (Hunan Province in China) where several grains of ancient rice with an age of around 12000-9000 cal BP were identified (Gong et al., 2007 and references therein; Fuller and Qin, 2009). At that time wild rice was mainly gathered in marshes (Fuller and Qin, 2009). The regular cultivation of wild rice started probably 8000-7000 cal BP on peripheral areas of wetlands (Fuller and Qin, 2009; Wang et al., 2010). The first typical wetland rice cultivation, which is comparable with the paddy management of today, was performed later in the Lower Yangtze region during the Majiabang culture (5000 cal BP) in Chuodun and Caoxieshan (Cao et al., 2006; Fuller and Qin, 2009; Wang et al., 2010).

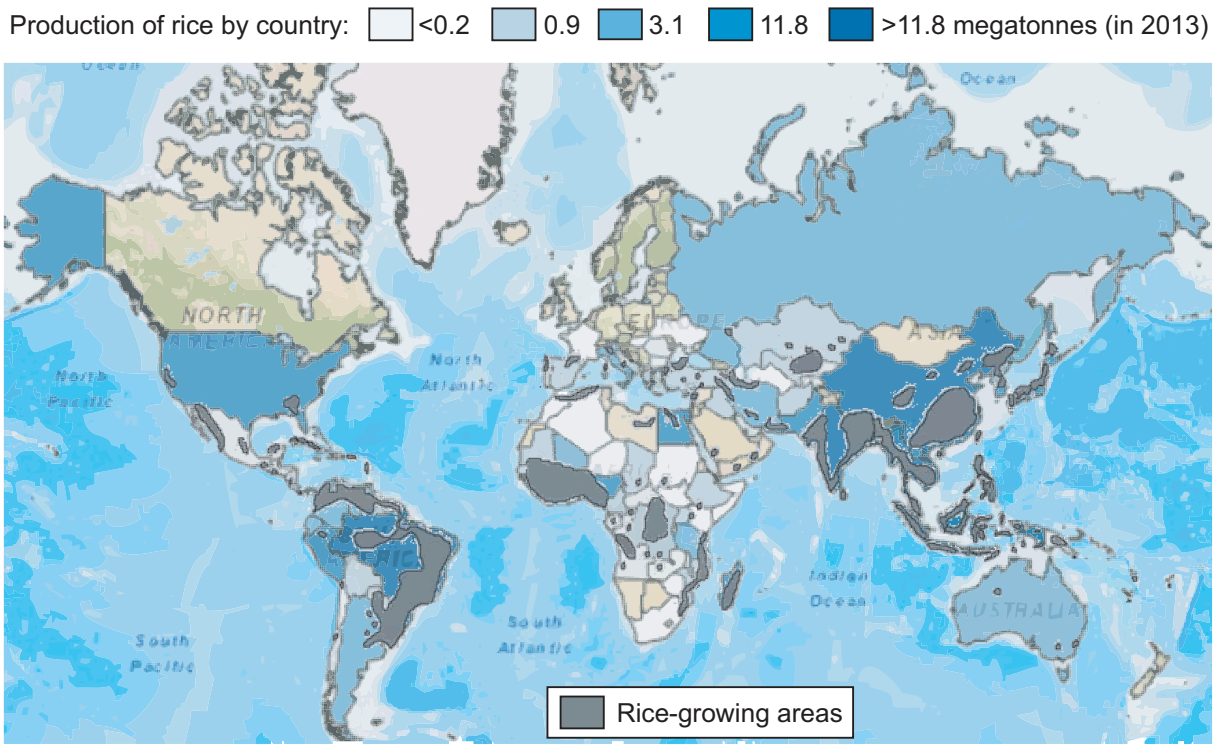


Fig. 1.1. Map of production quantities of rice that was cultivated in 2013 under paddy management according to FAO statistics (2013; http://faostat3.fao.org/home/E). Top producers were China (205 Mt), India (159 Mt) and Indonesia (71 Mt). Locations of rice-growing areas are marked according to Fernando (1993).

Long and intensive cultivation of wetland rice leads to specific soil properties, which affect the development of a so-called paddy soil. Ploughing and puddling, additionally to alternating artificial submerging and draining of fields are typical management practises during the cultivation of rice that control the development of diagnostic horizons in soil (Kögel-Knabner et al., 2010). Puddling is a kind of land preparation where the water saturated soil is tilled by ploughing and harrowing. This leads to the breakdown of soil aggregates, which reduces the bulk density and creates an uniform pedostructure (Gaunt et al., 1995). Paddy topsoils commonly comprise an approximately 20 cm thick anthraquic horizon that contains a puddled layer above and a dense plough pan at the base. The continuous paddy management leads to the slow infiltration of mobilized Fe-Mn into deeper soil layers. Furthermore, the successive and long-term wet cultivation of rice influences the development of a further characteristically

horizon below the anthraquic horizon, which is designated as hydragric horizon (Kölbl et al., 2014). This horizon may exhibit various distinctive features, as e.g. reduction in the pores (soil particles being coated or have halos with a chroma of 2 or less) and oxidative characteristics in the matrix (accumulation of Fe and/or Mn). Fully developed paddy soils were classified as Hydragric Anthrosols (IUSS Working Group WR, 2007). The investigation of paddy soils located on subtropical marshland (Kölbl et al., 2014), showed that the formation of these hydragric horizon may require several hundreds of years until full development (here at least 700 yr paddy management was needed). A typical paddy soil profile including specific pedogenetic horizons is shown in Fig. 1.2. (designation according to FAO, 2006).

Desalinization and decalcification are important processes in topsoils, with the latter exhibit higher rates in soils under paddy management compared to equivalent developed upland soils (Kölbl et al., 2014). In general, the process of decalcification is the dissolution of in water insoluble carbonates (e.g. calcium carbonate) through CO_2 and acidic soil water followed by the translocation to deeper soil layers (Borggaard, 1997).

$$CaCO_3 + CO_2 + H_2O \rightarrow Ca^{2+} + 2HCO_3^- \quad (1)$$

The accelerated loss of carbonates in paddy soils is particularly related to alternating oxic and anoxic conditions that increase the production of CO_2 via the aerobical decomposition of organic matter, methanogenesis and respiration from roots (van den Berg and Loch, 2000). High numbers of redox cycles, induced by higher numbers of rice cultivation cycles, intensifies mineral weathering, mineral transformation (Nanzyo et al., 1999) and leaching processes (Ponnamperuma, 1972; Kölbl et al., 2014). Redox reactions alter the biochemistry of soils.

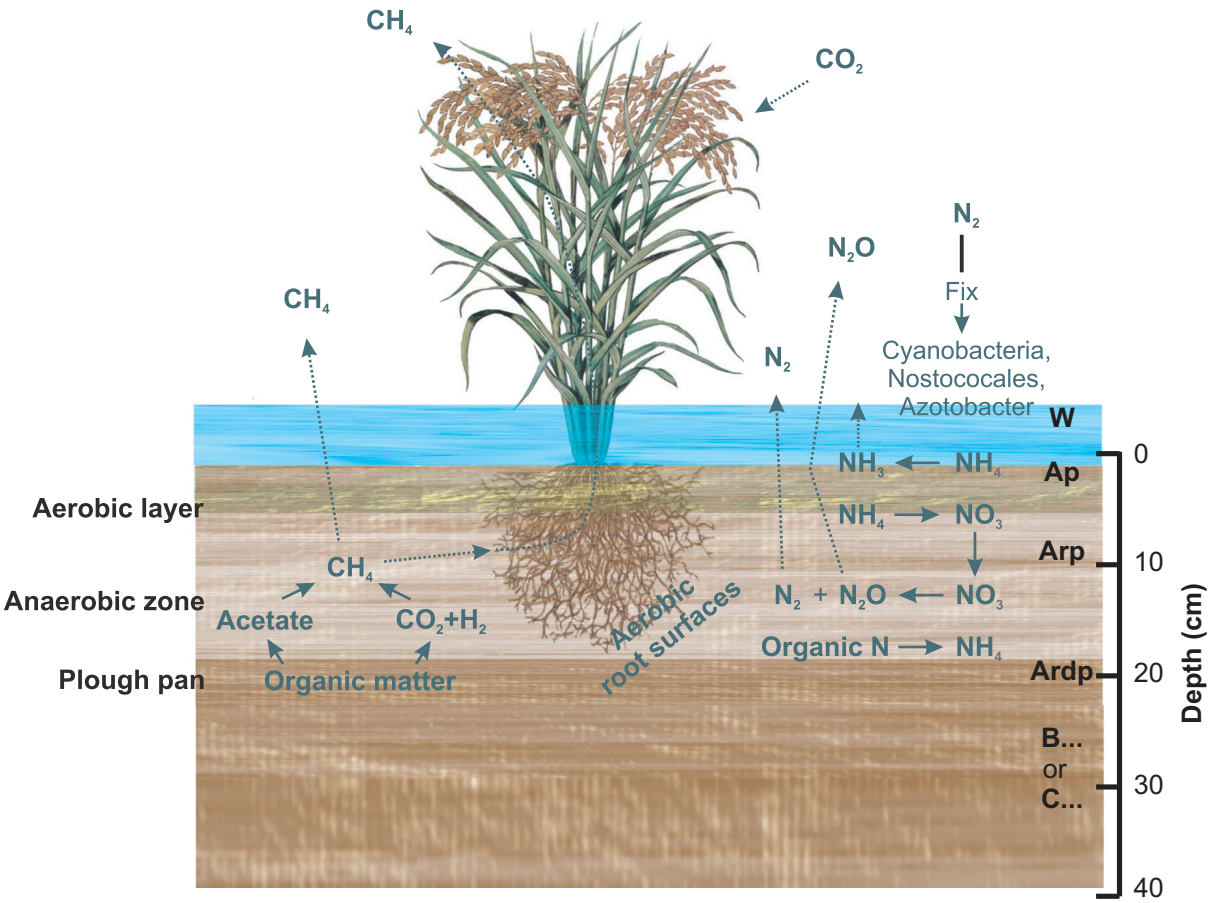


Fig. 1.2. Typical paddy soil profile (modified from Kögel-Knabner et al., 2010). **W:** oxic water layer, **Ap:** oxic and partially oxic zone during oxygen release from roots, **Arp:** reduced puddled layer as upper part of anthraquic horizon, **Ardp:** compact plough pan as lower part of anthraquic horizon with stagnic and reduced conditions as well as low hydraulic permeability, **B or C:** oxic or reduced subsurface horizon consist of either subsoil, which can have a 10 cm thick hydragric horizon, or low pedogenic affected parent rock material. Reduced conditions prevail in gleyic moisture regimes.

Primary regulators of the redox potential in wetlands are: temperature, pH, fluctuation of the water table as well as the presence of electron acceptors and donors, which are not in equilibrium within a biological system due to continuous addition and removal of electrons. Especially, the decomposition of organic matter coupled to the absence of O_2 in soils under submerged conditions initiates a decline of the redox potential (van den Berg and Loch, 2000), which drops toward anoxic conditions already 5 days after flooding of paddy soil (Kölbl et al., 2014). The principle here is that microbially mediated degradation of organic

compounds via aerobic respiration, using oxygen as electron acceptors, decreases until a stabilization of highly reduced conditions is reached. Then a switch-over of predominating metabolisms to the anaerobic fermentation (Sahrawat, 2004; 2005; Kögel-Knabner et al., 2010), which is thermodynamically less efficient for microbes than the aerobic degradation (Peters et al., 2005) commences. Continuous and successive input of organic matter and the slower decomposition of these substances, due to low energy yields during anaerobic decomposition (Table 1.1), favour the accumulation of organic matter in paddy soils (Lal, 2002; Sahrawat, 2005).

Table 1.1. Selected reaction pathways of microbial metabolism, their reduction potentials and their energy yields (modified from Valiela, 1995 and Nelson and Cox, 2005; Comeau, 2008).

Metabolism	**Reaction examples**	**Ranges of redox potential (mV)**	**Energy yield (kcal)**
Aerobic respiration	$C_6H_{12}O_6 + 6\mathbf{O_2} \rightarrow 6\ CO_2 + 6H_2O$	+800 to -420	686
Nitrate reduction and denitrification	$5\ C_6H_{12}O_6 + 24\ \mathbf{NO_3^-} + 24\ H^+ \rightarrow 30\ CO_2 + 12\ N_2 + 42\ H_2O$	+750 to -420	649
Fermentation	$\mathbf{C_6H_{12}O_6} \rightarrow 2\ CH_3CHOCOOH$	420 to -185	58
	$\mathbf{C_6H_{12}O_6} \rightarrow 2\ CH_2CH_2OH + CO_2$		57
Sulfate reduction	$C_6H_{12}O_6 + 3\ \mathbf{SO_4^{2-}} \rightarrow 6\ CO_2 + 3\ S^{2-} + 6\ H_2O$		
	$CH_3CHOHCOO\text{-} + ½\ \mathbf{SO_4^{2-}} + 3/2\ H^+ \rightarrow CH_3COO^- + CO_2 + H_2O + ½\ HS^-$	-250 to -420	8.9
	$CH_3COO\text{-} + \mathbf{SO_4^{2-}} \rightarrow 2\ CO_2 + 2H_2O + HS^-$		-
Methanogenesis	$H_2 + ¼\ \mathbf{CO_2} \rightarrow ¼\ CH_4 + ½\ H_2O$	-250 to -420	8.3
	$CH_3COO^- + 4\mathbf{H_2} \rightarrow 2CH_4 + 2H_2O$		39
	$CH_3COO^- \rightarrow CH_4 + CO_2$		6.6
Methane oxidation	$CH_4 + \mathbf{SO_4^{2-}} + 2H^+ \rightarrow CO_2 + H_2O + HS^-$		3.1
	$CH_4 + 2\ \mathbf{O_2} \rightarrow CO_2 + 2\ H_2O$		193.5

Wetland rice cultivation is known for increased emission of greenhouse gases methane and nitrous oxide (Scharpenseel et al., 1996; Neue et al., 1997; Liesack et al., 2000). Especially,

microbial metabolic reactions lacking oxygen as electron acceptor produce these gases as intermediates, e.g. through the further degradation of fermentation products under strong anoxia during acetoclastic methanogenesis (Liesack et al., 2000). Nitrate (NO_3^-) is one of the major plant nutrients that is strongly depleted in arable soils, because of the intense plant uptake during growth or due to loss via relocation after dissolution in water. The extracted amounts of NO_3^- are commonly restored by adding of nitrogen fertilizers (using nitrate fertilizer or ammonium fertilizer that is aerobically nitrified to nitrate). In paddy soils, if slight anaerobic conditions dominate, denitrification may occur, which is a microbially mediated stepwise process that converts nitrate to gaseous nitrogen during the degradation of organic substances by means of nitrate as electron acceptor. The greenhouse gas nitrous oxide (N_2O) is one of the important intermediates in denitrification, which is most intensive in paddy soils during changes in management: from flooded rice to non-flooded upland crops or reverse (Xiong et al., 2007). Denitrifiers are facultative anaerobic organisms that are able to switch between aerobic and anaerobic conditions (Jones et al., 2008). In addition to the pathway under anoxic condition, N_2O is also an intermediate of the reduction of NO_2^- to N_2 after the oxidation of ammonia, the first step of the aerobic nitrification (e.g. ammonium fertilizers). Here, denitrification by ammonia oxidizing bacteria (AOB) produces minor amounts of N_2O (Kool et al., 2011). In a study of Hwang and Hanaki (2000) the highest production of N_2O was noticeable, if 5% oxygen and 50% water content prevailed.

1.3 Soil organic matter

The soil organic matter (SOM) is a dynamic fraction in soils (influenced by continuous decomposing and transformation activities), which is composed of a complex and heterogeneous mixture of organic compounds that may affect the soil properties during pedogenesis. Simplified, the term SOM denotes the sum of organic carbon-containing substances that derive from microorganisms as well as from plant and animal remains in different stages of decomposition (Huang et al., 1996, Nieder and Benbi, 2008). According to this, in soils the organic matter occurs in unaltered form and/or as one of the intermediate products from the biological and chemical degradation (Nieder and Benbi, 2008; Simpson and Simpson 2012). If the decomposition of the organic material has advanced to an extent where

the former structural organization in the source material is no longer recognizable then it is called humus (Amundson, 2001). In general, a differentiation of various types of SOM can be achieved as in Table 1.2.

Table 1.2. Differentiation of SOM (modified from Baldock and Skjemstad, 2000).

Component	**Definition**
Soil biota	Organic materials associated with the tissues and cells of living plants, microorganism (bacteria, archaea, fungi) and soil fauna (nematodes, protozoa, earthworms).
Non-living components	
Humus	A mixture of altered or transformed materials.
Non-humic compounds	Contains identifiable biomolecules (amino acids, lipids, polysaccharides, lignin, nucleic acid).
Humic compounds	Contains molecules (fulvic acid, humic acid, humin) that are not related to a distinct precursor. Different theories for possible formation pathways exist that have lignin, cellulose or sugar and amine as source as origin.
Dissolved organic matter (DOM)	Water soluble organic materials (< 0.45 µm).
Particulate organic matter (POM)	Fragments that have a recognizable cellular structure (dominated by plant materials).
Inert organic matter (IOM)	Highly carbonized materials including charcoal, charred plant residues, graphite and coal.

Various mechanisms control the fate of SOM (Fig. 1.3.) such as: (i) the selective preservation of recalcitrant SOM, (ii) biological recycling of carbon and nitrogen via the microbial metabolism, (iii) protection of SOM from mineralization in micro- and nanopores, in aggregates or through encapsulation and hydrophobic surroundings and (iiii) chemically interaction with phenols, amides, metal ions and minerals (Amelung et al. 2008 and references therein).

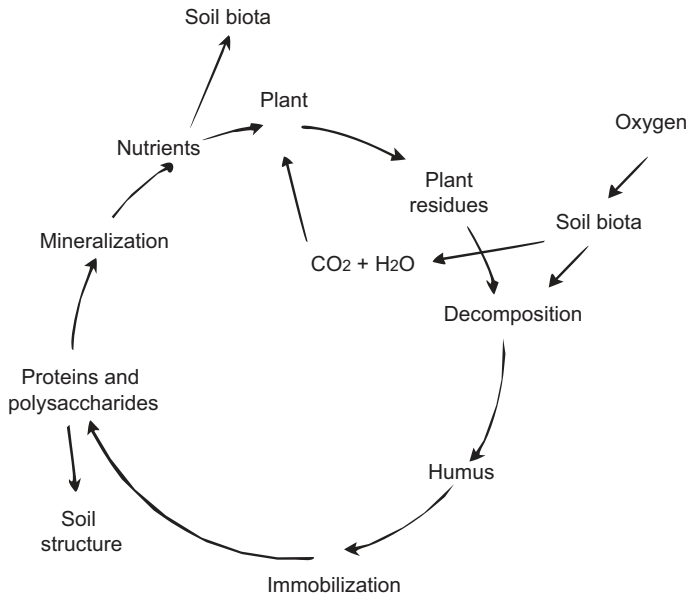


Fig. 1.3. Generalized cycle of soil organic matter including continues transformation of plants, micro- and macro-organism derived organic and inorganic carbon (modified from www.fao.org).

The vegetation cover and the agricultural management of soil are important factors for the storage of soil organic carbon (SOC). Thus, green manuring, agroforestry and usage of mulch have an enhancing effect on the amount of SOC (Lal et al., 2002). As observed for various Chinese soils, in paddy soils redox conditions prevail that may affect the enhanced preservation of SOM in contrast to permanently oxic upland soils (Lal et al., 2002 and references therein; Wu, 2011). The highest concentrations of organic matter are often detectable in topsoils (A horizon) with SOC contents ranging from 0.7% to 5% on a weight basis in forest and arable soils. Furthermore, exceptionally high SOC contents are typical in permanent grasslands (7.5% to 15%) and redoximorphic soils that comprise gleyic soils and peat bogs (15% to 50%) (Blume et al., 2010).

Decomposition, followed by mineralization of SOM (conversion into inorganic constituents) are important processes, which affect the formation of smaller organic compounds and the release of essential nutrients for plant growth. In addition to carbon (C) the soil organic matter consists of other non-metals as hydrogen (H), oxygen (O), nitrogen (N), sulphur (S) and phosphor (P) as well as a suite of metals (Blume et al., 2010). The progression of humification

(transformation of carbon from organic residues to humus) can be estimated by using the C/N ratio. The principle here is that organic matter derived from higher plants contains generally lower proportions of nitrogen than the organic matter of microorganism (Simpson and Simpson 2012). Depending on the type of plant materials the C/N can vary internally, with high ratios of 100-400 indicative for wood, lower ratios of 50-60 for fine roots and matured plants, as well as ratios of 30-50 for leaves, legumes and young plants (Yerima and van Ranst, 2005; Blume et al., 2010). The C/N ratio in plant tissue itself is variable, depending on species and age, but the end-product of plant decomposition is always humus with fairly constant C/N ratios particularly in a natural soil system (Yerima and van Ranst, 2005). Degradation of plant material is accompanied with a decline in the carbon content (linked with the release of CO_2) and an increase of nitrogen that results in lower C/N ratios than in unaltered plant materials. Expressed in numbers, soils with predominant mineralization processes exhibit low C/N ratios of < 10. These soils are commonly associated with a reduced heterotrophic activity owing to limited carbon availability and with the release of plant available nitrogen that is finally assimilated by plants or soil biota (Luce et al., 2011). In arable soil the C/N ratio commonly ranges between 8 and 15 (Yerima and van Ranst, 2005). If the C/N ratio is around 20, then mineralization and humification are in equilibrium such as noticeable in grassland and manure. By contrast, the excess of carbon after incorporation of fresh organic materials is associated with ratios > 50 that indicates slow or minor humification and mineralization in soil (Yerima and van Ranst, 2005). Microorganisms are key players in humification contributing 2-5% to the total SOM. Their growth is positively related to increased amounts of nitrogen-rich organic matter (Yerima and van Ranst, 2005).

In arable soil the intense utilization induces a decline of plant-available nutrients. Here, especially harvesting of crop plants leads to an increased loss of recyclable substances. Therefore, excessive manuring with mineral and organic fertilizer is conducted to recharge the soils with nutrients. The specific water management during the cultivation of rice affects the elevated loss of nitrogen via nitrate leaching, runoff and emission of nitrogen oxides, which exceeds the N-return via re-incorporation of rice straw after harvest (also often as combusted residues). Hence, the application of nitrogen fertilizer on paddy soils is indispensable and in

general accounts for 15.4% of all nitrogen fertilizer used in world (Heffer, 2013). Alternatively, but rarely utilized management practises of paddy soils exist that can improve the situation of plant nutrients without excessive manuring. For instance the concomitant growing of rice and Azolla, a plant that symbiotic cohabits with the nitrogen-fixing cyanobacterium *Anabaena azollae* (Watanabe, 1982; Cheng et al., 2010), or the simultaneous farming of fish in paddy fields (Tsuruta et al., 2011; Hu et al., 2013) are such methods that would lowering the application of fertilizer. But the prevailing cropping systems on paddy fields are dominated by the rice-upland crop rotation systems with growing of intercrops such as wheat, maize or vegetables, potato or legumes in the dry season. Alternatively the rice-rice rotation system is practised varying numbers of rice growing cycles (1-3 cycles are common) depending on climatic conditions and the demand on rice (Gaunt et al., 1995).

In paddy soils the primary input of unaltered organic materials are characterized by a broad diversity of different sources, such as plant litter, roots and exudates (rice plants, crop plants and weeds), epiphytic algae, management associated microbial consortia (bacteria, archaea and fungi), organisms from the micro- and macrofauna and residual organic matter from soil parent materials. Additionally, allochthonous inputs affect the SOM composition through the incorporation of substances during flooding the fields (input of e.g. riverine organic matter) or the accumulation of atmospherically transported organic matter. Paddy management in addition to fertilizers (e.g. urea, dung) may comprise the application of pesticides and fungicides to promote the growth of healthy plants. Other anthropogenic sources of SOM are e.g. lubricants and fossil fuel contaminations or combustion residues that reach the soil during e.g. ploughing and puddling with machines and the insertion of ashes into soil.

The complex organic matter is composed of a large number of compounds that are widely different in their chemical composition and their amounts, amongst others dependent of the type, nature and age of plants (Kögel-Knabner, 2002; http://agriinfo.in/). Typical substances originating from degraded residues of plants and soil biota (microbes, worms and insects) are: complex carbohydrates (cellulose, starch), simple sugars, lignin, pectin, proteins (amino acids), fats, oils, waxes, resins, alcohols, organic acids, phenols, nucleic acids etc. and others. Despite the highly diverse SOM composition and the occurrence of products in various stages

of decomposition/transformation, specific compounds can indicate their origin. These compounds are so-called biomarkers that can be traced to a former living organism (specific plant, microbial or anthropogenic source) because their carbon skeleton was preserved after abiotic and biotic degradation (Peters et al., 2005). Over time diverse methods have been developed, to detect these compounds in plants and soils. Table 1.3 gives an overview about the commonly used ones.

All of these compounds have various residence times in soils, depending on their chemical stability, the intensity of surface interactions with soil constituents and the environmental conditions prevailing. The stabilization of the SOM is also controlled by: their interaction with the chemical composition of the mineral fraction, the presence of multivalent cations, the availability of adsorbable mineral surfaces for organic materials and the construction of the soil matrix (Baldock and Skjemstad, 2000). In general, the SOC may be differentiated in three pools depending on residence time: in an active pool (~1 year) including plant litter and root exudates, a slow pool (10-100 years) including intermediate degraded compounds and a passive pool (≥100-1000 years) including physically or chemically protected organic compounds (Amundson, 2001).

The combination of biomarker analyses with stable isotope allows the reconstruction of SOM pathways and time-scales (Amelung et al., 2008). For instance, incubation studies with stable isotope labelling (commonly with ^{13}C, ^{15}N) of plant material, of soil substrate or using CO_2 from artificial sources, allows the tracing of the carbon or nitrogen pathway mainly for short-term mechanisms (Amelung et al., 2008 and references therein). Especially compound-specific stable isotope analyses may identify the participation of specific decomposer organism or the increased degradation of biomarkers via abiotic processes or *in situ* formation of compounds (e.g. *n*-alkanes, carbohydrates, amino sugars and PLFA). Fig. 1.4 shows an overview about mean residence times of several compounds such as sugars and starches that are degraded first, followed by phospholipid fatty acids (PLFA) and proteins (amino acids).

Table 1.3. Overview of different biomarker methods commonly used to elucidate the molecular composition of SOM (modified from Simpson and Simpson, 2012 and references therein). GC-MS stands for gas chromatography-mass spectrometry and LC-MS stands for liquid chromatography-mass spectrometry.

Substances	**Biomarker methods (GC-MS or LC-MS)**	**Potential Source(s)**
Carbohydrates/sugars	Free simple sugars can be extracted by organic solvents.	All organism
	Acid hydrolysis can be used to break up carbohydrates into simple sugars.	All organism
Biochar/black carbon	Levoglucosan can be extracted by organic solvents.	Cellulose (biomass burning)
Cutin and suberin	Base hydrolysis release the bonded *n*-alkanes, *n*-alkanol, *n*-alkanoic acids and hydroxyalkanoic acids.	Plants: cutin often in aboveground parts and suberin in underground parts and woody stems.
Lignin	Lignin-derived phenols (monomers and dimers) from intact lignin molecules using alkaline copper (II) oxidation.	
	Free phenols in solvent extracts can also be from suberin (i.e., ferulic acid) and are not necessarily lignin-derived.	
Microbial-derived compounds	Amino sugars such as: glucosamine, galactose-amine, muramic acid.	Bacteria and fungi
	Bacteriohopanepolyols (BHPs)	Bacteria
	Branched glycerol dialkyl glycerol tetraethers	Bacteria
	Isoprenoid glycerol dialkyl glycerol tetraethers	Archaea
	Ergosterol	Fungi
	Hopanoids	Bacteria
	Phospholipid fatty acids (PLFAs)	Bacteria, plants and fungi
Peptides	Free amino acids can be extracted in organic solvents.	
	Amino acids can also be measured after acid hydrolysis.	
Lipids/waxes	Phytosterols and free simple alkyl lipids as hydrocarbons and carboxylic acids can be extracted by organic solvents.	Plants

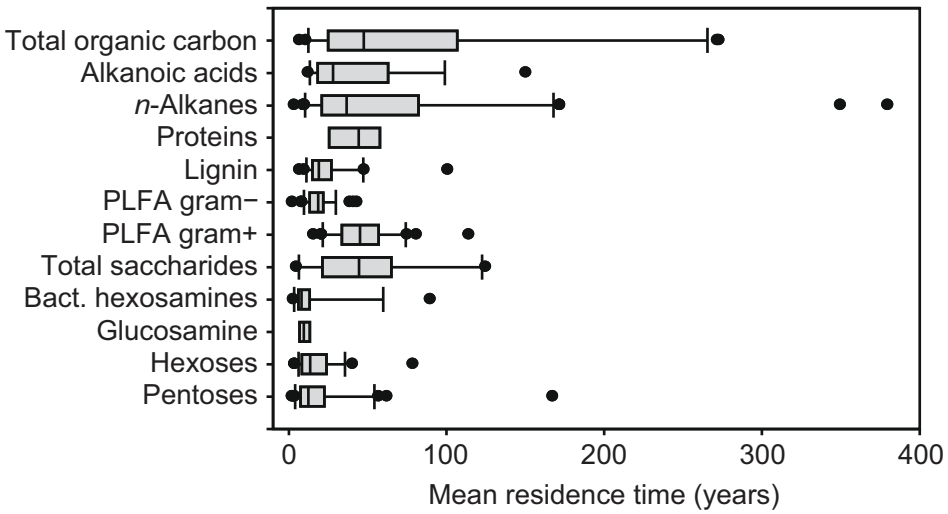


Fig. 1.4. Box plot of MRT (mean residence time) of organic compounds and biomarkers occurring in soils (Amelung et al., 2008).

Although, proteins and carbohydrates are the most abundant substances within living cells (plants, microorganism), these components have only low residence times in soils as many microorganism easily metabolize them (Kögel-Knabner, 2002). Lipophilic compounds as lignin, fats, *n*-alkanes, carboxylic acids and waxes (especially the long-chains) are potentially better preserved than others (Marschner et al., 2008; Wiesenberg et al., 2004; Amelung et al., 2008). The turnover rates of individual compounds are mainly influenced by intrinsic stability itself, however transformation reactions, as e.g. the reduction of alkenes and alcohols, the decarboxylation of fatty acids or the degradation of aliphatic biopolymers, also have an effect on the summarized dwell time of *n*-alkanes. (Lichtfouse et al., 1998; Amelung et al., 2008).

1.4 Lipids

Within soils lipids may occur in functionalized or de-functionalized form depending on their origin and the degree of decomposition of organic material. The stronger stability of some

lipids and their allocation to specific source organisms allow reconstructing accumulation and decomposition activities in soils. However, what are lipids? Lipids have broadly been defined as a heterogeneous group of organic compounds, which are soluble in organic solvents but insoluble in water because of their hydrophobic or amphipathic properties (Dinel et al., 1990; Fahy et al., 2005). This definition is unfortunately misleading, because in nature numerous exceptions exist, such as substances with lipid-like properties (e.g. hydrophobic proteins) or lipids that are more soluble in water than in organic solvents, as gangliosides a sphingolipid (http://lipidlibrary.aocs.org/). Based on distinctive moieties serving as building blocks, Fahy et al. (2005) classified the lipids into eight categories (Table 1.4 and Fig. 1.5) such as fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, saccharolipids and polyketides (derived from carbanion-based condensations of thioesters), and sterol as well as prenol lipids (derived from carbocation-based condensation of isoprene units). The generalized overview of the biosynthetic network of lipids is shown in Fig. 1.5.

Table 1.4. Lipid categories and examples that occur in eukaryotes and prokaryotes (modified from Fahy et al., 2005 and http://lipidlibrary.aocs.org/).

Category of lipid	**Examples for subclasses**
Fatty acyls	Fatty acids, fatty alcohols and fatty esters, wax monoesters
Glycerolipids	Diacyl- and triacylglycerols
Glycerophospholipids	Phosphatidylcholines and phosphatidylethanolamines
Sphingolipids	Ceramides and phosphosphingolipids
Saccharolipids	Acylaminosugars
Polyketides	Flavonoids and aromatic polyketides
Sterol lipids	Cholesterol, bile acids and steroids
Prenol lipids	Isoprenoids, Hopanoids and Quinones

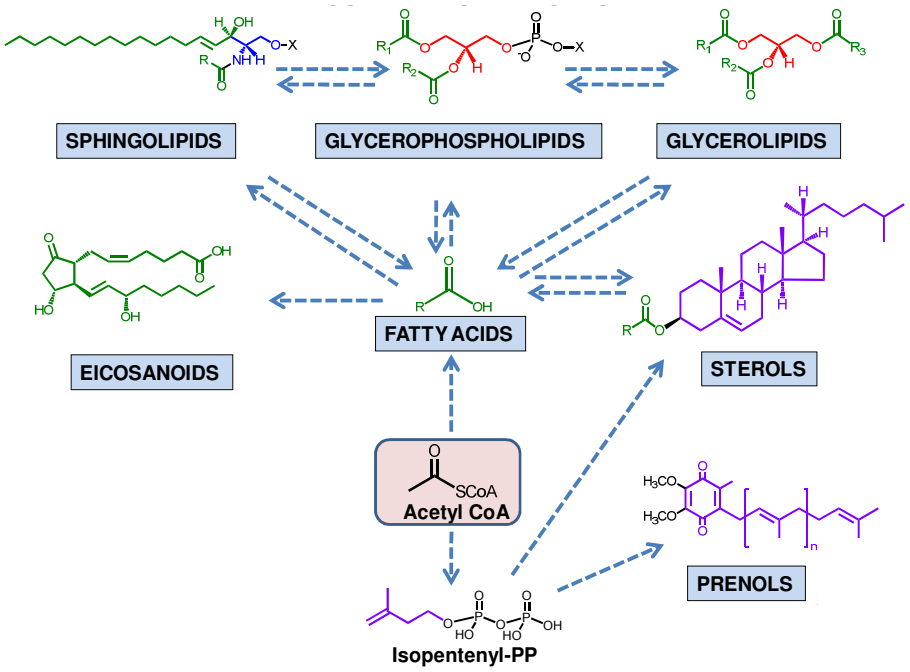


Fig. 1.5. Overview of the biosynthetic network of lipids and their structural characteristics. Various colours depict different moieties/functional groups in chemical structure: fatty acid moieties (green), glycerol group (red), amide group (blue), phosphate group (black) and isoprene moieties (purple) (http://lipidmaps.org/).

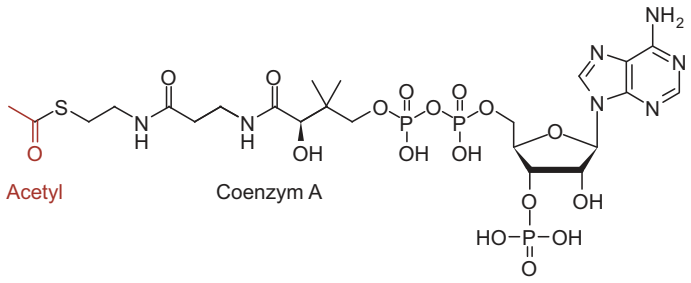


Fig. 1.6. Molecular structure of acetyl coenzyme A.

The biosynthesis of all these lipids generally starts with the formation of the molecule acetyl coenzyme A (acetyl-CoA), a coenzyme consisting of a CoA unit that is linked over a thioester bond to an acetyl moiety (Fig. 1.6). Thus, for instance, the synthesis of fatty acids, which are often construction units in composite lipids as e.g. glycerolipids and glycerophospholipid,

starts with the irreversible carboxylation of acetyl-CoA to malonyl-CoA by the catalyzing enzyme acetyl-CoA carboxylase (ACC).

The ACC have distinctive composition-forms depending on organism type, with a multiple subunit form [composed of biotin carboxylase (BC), the biotin carboxyl carrier protein (BCCP) and α- as well as β-carboxyltransferase (CT)] occurring in prokaryotes and a single multifunctional polypeptide with 4 subunit domains being present in the cytosol of eukaryotes (Sasaki and Nagano, 2004; Gerhardt et al., 2015). In plants often both forms occur, with the multiple subunit form inside plastids and the single multifunctional polypeptide form in cytosol, except for grasses as e.g. wheat and rice that have only the multifunctional polypeptide form in the cytosol as well as in the plastids (Ohlrogge and Browse, 1995; Sasaki and Nagano, 2004 and references therein). In addition, the activity of ACC is differently regulated, as for instance by the citrate cycle in mammals or by a light-dependent increase of the pH and [Mg^{2+}] in the stroma of plants (Sasaki and Nagano, 2004; Nelson and Cox, 2005). Recently, Gerhardt et al. (2015) discussed that the signal transduction protein GlnB, a nitrogen regulatory protein (e.g. for glutamine synthesis), also contributes to the regulation of the ACC in bacteria, with a reduced activity if GlnB interacts with the BCCP through the presence of a higher amount of available ammonium.

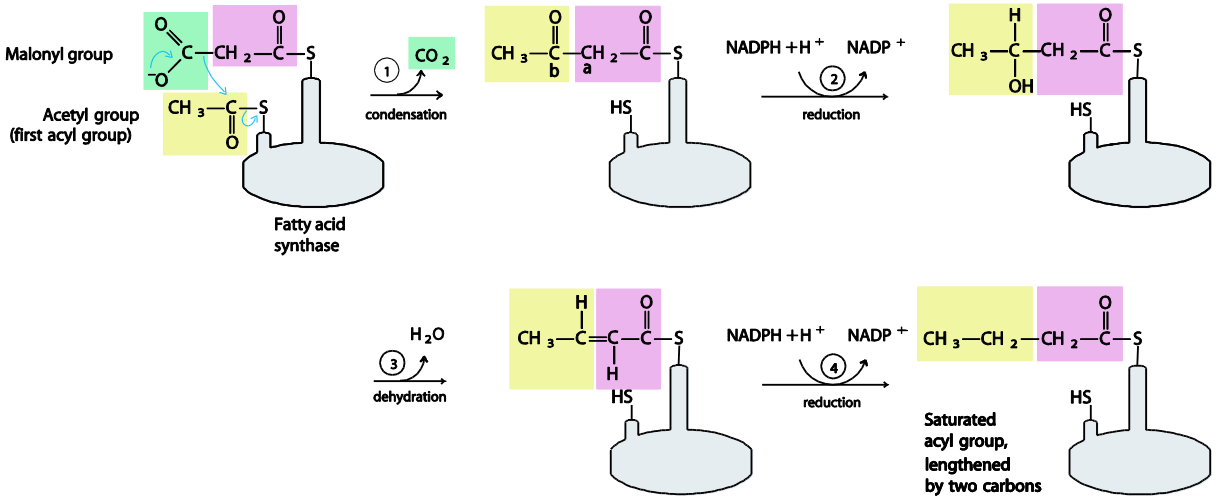


Fig. 1.7. Depiction of long chain fatty acid synthesis in a repetitive reaction sequence. Modified after Nelson and Cox (2005).

After the transfer of the malonyl group and a further acetyl group to the fatty acid synthase (FAS), a set of reactions with four main steps (condensation, reduction, dehydration and reduction) follows to extend the fatty acyl chain through each cycle by 2 carbon units (Fig. 1.7; Nelson and Cox, 2005). Thus, the most common fatty acids have an even number of carbon atoms (Peters et al., 2005). As described for the ACC before, the enzyme structures of various organisms are different, with FAS-I acting as one multiple-enzyme complex (e.g. in mammals, fungi and some bacteria) and FAS-II catalyzing the reactions with six different active enzymes and one separate acyl carrier protein (e.g. in plants, fungi and prokaryotes). Nevertheless, the four-step cycle is the same in all organisms (Nelson and Cox, 2005) and the product after seven cycles is the saturated fatty acid palmitate ($C_{16:0}$), which can undergo further modifications as desaturation and/or elongation (catalyzed by elongases and desaturases). In contrast to mammals, that in principal produce saturated fatty acids, plants and microorganisms synthesize in addition unsaturated fatty acids, waxes and hydrocarbons e.g. within the chloroplast (Nelson and Cox, 2005).

Lipids are major constituents in eukaryotes and in prokaryotes, where they exhibit a wide diversity in their chemistry and biological functions. Lipids have different properties and the individual composition depends on application conditions. According to this, they may serve as storage and insulation lipids (triglycerides, derivatives of fatty acids), as structural lipids to protect the cell from external influences (glycerophospholipids, sterols) or they facilitate the transmission of signals as so-called lipid messenger via the binding on a protein (steroid hormones, ceramides, enzyme cofactors and pigments). Some of them are specific for individual organism groups (have a chemotaxonomic potential) and others are universally findable in most organism. Thus, for instance, the cell membranes of bacteria differ strictly from those of the archaea with straight or branched alkyl chains that are either ester-linked or ether-linked to glycerol backbones in bacteria and isoprenoid alkyl chains that are connected by ether bonds to the glycerol backbones in archaea (Fig. 1.8). The latter ones as well as sterol lipids are also synthesized from acetyl-CoA, but the assembly plan is different from those of fatty acids. Here, one pathway is the formation of the six-carbon intermediate mevalonate

through the condensation of three acetate molecules, followed by further conversion to the activated isoprene isopentenyl-diphosphat (Nelson and Cox, 2005; Peters et al., 2005).

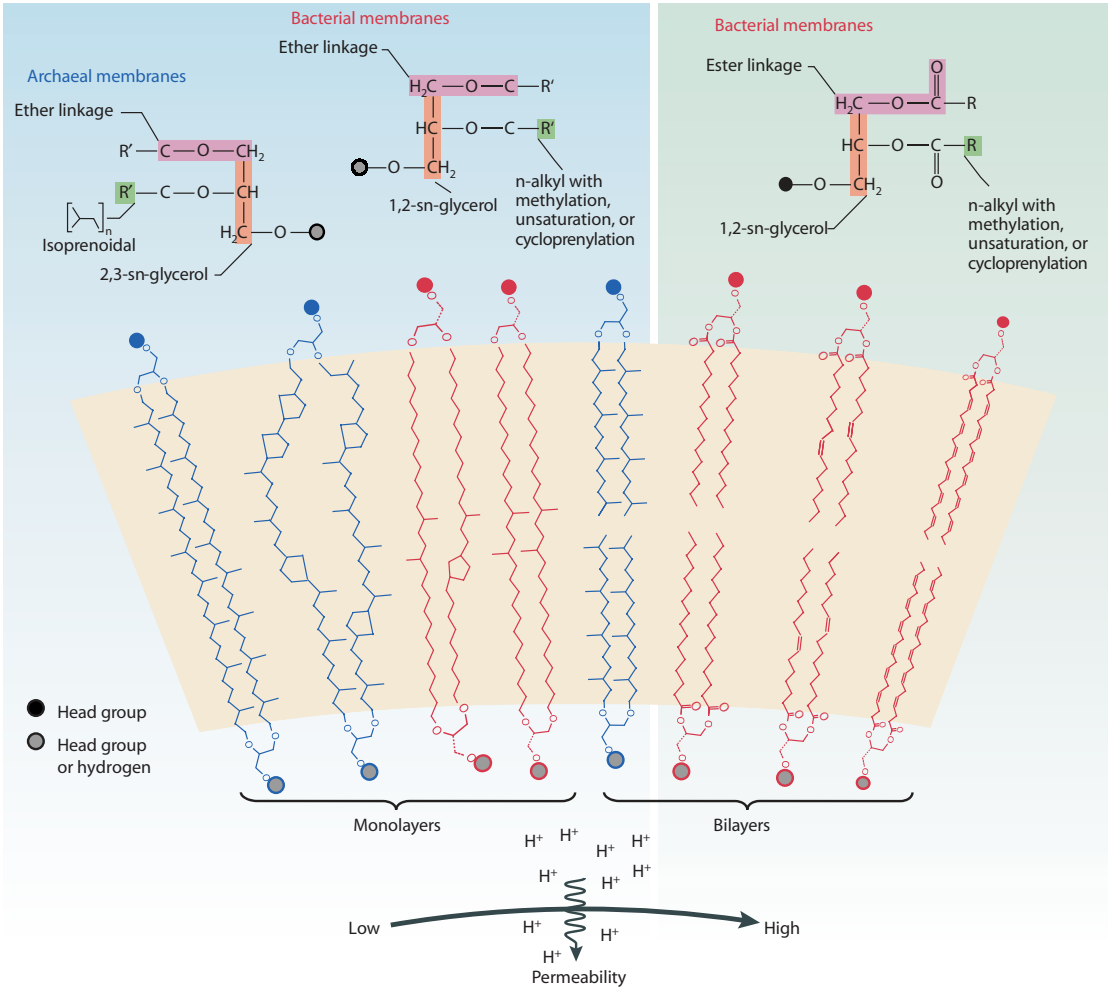


Fig. 1.8. Chemical structures of archaeal (left) and bacterial (right) membrane lipids. Archaeal lipids consist of isoprenoidal alcohols that are ether-linked to glycerol backbones and bacterial lipids dominantly consist of fatty acids that are either ester-linked or ether-linked to glycerol. The stereochemical configuration of glycerol backbone is different in archaea and bacteria, with archaea possessing a 2,3-di-*O*-alkyl-*sn*-glycerol configuration and bacteria a 1,2-di-*O*-alkyl-*sn*-glycerol configuration. The arrow indicates the membrane permeability to ions. Modified after Valentine (2007) and Weijers et al. (2006a).

In addition to structural membrane lipids wax components exist that serve as water repellents and regulators of evapotranspiration (Eglinton and Hamilton, 1967; Lockey, 1988, 1991). Here, especially ester of long-chain saturated and unsaturated fatty acids with alcohols or long-chain alkanes serve as protective coatings on leaves and minimize the water loss for plants.

Characteristic distribution patterns of *n*-alkanes in soils/sediments allow for differentiation between marine and terrestrial plant OM inputs as well as for discrimination of vascular plants from submerged plants or algae (Poynter et al., 1989; Ficken et al. 2000). In this connection, long chain homologues > C_{20} may originate from cuticle waxes as well as from cell membranes of higher land plants (Eglington and Hamilton, 1967; Kolattukudy, 1976; Bianchi, 1995) and short chains < C_{20} predominantly derive from algae or bacteria (Cranwell, 1982). Long-chain homologues (C_{25} to C_{35}) of *n*-alkanes with an odd over even predominance (Eglinton and Hamilton, 1967) and a prevalence of either C_{31} or C_{29} were commonly associated with crop plant and grass inputs into soils (Wiesenberg et al., 2004; Rommerskirchen et al., 2006).

1.5 Aims and thesis outline

This PhD thesis was part of the Research Unit FOR 995 "Biogeochemistry of paddy soil evolution" that was funded by the German Research Foundation (DFG). In order to improve the knowledge concerning the biogeochemistry of paddy soil evolution and the associated main processes, a collaborative approach of soil chemistry, mineralogy and microbiology was conducted. Rice paddy soils are characterized by preservation of organic matter predominantly derived from the rice plant itself and from associated microbial consortia under at least periodically anaerobic conditions due to submergence. Depending on human management practices, paddies may also receive exogenous organic inputs via manuring or accumulation of combustion residues. Lipids as molecular markers of rice and other plants (crop plants including wheat or plants inhabiting natural wetlands) and as markers for microbial and anthropogenic input into paddy soil were investigated in diverse sampling sites located in tropical (Indonesia, Vietnam and Philippines) as well as in subtropical (China and

Italy) climate zones. To this, soil samples including 119 paddy, 37 upland, 9 forest, 2 bushland and 3 marsh soils as well as diverse reference substrates from tidal wetland to freshwater sediments and plant samples were collected and analysed for their lipid composition. In the following chapters, the main findings of this thesis are presented.

Chapter 2 describes the variability of numerous biogeochemical soil parameters detected in paddy soils from the 2000 yr old chronosequence in the Zhejiang province, China. Results of descriptive, explorative and non-parametric statistical data analyses were used to confirm the low intrinsic heterogeneity for the more conservative and stable parameters (e.g. SOC and lipid) within a paddy rice field. In addition, long-term rice paddy management indicates a decline of the variability according to progression of soil evolution. Individual parameters exhibit first an insufficient biogeochemical signature in the younger paddy soils ($\leq$ 300 yr), which develops with the duration of management to characteristic and stable biogeochemical conditions. In addition, the management of paddy soils over centuries is characterized by numerous flooding and drying cycles of soils, which influences the presence of redox conditions and therefore resulted in slower decomposition of organic matter.

One part of **Chapter 3** deals with the characterization of plant wax alkane patterns of rice plants and several intercrop plants that were all subdivided for plant tissues and for growth stages. In particular, the predomination of the *n*-C_{33} alkane was a distinctive feature in the *n*-alkanes composition in all rice plant tissues compared to crop plants. Furthermore, an adaption of the *n*-alkane distributions to the water level above paddy soils was detected. Submerged plant tissues as roots from adult rice as well as the entire young rice plants contained higher amounts of mid-chain alkanes. The *n*-alkane distributions of the agricultural topsoils investigated reflected the pattern of rice or crop plant residues that remained on the field after harvesting. Soils here all evolved on former tidal wetland sediments. Compositional differences between subsoils and topsoils indicated different botanical origins. Independent of the agricultural usage, the substrates alkyl lipid signature was preserved effectively in the deeper soil horizons and allowed the reconstruction of past land use change.

Chapter 4 describes the composition and site-dependent variability of microbially derived glycerol dialkyl glycerol tetraether (GDGTs) lipids, which occurred in different soils.

Both, branched (bacterial) as well as isoprenoid (archaeal) GDGTs indicated that their compositions are preferentially controlled by the management type followed by climatic exposition. Paddy soils exhibited primarily a signature of anaerobic bacterial GDGT lipids as indicated by the dominance of branched GDGTs. Exceptionally high proportions of isoprenoid GDGT-0, furthermore confirmed the increased presence of methanogens appearing commonly in flooded rice paddy soils. Differences in agricultural management amongst others comprise differences in soil water content, redox conditions, soil pH, nutrient content and cultivated plants, which separately and/or in combination affect the associated microbial community composition. The comparison of adjacently located paddy and upland soils, excluding climate as an influencing factor, revealed a management depend behaviour of microbial cell membrane lipids. Noticeably different results for indices that were calculated based on isoprenoidal and/or branched GDGTs (TEX_{86}, MBT', CBT, BIT) were obtained. Especially, varying results in periodically flooded paddy soils compared to non-flooded upland soils suggested soil moisture to affect the soil temperature, which additionally to the soil pH as a further factor controlled the bacterial tetraether lipid composition.

Chapter 5 provides an insight into the composition of ammonia oxidizing archaea in flooded paddy soils using a combined approach with lipidomics and genomics. The abundance of the functional gene *amoA*, which encodes the key enzyme ammonium monooxygenase in ammonia-oxidizing archaea (AOA), correlates with the sum of isoprenoid GDGTs. In contrast to *amoA* AOA gene copies, reflecting only a snapshot of the archaea present in soil during sampling time, the long-time paddy management lead to the successive increase of ammonia-oxidizing *Thaumarchaeota* as deduced by high relative abundances of crenarchaeol. The inconsistent relationship between gene abundances and crenarchaeol in some paddy soils either suggests a higher content of fossil thaumarchaeal lipids or indicates the existence of *Thaumarchaeota* (e.g. Group 1.1c) that do not contain the functional gene *amoA*.

In **Chapter 6** it is shown that combustion residues, detected on molecular level as polycyclic aromatic hydrocarbons (PAHs) and black carbon (BC) are usable to trace the input of plant combustion residues into soils. In-field combustion of crop residues is a commonly

used post-harvest practice to remove not required crop residues and to release nutrients into soils. PAHs and BC were detected in paddy and uplands topsoils, with the highest abundances found in the former. In addition, the increased enrichments in soils under long-term paddy management suggest a higher accumulation and preservation compared to upland soils. Despite of that, the distribution patterns of PAHs showed a higher compositional variability than those from BC derived benzenepolycarboxylic acids (BPCAs). In paddy soils the PAH patterns indicate the predominance of rice straw ash. The PAH patterns in deeper soil horizons reflect the signature of the tidal wetland substrate that was mainly derived from Yangtze River sediments and East China Sea sediments. Here, especially the high proportion of perylene, a non-combustion derived PAH, confirmed this assumption. Buried layers in the oldest soils seem to have been former topsoils, as suggested by PAH patterns comparable to those of rice ash.

Chapter 7 gives a summary of the major results of this thesis.

2. Intra-versus inter-site macroscale variation in biogeochemical properties along a paddy soil chronosequence

C. Mueller-Niggemann[1], A. Bannert[2], M. Schloter[3], E. Lehndorff[4], L. Schwark[1]

[1]Institute of Geosciences, Christian-Albrechts-University of Kiel, Kiel, Germany

[2]Chair of Soil Ecology, Technische Universität München, Neuherberg, Germany

[3]Research Unit for Environmental Genomics, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany

[4]Institute of Crop Science and Resource Conservation, Soil Sciences, Bonn University, Bonn, Germany

Published in *Biogeosciences* **9** (2012) 1237-1251. doi:10.5194/bg-9-1237-2012

Abstract. In order to assess the intrinsic heterogeneity of paddy soils, a set of biogeochemical soil parameters was investigated in five field replicates of seven paddy fields (50, 100, 300, 500, 700, 1000, and 2000 yr of wetland rice cultivation), one flooded paddy nursery, one tidal wetland (TW), and one freshwater site (FW) from a coastal area at Hangzhou Bay, Zhejiang Province, China. All soils evolved from a marine tidal flat substrate due to land reclamation. The biogeochemical parameters based on their properties were differentiated into (i) a group behaving conservatively (TC, TOC, TN, TS, magnetic susceptibility, soil lightness and colour parameters, $\delta^{13}C$, $\delta^{15}N$, lipids and *n*-alkanes) and (ii) one encompassing more labile properties or fast cycling components (N_{mic}, C_{mic}, nitrate, ammonium, DON and DOC). The macroscale heterogeneity in paddy soils was assessed by evaluating intra- versus inter-site spatial variability of biogeochemical properties using statistical data analysis (descriptive, explorative and non-parametric). Results show that the intrinsic heterogeneity of paddy soil organic and minerogenic components per field is smaller than between study sites. The coefficient of variation (CV) values of conservative parameters varied in a low range (10% to 20%), decreasing from younger towards older paddy soils. This indicates a declining

variability of soil biogeochemical properties in longer used cropping sites according to progress in soil evolution. A generally higher variation of CV values (> 20–40%) observed for labile parameters implies a need for substantially higher sampling frequency when investigating these as compared to more conservative parameters. Since the representativeness of the sampling strategy could be sufficiently demonstrated, an investigation of long-term carbon accumulation/sequestration trends in topsoils of the 2000 yr paddy chronosequence under wetland rice cultivation restricted was conducted. Observations cannot be extrapolated to global scale but with coastal paddy fields developed on marine tidal flat substrates after land reclamation in the Zhejiang Province represent a small fraction (< 1%) of the total rice cropping area. The evolutionary trend showed that the biogeochemical signatures characteristic for paddy soils were fully developed in less than 300 yr since onset of wetland rice cultivation. A six-fold increase of topsoil TOC suggests a substantial gain in CO_2 sequestration potential when marine tidal wetland substrate developed to 2000 yr old paddy soil.

2.1 Introduction

On global scale rice (*Oryza sativa*) is the most important staple crop feeding more than 50% of the World's population. Cultivation of rice thus affords large proportion of arable land, amounting to approx. 156×10^6 hectare, of which > 90% is used for wetland or paddy rice cultivation in 2008 (IRRI, 2010). A critical factor in paddy rice cropping is the periodic flooding of soils and the associated fluctuations in soil redox conditions, biogeochemical cycling of essential and trace elements, and microbial community structure. Rice paddy fields are assumed to contribute significantly to the emission of potent greenhouse gases CH_4 and N_2O (e.g. IPCC, 2007; Conrad, 2009) and to the loss of nitrate into the environment (Shrestha and Ladha, 1998; Zhu et al., 2000; Kögel-Knabner et al., 2010). Consequently, the investigation of biogeochemical processes in paddy soils is of critical importance in order to assess environmental impact and initiate reduction strategies.

A major problem in the design of biogeochemical studies of paddy fields is the intrinsic heterogeneity of paddy soils in the spatial as well as the temporal realm. Spatial variability may occur on the micro (nm–mm), meso (cm) and macro (m–tens of m) scale level. In paddy

soils microscale variability has been described for soil aggregates (Li et al., 2007) and within the rhizosphere whereas mesoscale variations occur within paddy soil profiles on cm or decimetre scale and can be related preferentially to changes in redox conditions (Kögel-Knabner et al., 2010). Macroscale heterogeneity in paddy soils occurs over distances of meters or tens of meters and is less well studied than micro or mesoscale variability. The focus on such heterogeneity investigations has been placed on soil fertility, crop yields and nutrient levels in paddy fields (Tatsuya et al., 2004; Wang et al., 2009; Wei et al., 2009; Yanai et al., 2001; Zhao et al., 2009). However the heterogeneity of bulk organic, molecular and isotopic biogeochemical parameters used to interpret paddy soil processes has not yet been investigated on the macroscale. The objectives of this study thus were, first to evaluate intra and inter-site spatial variability of geochemical properties indicative for soil organic matter (SOM), minerogenic substrate and nutrients in paddy fields. Hereby, a differentiation of parameters assumed to behave conservatively by reflecting time-integrated properties (averaging over years or decades) versus fast reacting or labile parameters (reflecting daily, weekly or seasonal changes) was performed. Biogeochemical properties assumed to behave conservatively comprised soil TOC, TN, $\delta^{13}C$, $\delta^{15}N$ and extractable lipid content and composition (Zhou et al., 2009; Wiesenberg et al., 2004, 2008), reflecting time-averaged influx and composition of crop biomass and microbial re-mineralization. Further parameters reflecting time-averaged soil mineralogy and redox conditions were soil magnetic susceptibility and spectral soil colour (Mullins, 1977; Viscarra Rossel et al., 2006). As labile parameters microbial biomass carbon and nitrogen (Chantigny et al., 1996; Bai et al., 2000), nitrate, ammonium (Myrold et al., 1986; Davidson et al., 1992; Stark and Hart, 1997), dissolved organic N, and dissolved organic C (Kalbitz et al., 2003; Zhang et al., 2011) were considered.

Factors influencing macroscale paddy soil heterogeneity can be either linked to natural variability of the substrate on which paddy soils evolved and/or management practices (Rüth and Lennartz, 2008) that locally affect influx and efflux of various components into soils, which in turn regulates the soil microbial community (Dupuis and Whalen, 2007). Management practices can cause very localized and arbitrary enrichment (spots of 1–3m

diameter) of fertilizers, pesticide application, vegetation waste, or biomass combustion residues (heaps of burning rice straw) on paddy fields. Additionally, more systematic in-field variations in soil properties may result from flow pathways of irrigation water and its suspended materials load (Schmitter et al., 2010). Puddling of rice fields (ploughing under flooded conditions) is considered a key factor in the homogenization of locally constrained inputs and when repeated oftentimes may finally lead to the establishment of homogeneous distribution of conservative soil parameters, whereas the labile components may still exhibit severe spatial variability on the field scale.

Depending on the methodological approach applied, challenges to obtain representative paddy soil samples may vary considerably. This may lead to incompatible results, if e.g. microbial ecology conducted by genomics or proteomics targeting labile components is compared to lipidomics (analysis of phospholipid fatty acids or other microbial membrane lipids) employing conservative components.

Secondly, if it could be proven that inter-site variations exceed intra-site variability for specific parameters, the biogeochemical trends over up to 2000 yr of rice cultivation could be evaluated for a chronosequence (Walker et al., 2010) from the Zhejiang Province, China. Here rice cultivation started in coastal regions following land reclamation after dyke building at well dated times (Cheng et al., 2009; Chen et al., 2011), which allows for investigation of long-term evolutionary trends in rice paddy biogeochemistry. It is postulated that ongoing paddy soil evolution will continuously diminish the intrinsic heterogeneity of young paddy soils and ultimately establish homogeneous soil biogeochemical conditions as a consequence of destruction of aggregates and macropores (Ringrose-Voase et al., 2000) and by formation of an efficient plough pan (Lennartz et al., 2009) via puddling. Verification of paddy soil homogeneity in this investigation will contribute to validating pedogenic and biogeochemical studies of the same chronosequence conducted previously (Cheng et al., 2009; Bannert et al., 2011a, b; Chen et al., 2011; Roth et al., 2011; Wissing et al., 2011). All biogeochemical investigations were carried out using 5 field replicates that were treated statistically and allow assessing whether a composition in one field or a trend over several fields is robust and representative.

2.2 Material and methods

Study sites

The sites are located in the coastal Cixi area (Hangzhou Bay) in the north-east of the Zhejiang Province, China, as shown in Fig. 2.1. The Bay is affected by river runoff and tide from the East China Sea. The Yangtze (Changjiang) River with an average water runoff of 925 x 10^9 m^3 yr^{-1} and sediment load of 480 x 10^9 kg yr^{-1} supplied the dominant amount of sediment to the Hangzhou Bay (Li et al., 2009; Wang et al., 2008a), where it is re-deposited by southward coastal currents and tides (Jilan and Kangshan, 1989; Xi et al., 2009). The climate is subtropical monsoonal with annual average temperature and rainfall of 16.3°C and 1418 mm, respectively. The coastal plain of Cixi is densely covered by rivers, lakes, as well as urban and agriculture areas with main crops being wetland rice, rape, barley, and cotton (Hua and Zhu, 2000).

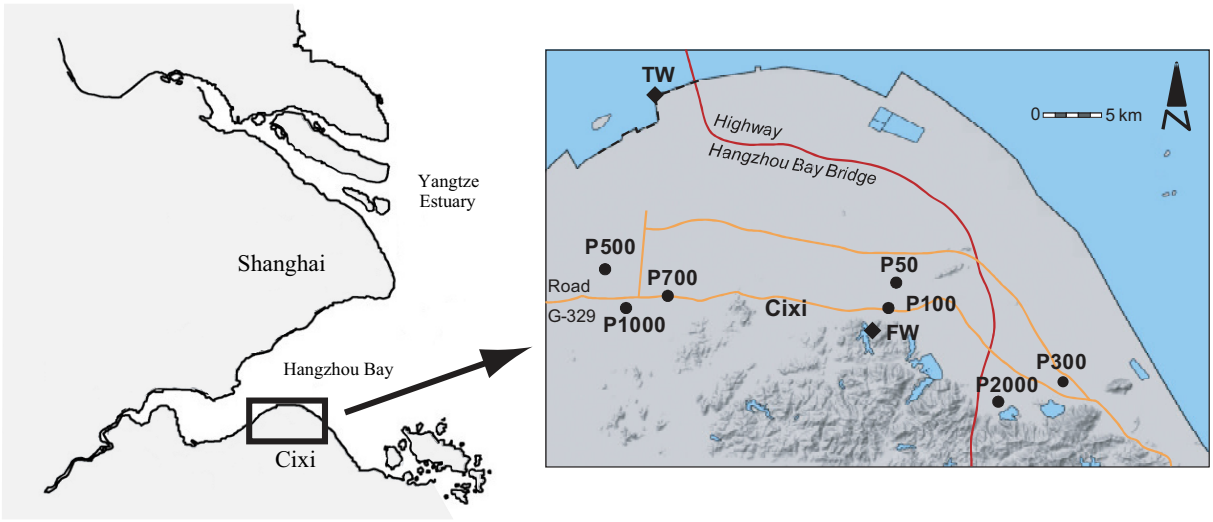


Fig. 2.1. Map of study area, depicting generations of dykes constructed for land reclamation purposes and sampling locations (courtesy of R. Jahn, University Halle).

Step by step land reclamation on marine tidal mudflat sediments (continuous alluvial plain of Andong Beach) through the building of protective dykes over the past 2000 yr, allows the investigation of a soil chronosequence with different stages of development and well known

starting dates of cultivation (Edit Committee of Chorography of Cixi County, 1992; Yu et al., 2003; Cheng et al., 2009; Chen et al., 2011). Wetland rice cultivation generally started when salt concentration decreases to tolerance levels, commonly after < 5 yr. Based on the time of dyke construction and information of the Edit Committee of Chorography of Cixi County (1992) sites with ongoing wetland rice cultivation for 50, 100, 300, 500, 700, 1000 and 2000 yr were identified.

In this region the generally cropping system constitutes one wetland rice season and one dry inter-crop (vegetables, cotton or cereals) season per year, called paddy-upland rotation. Soils with wetland rice cultivation represent Anthraquic Anthrosols, or briefly paddy soils. These are exposed to longer phases of irrigation influenced by oxygen deprivation and establishment of reducing conditions. The sampled paddy soils can be differentiated into a Stagnic Endogleyic Cambisol (50 yr old paddy), an Endogleyic Cambisol (100 to 500 yr old paddy soils) and an Endogleyic Stagnosol (700 to 2000 yr old paddy soils).

Sampling

Sampling was conducted in June 2008 after the harvest of the upland crop from seven paddy sites (P50, P100, P300, P500, P700, P1000, P2000) before flooding. In addition, sediment from a flooded paddy nursery site (P50N), a marine site (TW for tidal wetland), and a lacustrine site (FW for freshwater sediment) were analysed. From each site the top soil/ sediment down to the plough pan (roughly 0–20 cm) was sampled by a soil auger. The sample representativeness was ensured by collecting five field replicates at each site. The sampling strategy in Fig. 2.2 shows that each field replicate constitutes a composite sample of seven subsamples. After storage of the soil and sediment samples in a freezer, soil microbial biomass analyses followed within the next 2 weeks. The remaining sample aliquots were freeze dried and homogenized by grinding to fine powder prior further analyses.

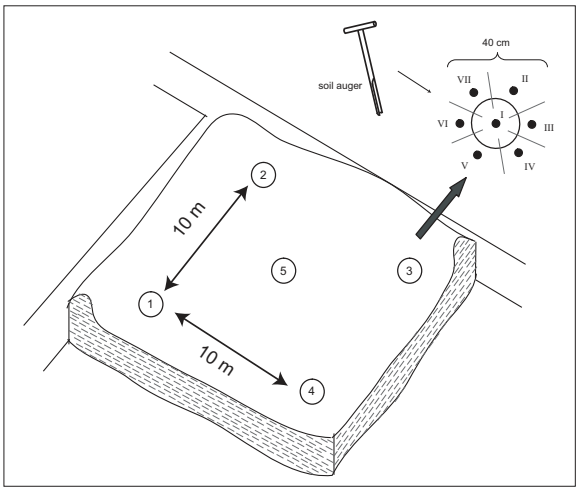


Fig. 2.2. Design for recovery of field replicates at 10m regular spacing, each of which is a composite of 7 subsamples taken at 40 cm regular spacing.

Laboratory analysis

Total organic carbon (TOC) concentrations of the soils and sediments were determined with a LECO CS-225 analyser after decarbonatisation with HCl (10%) and neutralization with distilled water. The total carbon (TC), total nitrogen (TN) and total sulphur (TS) were measured directly with a CNS analyser (Elementar Vario EL-III). Bulk magnetic susceptibility was analysed at room temperature with a Kappabridge (KLY-2, noise level 4×10^{-8} SI) to characterize the magnetizability of ferromagnetic particles in the sample. Soil colour was quantified using a Minolta (CM-700d/600d) spectrophotometer by measuring the colour parameters on the surface of air-dried samples as described in Wiesenberg et al. (2006). Determinations of water extractable organic carbon (DOC) and nitrogen (DON) were conducted after extraction of the samples with 0.01 M $CaCl_2$ (solid to liquid ratio of 1:3) with a total organic carbon analyser DIMATOC 100 (Jörgensen and Brookes, 1991). For the detection of microbial biomass carbon (C_{mic}) and nitrogen (N_{mic}) content, aliquots of soils/sediments were fumigated with chloroform (24 h) prior to $CaCl_2$ extraction. The nitrate and ammonium concentrations were measured in $CaCl_2$ extracts by a photometric autoanalyser (CFA-SAN Plus/Skalar Analytic) using the commercial kits NANOCOLOR Nitrat-50 and Ammonium-3.

Bulk elemental analysis-isotope ratio mass spectrometry (EA-IRMS) was conducted with an elementary analyser (FlashEA1110, ThermoFisherScientific) coupled to a mass spectrometer

(DeltaV Advantage, ThermoFisher Scientific). The isotopic compositions were expressed as $\delta^{13}C$ and $\delta^{15}N$ permil units in relation to internal standards V-PDB and air nitrogen.

The extractable lipids of soils and sediments were obtained by pressurized solvent extraction (Dionex ASE 200). Briefly, lipids from ca. 8 g dry soil were extracted with a dichloromethane/methanol (3/1; v:v) solvent mixture for 20 min at 100°C and 7 x 10^6 kg m^{-1} s^{-2}. Elemental sulphur was removed from the total lipid extracts by addition of activated copper. For quantification known amounts of perdeuterated *n*-tetracosane were added as internal standard. Total extracts dissolved in *n*-hexane were separated into apolar and polar compounds using small scale chromatography (Bastow et al., 2007). The aliphatic hydrocarbons were eluted with n-hexane after passing the glass column (4 mm×8 cm) filled with activated silica gel (2 h at 130°C). The sample volume was reduced via evaporation prior to transfer to GC-MS vials. Aliphatic hydrocarbon fractions were analysed on a 30 m, ZB-1ms fused silica capillary column (0.25 mm internal diameter, film thickness 0.25 µm; Phenomenex) in a HP 5890 Series II gas chromatograph (oven programme: 70°C, 3 min isothermal, 10°C/min to 140°C, 3 °C/min to 340°C, 3.5 min isothermal) equipped with a split/splitless injector coupled to a HP 5971A mass spectrometer operated in EI-mode at 70 EV.

Statistical analysis

All individual data sets were subjected to a statistical evaluation including calculation of various descriptive statistics such as the average (AV), the standard deviation (SD) and the coefficient of variation (CV), whereby the latter describe the spread and relative proportion of variation in the data set. Non-parametric statistical analyses were applied to compare all measured soil parameters among the different sampling sites. For identification of significant variations between the sites a Kruskal-Wallis Test, suitable for non-Gaussian distributed populations, was operated (null hypothesis was all medians are equal), where the asymptotic significances (*p*-values) < 0.05 indicate one or more medians are different. Additionally, multivariate techniques such as a principal component analysis and a cluster analysis were conducted to identify inter-site variability. The analyses were performed using MS Office Excel and PASW Statistics 18.

2.3 Results and discussion

Soil parameters

For the control of sample representativeness five field replicates (each a composite sample of seven individual subsamples covering an area of 2 square meters) were taken from every site and investigated for 23 different biogeochemical parameters. The summary of the descriptive statistics with average, standard deviation and the coefficients of variations of all parameters determined for the individual sites are listed in Table 2.1.

Macroscale intra-site variability

The investigation of intra-site variability on macroscale level provides information about heterogeneity and distribution of soil parameters within a single field. According to their properties the parameters chosen could be pre-differentiated in two groups. The first group termed as "conservative" parameters contained the soil properties that were assumed to represent time-averaged and well homogenized (puddled) soil properties including TC, TOC, TN, TS, lipids, *n*-alkanes, stable carbon and nitrogen isotopes ($\delta^{13}C$, $\delta^{15}N$), magnetic susceptibility (χ), soil lightness (L^*) and colour parameters (a^*, b^*). The second group termed "labile" parameters encompassed all fast reacting properties (reflecting daily, weekly or at maximum seasonal changes) included microbial biomass nitrogen and carbon (N_{mic}, C_{mic}), nitrate, ammonium, DON and DOC (Table 2.1).

Bulk organic and minerogenic parameters

The results show much lower variation of individual parameters in the conservative group with non-uniform distribution pattern over different paddy soil cultivation times (Table 2.1, Fig. 2.3a). A coefficient of variation (CV) for soil parameters lower than 20% generally indicates insignificant variability of these soil properties. The lowest spreading of CV values (< 1.4%) within all soils and sediments was observed for the bulk soil organic matter (SOM) $\delta^{13}C$ isotope values and the lightness parameter L^* (CIE axis ranging from black to white). A slightly higher but generally low variation of CV values (between 1 to 13%) was detected for the $\delta^{15}N$ values of bulk SOM, magnetic susceptibility (χ), and soil colour parameters a^* and b^* (Fig. 2.3a). These parameters and their marginal variation within a sampling site,

comparable in soils and sediments provide information about the homogeneity of the parent material on which the respective paddy soils developed. Only the 1000 yr old paddy soil presents an exception (Fig. 2.3a) with a higher variation in magnetic susceptibility (18.4%), which can be explained by a mixture of the paddy soil with adjacent upland soils of different mineralogy caused by topsoil removal and mixing in the course of dyke maintenance work (Roth et al., 2011; Wissing et al., 2011).

Table 2.1. Descriptive statistics of all biogeochemical parameters determined for the 10 study sites. P50N designates the seedling nursery paddy, TW designates the marine tidal flat substrate and FW designates the freshwater limnic substrate. AV=average value, SD=standard deviation, CV=coefficient of variation. Conservative parameters were grouped TC to *b** (D65), labile parameters were grouped N_{mic} to DOC.

	P50			P50N			P100			P300			P500		
	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV
TC [%]	1.49	0.17	11.3	1.33	0.11	8.1	1.40	0.27	19.2	2.41	0.10	4.1	1.39	0.07	4.8
TOC [%]	0.99	0.14	14.2	1.04	0.07	6.3	1.39	0.28	20.0	2.25	0.12	5.3	1.33	0.11	8.4
TIC [%]	0.50	0.06	12.8	0.29	0.08	26.6	0.02	0.02	98.2	0.17	0.10	63.0	0.06	0.08	140.1
TN [%]	0.118	0.017	14.5	0.122	0.014	11.1	0.157	0.026	16.7	0.261	0.010	3.9	0.158	0.012	7.5
TS [%]	0.032	0.006	19.6	0.031	0.001	4.3	0.030	0.005	16.9	0.056	0.010	17.7	0.030	0.004	14.1
C/N	8.37	0.19	2.3	8.54	0.78	9.1	8.85	0.59	6.6	8.63	0.40	4.6	8.44	0.37	4.4
Lipids [mg kg^{-1} dw]	553	115	20.8	544	42	7.7	735	154	21.0	1210	51	4.3	651	50	7.6
Lipids [g kg^{-1} TOC]	56	5	9.0	52	3	6.1	53	4	8.4	54	4	6.7	49	3	5.5
n-Alkanes [µg kg^{-1} dw]	1531	310	20.3	1271	96	7.6	2141	446	20.8	3440	221	6.4	2346	186	7.9
n-Alkanes [mg kg^{-1} TOC]	155	30	19.0	123	15	12.0	154	16	10.5	153	13	8.6	176	10	5.7
CPI_{short} [b]	1.8	0.2	11.9	1.6	0.2	10.1	1.4	0.1	3.6	1.3	0.1	9.2	1.4	0.2	12.9
CPI_{long} [c]	5.7	0.8	14.6	4.2	0.9	21.4	5.8	1.0	18.1	6.1	0.5	8.4	10.4	1.8	17.0
P_{aq} [d]	0.24	0.05	19.2	0.22	0.05	22.2	0.19	0.03	16.1	0.19	0.01	3.4	0.07	0.01	17.6
$\delta^{13}C$ [‰]	-28.0	0.4	-1.4	-27.1	0.3	-1.2	-28.2	0.2	-0.6	-28.5	0.1	-0.4	-28.0	0.1	-0.4
$\delta^{15}N$ [‰]	3.1	0.1	2.9	5.7	0.3	4.4	4.6	0.1	2.9	2.7	0.3	11.4	5.6	0.1	2.6
χ [10^{-8} m^3 kg^{-1} dw]	26.1	1.1	4.1	24.7	0.3	1.4	14.8	1.1	7.2	11.8	0.9	7.7	21.3	0.8	3.9
*L** (D65)	55.6	0.4	0.7	56.3	0.6	1.0	55.3	0.4	0.7	54.1	0.6	1.1	55.0	0.6	1.0
*a** (D65)	3.7	0.1	2.8	3.6	0.1	4.1	3.7	0.1	3.5	3.3	0.1	2.0	3.2	0.0	1.3
*b** (D65)	15.3	0.2	1.3	15.1	0.5	3.0	14.8	0.2	1.1	14.0	0.3	2.1	13.6	0.4	2.6
N_{mic} [µg g^{-1} dw]	33.4	8.5	25.5	45.1	5.8	12.9	37.6	3.3	8.8	4.1	1.1	27.9	15.6	5.3	33.7
C_{mic} [µg g^{-1} dw]	293	103	35.0	558	61	11.0	522	108	20.6	167	33	19.5	490	93	18.9
Nitrate [µg N g^{-1} dw]	2.7	1.8	67.1	3.6	1.1	30.8	7.5	3.9	52.1	27.5	12.0	43.5	12.0	3.2	27.0
Ammonium [µg N g^{-1} dw]	0.4	0.1	33.3	0.5	0.4	83.8	0.2	0.1	30.8	0.6	0.3	59.0	0.1	0.0	13.1
DON [µg g^{-1} dw]	2.2	0.9	42.1	3.0	0.6	19.9	5.2	1.3	24.3	5.8	1.4	23.3	12.0	2.6	21.8
DOC [µg g^{-1} dw]	21.5	7.7	36.0	9.3	2.8	29.9	14.7	4.6	31.2	16.4	5.2	31.5	23.6	4.2	17.9

	P700			P1000			P2000			TW			FW			soils		
	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV	AV	SD	CV
TC [%]	2.10	0.19	8.9	1.16	0.11	9.4	3.05	0.13	4.3	1.28	0.02	1.6	1.62	0.17	10.3	1.80	0.65	36.2
TOC [%]	2.00	0.09	4.7	1.10	0.10	8.9	2.85	0.10	3.5	0.49	0.06	12.4	1.51	0.21	14.0	1.71	0.70	41.0
TOC [%]	0.99	0.14	14.2	1.04	0.07	6.3	1.39	0.28	20.0	2.25	0.12	5.3	1.33	0.11	8.4			
TIC [%]	0.12	0.12	98.8	0.06	0.03	60.7	0.20	0.13	67.7	0.79	0.05	6.3	0.10	0.11	101.2	0.09	0.12	126.2
TN [%]	0.208	0.022	10.5	0.128	0.017	13.2	0.361	0.018	4.9	0.052	0.011	21.4	0.148	0.028	19.1	0.191	0.082	42.9
TS [%]	0.050	0.006	12.2	0.026	0.004	17.0	0.057	0.005	9.6	0.047	0.004	8.0	0.035	0.007	20.1	0.039	0.013	34.1
C/N	9.68	0.77	8.0	8.67	0.57	6.5	7.93	0.61	7.7	9.63	1.99	20.6	10.38	1.34	12.9	8.99	0.74	8.2
Lipids [mg kg-1 dw]	1358	143	10.5	639	54	8.4	1764	73	4.2	862	248	28.7	1438	177	12.3	941	440	46.8
Lipids [g kg-1 TOC]	68	8	11.7	59	10	16.7	62	2	3.4	176	49	28.1	95	6	5.9	54	8	13.9
n-Alkanes [µg kg-1 dw]	6557	1456	22.2	1690	146	8.7	5806	961	16.6	975	102	10.5	4551	1341	29.5	3.138	2.017	64.3
n-Alkanes [mg kg^{-1} TOC]	330	80	24.4	153	8	5.1	204	34	16.9	200	11	5.6	299	74	24.8	175	65	37.1
CPIshort b	1.6	0.1	5.7	1.5	0.2	16.3	1.7	0.2	10.4	1.3	0.1	6.6	2.1	0.2	10.5	1.5	0.2	13.5
CPI long c	2.5	0.1	5.0	4.2	0.9	20.3	4.1	0.2	4.1	3.4	0.3	9.5	4.0	0.4	9.6	5.4	2.4	44.9
Paq d	0.24	0.01	5.8	0.24	0.04	17.5	0.25	0.01	5.2	0.26	0.02	6.0	0.34	0.03	7.6	0.21	0.06	30.2
$\delta^{13}C$ [‰]	-28.0	0.0	-0.2	-25.9	0.2	-0.6	-29.4	0.1	-0.3	-24.4	0.1	-0.2	-22.6	0.4	-2.0	-27.9	1.0	-3.6
$\delta^{15}N$ [‰]	4.9	0.2	4.0	5.3	0.2	3.7	2.1	0.3	13.4	4.6	0.2	3.7	1.6	0.2	11.6	4.3	1.3	31.3
χ [10^{-8} m^3 kg^{-1} dw]	15.5	0.8	4.9	15.5	2.9	18.4	12.0	1.0	8.7	62.8	0.7	1.1	16.0	2.1	13.1	17.5	5.3	30.3
L* (D65)	55.1	0.7	1.2	57.1	0.8	1.4	57.2	0.2	0.4	56.4	0.4	0.7	58.7	0.2	0.3	55.7	1.1	2.1
a* (D65)	2.8	0.1	2.9	2.9	0.1	4.8	1.8	0.2	8.5	3.7	0.2	4.1	2.7	0.3	12.4	3.1	0.6	19.9
b* (D65)	14.5	0.2	1.2	15.2	0.5	3.4	12.4	0.7	5.7	14.5	0.3	2.2	14.3	0.7	5.2	14.3	1.0	6.9
Nmic [µg g-1 dw]	30.0	6.2	20.8	37.7	3.8	10.1	32.6	12.4	38.0	27.4	5.6	20.3	57.2	26.6	46.4	29.4	14.1	47.8
Cmic [µg g-1 dw]	535	43	8.1	331	71	21.4	4043	1591	39.4	255	41	16.2	1006	616	61.3	882	1340	151.9
Nitrate [µg N g-1 dw]	7.0	2.8	40.0	6.5	0.8	12.3	15.1	5.0	33.5	3.9	2.3	60.9	21.9	20.9	95.5	10.4	9.0	85.9
Ammonium [µg N g-1 dw]	0.2	0.0	19.3				6.6	5.8	87.9	0.6	0.2	32.6	12.4	8.4	68.2	1.1	2.9	262.7
DON [µg g-1 dw]	9.7	6.4	65.4	4.4	0.9	20.6	10.7	3.0	27.7	0.7	0.5	72.8	21.0	9.5	45.5	6.8	4.3	63.5
DOC [µg g-1 dw]	21.1	4.0	18.9	8.1	2.1	26.0	34.7	7.7	22.1	13.6	2.8	20.3	30.7	20.4	66.3	18.6	9.4	50.5

[a] $\sum$ *n*-Alkanes = $n\text{-}C_{13\text{-}40}$

[b] CPI_{short}: 0,5 * ([($n\text{-}C_{15}$+$n\text{-}C_{17}$+$n\text{-}C_{19}$+$n\text{-}C_{21}$)/($n\text{-}C_{14}$+$n\text{-}C_{16}$+$n\text{-}C_{18}$+$n\text{-}C_{20}$)]+ [($n\text{-}C_{15}$+$n\text{-}C_{17}$+$n\text{-}C_{19}$+$n\text{-}C_{21}$)/($n\text{-}C_{16}$+$n\text{-}C_{18}$+$n\text{-}C_{20}$+$n\text{-}C_{22}$)])

[c] CPI_{long}: 0,5 * ([($n\text{-}C_{25}$+$n\text{-}C_{27}$+$n\text{-}C_{29}$+$n\text{-}C_{31}$)/($n\text{-}C_{24}$+$n\text{-}C_{26}$+$n\text{-}C_{28}$+$n\text{-}C_{30}$)]+ [($n\text{-}C_{25}$+$n\text{-}C_{27}$+$n\text{-}C_{29}$+$n\text{-}C_{31}$)/($n\text{-}C_{26}$+$n\text{-}C_{28}$+$n\text{-}C_{30}$+$n\text{-}C_{32}$)])

[d] P_{aq}: ($n\text{-}C_{23}$+$n\text{-}C_{25}$)/($n\text{-}C_{23}$+$n\text{-}C_{25}$+$n\text{-}C_{29}$+$n\text{-}C_{31}$)

Other conservative parameters (TOC, TN, C/N) influenced predominantly by organic matter input at a given site also show minor dispersion with coefficients of variation generally < 10% in soils with more than 100 yr of paddy cultivation. Only in younger paddy soils and in the reference substrates the CV values of these parameters varied in a slightly broader range up to approximately 20% (Table 2.1).

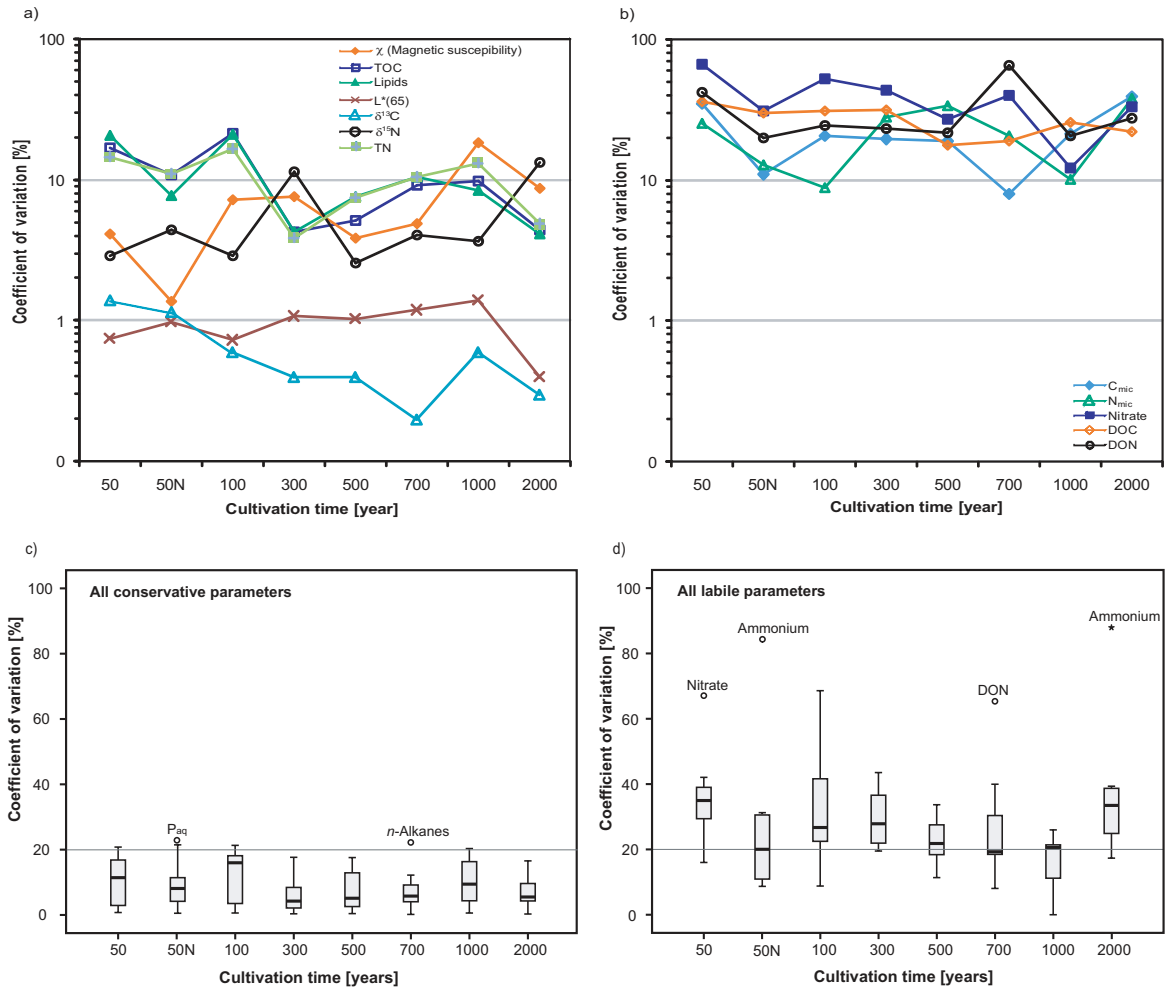


Fig. 2.3. Coefficients of variation for paddy soil sites sampled in pentuplicate with **(a)** conservative parameters, **(b)** labile parameters, **(c)** box and whisker-plots showing median value, 75 percentile, 90 percentile and outliers for conservative parameters, **(d)** same as **(c)** but for labile parameters.

A range in the CV close to or less than 10% at older paddy sites is better than expected and indicates a decreasing variability in soil organic matter parameters in longer used cropping sites. In reference substrates the CV values of conservative parameters are generally higher

than in paddy soils, exemplified by variation of TN and C/N in marine sediments and of TN and TS in lacustrine sediments (Table 2.1).

Lipid and alkane concentration and composition

The highest variations of conservative parameters within a sampling site were noted for the concentration of total extractable lipids and the lipid class of *n*-alkanes. The latter is derived from land plant wax coatings, limnic macrophytes, marine/limnic algae and cyanobacteria or may arise from anthropogenic contamination. As the compound class of the alkanes is most widely studied in soils and it was chosen for compositional analysis. Further lipid classes were excluded from detailed compositional determination. The *n*-alkane distributions reveal a maximum CV of 22% in paddy soil sites and of 10% or 30% in the TW and FW reference substrates, respectively (Table 2.1). The reason for the higher variation in the limnic environment could be attributed to sampling in a shallow water environment. This favoured mixed organic matter input from submerged aquatic macrophyte biomass and terrestrial plant matter supplied by the catchment to the near-shore limnic setting. The range of alkane concentrations in paddy soils is caused by diverse organic matter input from actual crop or weed vegetation, products from incomplete biomass combustion, or fossil fuel contaminants at different "hot-spots" on a site. In general, total extractable lipids in paddy soils represent 5.6% of the total soil organic carbon and are mainly composed of *n*-fatty acids, *n*-alcohols, sterols, long-chain wax esters, sugars and other functionalized lipid classes. On average 3.1 mg kg^{-1} of these soil lipids are composed of source-diagnostic *n*-alkanes (Table 2.1) but in the 700 yr old paddy soil substantially higher proportions of *n*-alkanes (6.5 mg kg^{-1}) were observed, which could be attributed to fossil fuel contamination.

A partial origin of *n*-alkanes from fossil contamination is evident from the presence of a pronounced "unresolved complex mixture" (UCM) and a high abundance of thermally mature tricyclic and pentacyclic triterpenoids (hopanoids) dominating over the recent microbial triterpene diploptene (Fig. 2.4). Recent bacteria biosynthesize the unsaturated 17β(H)-, 21β(H)-hope(22,29)ene also termed diploptene, which is only stable under near-surface conditions (Ourisson et al., 1987). This compound is diagenetically transformed into saturated analogues upon sediment burial when reaching thermal maturity (Peters et al., 2005). In

petroleum thermally stable hopanoids with 17α(H)-, 21β(H)-isomer configuration and a predominance of the 22S over 22R isomers are found. Such a petroleum derived hopanoid distribution has been encountered in the P700 topsoils (Fig. 2.4).

Fossil fuel contamination in a paddy field could originate from a point source in the field, e.g. caused by breakdown of motorized farming machinery associated with spillage of lubricants or fuels. In such a case, only a small area of a few square meters would be contaminated, due to hydrocarbon hydrophobicity preventing further dispersal. The spatially continuous presence of fossil fuel derived hydrocarbons in the P700 field argues against such a localized point spill, but points towards a diffuse contamination, e.g. by inflow of contaminated irrigation canal waters.

The compositional variation of *n*-alkanes in the paddy soils can be evaluated using standardized parameters describing the preferential enrichment of individual alkanes. The carbon preference index (CPI) established by Bray and Evans (1961) is used to highlight the predominance of odd-over even numbered *n*-alkane homologues. High CPI_{long} values for long-chain components (> nC_{23}) derive from fresh plant waxes and tend to decline with increasing biodegradation and thermal maturity. The same accounts for short chain *n*-alkanes (< nC_{22}) derived from algae or cyanobacteria. Fossil fuels exhibit CPI values close to unity. Variation in CPI values thus reflects recent diagenetic progress or fossil fuel origin. The short chain alkanes for paddies and TW reveal CPI values < 1.7 indicative of minor algal and/or cyanobacterial input with only the FW site giving a higher CPI_{short} of 2.1 pointing to more enhanced algal/cyanobacterial contributions. In conjunction with a small average CV of < 10% the overall proportion of aquatic microbial biomass has been low. An even-over odd predominance of short chain (< nC_{21}) *n*-alkane homologues as marker of incomplete combustion products (Wiesenberg et al., 2009, Kuhn et al., 2010) was not observed in topsoils and sediments. More substantial variation was observed for CPI_{long} values around 6.0 for the young paddy sites (P50– P300) and around 4.0 for the older paddies and reference sites, indicative of progress in diagenetic overprinting. Exceptionally high CPI_{long} values were observed for P500 (Table 2.1) and indicate an origin of plant waxes from crops other than rice. Based on comparison with recent crop plants (data not shown), the *n*-alkane distribution

at this site is governed by input of wax lipids from the upland crop rape (*Brassica napus*). Very low CPI_{long} values of 2.5 for P700 are due to admixture of fossil fuel with a CPI near unity and thus support the UCM and hopanoid data. For all CPI_{long} values CV values are between 5 to 20% showing no preference for source or degree of diagenesis.

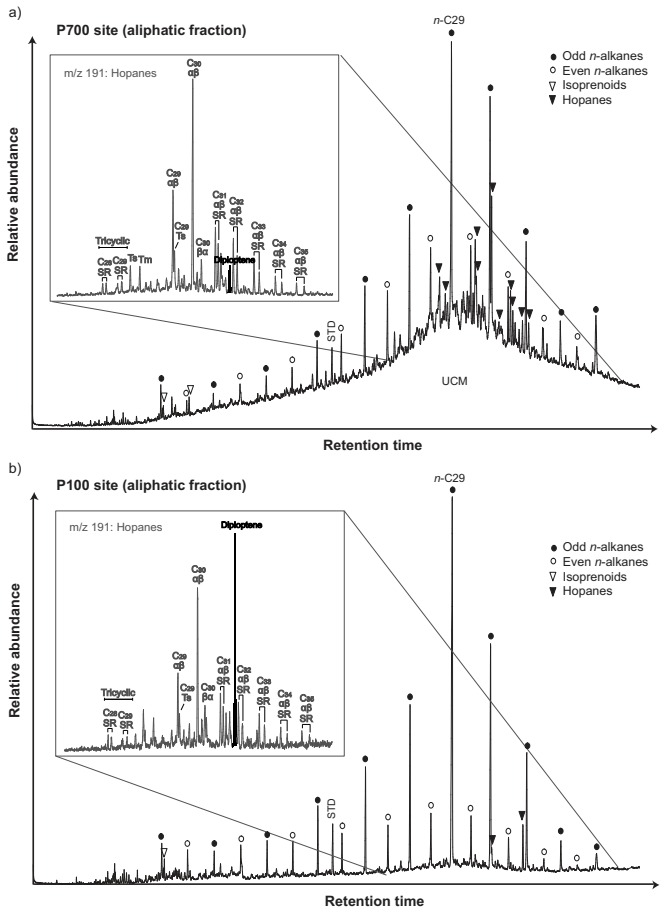


Fig. 2.4. GC/MS total ion traces obtained from aliphatic hydrocarbon fractions, with major peaks labelled for identification. The inset shows the extracted mass fragmentogram of *m/z*=191, indicative for tri- and pentacyclic triterpenoids. Diploptene marked black is indicative of recent bacteria, hopanes and tricyclic hydrocarbons in grey derive from fossil fuel contamination. Peaks are labelled according to number of carbon atoms per molecule and isomerisation at position C_{17}, C_{21} and C_{22}.
Ts=trisnorhopane, Tm=Trisnorneohopane.
Note high abundance of fossil fuel hopanes vs. recent diploptene in P700.

The P_{aq}-ratio established by Ficken et al. (2000) has been used to determine the relative contribution of submerged aquatic macrophytes to the sedimentary *n*-alkane load. Values for all paddies except for P500 are close to that of the TW substrate and only the FW reference sites reveal enhanced values of 0.34 (Table 2.1). The exceptionally low P_{aq} of P500 is due to an origin of alkanes from non-aquatic plants and emphasizes an origin from upland crops growing under dry conditions. During time of sample acquisition the P500 site was used as a paddy soil and had been under this utilization for more than 3 yr. The molecular composition

of lipids, in particular *n*-alkanes, from this site however clearly reflects its previous long-term use as a non-paddy upland cropping site. The time-integrative manner of conservative molecular biogeochemical indicators unravelling the temporally dominating land-use of soils is well illustrated in the P500 case. Despite the coexistence of older (< 3 yr) and fresh lipids, the CV for various molecular parameters is only 12 to 17%.

For all sites the CV values are below 20% and confirm that application of molecular proxies for source identification, degree of diagenetic overprinting and fossil fuel contamination are very robust and reliable.

Integrating conservative and labile parameters over cultivation times

Calculation of averages for the most important conservative parameters (TC, TOC, TN, TS, lipid yield, alkane yield, $\delta^{13}C$, $\delta^{15}N$, χ, L^*, a^*, and b^*) resulted in CV values of < 20% for all cultivation times as shown in Fig. 2.3c. A prominent outlier (exceeding the 95 percentile) was the *n*-alkane yield at P700 that is controlled by addition of fossil fuel contaminants to this site.

A comparison of the P50 and P50N site reveals differential behaviour that can be explained by management practices. Soil at the P50 site has experienced a long series of redox cycles like all other paddy soils, whereas the P50N site as a rice seedling nursery is kept under flooded conditions for longer times and thus has been going through less frequent and less dramatic redox cycles. As a result paddy soil evolution at the P50N site proceeded further compared to other sites of comparable overall cultivation time. A notable exception is the P_{aq} ratio that shows a CV comparable to the P50 or P100 sites (Fig. 2.3c), indicating that water table fluctuations at the P50N site affect wax lipid composition of rice seedlings, even if the soils do not pass through completely dry cycles.

Except for the P1000 site where interferences due to dyke maintenance have been reported (Roth et al., 2011; Wissing et al., 2011), a decline in the CV values over cultivation time is noted (Fig. 2.3c). This can be explained by an increasing degree of paddy soil evolution and homogenization of SOM and minerogenic composition, accompanied by stabilization in soil microbial community structure. This establishment of quasi-continuous composition in conservative paddy soil biogeochemical parameters was established after only 300 yr of cultivation time.

The averages for labile parameters according to cultivation age depicted in Fig. 2.3d show a much higher degree of variability, with the lower and upper boundaries of CV values for the 75 percentile ranging between 20 and 40% (Fig. 2.3d). Outliers exceeding the 95 percentile are nitrate for the P50, DON for the P700, and ammonium for the P2000 site. All of these spatially highly variable parameters are associated to reactive compounds of the nitrogen cycle and are highly influenced by spatially non-systematic human manipulation such as fertilization. Additionally, the presence of biopores and cracks in the plough pan could contribute to irregular leaching processes coupled with a high variability of these water soluble parameters within a field (Sander and Gerke, 2007). The approximately 20 times higher ammonium content in P2000 could have been induced by uneven manual application of nitrogen fertilizer and an inefficient field management practices just prior to sampling (see also Roth et al., 2011). Other indicators of nitrogen cycling in paddy soil including microbial N (CV_{avg} = 33%), C/N (CV_{avg} = 5%), or $\delta^{15}N$ (CV_{avg} = 3%) behave stable and demonstrate the establishment of a well controlled nitrogen cycle in paddy soils.

Inter-site variability

Reliable identification of differences in biogeochemistry between individual paddy fields and interpretation of evolutionary trends according to cultivation time, physiogeographical properties, management practises, etc. can only be achieved, if the intra-site heterogeneity is smaller than inter-site differences in biogeochemistry. We thus tested individual parameter relationships and applied statistical approaches to the entire data set employing PCA and non-parametric tests for variance analysis as well as the Kruskal-Wallis test to verify that inter-site exceeds intra-site variation. Finally, a cluster analysis was performed to elucidate, if duration of paddy soil management and associated soil evolution leads to establishment of robust clusters of comparable soil properties for the different paddies.

Examination of pairs of individual parameters revealed that in general values for each site clustered closely and only moderate overlap between site clusters occurred. This is exemplarily shown in Fig. 2.5a, b and c for $\delta^{13}C$ vs. $\delta^{15}N$, χ vs. a^* and $\delta^{15}N$ vs. P_{aq}, respectively.

Table 2.2. Non-parametric variance analysis by Kruskal-Wallis test, suitable for non-normal distributed data sets, performed for all sites (n = 49) and for paddy sites only (n = 34). Significant variation between sites is indicated, *H* values from Chi-squared test exceed the critical *H*-values of the null-hypothesis. Parameters indistinguishable between sites because intra-site variance exceeds inter-site variance are plotted in italic.

	All sites				**All paddy sites**			
	H	***df***	***p***	***Hcrit***	***H***	***df***	***p***	***Hcrit***
TC [%]	40.824	9	0.000	28.992	29.524	6	0.000	27.800
TOC [%]	43.880	9	0.000	39.396	27.995	6	0.000	29.200
TIC [%]	33.971	9	0.000	28.191	22.712	6	0.001	14.800
TN [%]	43.005	9	0.000	36.195	29.282	6	0.000	29.200
TS [%]	38.561	9	0.000	37.995	26.502	6	0.000	27.800
C/N	24.233	9	0.004	17.187	16.390	6	*0.012*	16.400
TOC/S	30.601	9	0.000	23.389	19.565	6	0.003	16.400
Lipids [mg kg^{-1} dw]	41.197	9	0.000	40.997	28.920	6	0.000	29.200
Lipids [g kg^{-1} TOC]	37.719	9	0.000	29.992	17.955	6	0.006	18.200
n-Alkanes [µg kg^{-1} dw]	44.868	9	0.000	39.396	30.076	6	0.000	29.200
n-Alkanes [mg kg^{-1} TOC]	41.218	9	0.000	36.195	26.051	6	0.000	18.200
CPI_{short}	30.716	9	0.000	28.191	16.890	6	*0.010*	18.000
CPI_{long}	39.956	9	0.000	36.195	25.216	6	0.000	22.800
P_{aq}	35.325	9	0.000	25.190	14.853	6	*0.021*	14.000
$\delta^{13}C$ [‰]	45.219	9	0.000	37.995	30.028	6	0.000	23.600
$\delta^{15}N$ [‰]	46.165	9	0.000	40.997	31.663	6	0.000	29.200
χ [10^{-8} m^3 kg^{-1} dw]	44.184	9	0.000	36.195	29.069	6	0.000	21.200
*L** (D65)	42.415	9	0.000	40.997	27.689	6	0.000	23.600
*a** (D65)	44.108	9	0.000	39.396	30.311	6	0.000	29.200
*b** (D65)	37.053	9	0.000	28.191	26.286	6	0.000	21.200
N_{mic} [µg g^{-1}dw]	33.908	9	0.000	24.190	21.355	6	0.002	16.600
C_{mic} [µg g^{-1} dw]	41.157	9	0.000	36.195	28.730	6	0.000	26.000
Nitrate [µg N g^{-1} dw]	35.111	9	0.000	28.191	25.720	6	0.000	19.600
Ammonium [µg N g^{-1} dw]	41.891	9	0.000	37.995	26.021	6	0.000	15.000
DON [µg g^{-1} dw]	43.210	9	0.000	39.396	27.202	6	0.000	21.200
DOC [µg g^{-1} dw]	31.265	9	0.000	28.992	25.586	6	0.000	21.400

The binary plots demonstrate that individual single parameters often do not show a clear separation between sites, whereas addition of a second dimension allows full discrimination of the site clusters. The bulk isotope parameters show no overlap between site clusters and clearly separate the soils from the marine and limnic substrates (Fig. 2.5a). Soil magnetic susceptibility and colour parameters depending on minerogenic composition and expression of redoximorph features also show clear separation of site clusters with little to no overlap (Fig. 2.5b), except for one outlier in soil colour at the FW site. Combinations of molecular compositional and isotope parameters are suitable for site differentiation also exhibiting less

intra-site versus inter-site variability (Fig. 2.5c). Hereby the variance in P_{aq} ratios is substantially higher than for $\delta^{15}N$ signatures. In general, overlap due to spread in one parameter is more frequent in young, less well developed or in disturbed soils (P1000).

The Kruskal-Wallis test is employed in ecological, biogeochemical, and environmental quality studies to evaluate, whether variance between sites exceeds variance within sites (Gratton et al., 2000; Katsaounos et al., 2007; Lehndorff and Schwark, 2008). Including all sites and parameters the asymptotic significance gave $p < 0.001$, except for the C/N ratio where a value of 0.004 was reached (Table 2.2), indicating that sites are less well distinguishable. Nevertheless the critical H values for all sites did not exceed the H-values proposed for the null-hypothesis (Table 2.2), implying that a full discrimination of all sites using median values of any of the selected parameters was possible. If the data set was reduced to contain only the paddy sites, i.e. P500, TW, and FW excluded (Table 2.2), asymptotic significance values for the C/N, CPI_{short} and P_{aq} -ratios for $p > 0.01$ could not be met. Furthermore, the median-referred critical H-values exceeded the H-values for the following parameters: TS, C/N, TOC, extract yield and CPI_{short} (Table 2.2), indicating that the intra-site variance was comparable to or exceeded inter-site variance. As most of these parameters represent concentrations that are primarily related to the absolute amount of soil organic matter rather than its compositional differences, a discrimination of sites using these such indicators is not feasible.

Application of principal component analysis allows evaluating the entire data set and was carried out on all paddy soils using all parameters determined (Fig. 2.5d), and for all paddy soils employing conservative parameters only (Fig. 2.5e). A full discrimination of all sites was achieved (Fig. 2.5d), when using the 1st and 2nd regression factors of all parameters, explaining 39.8 and 17.3% of variance in the data set (Table 2.3). Factor 1 exhibits highest loading by parameters associated with soil organic matter concentration, whereas the 2nd factor is primarily controlled by minerogenic composition parameters (Table 2.3). No overlap of parametric values between the individual sites was observed, which was taken as direct evidence that the intrinsic heterogeneity of paddy soil does not exclude detailed interpretation of biogeochemical differences between sites.

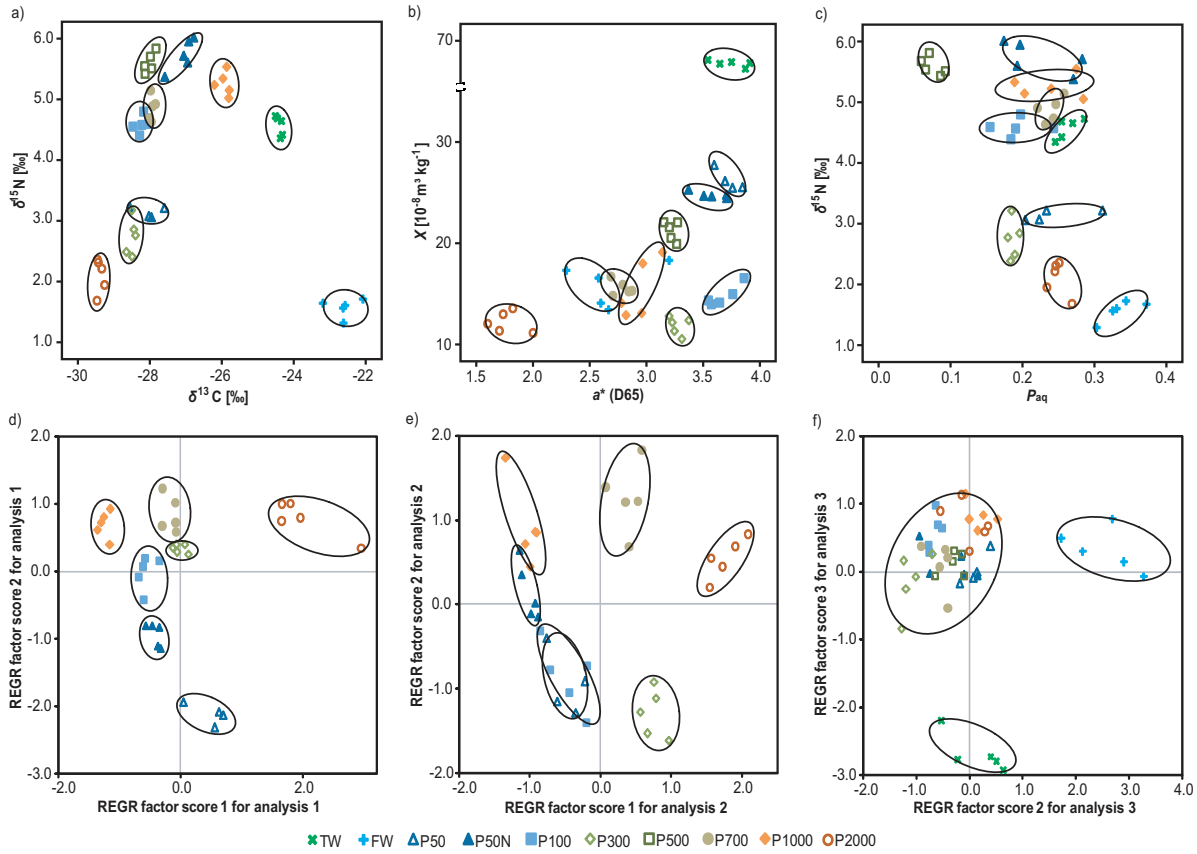


Fig. 2.5. Discrimination of variance between sites versus in-site using descriptive statistics **(a)** to **(c)** and non-parametric and multivariate methods **(d)** to **(f)**. The factor plots obtained from PCA are shown for **(d)** all paddy soils using all parameters, **(e)** all paddy soils using exclusively conservative parameters, and **(f)** using all paddy soils, non-paddy soil P500 and substrates FW and TW. Note that discrimination of substrates was achieved best, when using the 2nd and 3rd factor rather than 1st and 2nd factor as in **(d)** and **(e)**.

If the data set was reduced to the conservatively behaving parameters, the discriminative power was reduced (Fig. 2.5e), whereby preferentially overlap for younger sites was observed. The 1st and 2nd regression factors for the data restricted to conservative parameters explain 50.6 and 14.2% of the variability and are controlled by organic matter concentration and n-alkane compositions, respectively (Table 2.4). The similarity in biogeochemical properties concerning the conservative parameters, in particular for the P50 and P100 sites, can be attributed to the low evolutionary stage of the paddy soils. All sites under paddy cultivation for 300 yr and more have developed individual soil characteristics as mentioned above when discussing CV for individual age classes. Not only the duration of paddy soil

utilization is of critical importance but also the individual management practice. Two sites used for 50 yr of rice cultivation were investigated, whereby one of these sites was used as nursery (P50N) for growing rice seedling prior to transplantation. As the P50N site is consistently kept under flooded conditions, soil evolution proceeds differently from the P50 site. This allows full discrimination of the P50N from the P50 and P100 sites, whereas the latter two do show considerable overlap, when PCA is conducted (Fig. 2.5e).

Table 2.3. Factor loadings table obtained from PCA performed with all paddy soils and all parameters illustrated in Fig. 2.5d.

Analysis 1	**Factor loadings**			
	1	**2**	**3**	**4**
TC	0.932	0.306	-0.054	-0.061
TN	0.882	0.427	-0.062	0.127
TS	0.836	0.267	-0.102	-0.336
C/N	-0.028	-0.122	0.108	-0.790
TOC/S	0.373	0.611	-0.041	0.396
TOC	0.885	0.437	-0.063	-0.056
CPI_{short}	0.274	-0.488	0.509	0.005
CPI_{long}	-0.028	-0.264	-0.842	0.265
P_{aq}	0.100	-0.155	0.687	0.017
$\delta^{15}N$	-0.860	0.085	0.317	-0.127
$\delta^{13}C$	-0.874	0.021	0.222	-0.035
*L**(D65)	-0.075	0.044	0.651	0.556
*a**(D65)	-0.646	-0.522	-0.439	-0.157
*b**(D65)	-0.817	-0.369	-0.019	-0.149
χ	-0.370	-0.839	0.263	-0.096
Lipids	0.835	0.502	0.122	-0.117
n-Alkanes	0.677	0.460	0.320	-0.389
TIC	-0.170	-0.926	0.078	0.002
N_{mic}	-0.410	-0.183	0.571	0.415
C_{mic}	0.711	0.143	0.318	0.541
Nitrate	0.464	0.370	-0.634	-0.033
Ammonium	0.648	0.065	0.186	0.533
DON	0.529	0.507	0.205	0.038
DOC	0.845	-0.026	0.218	0.155

Table 2.4. Factor loadings table obtained from PCA performed with all paddy soils and all conservative parameters illustrated in Fig. 2.5e.

Analysis 2	**Factor loadings**			
	1	**2**	**3**	**4**
TC	0.987	0.009	0.006	-0.018
TN	0.964	0.028	-0.191	-0.111
TS	0.897	0.001	0.304	-0.051
C/N	-0.019	0.137	0.815	0.035
TOC/S	0.533	0.093	-0.512	-0.383
TOC	0.983	0.052	-0.018	-0.133
CPI_{short}	0.094	0.195	-0.025	0.812
CPI_{long}	-0.064	-0.909	-0.173	-0.101
P_{aq}	-0.008	0.608	-0.133	0.532
$\delta^{15}N$	-0.799	0.383	0.170	-0.249
$\delta^{13}C$	-0.840	0.322	0.006	-0.158
*L**(D65)	-0.106	0.579	-0.617	0.233
*a**(D65)	-0.748	-0.543	0.283	0.047
*b**(D65)	-0.899	-0.081	0.233	0.036
χ	-0.614	-0.054	0.230	0.654
Lipids	0.950	0.247	0.031	-0.141
n-Alkanes	0.796	0.445	0.306	-0.106

Including the reference substrates in PCA for the conservative parameters reveals a more pronounced differentiation of the substrates from the soils that develop on them when 2nd factor and 3rd factor extracted by PCA, explaining 18.3% and 17.0 percent of the variance

(Table 2.5) are used for discrimination (Fig. 2.5f). The abandonment of factor 1 explaining 37.93 percent of variance in the regression analysis leads to incomplete separation of the individual paddy sites emphasizing the importance of this factor in discriminant analysis. The properties exhibiting the highest loading scores on the 1st factor are organic matter concentration-related, those for the 2nd factor are governed by alkane composition, isotope signature and soil brightness, those for the 3rd factor are dominated by properties linked to redox-conditions (Table 2.5).

Table 2.5. Factor loadings table obtained from PCA performed with all sites and all conservative parameters illustrated in Fig. 2.5f.

Analysis 3	**Factor loadings**			
	1	**2**	**3**	**4**
TC	0.942	-0.161	0.207	0.061
TN	0.829	-0.245	0.477	-0.022
TS	0.871	-0.211	-0.345	0.174
C/N	0.055	0.629	-0.391	0.075
TOC/S	0.271	0.050	0.894	-0.200
TOC	0.871	-0.139	0.451	0.001
CPI_{short}	0.131	0.758	0.302	0.061
CPI_{long}	-0.154	-0.207	0.152	-0.910
P_{aq}	0.093	0.568	-0.085	0.773
$\delta^{15}N$	-0.625	-0.465	-0.114	-0.131
$\delta^{13}C$	-0.307	0.771	-0.365	0.162
*L**(D65)	0.003	0.798	0.089	0.279
*a**(D65)	-0.750	-0.293	-0.406	-0.047
*b**(D65)	-0.813	-0.038	-0.039	0.416
χ	-0.261	0.067	-0.904	0.033
Lipids	0.905	0.260	0.100	0.187
n-Alkanes	0.797	0.124	0.286	0.170

Cluster analysis performed to evaluate whether the statistical approach would group individual sites in clusters based on the entirety of all biogeochemical parameters determined is presented in Fig. 2.6. All older paddy soils not affected by secondary alteration (P500: extended non-paddy use, P1000: removal of surface soil for dyke maintenance) are clustered appropriately, whereas the younger soils exhibit insufficient development of individual biogeochemical paddy soil characteristics.

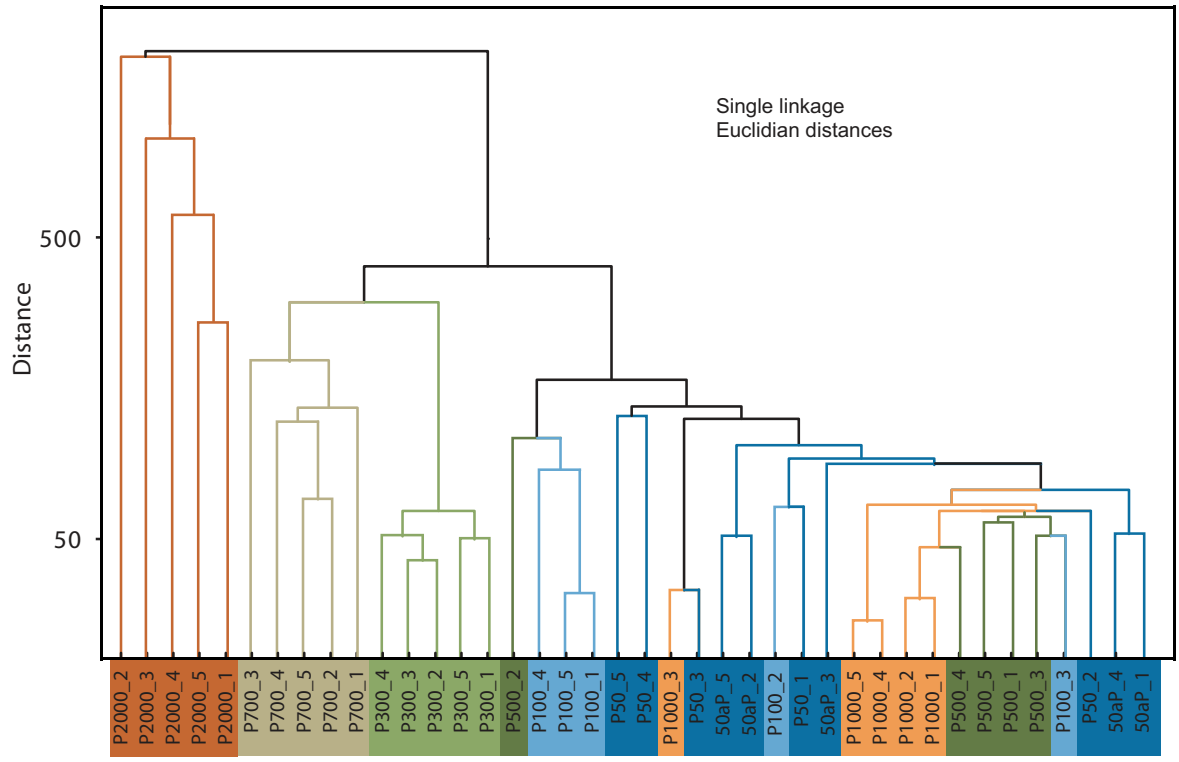


Fig. 2.6. Cluster diagram for paddy soils and non-paddy soil P500, constructed using all parameters.

Contamination of the P700 site did not lead to a significant change in time-integrated basic soil biogeochemical parameters but preferentially affected the aliphatic hydrocarbon composition. This indicates that the addition of the petroleum contaminants did not detrimentally affect the soil microbial community or inhibited plant growth by adding toxic substances or providing alternative substrates for microbial utilization. Seen from a temporal perspective, the P700 site acquired its biogeochemical profile over a time span of about 700 yr, whereas the minor petroleum contamination is assumed to have occurred only a few years ago and thus has negligible influence on the overall biogeochemical status. In a similar but opposite fashion, the P500 site reveals the cumulative biogeochemical characteristics of almost 500 yr of use as upland cropping site and only recently (approx. 3 yr before sampling) was converted to a paddy field. Consequently, this site still exhibits the time-integrated features of a non-paddy land management and clusters with the very young paddies (P50,

P50N, P100) developed on a marine tidal substrate (Fig. 2.6). Similarly, the P1000 site, though continuously utilized as paddy field, groups with the young soils due to repetitive removal of surface soils and dilution with soil material of non-paddy origin. Both of these sites, the P500 as well as P1000 exhibit four subsamples of close similarity and one subsample of largely deviating character, indicating the large intra-site variation caused by human interference. Although several sites of the chronosequence studied have been affected by anthropogenic perturbations, the cluster analysis indicates that paddy soil evolution over time led to the establishment of stable biogeochemical properties and conditions, even if permanent human intervention via soil management and utilization prevailed.

Organic matter accumulation trends

The environmental budget of paddy soils is under debate and considered to show a negative balance due to the emission of greenhouse gases and intensive nitrate loss from paddies (IPPC, 2007; Conrad, 2009; Kögel-Knabner et al., 2010). On the other hand a positive balance could be attributed to paddy soils based on atmospheric CO_2 sequestration via surface soil accumulation and preservation of fresh photosynthate. Organic carbon accumulation in Chinese paddy ecosystems has been investigated for different locations, climates, soil substrates, and management systems (Cheng et al., 2009; Pan et al., 2003; Wissing et al., 2011; Wu, 2011). The results obtained in this study cannot be generalized to global scale or taken to represent a Chinese paddy soil carbon budget as the Zhejiang land reclamation area with about 1500 km^2 of paddy rice cropping only represents a very minor fraction of the total wetland rice cropping area of 156×10^6 ha (R. Jahn, personal communication, 2011). Wu (2011) identified generally higher organic carbon stocks in paddy soils than in other arable land-use types in different landscapes of China, attributed to slower organic matter degradation rate under anaerobic conditions (Lal et al., 2002). The chronosequence studied here offers the opportunity to evaluate CO_2 sequestration in paddies, comparison with non-paddy sites (P500) and interferences via intentional management (P1000) or unintentional contamination (P700). At all sites paddy soils developed on the same tidal flat sediments and land management practices were comparable. The accumulation trends for TOC, lipids and *n*-alkanes over cultivation time are shown in Fig. 2.7a–c, complemented by the accumulation of

lipids and alkanes normalized to TOC (Fig. 2.7d and e). The TOC concentrations of paddy soil reach approx. 1% after 50 yr of cultivation, i.e. more than double the concentration of the parent substrate (Table 2.1).

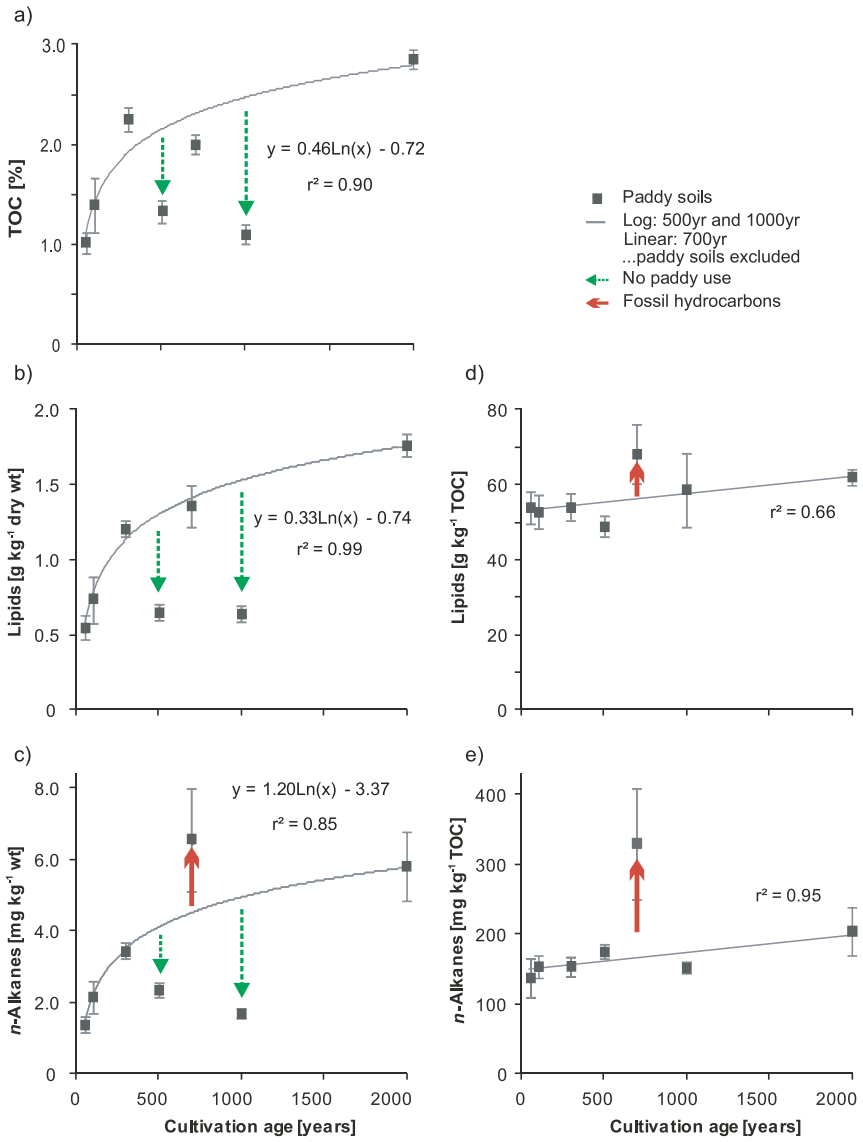


Fig. 2.7. Accumulation trends of **(a)** TOC, **(b)** lipid yield and **(c)** *n*-alkane yield normalized to dry sample weight and **(d)** lipid yield and **(e)** *n*-alkane yield normalized to TOC over cultivation time. Arrows denote deviations from natural accumulation trends due to human disturbance of the paddy soil system. The P500 was used as upland field and only recently converted to paddy soil use, the P1000 site experienced topsoil removal and admixture of other soil material in the course of dyke maintenance work, the P700 site suffers from petroleum contamination.

Increase in TOC continues to be rapid (Targulian and Krasilnikov, 2007) until about 300 yr and levels off to reach maximum concentrations after 2000 yr of cultivation. Severely lower TOC concentrations are noted for the P500 site, which might indicate a use as non-paddy for a longer period and thus much less TOC compared to a paddy soil has accumulated. Comparison of paddy and non-paddy soils from the same chronosequence revealed significantly higher organic carbon contents in paddy soils, whereas in non-paddy soils no increase in organic matter accumulation upon long-term utilization was observed (Cheng et al., 2009; R. Jahn, personal communication, 2011). The P1000 site is assumed to have accumulated TOC continuously but has lost about 50 to 70% of this TOC due to human interference as also observed in Wissing et al. (2011).

The lipid concentration of the paddy soils reveals a similar accumulation pattern though solvent extractable lipids are less stable than recalcitrant TOC (Wiesenberg et al., 2004; Marschner et al., 2008) that includes non-extractable fractions, e.g. black carbon from incomplete combustion of rice straw. Normalization of lipid yield to soil TOC indicates that these constitute approx. 5 to 6% of the organic carbon (Fig. 2.7d) and that the relative concentration increases with cultivation age. This indicates that the labile proportion of organic input into paddy soils and/or the standing microbial biomass is better preserved in long cultivated paddy soils.

The *n*-alkane accumulation trends exhibit not only negative perturbations as did TOC and lipid yield but also a pronounced increase at the P700 site due to fossil fuel contamination (Fig. 2.7c). This externally added *n*-alkane fraction even increases in proportion, if normalized to TOC concentrations (Fig. 2.7e). Exclusion of the P700 site still results in an increase of n-alkanes over time, which is explained by the lower mineralization of *n*-alkanes compared to functionalized lipids, e.g. fatty acids or alcohols. These components undergo oxidation and decarboxylation reactions upon diagenesis, finally leading to generation and accumulation of *n*-alkanes. The stability and turnover times of *n*-alkanes in paddy soils have not yet been studied and due to intensive redox cycling may deviate from upland arable and grassland soils. The former were shown to have turnover times for *n*-alkanes slightly shorter

than total organic carbon (Wiesenberg et al., 2004) whereas under the latter conditions the *n*-alkanes may remain even more stable than total soil organic matter (Wiesenberg et al., 2008). Aberrations in the accumulation of organic matter and organic matter fractions can be sufficiently explained by applications of molecular proxies, unravelling deviating sources of organic input or fossil fuel contamination. A different input of plant material, preferentially rape based on comparison of *n*-alkane distributions with recent reference crops, at the P500 site revealed this site to have been used as upland field for prolonged periods. Exceptional concentrations and compositional differences in aliphatic hydrocarbons, in particular *n*-alkanes, hopanes and UCM identify human perturbations of the soil ecosystem by petroleum contamination.

2.4 Conclusions

Biogeochemical proxies determined on five field replicates of paddy soils differing in cultivation age and two substrates on which these paddy soils evolved showed that the intrinsic heterogeneity of paddy soil organic and minerogenic components is smaller than differences in biogeochemical properties between study sites. This conclusion was drawn based on interpretation of individual parameters, descriptive and non-parametric statistical analysis, PCA and cluster analysis. The coefficient of variation for conservative parameters determined in pentuplicate and reflecting time-integrated evolution of soil properties in general was 10% or lower. Based on this study, collection of one composite field sample is considered sufficient for generation of representative biogeochemical data in paddy soils. In field heterogeneity of fast cycling and anthropogenically amended nutrients was found very high with coefficients of variation usually between 20 and 40% and frequent outliers. Sampling strategies covering the heterogeneity of such parameters will require much higher sampling frequency and spatial resolution.

Biogeochemical properties acquired by paddy soils over centennial periods of time behave conservatively and do adapt not rapidly, if management conditions or practices are altered. Hence, previous historic land use or management practices can be reconstructed, even after new utilization has been established.

Duration of cultivation as paddy soil leads to establishment of specific soil characteristics that become increasingly stable with cultivation time. For paddies evolving on marine tidal substrates as in this study, the full development of paddy biogeochemical signatures was completed in less than 300 yr.

The environmental/ecological budget of paddy soils in the coastal Zhejiang land reclamation area studied here revealed a positive balance, when sequestration of atmospheric CO_2 was considered. Paddy soils show a strong organic matter accumulation trend during 2000 yr paddy cultivation, which indicates the sequestration of atmospheric CO_2. The environmental/ ecological budget is only of local importance, due to soils evolving in a highly specific environment. Perturbation of paddy soils leading to severely reduced sequestration potential can be identified by application of molecular source proxies. Thus the integrity of the carbon accumulation history of paddy soils in unknown areas can be critically evaluated.

2.5 Acknowledgements

We thank the German Research Foundation (DFG) for financial support (Schw554/20). Chinese and German partners of Research Initiative FOR 995 are thanked for field work collaboration. We appreciate analytical assistance by laboratory staff at Cologne University. We thank R. Jahn, for support and providing information on sample location.

3. Chemotaxonomy and diagenesis of aliphatic hydrocarbons in rice plants and soils from land reclamation areas in the Zhejiang Province, China

C. Mueller-Niggemann[1], L. Schwark[1, 2]

[1]Institute of Geosciences, Christian-Albrechts-University of Kiel University, Kiel, Germany
[2]WA-OIGC, Curtin University, Perth, Australia

Published in *Organic Geochemistry* **83-84** (2015) 215-226.
doi:10.1016/j.orggeochem.2015.03.016

Abstract. Rice is the most important staple food globally and requires large growth areas kept under flooded paddy conditions, contributing significantly to microbial greenhouse gas emissions. Biogeochemical cycling in such agroecosystems has been investigated intensively but molecular biomarker studies are scarce. We conducted a chemotaxonomic investigation of wax alkanes differentiated for plant tissues (leaf, stem, root) of rice and intercrop plants (maize, sorghum, rape, mustard, bean, cotton) and studied the incorporation of these lipids into soil under paddy compared with upland management forms. Soil chronosequences reflecting paddy and upland agroecosystem changes compared with natural soil substrates over the past two millennia were studied in land reclamation areas of the Zhejiang Province, China. Soils evolved on tidal wetland sediments contained predominantly lipids derived from terrigenous supply by the Yangtze River (YR = Chang Jiang) and to a lesser extent from marine sources via the East China Sea (ECS). Agricultural usage converted lipid composition of topsoil within 50 yr to reflect the *n*-alkane patterns of crops with their relative proportion increasing with cultivation time. Alkyl chain length distribution of rice was broad compared with upland plants, due to the water regime changing over the growth period. This separated paddy from upland managed soils on the basis of alkyl lipids, allowing reconstruction of past land use change. Combustion of crop biomass after harvest is common practice and generates

alkenes upon incomplete combustion which, due to their high reactivity are immediately converted to alkanes with lower carbon preference index (CPI) in topsoil. The storage of lipids and organic matter over time is greater in paddy than in upland managed agroecosystems and contributes to CO_2 sequestration from the atmosphere.

3.1 Introduction

Acyclic alkanes and functionalized alkyl lipids serve as important constituents in plant wax and insect cuticular wax, regulating evapotranspiration (Eglinton and Hamilton, 1967; Lockey, 1988, 1991; Samuels et al., 2008; Schwark, 2010). Their relative abundance and distribution in plant wax have been employed for chemotaxonomic classification (e.g. Eglinton et al., 1962; Herbin and Robins, 1969; Gülz, 1994; Bianchi, 1995; Schwark et al., 2002; Cameron et al., 2002; Rommerskirchen et al., 2006; Vogts et al., 2009; Diefendorf et al., 2011; Bush and McInerney, 2013; Carr et al., 2014), often complemented by carbon and hydrogen compound-specific isotope analysis. Alkyl lipid properties differentiate C_3 from C_4 plants and reflect biosynthetic dependence on hydrological conditions (Collister et al., 1994; Lichtfouse et al., 1994; Lockheart et al., 1997; Bianchi and Bianchi, 1990; Chikaraishi and Naraoka, 2003; Sachse et al., 2004; Rommerskirchen et al., 2006; Diefendorf et al., 2011; Kahmen et al. 2013). Therefore, plant wax composition and corresponding compound specific isotopes have been employed in paleoclimate reconstruction (Meyers, 1997; Xie et al., 2000; Schwark et al., 2002; Schefuß et al., 2003; Liu et al., 2005; Zhou et al., 2005; Sachse et al., 2012). Plant wax lipids have been applied in agroecology, soil and crop science to issues including, among others, crop differentiation (e.g. Bianchi 1995; Wiesenberg et al. 2004; Wiesenberg and Schwark, 2006), forage assessment (Dove, 1992; Dove et al. 1996), plant soil lipid transformation pathways and lipid degradation in soils (Dinel et al 1990; Amblès et al., 1994) and the calculation of biomass turnover rate in soil (Balesdent and Mariotti, 1996; Wiesenberg et al., 2004; Amelung et al., 2008). Studies aimed at the characterization of agroecosystems are dominated by investigations of upland soils, i.e. soils without episodic water flooding and associated crop plants, including corn/maize, wheat, barley, rape and other

grasses. However, > 50% of the world's population depends on rice as a staple food. Only a small portion of the rice grown is in upland fields, with > 95% from cultivation in rice paddies (Kölbl et al., 2014), which are episodically water-flooded fields. Despite its economic and ecological importance, the lipid composition of rice plants, their combustion products and paddy soils has rarely been investigated and if so, only–single sample data were provided (Hays et al., 2005; Atahan et al., 2007; Islam et al., 2009; Mueller-Niggemann et al., 2012; Kölbl et al., 2014). There is only one short report (Bianchi et al., 1979) about the composition of rice plant wax. This indicates the need for a more systematic analysis of rice plant lipids and their mode of incorporation into rice paddy soil. When addressing plant lipid constitution, variation in lipid composition according to plant organ and growth stage should be taken into account. In agroecosystems with sufficient water supply, up to three rice harvests per annum are feasible (Kölbl et al., 2014) but, under restricted water supply, only one or two harvests per annum can be achieved. Farmers then use the fields outside of the rice growth period for planting other cereal crops, vegetables or fallow N fertilizer plants (beans, rape) under upland conditions. Consequently, soil under this management type should receive a plant wax input not exclusively from rice but also from intercrop plants. This requires a parallel study of intercrop plants to account for their lipid contribution to paddy soils. Soil *n*-alkanes possess turnover times > 100 yr (Wiesenberg et al., 2004), causing soil having been under paddy management for decades to still contain *n*-alkanes derived from the original soil substrate. These *n*-alkane fractions would have to be determined by using reference substrates for paddy soils. Despite input via subaerial crop biomass, agricultural soils should receive contributions from crop root exudates, weeds, epiphytic algae, microbial organic matter (OM), organic fertilizer (manure), biomass combustion residues and riverine influx during flooding. In Chinese arable land, the accumulation behaviour of soil organic carbon (SOC) in various locations, climates, soil substrates and management systems has been investigated (Pan et al., 2003; Cheng et al., 2009; Wissing et al., 2011; Wu, 2011; Mueller-Niggemann et al., 2012). Generally higher SOC content in paddy soil than in other arable land use types (Wu, 2011; Mueller-Niggemann et al., 2012; Kalbitz et al., 2013; Kölbl et al., 2014) has been reported and may be attributed to slower degradation and enhanced preservation of OM under

anaerobic conditions, caused by management-induced redox cycling (Lal, 2002; Mueller-Niggemann et al., 2012; Kölbl et al., 2014) and input of black carbon (BC) originating from crop residue burning (Lehndorff et al., 2014).

In the present study we focus on *n*-alkanes in rice and intercrop plants, as well as aliphatic hydrocarbons in rice paddy soils and potential substrates, including anthropogenic pollution in rice paddy fields in the Zhejiang Province (China). In this region, soils with a well documented 2000 yr history of land reclamation from marine tidal flats after dyke construction along the shoreline of the southern Hangzhou Bay, allowed investigating specific biogeochemical soil dynamics involved in long term paddy or upland soil development. Previous geochemical investigations in this area concentrated predominantly on chemical properties such as desalinisation, decalcification and nutrient content in marsh soils (Shi et al., 2005; Iost et al., 2007). Furthermore, biogeochemical studies of the same chronosequence have addressed soil pedogenesis and heterogeneity via investigations of SOC, minerogenic content, texture and microbial composition (Cheng et al., 2009; Bannert et al., 2011a,b; Chen et al., 2011; Roth et al., 2011; Wissing et al., 2011; Mueller-Niggemann et al., 2012; Kölbl et al., 2014; Huang et al., 2015). Sediments from the Yangtze Estuary and the ECS, invoked as parent substrates of soils investigated here (Mueller-Niggemann et al., 2012; Kölbl et al., 2014), were characterized for stable C and N isotopes (Kao et al., 2003; Zhang et al., 2007) as well as lipids (Zegouagh et al., 2000; Jeng and Huh, 2004; Wang et al., 2008b) to provide information about sources and preservation of OM and biogeochemical processes.

The objectives of the study were to answer the following questions: (i) What is the rice wax lipid composition, differentiated according to plant organs (leaf, stem, root) and growth stage? (ii) What is the epicuticular wax composition of intercrop plants (maize, sorghum, rape, mustard, beans and cotton)? (iii) What is the soil aliphatic hydrocarbon composition after different time intervals of paddy management (chronosequence of 50, 100, 300, 500, 700, 1000, 2000 yr)? (iv) What is the soil aliphatic hydrocarbon composition after different times of upland crop management (chronosequence of 50, 100, 300, 500, 700 yr)? (v) Does the management under at least an episodically anoxic paddy redox regime vs. aerobic upland soil management lead to different soil alkane pools in the two chronosequences? (vi) What is the

aliphatic hydrocarbon pattern in substrates of rice paddy soils (freshwater sediment, marine tidal flat sediment, desalinated marshland)? (vii) Can an input of crop residue burning and anthropogenic pollution be detected in paddy soils?

3.2 Material and methods

Study area

The study sites are in a land reclamation area in Hangzhou Bay near Cixi (30° 10′N, 121° 14′E), northeast of Zhejiang Province (Fig. 3.1). The bay is affected by river runoff and tide from the ECS. The YR, with an average water runoff of 925×10^9 m^3/yr and sediment load of 480×10^9 kg/yr supplies the dominant amount of sediment to the bay (Wang et al., 2008a, Li et al., 2009), where it is re-deposited by southward coastal currents and tides (Jilan and Kangshan, 1989; Xie et al., 2009). The modern climate is subtropical monsoonal with average temperature 16.3 °C and annual rainfall 1325 mm (Cheng et al., 2009). The coastal plain area varies in altitude from 2.6 to 5.7 m above sea level (Zhang et al., 2004) and is densely covered by rivers and lakes, as well as by urban and agriculture areas, with the main crops being wetland rice, rape, barley and cotton (Hua and Zhu, 2000). A special ecosystem property in the Cixi region is successive land reclamation via dyke construction on marine tidal flats since > 1000 yr ago, which has led to an artificial coastline relocated 18 km northeastwards (Feng and Bao, 2005). The sediment was supplied marginally from the Qiantang and Cao'e River and predominantly from the YR, the major contributor of terrigenous sediment to the ECS (Wang et al., 2008a, 2011), in particular before construction of dams along the river, which act as sediment traps. High sediment supply to the coastal areas of Cixi and the drop in sea level allowed intensive land reclamation. Dyke construction led to artificially initiated and accelerated pedogenesis and, depending on agriculture utilization, to the formation of either paddy or upland soils on the marine tidal flat substrate (primary YR suspended sediment).

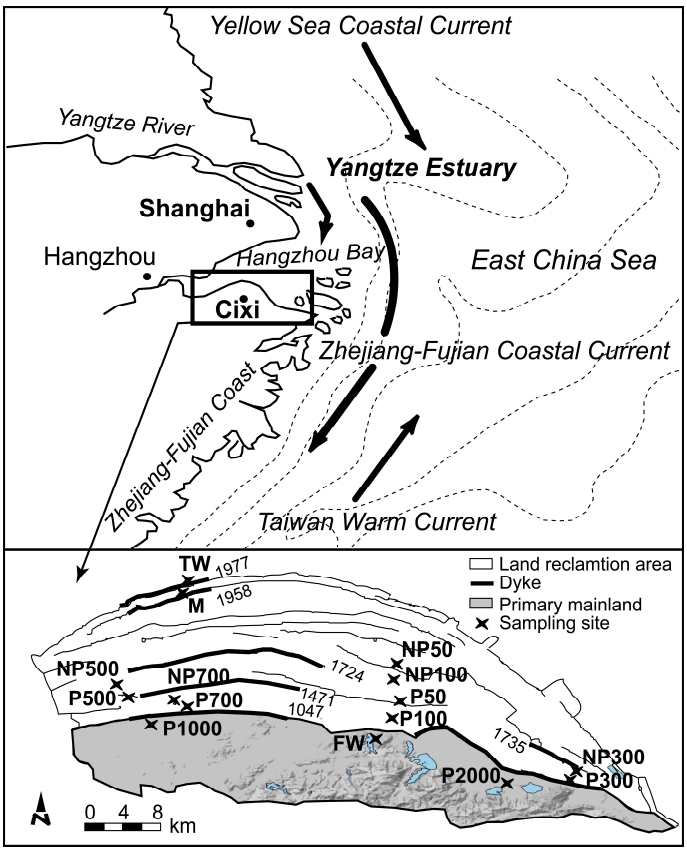


Fig. 3.1. Location of study area in the northeast Zhejiang Province. Dark lines in inset map depict position of protective dykes and numbers denote their year of construction. P50-P2000, sites with submerged paddy rice cultivation; NP50-NP700, sites with non-paddy upland use; TW and FW, potential soil substrates as tidal wetland sediment and freshwater sediment; M, for marshland soil/sediment.

Soil development started after land reclamation through the building of protective dykes on marine tidal mudflat sediments (continuous alluvial plain of Andong Beach). Based on the time of dyke construction and land use information from the Edit Committee of Chorography of Cixi County (1992), agricultural soils under continuous non-flooded upland use (50, 100, 300, 500 and 700 yr) as well as wetland rice cultivation (50, 100, 300, 700, 1000 and 2000 yr) were identified. Soil development from similar parent material has been supported by similar soil texture (Kölbl et al., 2014) and by radiocarbon analysis, indicating that soil formation evolved from homogeneous sediment. The ^{14}C activity of redeposited OM was around 50 pMC (percent modern carbon) in recent tidal wetland substrates as well as in all paddy subsoils > 70 cm (Bräuer et al., 2013).

Sampling

In June 2008 seven paddy (P50, P100, P300, P500, P700, P1000 and P2000) and five upland (non-paddy, NP50, NP100, NP300, NP500 and NP700) soils were sampled after harvest. The local cropping system on paddy fields is paddy-upland rotation, with one wetland rice season and one intercrop (vegetables or cereals) season per year (Cheng et al., 2009). In contrast to upland agricultural management, paddy management leads to formation of a puddled layer and a dense plough pan in the soil profile. Detailed soil description of all sampling sites according to FAO (2006) is available in Kölbl et al. (2014). In a previous study, it was demonstrated that repetitive puddling paddy management leads to a homogeneous distribution of particulate OM and soil lipids (Mueller-Niggemann et al., 2012) on an intra-field scale (macroscale intrinsic heterogeneity < 20% variation). This enables reliable differentiation of inter-field differences in biogeochemical and ecosystem properties based on only one representative profile at each location.

In order to identify the starting conditions in the chronosequence pedogenetic evolution, the potential soil substrates (day zero of soil development), a marine tidal wetland sediment (TW), a limnic freshwater sediment (FW) and a 30 yr old marshland soil (M) were sampled. To discriminate the main wax input source in soils, the dominant intercrop plants (maize, sorghum, rape, mustard, beans and cotton) in the Cixi area were collected during the field excursion and separated into plant tissues - roots, stems and leaves. Samples were air-dried, homogenized by grinding and sieved to < 2 mm prior to analysis.

Lipid analysis

Lipids were extracted from 10-12 g soil and 0.2-1 g plant samples using accelerated solvent extraction (ASE 200, Dionex) with CH_2Cl_2/MeOH (3:1 v/v) at 100 °C and 7×10^6 Pa for 20 min. The total lipid extract (TLE) was dried and separated into neutral and acid fractions via solid phase extraction (SPE) using silica gel impregnated with KOH in MeOH (10%, w/w). The fractions were obtained by successive elution with CH_2Cl_2 and CH_2Cl_2/HCO_2H (99:1 v/v). Neutral components were further separated into three fractions using small scale chromatography (Pasteur pipette filled with silica gel) and a solvent polarity gradient.

Aliphatic and aromatic hydrocarbons and polar compounds were eluted with *n*-hexane, *n*-hexane/CH_2Cl_2 (7:3 v/v) and MeOH, respectively. Elemental S was removed from the aliphatic fraction by addition of activated Cu (Blumer, 1957). For quantification of *n*-alkanes, a known amount of perdeuterated d_{50}-*n*-tetracosane was added as internal standard prior to separation.

The aliphatic hydrocarbons were analysed using gas chromatography-mass spectrometry (GC-MS) with an Agilent 7890A GC instrument equipped with a split/splitless injector and ZB-5HT fused silica column (30 m × 0.25 mm i.d., 0.25 μm film thickness; Phenomenex). The oven programme was: 70 ^{o}C (3 min) to 140 ^{o}C at 10 ^{o}C/min and then to 340 ^{o}C (held 13 min) at 3 ^{o}C/min. He was the carrier gas at a constant 1 ml/min. The GC instrument was coupled to an Agilent 5975B mass spectrometer operated in scan mode (*m/z* 50-650) with electron ionization (EI) at 70 eV. Assignment of individual compounds was based on mass spectra or mass chromatograms and comparison with retention times of standards. Quantification was conducted by way of normalization of peak areas vs. internal standard.

3.3 Results and discussion

n-Alkanes in crop plants

The *n*-alkane pattern in various crop plants with a differentiation according to tissues (leaf, stem, roots) indicated characteristic lipid composition within plants and dependence on growth stage (Fig. 3.2; Supplementary material, Table S3.1). As expected, in all plants long chain homologues (> C_{20}) with an odd/even predominance were detected. In crop plants the distributions for stems and roots differed from those of leaves and maximized at C_{29} except for maize leaves which were dominated by C_{31}, as observed previously (Wiesenberg et al., 2004). A substantial difference was seen in the distribution in maize leaves, which were enriched in C_{33}, whereas stems and roots lacked this component. Rape and mustard were particularly enriched in $C_{29,}$ as noted by Conte et al. (2003), whereas sorghum showed a proportionally higher content of C_{27} (Fig. 3.2, Table S3.1). In general, the distribution pattern of maize and sorghum was rather broad (Fig. 3.2), whereas that of rape was exceptionally narrow, with C_{29} contributing > 80% to the sum of wax *n*-alkanes.

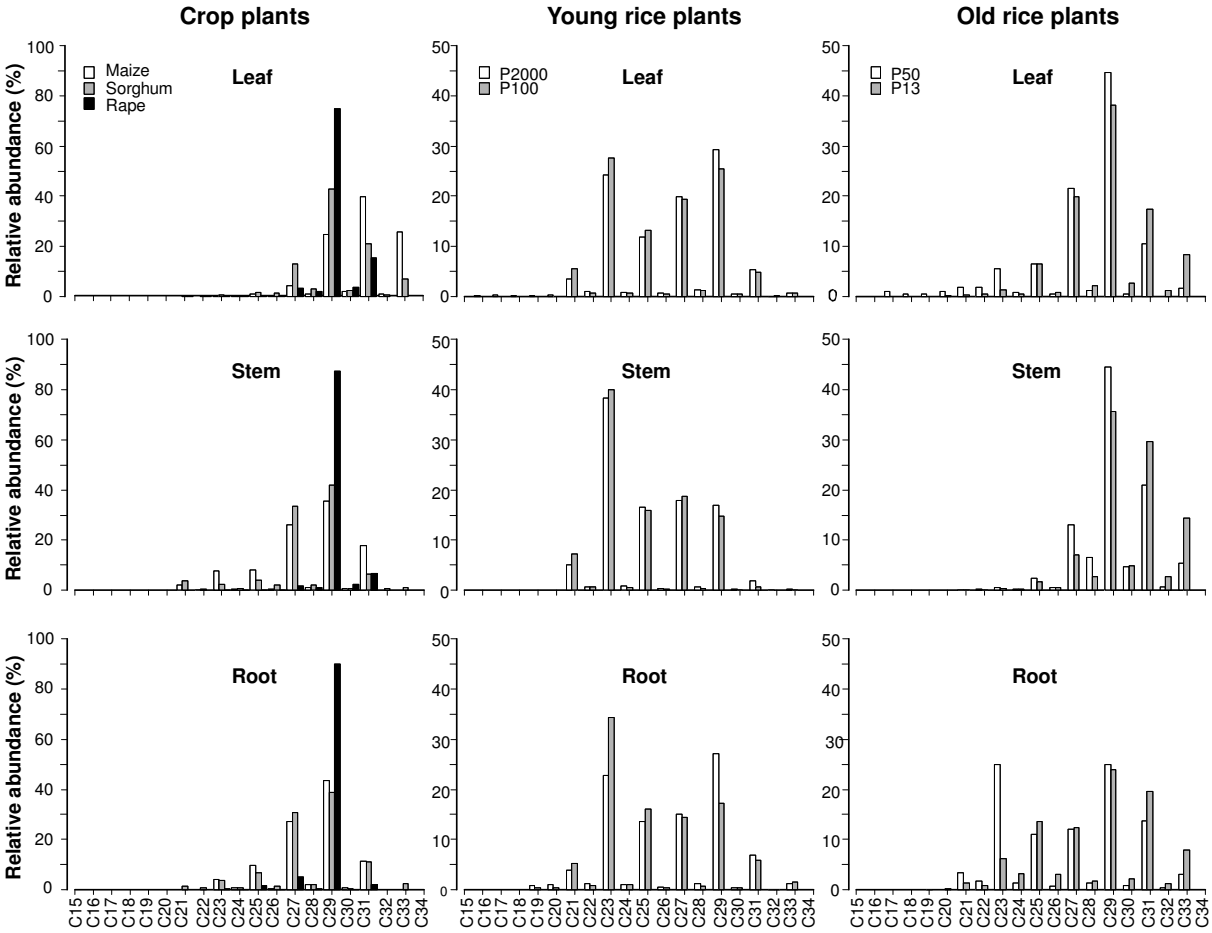


Fig. 3.2. *n*-Alkane distributions in upland crop plants (maize, sorghum, rape), young rice plants (several days old) and adult rice plants collected during the field excursion in Cixi. All plants were separated into leaf, stem and root tissues.

Wax *n*-alkane composition revealed substantial differences between intercrop and rice plants. Adult rice plants had similar *n*-alkane patterns for leaves and stems. Rice plants grow under variable water regimes and thus biosynthesize wax lipids that change over the growth period. This was reflected in *n*-alkane distributions of young vs. old rice plants (Fig. 3.2). During water flooded development stages, mid-chain (C_{23} to C_{25}) alkanes were prominent in all plant tissues, whereas in older plants growing under dry conditions a shift towards a higher abundance of C_{29} and C_{31} occurred (Fig. 3.2). Root alkanes of old rice differed from those of subaerial plant parts by showing proportionally more mid-chain (C_{23} and C_{25}) alkanes,

comparable with those of young rice plants. This can be attributed to paddies being kept at soil water saturation level until a day before harvest, leading to root alkanes mimicking the distribution of young rice growing immersed in water. Despite this high abundance of mid-chain alkanes, root alkanes of mature rice, as well as subaerial plant parts, revealed a higher proportion of C_{33}, a feature that differentiated rice from all other crops, except for maize leaves. The abundant C_{33}, in combination with a significant amount of C_{27} and even C_{25} made rice the crop plant with the broadest *n*-alkane distribution, a feature less pronounced in stems than in leaves or roots (Fig. 3.2, Table S3.1). In summary, rice wax alkanes were characterized by a broad distribution with abundant C_{33}, maximising at C_{29} and showing a shift towards mid-chain alkanes in young, water-immersed plants.

Lipids in topsoil

Only in topsoil samples did extractable lipid and *n*-alkane concentrations correlate significantly with SOC (Fig. S3.1, $r^2 = 0.92$; $p < 0.01$), serving as an indicator of fresh OM input to the upper soil horizon. The potential soil substrates showed much higher OC and extractable lipid or *n*-alkane concentration for the limnic sediment compared with the tidal flat sediment. The latter, derived mainly from erosion in the YR hinterland, followed by long distance riverine transport and re-deposition, is heavily degraded but a further decline in extractable and labile OM occurs after dyke installation in marshlands behind the dyke. The lipid concentration in the marsh sediment/soil must be regarded as a starting point for lipid accumulation during upland or paddy soil development. Before land reclamation along the Zhejiang coast, the shoreline in the area was more rugged and several embayments, coastal brackish lagoons and nearshore lakes existed (Feng and Bao, 2005). These coastal environments were incorporated into reclaimed land areas, with some of the former coastal lakes or lagoonal embayments then forming landlocked lakes in the reclaimed marshland. Soil substrates in such settings may have preserved parts of their original freshwater lipid signature.

Though extractable lipids are less stable than recalcitrant SOC (Wiesenberg et al., 2004; Marschner et al., 2008), substantially higher OM content in paddy soil than in other arable

land-use types agreed with comparative studies from the same chronosequence (Cheng et al., 2009) and in different landscapes of China (Wu, 2011). Compared with the TW substrate, higher amounts of OM and free lipids were noted in paddy topsoils, where SOC content (Table 3.1) tripled after only 50 yr cultivation. Oxygen limitation under submergence causes reducing conditions, inducing enhanced storage of soil OM, including lipids, in paddy topsoil (Sahrawat, 2005). In contrast to initial enrichment in paddies, the annual accumulation rate of lipids in soil under paddy cultivation for > 100 yr decreased to a rate < 1%/yr and in soil > 1000 yr approached steady state conditions, with a rate < 0.15%/yr (Table 3.1). In upland topsoil, enhanced SOC accumulation was not observed. Annual accumulation rate < 1%/yr within the first 50 yr and < 0.1%/yr in soil with > 300 yr cultivation was low and demonstrated that steady state conditions levelled out faster in permanently oxic than periodically anoxic soil (Table 3.1).

The concentration of *n*-alkanes (sum of C_{12} through C_{40}) was notably higher in topsoil than in subsoil (Table S3.2). In upland soil generally a low amount of extractable *n*-alkanes [0.7 to 1.2 µg/g dry wt. (dw)] with no apparent age trend was noted (Table S3.2). The *n*-alkanes, originating mainly from soil-introduced crop plant wax (Lichtfouse et al., 1998; Wiesenberg et al., 2004; Jansen et al., 2006; Rommerskirchen et al., 2006), exhibited a particularly rapid change within the first 50 yr of rice cultivation. Compared with the TW substrate, the twofold increase reflected a high preservation potential under these specific management-induced conditions and long term paddy soil use. The successive deposition and preservation of lipids led to an 8 × higher content of *n*-alkanes after 2000 yr paddy use (Fig. S3.1b).

Among the paddy soils, two locations deviated from the common trend by showing exceptionally low SOC and extractable organic concentration (Fig. S3.1, Table 3.1). The P1000 site, although developed as a paddy over the classified period, had received some very recent anthropogenic addition of non-paddy material, which was thoroughly mixed with the paddy soil (Mueller-Niggemann et al., 2012). This accounts for the low mass-normalized SOC, TLE and alkane yields. The second site, P500, though documented to have been cultivated as paddy field for several years before, must have been cropped in upland management practice for several decades, leading to reduced SOC and TLE accumulation.

This could be deduced from the lipid profiles as discussed below. Soil OM concentration at these two paddy sites equalled those of upland locations (Fig. S3.1), whereas all other results demonstrate a clear offset between paddy and upland managed fields. Similar observations have been reported for other compound classes (e.g. amino sugars) investigated at the same chronosequence (Roth et al., 2011; Wissing et al., 2011; Kölbl et al., 2014).

Table 3.1. Average concentration and annual accumulation rate for SOC, lipids and *n*-alkanes in topsoil.

	Average concentration in topsoil			**Accumulation rate[a] in topsoil**		
Site	**SOC [%]**	**Lipids [µg/g dw]**	***n*-Alkanes [ng/g dw]**	**SOC [%/yr]**	**Lipids [%/yr]**	***n*-Alkanes [%/yr]**
P50	1.79	608	1547	2.12	3.22	1.45
P100	1.77	931	2528	1.04	2.03	1.22
P300	2.51	1083	3664	0.46	0.72	0.53
P500	1.05	275	1028	0.10	0.16	0.06
P700	2.45	1330	9591	0.19	0.34	0.36
P1000	1.12	400	1358	0.06	0.12	0.06
P2000	3.64	1433	6846	0.09	0.12	0.11
NP50	0.93	215	743	0.80	1.10	-0.03
NP100	0.74	177	1191	0.16	0.35	0.46
NP300	1.17	363	1743	0.21	0.36	0.28
NP500	0.76	190	612	0.04	0.08	-0.04
NP700	0.77	324	1136	0.03	0.14	0.06
TW	0.61	639	846			
Marsh	0.63	125	753			
FW	0.98	615	5653			

[a] $[(X_t/X_0)^{(1/t)}-1] \times 100$, < 0.15%/yr for steady state conditions, where
X_t = mean concentration at cultivation time t (yr) and
X_0 = mean concentration in marsh soil developed from tidal wetland (t = 0).

Compared with all other topsoil samples, surface horizons of the 700 yr old paddy soil contained a substantially higher proportion of *n*-alkanes (Fig. S3.1b). This to some extent is attributed to fossil fuel contamination, which was confirmed by the presence of a pronounced unresolved complex mixture (UCM), as well as a high abundance of tricyclic and pentacyclic triterpanes (Fig. S3.2). These triterpenoids revealed a characteristic predominance of hopanes with a mature isomer distribution (Fig. S3.2a) dominating the unsaturated diploptene

triterpene indicative of the presence of recent bacteria (Rohmer et al., 1984). Minor amounts of contaminant hopanes were also present in other paddy soils (Fig. S3.2b) and can be attributed to ubiquitous potential anthropogenic sources, e.g. lubricants used in agricultural machinery, aerosol input or contaminated irrigation water (Peters et al, 2005).

In all topsoil samples, long chain *n*-alkanes with a unimodal distribution and an odd predominance prevailed. Alkanes in both paddy and upland topsoils were dominated by C_{29} and C_{31} (Fig. 3.3 and Fig. S3.3). In soil profiles the C_{29}/C_{31} ratio (Schwark et al., 2002; Jansen et al., 2006) alone was unsuitable for differentiating between paddy and upland development (Fig. 3.4a). Whereas subsoils and substrates exhibited similar C_{29}/C_{31} ratio values around 0.8, a severe shift to 1.5 was noted for topsoils. The marked difference in lipid composition between topsoils and subsoils in paddies argued for a high efficiency of the plough pan barrier, preventing downward translocation of hydrophobic alkyl lipids. Paddy soils possessed a higher amount of intermediate chain length alkanes, i.e. C_{23} and C_{25}. Soil *n*-alkane distributions in paddy vs. upland soils differed as shown in Fig. 3.4b, with an enhanced contribution of C_{25} and C_{33} vs. C_{29} and C_{31} under rice paddy management, a feature also noted in rice roots. The relative proportion of mid- to long chain homologues, also known as P_{aq} index (Ficken et al., 2000) was calculated using the following equation:

$$P_{aq} = \sum C_{23} + C_{25} / \sum C_{23} + C_{25} + C_{29} + C_{31} \quad (1)$$

This revealed a predominant input of emergent aquatic macrophytes (i.e., rice plants) in paddy soil (0.1-0.2) and of land plants in upland soil (< 0.1; Table S3.3).

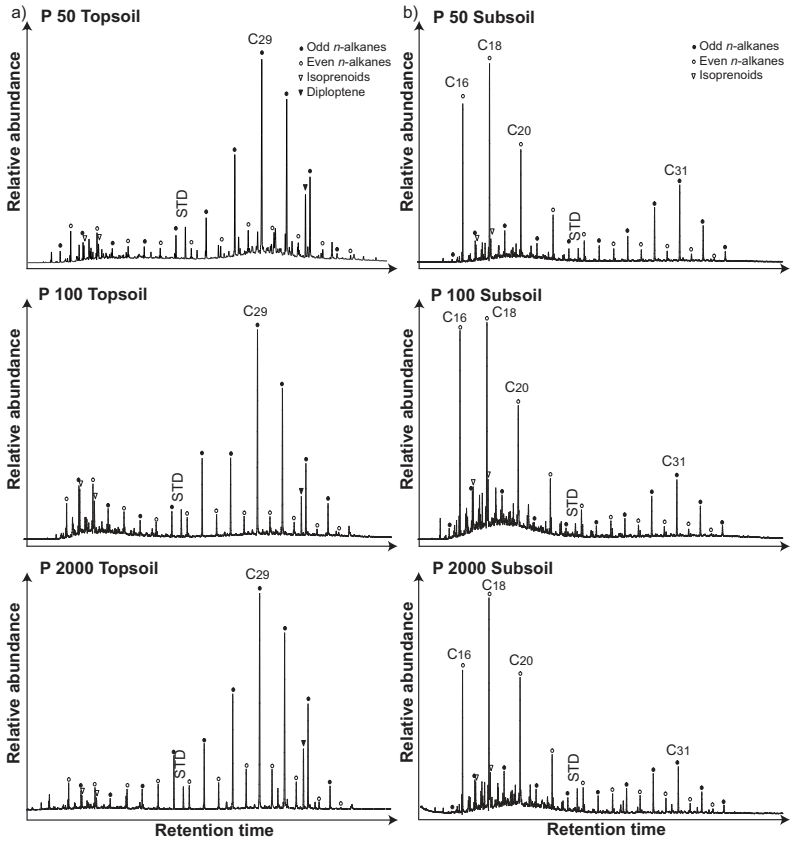


Fig. 3.3. GC-MS total ion chromatograms of aliphatic hydrocarbon fraction from several representative samples: **(a)** paddy, **(b)** paddy subsoil. STD, internal standard (d_{50}-*n*-tetracosane). Open triangles denote pristane and phytane.

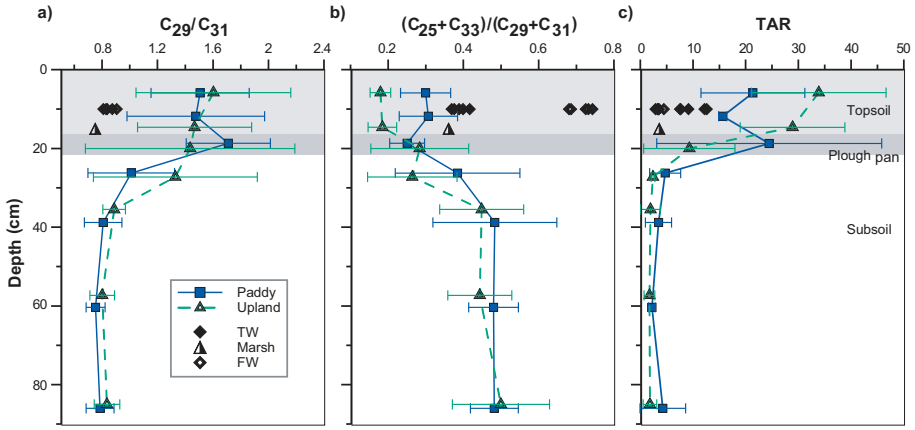


Fig. 3.4. Depth profiles of *n*-alkane ratios: **(a)** C_{29}/C_{31} discriminating topsoil from subsoil as well as potential soil substrates; **(b)** $(C_{25}+C_{33})/(C_{29}+C_{31})$ differentiates the management types (paddy and upland cultivation); **(c)** depth profile of *n*-alkane based proxy, the terrestrial to aquatic ratio (TAR) for the paddy and upland soil chronosequence. TAR = $\sum C_{27}+C_{29}+C_{31}/\sum C_{15}+C_{17}+C_{19}$. Mean values for 50, 100, 300, 500 and 700 yr old sites are shown (error bars = standard deviation).

Source of *n*-alkanes in topsoil – influence of agricultural management

The compositional variation in chemotaxonomically distinct *n*-alkanes in plant tissues and their preservation in topsoil allowed estimation of the primary lipid input to differently managed soils. The *n*-alkane distribution, as shown in a discrimination plot based on C_{25}, C_{27} and C_{31}, differed significantly between paddy and upland topsoils (Fig. 3.5a). In contrast to C_{31} enriched upland topsoil, the paddy surface horizons revealed a distribution similar to the soil substrates (FW, M, TW). The pattern in upland soil samples agreed with a dominant input of wax lipids originating from cereals such as maize and rape, as well as vegetables. The distribution in rice plants proved highly specific and useful for identification of rice derived OM in soil. The distribution in paddy soil matched the wax composition of adult rice plants (Fig. 3.5b). Young rice plants and seedlings grown immersed in water exhibited an *n*-alkane pattern shifted towards shorter chain length, being incompatible with the pattern in paddy soil (Fig. 3.5b). Among different rice plant tissues, subaerially-grown stem-derived wax alkanes were most similar to upland plant lipids (Fig. 3.5a and b), whereas root-derived alkanes, except for cotton, in general showed an enhanced C_{25} proportion of around 40% (Fig. 3.5b). Leaf alkanes did not exceed 20% in C_{25} proportion but revealed the broadest range in relative abundance of C_{31}. The mature rice plant patterns averaged from stem, leaf and root signatures fell within the field of paddy soil alkanes (Fig. 3.5a; $C_{25} > 15\%$, $C_{27} > 25\%$, $C_{31} < 60\%$). The much larger spread in land plant wax was not fully represented in upland soil samples due to significant mixing and averaging of wax alkanes upon crop rotation.

Discrimination plots based on C_{29}, C_{31} and C_{33} (Fig. 3.5c and d) separated the FW and TW substrates, with only marginal variation between TW and the desalinized marsh behind dyke site (Fig. 3.5c). The paddy soil sites exhibited a trend approaching the FW substrate end member, whereas the upland soils showed a higher affinity for the TW substrate (Fig. 3.5c). A systematic variation in *n*-alkane patterns according to the chronosequence evolution was not well expressed. Comparison of the soil and substrate $C_{29}/C_{31}/C_{33}$ pattern with the corresponding plant wax distribution showed that rice plants had an intermediate position in the paddy soil trend. The 100 yr and 500 yr samples and one 1000 yr old paddy soil sample gave a slightly higher proportion, all other samples having a lower proportion of C_{29}. The

decline in C_{29} was accompanied by a concurrent increase in C_{33}. The overall trend exhibited a dilution of the primary substrate signature by way of addition of fresh rice wax derived alkanes, shifting the distribution to higher relative C_{33}, rather stable C_{31} and highly variable C_{29} proportions. Though C_{29} was the most abundant *n*-alkane in rice plant wax (Fig. 3.5d), the differentiation between upland and paddy soils resulted primarily from the C_{33} proportion. The significant variability in the C_{29} proportion of paddy soils, ranging from ca. 40-60%, may derive largely from intercrops planted in rice fields during the dry season. All upland sites were characterized by a lower relative contribution of C_{33}, as also observed for the TW and marsh substrates. The shift away from the substrate *n*-alkane composition (Fig. 3.5c) for upland soils was caused by a lower contribution of the only C_{33}-rich upland crop (maize) and enhanced contributions of particularly C_{29}-rich rape and mustard wax to the upland soil *n*-alkane pool. The range in C_{29} was similar to that of paddy soil samples, ranging from ca. 35% to almost 60%. This crop-dependent range agreed with the paddy soil wax inventory and supported the conclusion of intercrop plant wax dominating the spread in C_{29} relative abundance in paddies. The low relative abundance of C_{33}, except for maize, in all the upland crop plants, irrespective of plant tissue type (Fig. 3.5d), was the main factor in discriminating upland crops from rice. The variability in C_{29} vs. C_{31} was governed by crop type, with beans and cotton being C_{31} enriched, without significant influence by tissue type for the latter (Fig. 3.5d).

The soil *n*-alkane inventory allowed us to identify the input of rice plants in arable soils because of the broad range in rice wax *n*-alkanes leading to an enhanced contribution of C_{25} and C_{33}, both bracketing the C_{27}, C_{29} and C_{31} alkanes, which were more abundant in upland crop plants. The $(C_{25}+C_{33})/(C_{29}+C_{31})$ ratio deviated in paddy and upland topsoils as a consequence of a predominant wax alkane input from rice vs. intercrop plants (Table S3.3). Intercrop plants had lower $(C_{25}+C_{33})/(C_{29}+C_{31})$ ratio values than, in particular, rice roots (Table S3.1), suggesting these to be a major contributor to the paddy soil alkane inventory. The broad distribution in rice wax alkanes is assumed to be caused by variable water availability affecting plants, in particular stems, during individual growth stages via evapotranspirative water loss. The typical rice wax alkane pattern or its absence can help in

determining past land use and in archaeo-botanical investigations, where it can be combined with other indicators, e.g. rice phytolith abundance (Cao et al., 2006; Itzstein-Davey et al., 2007).

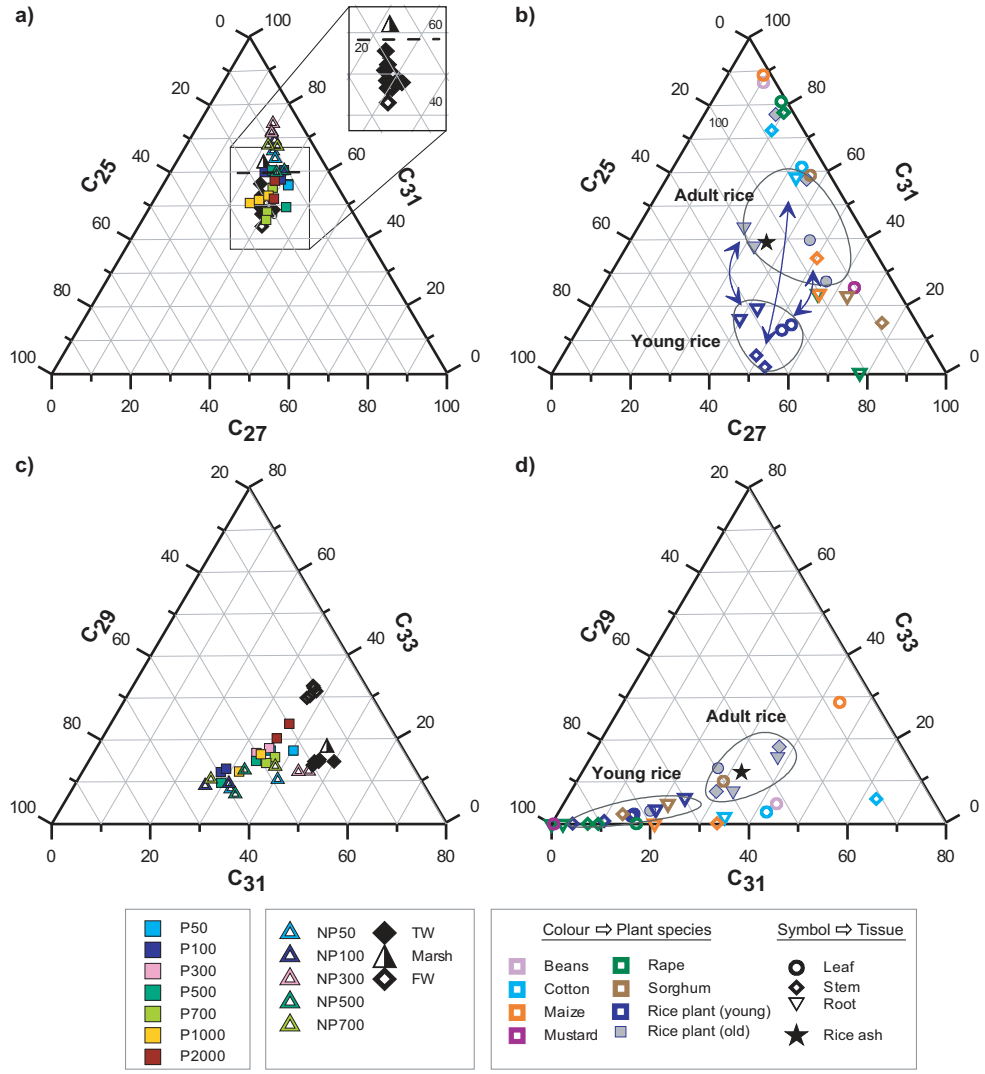


Fig. 3.5. Relative proportion of selected long chain *n*-alkanes in soil, reference sediment and crop plant: **(a)** paddy, upland and substrate soil (inset shows substrate only), **(b)** crop plant differentiated by tissue and growth stage, **(c)** paddy and upland soil vs. substrate, **(d)** crop wax alkanes differentiated by tissue and growth stage.

In general, increased lipid accumulation in surface horizons of rice paddy soils suggests an apparently larger input of plant-derived epicuticular wax via specific crop rotation management as well as slower decomposition of soil OM. The latter is favoured by longer periods under waterlogged, O_2 depleted conditions and more acidic pore water (Sahrawat, 2005; Kölbl et al., 2014). In addition, the *n*-alkane composition in surface horizons exhibited substantial differences according to agricultural management type. Molecular proxies such as average chain length (ACL) and carbon preference index (CPI) are used for estimation of biological sources and degree of lipid degradation (Bray and Evans, 1961; Eglinton and Hamilton, 1967; Poynter et al., 1989). ACL was calculated with **a** as the amount of compounds with **A** carbons and **n** number of alkanes as:

$$ACL = \sum(a \times A) \times n^{-1} \quad (2)$$

CPI was calculated for long chain *n*-alkanes as:

$$CPI_{long} = [(\sum C_{25}+C_{27}+C_{29}+C_{31}/\sum C_{24}+C_{26}+C_{28}+C_{30}) + (\sum C_{25}+C_{27}+C_{29}+C_{31}/\sum C_{26}+C_{28}+C_{30}+C_{32})]/2 \quad (3)$$

Soils and plant tissues contained long chain *n*-alkanes with highly variable ACL and CPI values (Fig. 3.6a-d, Table S3.3). In crop plants (rape, maize, cotton) CPI_{long} declined simultaneously with an increase of ACL_{long} (C_{26} to C_{35}) from roots and stems towards leaves (Fig. 3.6a). Young rice plant wax alkanes occurred in the same range as crop plants, but adult rice revealed higher ACL_{long} in addition to a lower CPI_{long} (Fig. 3.6a and b). The main *n*-alkane input to topsoil originated from plant residues such as root and stem tissues after harvesting. Therefore, compared with fresh plants a slight degradation in upland topsoils was noticeable (Fig. 3.6a and c) by way of higher ACL_{long} and lower CPI_{long} values. Paddy topsoil samples were in a similar range to adult rice plants except for the 2000 yr and 700 yr old soil, which exhibited a lower CPI_{long} (Fig. 3.6d). The deviation in the 700 yr old site is explained by fossil fuel contamination, adding *n*-alkanes with an approximate CPI_{long} of about unity. For the oldest soil from the 2000 yr old paddy, the initial degradation activity of wax lipids accumulated during the long term agricultural utilization is assumed to have slightly lowered the CPI_{long} value (Fig. 3.6d).

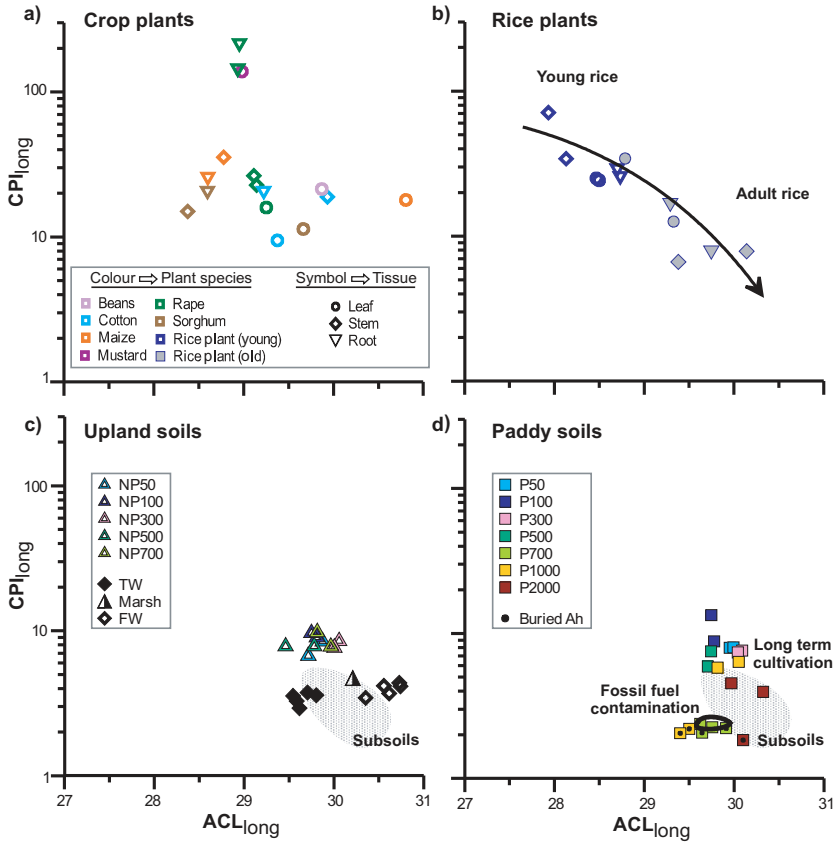


Fig. 3.6. Characteristics of average chain length (ACL) and carbon preference index (CPI) of long chain alkanes in several **(a)** crop plants, **(b)** rice plants, **(c)** upland topsoils, potential soil substrates, **(d)** paddy topsoils. Grey shaded background in (c) and (d) denotes subsoil.

n-Alkanes in subsoil – substrate characterization

The lowest concentrations of *n*-alkanes were in subsoils and ranged between 0.3 and 1.2 µg/g dw (Table S3.2). These values were comparable to those in the recent TW sediments. The proportion of *n*-alkanes varied from long chain dominated topsoils to a preference for short chain homologues in subsoils as a reflection of different OM origin. Most subsoil samples showed a bimodal distribution, with one mode in the short chain and one in the long chain homologues (Table S3.2, Fig. 3.3b). The proportion of long to short chain homologues expressed as terrigenous to aquatic ratio (TAR) can give an insight into the relative

contribution of land plant vs. algal/microbial sources to sediments (Meyers, 1997). Here a differentiation between topsoil and subsoil was possible, with values ranging between 13 and 60 in topsoil and < 8 in subsoil (Fig. 3.4c). The results indicate slightly higher aquatic/microbial OM content in subsoils than in modern TW reference sediments.

Higher abundances of short chain alkanes often appeared with an unresolved complex mixture (UCM) in the low molecular weight region (Fig. 3.3b). In addition, a specific characteristic of the distribution of short chain homologues was apparent for all subsoils. In contrast to topsoils, a strong even predominance with a maximum at C_{18}, followed by C_{16} and C_{20}, with a CPI_{short} (C_{14} to C_{22}) of 0.14 to 0.64, was observed. In buried A horizons the *n*-alkane patterns were equal to adjacent subsoils, except for the layers in 1000 yr old soil with a slightly higher abundance of C_{29}. In addition, regular isoprenoids like pristane and phytane were detected in generally uniform concentration. Algal/planktonic origins have been attributed to the presence of odd short chain *n*-alkanes (C_{15} and C_{17}) in sediments (Cranwell, 1982). In soils, the dominance of alkanes < C_{20}, coupled with no preference in odd or even chain homologues, is often associated with a bacterial contribution (Albro, 1976; Quénéa et al., 2006).

Previous studies described an analogous even alkane preference in suspended matter and sediments of the YR estuary (Qiu and Saliot, 1991; Wang et al., 2008a), pointing to a genetic relationship between these and the subsoil substrate. Xie et al. (2003) observed a bimodal distribution with maxima at C_{18} and C_{29} within paleosol profiles of the Xiushui River terraces near the catchment area of the YR in the middle-lower reaches. A study in central Queensland, Australia, suggests specific vegetation as a direct source of even short chain *n*-alkanes in soil (Kuhn et al., 2010). Other authors assumed that the preference of such homologues (maximizing at C_{18} or C_{16}) indicates thermal processes via breakdown of long chain *n*-alkanes or dehydration of alcohols upon biomass burning (Wiesenberg et al., 2009; Eckmeier and Wiesenberg, 2009). Laboratory combustion experiments established that a burning temperature > 350 °C leads to a shift towards a more balanced distribution of odd and even short chains (Wiesenberg et al., 2009; Knicker et al., 2013) and a higher temperature of around 500 °C initiates an increased production of even short chain *n*-alkanes (Wiesenberg et al., 2009). In top layers of sediment from the YR estuary an UCM, but no even predominance

of short chain alkanes, was observed (Bouloubassi et al., 2001). Alternative formation modes for short chain even *n*-alkanes, such as via pollution with fossil fuel-derived OM (Colombo et al., 1989) or bacterial byproducts during degradation of algal detritus (Venkatesan and Kaplan, 1982; Frysinger et al., 2003) have been postulated. Large area contamination, especially in soil deeper than 40 cm, with petrogenic or pyrogenic fossil fuels, was excluded due to low abundances of mature pentacyclic triterpenoids (hopanoids). Thus, combustion residues in former parent substrate come into consideration, which probably originate from wildfires or vegetation burning in agroecosystems in the YR hinterland area.

Incomplete combustion of rice straw after harvest is considered to be an important factor for adding OM to paddy topsoils, so a composite sample of rice straw ash and partially combusted rice stems was collected from heaps of burning rice straw in the Cixi area. These partially combusted biomass samples yielded a complex mixture of alkanes, alkenes and a highly branched isoprenoid (HBI) monoene (Fig. 3.7). At least under the conditions prevailing here, the above potentially combustion-generated short chain even *n*-alkanes were not detected. We postulate that the *n*-alkanes in the combustion residue comprised two fractions, one consisting of residual wax alkanes that survived partial combustion and still showed CPI values > 1, vs. another fraction represented by shorter chain *n*-alkanes with CPI close to unity. In addition to *n*-alkanes, short chain *n*-alk-1-enes were detected and interpreted as having been neo-formed during the incomplete combustion, e.g. via dehydration of *n*-alcohols (Wiesenberg et al., 2009). As no *n*-alkenes were detected in paddy soils, the lack of these combustion residues was attributed to either an insignificant overall input or more likely to rapid reduction of alkenes to alkanes under paddy soil conditions. Hence, *n*-alkenes in paddy soils are not suitable to serve as combustion indicators.

The presence and distribution of long chain *n*-alkanes in subsoil indicated a former wax lipid input of higher vascular land plants to the parent substrate (Eglinton et al., 1962; Kolattukudy et al., 1976). The distribution of the long chain *n*-alkanes was similar to modern reference sediments, with a low odd predominance (CPI_{long} 1.7 to 3.8) and a maximum at C_{31}. The *n*-alkane pattern exhibited a mixed composition of terrigenous (YR drainage area) as well as aquatic macrophyte sources, supported by P_{aq} values varying between 0.16 and 0.41 (Table

S3.3). A differentiation of preferentially marine tidal flat and lacustrine nearshore lake/lagoon depositional setting could be deduced from specific *n*-alkane ratios, clearly separating these potential soil substrates (Fig. 3.8). Subsoil horizons seemed marginally affected by more lacustrine conditions, visible in *n*-alkane patterns (Fig. 3.8) possessing higher relative amounts of intermediate homologues, potentially derived from aquatic macrophytes (Ficken et al., 1998). Additionally, FW sediments archive local input from surrounding land vegetation and may therefore be enriched in terrigenous C_{33} alkane. The latter has also been reported to derive from emerging aquatic plants, including *Carex* spp. or floating aquatic macrophytes, including *Potamogeton* spp. (Street et al., 2013). Differences in *n*-alkane pattern seen in Fig. 3.8 indicated that the deepest soil horizons of 2000 yr and 300 yr sites had a greater affinity for the FW substrate, whereas the soils of 500 yr and 1000 yr site more likely developed on TW substrates.

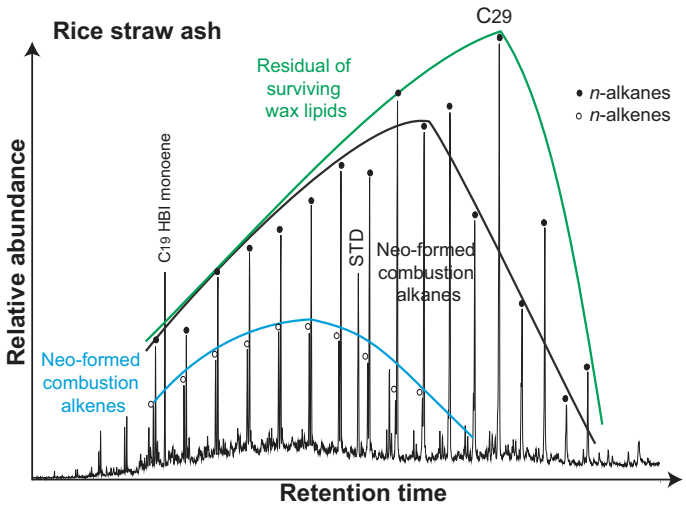


Fig. 3.7. Gas chromatogram of extract of partially combusted biomass from smouldering heaps of rice straw showing a complex mixture of residual wax alkanes accompanied by neo-formed *n*-alkanes of shorter chain length and neo-formed *n*-alkenes. The HBI monoene is derived from epiphytic diatoms. Note lack of isoprenoid alkanes or even-numbered short-chain alkanes, proposed to be formed via combustion, were observed.

The highly branched isoprenoid (HBI) derived from epiphytic diatoms on *Chara sp.* (Jaffé et al., 2001), a common weed in rice paddies (Ariosa et al., 2004). Here, the C_{20} HBI was present in the FW substrate (Fig. S3.4) but, in contrast to previous observations in fossil Chinese paddy soils (Atahan et al., 2007), could not be detected and hence had no power in differentiating upland from paddy soil.

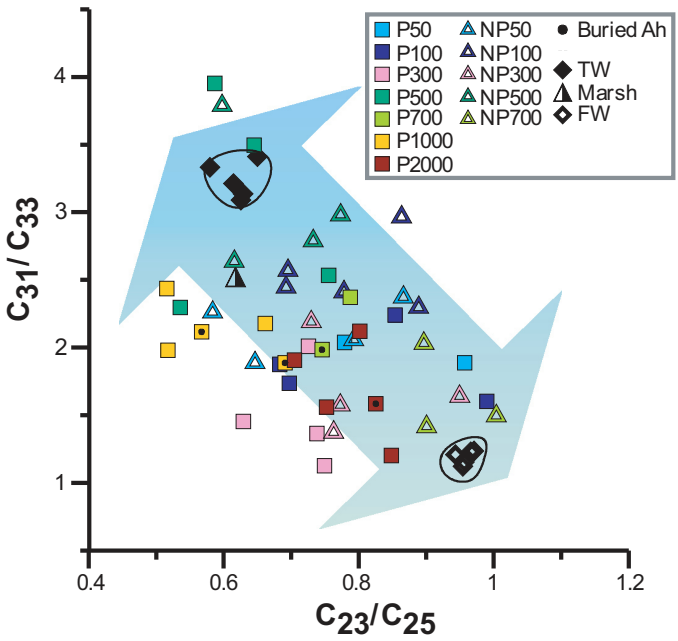


Fig. 3.8. Substrate discrimination plot based on *n*-alkane ratios in different subsoils and reference sediments serving as potential end members (circled).

In general, the aliphatic hydrocarbon patterns revealed a soil development from substrates with only a moderate variability, in agreement with rather invariable ^{14}C signatures (Bräuer et al., 2013). The lipid composition of the sediments of the YR estuary (Bouloubassi et al., 2001; Jeng and Huh, 2004; Wang et al., 2008b) agrees with the data from Cixi TW, marsh and subsoils studied here. Soil substrates contained a mixed composition of OM, including terrigenous compounds from YR suspended sediment and marine derived sources from the ECS.

3.4 Conclusions

This is the first study identifying the chemotaxonomic potential of *n*-alkanes in rice, the most important global staple food and the incorporation of these components into soil under paddy compared with upland management forms. The *n*-alkane composition of various crop plants (rice, maize, sorghum, rape, mustard, beans and cotton) differed according to plant tissue (leaf, stem, root) and growth stage. A distinction between rice plant and intercrop plant wax components was feasible, based on the prominent presence of C_{33} alkanes in rice. The *n*-alkane composition of rice plant tissues changed during growth and was adapted to the water

level on flooded fields. Therefore, in submerged plant tissue (young rice, roots) mid-chain (C_{23} to C_{25}) alkanes dominated vs. adult plants whose plant organs (leaf and stem) grew subaerially.

The *n*-alkanes in agricultural topsoil reflected the input of plant OM remaining in the field after harvesting, except for the 700 yr paddy soil, which was contaminated with fossil fuel. The soil *n*-alkane distributions in paddy and upland soil differed via enhanced contributions of C_{25} and C_{33} under rice paddy management as also noticed for e.g. rice root alkanes. In upland topsoil the alkane distribution patterns were in agreement with an input from other crop plants.

Alternate flooding of fields is a part of rice paddy management practices, leading to specific redox conditions, which result in slower decomposition of OM. In comparison with aerobic upland soils, higher amounts of lipids and aliphatic compounds, coupled with a successive enrichment until 2000 yr of cultivation, were observed in paddy topsoil. Simultaneously, the annual rate of *n*-alkane accumulation decreased slightly until soils reached 1000 yr in age. From then on, they approached steady state conditions. In upland soils this trend had already became apparent after only 300 yr of cultivation.

Lipid compositional changes from topsoil to subsoil reflected different botanical origins as well as advanced pedogenesis in tidal wetland sediment after land reclamation and associated development into lacustrine/limnic sediment, marsh soil and finally arable soil. The most intense changes in lipid composition were observed in topsoil (< 20 cm) after only 50 yr of agricultural use. Subsoil did not exhibit notable differences in lipid composition between sites. This suggests preservation of substrate lipid signatures derived from terrigenous substances, which may have contributed by YR suspended sediment and/or marine compounds from the ECS.

3.5 Acknowledgements

We acknowledge the German Research Foundation (DFG) for financial support (Schw554/20). Chinese and German partners of Research Initiative FOR 995 are thanked for field work collaboration. We appreciate analytical assistance by laboratory staff at Cologne University. The three anonymous reviewers are thanked for constructive comments.

3.6 Appendix A. Supplementary data

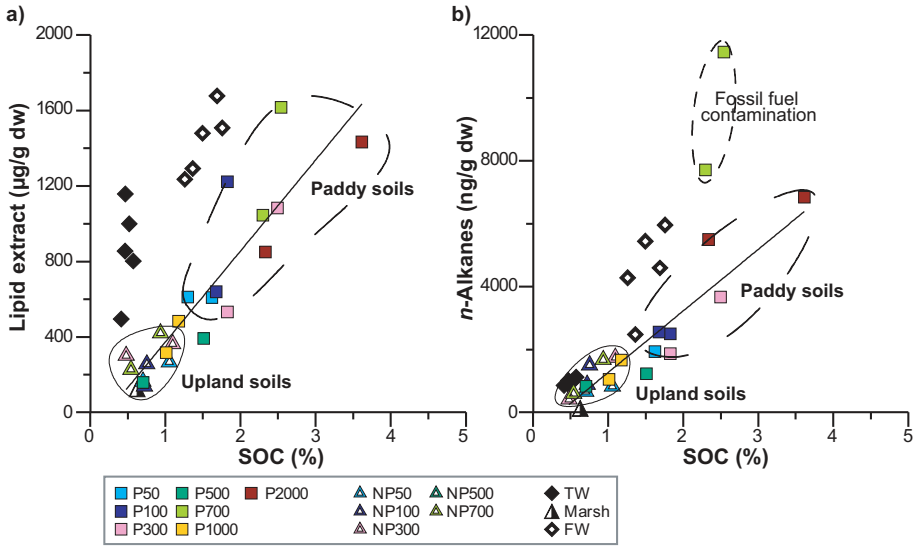


Fig. S 3.1. Scatter plots showing relationship between SOC content and concentration (a) lipids and (b) *n*-alkanes in topsoils and reference sediment. P, paddy soil; NP, upland soil; TW, tidal wetland sediment; FW, limnic freshwater sediment; marsh indicates desalinized marshland behind dyke, not yet in agricultural use.

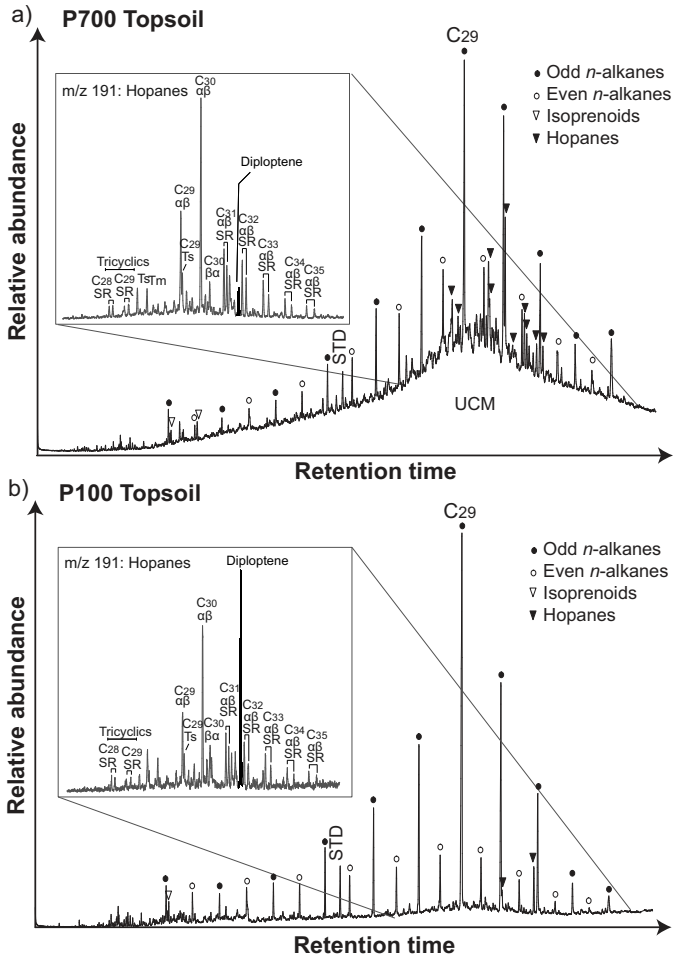


Fig. S 3.2. GC-MS total ion traces from aliphatic hydrocarbon fraction, with major peaks labelled. Inset shows *m/z* 191 chromatogram, indicative for tri- and pentacyclic triterpenoids. Diploptene marked black is indicative of recent bacteria; hopanes and tricyclic hydrocarbons in grey derive from fossil fuel contamination. Peaks are labelled according to number of carbon per molecule and isomerisation at C_{17}, C_{21} and C_{22}. Ts, trisnorhopane; Tm, trisnorneohopane. Note high abundance of fossil fuel hopanes vs. diploptene in P700.

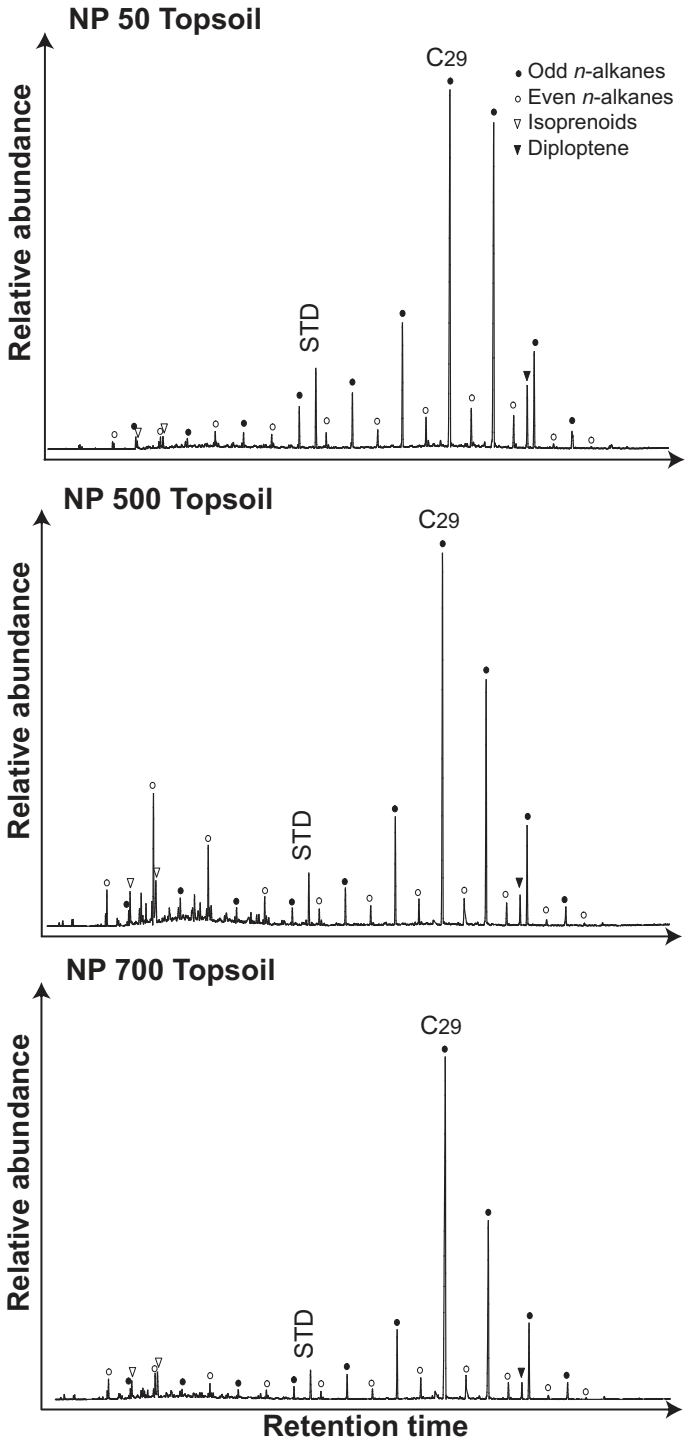


Fig. S 3.3. GC-MS total ion chromatograms of aliphatic hydrocarbon fraction from several representative samples: upland topsoil. STD, internal standard (*d50-n*-tetracosane). Open triangles denote pristane and phytane.

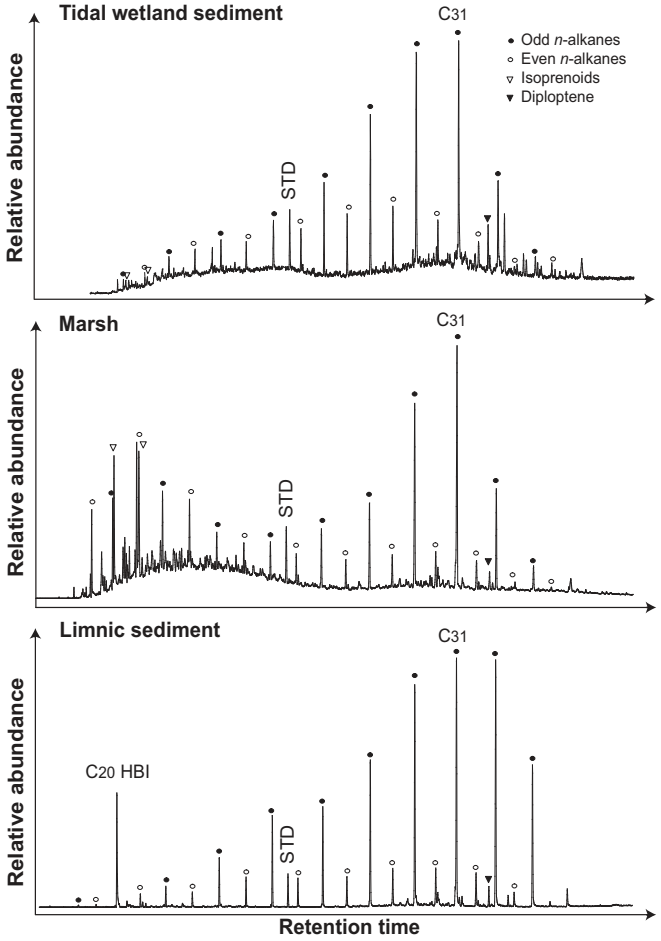


Fig. S 3.4. Total ion chromatograms of aliphatic hydrocarbons from parent substrate. STD, internal standard (*d50-n*-tetracosane).

Table S 3.1. Relative abundance of leaf, stem and root n-alkanes in upland crop and rice plants.

			Relative n-alkane abundance (%)														
Plant	**Species name**	**Tissue**	**C_{20}**	**C_{21}**	**C_{22}**	**C_{23}**	**C_{24}**	**C_{25}**	**C_{26}**	**C_{27}**	**C_{28}**	**C_{29}**	**C_{30}**	**C_{31}**	**C_{32}**	**C_{33}**	**$(C_{25}+C_{33})/(C_{29}+C_{31})$**
Beans	*Phaseolus vulgaris*	Stem	0.0	0.0	0.0	0.2	0.1	1.3	0.4	4.6	1.7	46.4	1.8	38.5	0.7	4.1	0.06
Cotton	*Gossypium*	Leaf	0.0	0.0	0.1	0.6	0.4	2.9	0.9	15.9	4.0	38.9	3.4	29.7	1.5	1.9	0.07
		Stem	0.0	0.0	0.0	0.3	0.1	5.3	0.9	13.0	1.7	23.8	1.5	47.7	1.4	4.3	0.13
		Root	0.0	0.0	0.2	1.0	1.0	3.9	1.8	14.5	1.5	48.4	0.6	25.8	0.0	1.2	0.07
Maize	*Zea mays*	Leaf		0.0	0.0	0.2	0.1	0.9	0.3	4.2	1.0	24.6	2.1	39.8	1.1	25.9	0.42
		Stem		2.1	0.0	7.6	0.4	8.2	0.4	26.1	1.2	35.4	0.7	17.8	0.0	0.0	0.15
		Root		0.0	0.0	4.1	0.8	9.9	0.4	27.1	2.0	43.4	0.7	11.4	0.0	0.0	0.18
Mustard	*Sinapsis arvensis*	Leaf	0.0	0.0	0.0	0.0	0.0	0.2	0.2	1.0	0.4	97.7	0.1	0.4	0.0	0.0	0
Rape	*Brassica napus*	Leaf	0.0	0.0	0.0	0.0	0.0	0.3	0.2	3.4	2.1	74.9	3.6	15.5	0.0	0.0	0
		Stem	0.0	0.0	0.0	0.0	0.0	0.2	0.0	1.8	1.1	87.4	2.4	6.9	0.0	0.0	0
		Root	0.0	0.0	0.0	0.3	0.0	1.8	0.2	4.9	0.4	90.2	0.1	2.1	0.0	0.0	0.02
Sorghum	*Sorghum sp.*	Leaf	0.1	0.2	0.2	0.8	0.5	1.8	1.3	13.0	2.9	42.9	2.2	21.2	0.8	7.1	0.14
		Stem	0.0	3.8	0.5	2.4	0.6	3.9	2.2	33.5	2.2	41.8	0.7	6.5	0.7	1.1	0.1
		Root	0.0	1.5	0.7	3.6	0.6	6.6	1.3	30.8	2.1	38.8	0.5	11.1	0.0	2.5	0.18
Rice (young)	*Oryza sativa*	Leaf	0.4	3.5	1.0	24.2	0.8	11.9	0.6	19.9	1.3	29.2	0.5	5.3	0.0	0.6	0.36
		Leaf	0.0	5.5	0.7	27.6	0.6	13.2	0.4	19.3	1.2	25.4	0.6	4.8	0.1	0.7	0.46
		Stem	0.1	5.1	0.7	38.3	0.9	16.6	0.4	17.9	0.7	16.8	0.2	1.9	0.0	0.1	0.89
		Stem	0.2	7.2	0.7	40.0	0.5	15.9	0.2	18.7	0.3	14.8	0.0	0.6	0.0	0.0	1.03
		Root	1.0	3.8	1.1	22.7	1.0	13.6	0.5	15.0	1.2	27.2	0.3	6.8	0.0	1.2	0.43
		Root	0.4	5.2	0.9	34.4	1.0	16.1	0.4	14.4	0.7	17.3	0.3	5.9	0.1	1.5	0.76
Rice (adult)	*Oryza sativa*	Leaf	1.0	1.8	1.9	5.5	0.8	6.5	0.4	21.5	1.1	44.7	0.5	10.5	0.0	1.7	0.15
		Leaf	0.2	0.3	0.5	1.3	0.5	6.5	0.8	20.0	2.2	38.1	2.6	17.3	1.2	8.4	0.27
		Stem		0.0	0.0	0.4	0.1	1.8	0.5	7.0	2.7	35.7	4.9	29.6	2.6	14.5	0.25
		Stem	0.0	0.1	0.2	0.5	0.2	2.4	0.5	13.0	6.5	44.6	4.6	21.0	0.8	5.4	0.12
		Root	0.2	1.4	0.8	6.2	3.2	13.6	3.1	12.4	1.7	23.9	2.2	19.6	1.2	7.9	0.49
		Root	0.0	3.3	1.6	24.9	1.3	11.1	0.7	12.1	1.4	24.9	0.9	13.8	0.4	3.0	0.36
Rice ash			4.8	4.6	5.6	5.7	5.7	6.4	4.7	8.6	4.4	16.4	4.5	9.6	1.3	3.6	0.39

Table S 3.2. Concentration of n-alkanes (C_{13}–C_{33}) in paddy soil and upland soil horizons and potential soil substrates.

Site	Depth	*n*-Alkane abundance (ng/g dw)																					Sum *n*-alkanes	
	(cm)	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30	C31	C32	C33	(µg/g dw)	(µg/g SOC)
P50	0-7	0.0	21.1	9.0	54.8	31.1	53.2	18.6	24.1	22.3	24.2	47.8	22.0	83.6	27.1	218.7	47.5	511.1	50.0	387.6	31.2	188.5	1.9	119.1
	7-14	0.0	0.8	0.0	3.4	13.9	15.0	8.9	12.5	23.2	13.6	39.8	13.1	58.4	18.1	153.6	29.9	281.2	29.5	267.9	21.8	114.9	1.2	89.4
	14-23	0.0	6.0	3.2	20.6	22.5	32.2	16.2	14.7	12.8	13.0	27.8	13.4	43.9	16.0	97.4	26.3	388.4	24.5	192.3	16.5	87.0	1.1	337.6
	23-38	0.0	6.3	4.6	259.1	29.6	338.8	54.8	214.2	45.8	92.0	29.6	40.9	22.0	21.9	29.2	15.9	65.8	14.9	91.2	10.6	70.8	1.5	458.1
	38-50	0.0	3.0	1.9	168.5	21.7	282.9	47.0	187.3	44.2	86.3	27.9	39.3	16.0	20.0	18.2	12.5	37.4	9.9	51.4	6.9	31.7	1.1	364.6
	50-70	0.0	6.4	4.5	201.1	25.8	303.3	40.4	199.3	31.8	82.0	22.5	36.8	23.5	21.2	32.5	16.7	65.5	13.2	88.1	10.8	46.7	1.3	416.6
	70-100	0.0	5.0	2.8	172.8	24.7	258.3	35.5	149.1	19.7	60.9	19.1	30.2	24.6	20.8	37.5	18.2	76.9	17.0	106.3	12.3	52.1	1.2	322.7
P100	0-9	0.0	0.0	0.0	0.0	6.4	6.1	5.1	11.7	19.3	19.6	52.4	17.7	93.9	24.2	243.7	47.9	1090.4	49.6	514.7	29.6	224.0	2.5	136.6
	9-15	0.0	14.0	7.4	47.8	52.4	74.9	37.6	34.2	29.8	30.2	64.6	31.2	102.2	37.3	226.5	61.1	903.3	56.9	447.2	38.4	202.3	2.6	152.6
	15-21	0.0	1.2	0.9	37.8	30.0	76.8	54.1	63.9	47.4	40.4	39.5	26.2	43.3	20.9	64.4	23.6	241.2	19.5	160.2	12.6	62.0	1.1	156.7
	21-30	0.0	9.9	18.8	64.0	100.6	125.5	81.4	62.1	34.7	24.3	22.6	19.9	26.5	16.9	34.1	16.2	96.9	16.1	87.8	12.1	39.2	0.9	150.5
	30-50	1.4	2.8	3.8	40.2	41.2	98.0	69.7	54.2	30.1	27.9	21.0	17.4	21.3	17.2	25.5	13.0	46.7	12.9	60.6	9.2	37.8	0.7	216.4
	50-75	0.0	17.1	18.3	331.3	64.4	415.3	67.4	216.5	29.3	75.2	25.7	37.5	37.6	28.1	50.9	25.2	98.7	22.2	140.6	16.3	75.0	1.8	550.7
	75-100	0.0	22.6	14.7	255.9	43.2	326.1	43.3	202.2	12.5	77.0	13.3	36.9	19.1	27.8	35.8	22.0	56.3	14.5	79.2	12.0	45.6	1.4	427.6
P300	0-18	0.0	2.8	7.9	27.5	37.3	49.9	18.9	30.8	42.4	50.0	126.2	53.1	169.5	53.3	337.6	90.0	1065.4	105.2	796.4	71.5	404.9	3.7	146.9
	18-24	0.0	2.9	1.7	5.6	9.7	12.6	7.5	16.0	25.6	27.2	65.3	28.5	91.1	31.6	180.0	48.3	574.3	53.0	379.4	36.7	191.6	1.9	102.2
	24-30	0.0	1.2	0.0	8.0	17.6	41.4	24.3	17.2	9.0	10.0	13.2	11.5	18.3	10.9	30.0	13.1	63.9	13.7	79.2	12.3	39.4	0.5	101.6
	30-50	0.0	1.5	1.6	100.3	16.0	200.5	14.1	127.7	8.5	54.3	8.0	24.9	10.6	14.7	14.6	10.5	26.1	9.4	38.3	8.1	34.0	0.7	233.4
	50-70	0.0	0.0	0.1	11.4	15.1	44.3	20.5	15.7	7.7	6.9	8.6	8.2	11.7	8.5	21.2	9.5	33.2	8.9	46.4	7.1	34.0	0.3	92.6
	70-100	0.0	5.0	4.3	176.2	31.7	270.7	34.3	152.3	11.4	62.5	9.5	30.5	15.1	19.0	24.6	14.5	39.9	11.4	54.3	8.8	37.4	1.0	340.3
P500	0-15	5.0	13.1	18.7	63.5	32.9	43.5	9.6	19.7	19.4	18.5	42.1	19.5	64.8	20.8	140.6	29.8	302.1	26.0	200.6	15.5	87.6	1.2	81.1
	15-19	11.0	13.2	2.9	37.3	8.8	45.0	4.8	14.9	9.4	14.6	17.8	19.6	30.4	17.2	55.5	22.8	268.9	24.5	130.4	15.6	42.6	0.8	117.6
	19-25	4.9	7.4	5.1	40.2	10.3	47.6	8.7	23.8	5.9	11.3	8.2	9.8	14.0	8.3	21.3	8.6	98.5	6.6	69.9	4.0	17.7	0.4	92.9
	25-48	4.8	5.8	2.8	38.2	10.4	57.5	8.8	30.1	5.5	13.7	8.3	10.3	12.9	7.9	18.1	7.6	45.2	5.3	54.1	3.0	15.5	0.4	108.0
	48-75	7.1	10.0	4.1	79.4	21.0	134.4	37.5	72.3	14.0	26.3	11.4	17.8	15.2	12.0	17.8	9.5	33.8	6.7	39.6	3.4	15.6	0.6	190.2
	75-100	5.5	3.9	1.2	8.1	2.8	16.3	4.1	9.2	4.1	7.2	7.6	7.8	14.1	7.6	17.6	6.9	35.0	4.3	43.3	1.7	18.8	0.2	112.2
P700	0-10	10.3	35.5	35.9	87.7	113.1	95.7	77.3	98.6	130.9	167.9	311.6	430.6	759.0	682.6	1079.7	727.0	2120.0	656.3	1689.8	444.1	709.7	11.5	451.4
	10-16	25.3	38.1	31.0	67.9	97.1	111.9	88.8	91.3	104.0	133.5	255.2	297.6	539.2	451.5	738.7	473.2	1458.1	392.9	1072.5	263.1	422.5	7.7	336.0
	16-22	14.3	15.2	11.5	28.1	15.4	35.3	16.9	28.5	27.2	29.6	38.5	30.9	58.4	34.0	100.4	42.4	284.1	37.3	195.8	24.2	78.6	1.2	127.8
	22-45	6.8	7.4	5.3	34.3	8.3	37.7	9.1	17.9	12.3	16.0	14.2	14.2	18.1	12.4	27.2	12.0	64.2	10.4	62.4	7.2	26.3	0.4	115.3
	45-69	5.7	8.4	9.0	57.1	15.4	52.7	12.2	23.3	17.2	22.2	22.7	22.7	30.4	25.0	41.3	26.1	59.3	22.3	57.1	13.5	28.8	0.6	100.5
	69-106	3.9	12.7	10.5	83.0	17.6	74.5	35.6	66.3	59.9	57.5	46.8	38.3	38.5	26.5	40.7	22.3	59.2	17.6	62.2	9.7	26.6	0.8	320.9
P1000	0-10	0.0	1.5	8.2	55.0	69.5	69.4	38.2	37.6	28.1	28.1	46.5	35.8	125.3	39.8	126.9	34.3	374.7	35.2	259.1	24.7	124.6	1.7	141.4
	10-16	4.7	10.3	11.0	34.3	41.1	53.6	29.5	30.5	21.8	21.1	31.2	22.5	62.0	26.9	76.1	26.1	258.7	22.9	147.1	13.1	57.2	1.1	103.9
	16-21	0.9	3.0	2.9	33.5	9.1	43.4	9.8	17.4	12.6	16.8	20.2	18.4	28.0	19.2	42.5	18.5	115.6	17.8	79.1	11.8	36.6	0.6	120.3
	21-40	2.0	3.1	5.2	37.3	13.3	61.7	11.7	24.4	15.3	23.4	28.0	28.4	40.5	33.3	55.3	27.4	71.4	23.9	64.2	14.3	34.0	0.6	134.0
	40-55	3.4	6.7	6.4	12.4	5.2	19.0	5.6	11.8	8.3	15.1	15.3	15.8	23.1	18.2	35.4	18.1	59.7	15.1	66.7	9.0	30.6	0.4	155.6
	55-80	5.3	9.6	7.8	41.7	14.4	54.2	12.1	19.4	13.1	20.2	19.9	23.9	38.5	23.7	63.7	26.0	94.6	19.6	98.2	11.1	40.3	0.7	186.4
	80-93	4.8	11.3	5.6	47.1	11.1	48.3	8.4	17.5	15.3	23.2	27.7	32.2	48.9	39.4	70.0	40.9	83.9	31.8	74.4	15.5	35.1	0.7	51.4
	93-100	6.7	11.3	3.4	16.4	14.2	49.3	20.7	30.7	12.0	17.7	12.5	16.8	24.2	18.0	29.2	19.3	41.7	18.4	48.1	10.0	24.3	0.5	154.6
P2000	0-15	45.1	82.5	28.7	135.9	95.1	119.0	60.0	78.3	122.2	146.8	318.2	149.7	403.9	167.3	688.7	241.2	1471.3	257.4	1180.0	173.2	672.2	6.8	189.4
	15-20	0.0	14.3	5.2	14.9	15.1	29.7	23.8	38.4	69.3	96.7	215.2	109.8	268.4	129.9	501.8	201.7	1135.3	246.9	1031.6	223.6	672.6	5.5	235.7
	20-27	5.5	26.7	15.2	248.7	67.6	329.3	46.0	160.2	23.6	69.8	30.8	41.5	40.9	34.9	67.4	41.0	122.0	45.2	148.0	35.3	94.8	1.7	354.9
	27-35	7.0	15.5	9.4	142.0	26.6	182.7	34.3	119.1	27.7	56.4	27.2	34.6	32.9	31.0	47.6	32.2	71.9	30.6	71.5	22.7	45.1	1.1	327.2
	35-50	6.3	14.4	9.9	88.8	26.1	103.6	44.9	43.8	21.8	17.7	17.5	18.1	24.8	19.1	33.2	21.6	48.9	18.5	51.9	12.5	27.2	0.7	303.6
	50-70	2.7	4.2	5.5	90.2	19.3	165.7	26.4	99.1	14.3	43.0	11.9	19.1	14.9	17.6	20.7	14.4	31.1	13.8	37.5	7.8	17.7	0.7	478.0
	70-100	5.4	9.9	3.8	38.9	30.7	222.3	59.6	143.4	21.7	47.6	12.0	23.5	14.2	17.0	18.9	15.8	33.5	13.3	39.6	8.7	32.9	0.8	684.5

Site	Depth	*n*-Alkane abundance (ng/g dw)																					Sum *n*-alkanes	
	(cm)	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30	C31	C32	C33	(µg/g dw)	(µg/g SOC)
NP50	0-9	1.3	3.8	1.6	5.3	4.2	9.2	6.0	9.7	10.0	9.7	25.3	10.7	35.3	12.5	74.4	18.7	259.1	23.0	214.3	19.4	54.9	0.8	77.9
	9-17	0.0	1.2	1.0	3.3	4.3	13.7	7.5	11.0	9.6	11.1	18.9	12.5	23.7	15.4	50.8	18.8	244.7	20.3	131.6	12.6	33.6	0.7	94.5
	17-24	0.0	1.9	0.3	6.8	5.8	89.1	22.0	76.1	9.8	30.9	10.1	16.1	15.7	13.8	23.7	11.1	49.5	11.5	62.9	9.4	33.3	0.5	100.9
	24-45	0.0	6.2	4.7	106.1	30.5	182.4	59.8	103.4	19.8	34.4	11.7	15.5	13.5	10.2	22.3	11.4	56.6	9.9	61.1	6.1	25.7	0.8	200.9
	45-70	0.0	4.5	4.3	91.8	36.5	228.5	61.6	134.8	17.8	50.2	13.4	26.4	16.9	18.3	24.5	13.2	54.1	12.8	73.6	8.2	35.8	0.9	228.2
	70-100	0.0	3.4	2.9	91.8	26.8	203.2	49.1	131.3	14.1	57.2	12.0	30.3	20.6	17.6	26.5	16.9	53.3	11.9	73.2	7.8	32.3	0.9	248.6
NP100	0-14	0.0	0.0	0.0	3.5	5.4	29.5	7.6	14.6	9.5	11.1	13.8	12.7	22.2	15.0	52.4	19.1	362.7	21.1	190.1	15.8	57.8	0.9	122.0
	14-25	0.0	4.3	5.3	115.0	74.3	271.1	75.9	157.7	20.8	60.2	14.7	26.0	17.0	18.6	37.8	16.9	189.3	14.2	108.4	13.8	36.6	1.3	155.9
	25-30	0.0	3.5	6.8	73.2	68.4	121.4	77.1	63.3	31.3	20.5	20.3	15.4	28.1	15.1	69.1	24.9	530.2	29.3	218.8	21.3	74.3	1.5	204.0
	30-38	0.0	1.5	1.3	81.9	70.6	295.8	130.6	169.0	32.5	47.1	13.5	22.4	17.4	14.7	26.2	12.7	65.7	12.3	71.1	9.3	29.5	1.1	269.5
	38-70	0.0	2.7	4.4	121.6	39.0	245.8	54.6	148.3	19.0	51.4	10.8	24.8	15.6	15.2	25.1	13.7	61.3	12.2	66.4	9.2	27.1	1.0	241.1
	70-100	0.0	15.2	11.3	159.4	73.2	247.5	94.3	139.7	33.9	62.4	49.5	72.1	71.2	66.2	62.9	40.0	70.0	21.6	71.8	11.8	27.9	1.4	387.5
NP300	0-11	2.5	6.5	3.9	23.0	18.2	59.1	19.2	37.0	18.2	23.7	21.4	17.2	48.2	19.1	133.0	50.2	475.8	46.6	522.2	28.7	140.7	1.7	159.0
	11-22	0.4	0.9	0.5	2.9	2.5	4.7	3.9	6.0	6.4	8.3	9.0	6.6	14.0	7.2	32.6	11.3	117.0	10.6	116.8	8.5	32.6	0.4	85.9
	22-32	0.1	1.1	2.6	85.7	16.9	147.2	31.7	93.8	12.2	31.9	9.1	15.1	12.5	10.0	21.1	9.4	54.7	9.8	60.1	7.0	27.4	0.7	127.8
	32-50	2.4	4.8	3.4	28.6	28.4	84.8	60.3	54.8	20.8	13.9	12.0	14.4	15.7	15.1	21.9	12.7	41.6	13.2	51.3	9.9	37.4	0.6	189.5
	50-70	2.2	7.5	8.7	110.2	43.2	225.3	63.6	137.6	20.6	49.2	12.5	24.1	16.1	20.9	23.2	15.4	42.9	13.7	59.3	9.6	37.7	1.0	386.2
	70-100	0.1	0.6	0.2	23.3	30.1	100.6	59.4	54.9	23.7	18.6	16.4	12.8	17.3	10.7	22.1	10.0	35.5	8.0	42.9	4.9	26.2	0.5	165.2
NP500	0-12	4.1	5.7	1.9	21.7	9.7	68.9	15.1	46.4	10.3	20.0	13.8	13.0	26.8	14.7	70.0	17.0	244.3	17.5	146.4	13.3	56.7	0.9	109.8
	12-17	4.6	4.9	4.7	11.4	5.4	13.6	3.8	9.3	7.2	8.8	8.3	9.2	15.2	8.8	31.1	8.3	120.5	7.0	68.4	3.4	14.1	0.4	57.4
	17-23	1.9	3.6	2.0	30.0	8.3	64.3	18.4	51.6	14.7	19.5	7.9	10.2	10.8	6.7	14.6	6.5	75.3	5.8	46.5	3.4	16.7	0.4	105.9
	23-45	0.0	0.0	0.0	0.0	0.0	14.3	9.3	32.4	11.1	12.5	5.9	8.3	9.6	6.2	11.2	4.4	18.8	2.6	19.1	1.0	7.2	0.2	66.5
	45-70	0.0	0.0	0.0	2.8	3.6	54.4	21.3	55.1	18.8	25.4	12.3	16.4	15.9	10.6	18.7	7.5	32.1	5.9	37.0	2.5	12.4	0.4	159.1
	70-100	1.6	2.7	0.4	24.4	5.5	53.5	13.4	34.4	11.8	16.8	8.8	10.8	14.7	9.5	15.8	7.7	28.6	5.5	33.2	2.2	8.8	0.3	87.2
NP700	0-12	5.0	8.6	6.5	28.7	16.5	37.5	12.9	22.4	16.4	16.9	22.4	15.0	41.6	19.1	106.6	38.2	711.9	40.8	306.9	27.9	120.3	1.7	179.1
	12-17	0.0	0.0	0.0	1.0	7.4	22.3	6.3	9.2	8.7	9.4	13.5	11.1	24.3	11.3	44.9	14.1	181.2	14.7	145.5	10.3	51.2	0.6	110.2
	17-23	2.4	3.6	1.4	4.6	1.8	6.0	4.6	6.6	6.8	7.9	9.3	8.3	14.6	8.8	25.0	10.2	70.8	10.4	78.0	8.8	35.6	0.3	87.3
	23-45	1.7	4.5	6.0	120.9	48.9	191.8	52.2	102.1	15.9	38.5	11.4	17.5	12.7	11.9	19.1	10.4	35.9	8.8	45.2	6.0	31.9	0.8	246.3
	45-70	0.5	1.1	3.3	104.8	50.2	203.0	50.6	117.3	18.3	41.5	13.8	21.2	13.7	14.9	21.0	12.3	38.6	12.2	51.0	8.3	34.1	0.9	279.5
	70-100	3.8	15.6	27.9	137.7	63.5	225.6	88.1	131.0	36.1	43.2	21.0	25.7	23.4	20.9	31.5	16.8	60.6	18.3	77.1	14.6	38.0	1.1	326.7
TW		0.0	3.1	2.5	51.3	13.9	85.2	13.8	46.1	13.2	20.7	20.1	18.2	36.5	19.0	59.9	19.3	130.7	21.0	171.9	13.3	67.8	0.8	138.6
Marsh		0.0	0.0	5.0	36.0	35.3	61.8	38.4	35.5	19.2	15.9	19.7	16.0	31.9	15.9	47.3	18.2	101.9	18.0	135.6	13.6	53.7	0.8	120.0
FW		0.0	12.9	20.6	117.2	54.5	89.2	30.2	49.6	93.8	118.0	282.4	141.9	366.1	174.7	576.7	234.2	882.8	256.5	947.5	202.1	668.7	5.7	576.9

Table S 3.3. Aliphatic hydrocarbon source proxies for paddy and upland soil profiles.

Site	Depth (cm)	Horizon FAO[a]	P_{aq}[b]	CPI[c]	ACL^{d}_{all}	$ACL^{d}_{long\ (26\text{-}35)}$	C_{23}/C_{25}	$(C_{25}+C_{33})/(C_{29}+C_{31})$
P50	0-7	Alp	0.13	7.95	27.9	30.0	0.57	0.30
	7-14	Arp	0.15	8.03	28.6	30.0	0.68	0.32
	14-23	Ardp	0.11	8.84	27.9	29.8	0.63	0.23
	23-38	Bwg1	0.25	2.75	21.7	30.7	1.35	0.59
	38-50	Bwg2	0.33	2.00	21.1	30.3	1.75	0.54
	50-70	Bwg3	0.23	2.88	21.7	30.2	0.96	0.46
	70-100	Blg	0.19	3.22	22.4	30.3	0.78	0.42
P100	0-9	Alp	0.08	13.39	29.1	29.7	0.56	0.20
	9-15	Alp2	0.11	8.84	27.9	29.8	0.63	0.23
	15-21	Ardp	0.17	6.15	25.7	29.8	0.91	0.26
	21-30	Bwg1	0.21	3.77	22.9	30.0	0.85	0.36
	30-50	Bwg2	0.28	2.75	23.4	30.3	0.99	0.55
	50-75	Bwlg1	0.21	3.23	21.7	30.4	0.68	0.47
	75-100	Bwlg2	0.19	2.19	21.2	30.1	0.70	0.48
P300	0-18	Alp	0.14	7.63	28.7	30.1	0.74	0.31
	18-24	Ardp	0.14	7.41	28.9	30.0	0.72	0.30
	24-30	Bwdl	0.18	3.86	26.4	30.3	0.73	0.40
	30-50	Bwl	0.22	1.80	21.8	30.6	0.75	0.69
	50-70	Bwlg1	0.20	3.26	25.7	30.6	0.74	0.57
	70-100	Bwlg2	0.21	2.14	21.3	30.3	0.63	0.56
P500	0-15	Alp	0.18	7.53	26.8	29.7	0.65	0.30
	15-19	Ardp	0.11	5.91	27.0	29.7	0.58	0.18
	19-25	Brdp	0.12	6.79	24.9	29.7	0.59	0.19
	25-48	Bwg1	0.18	4.83	23.8	29.8	0.65	0.29
	48-75	Bwg2	0.27	2.84	21.5	29.7	0.76	0.42
	75-100	Bwlg	0.22	4.76	26.3	29.9	0.54	0.42
P700	0-10	Alp1	0.22	2.26	28.2	29.8	0.41	0.39
	10-16	Alp2	0.24	2.38	27.7	29.6	0.47	0.38
	16-22	Ardp	0.17	4.52	27.2	29.8	0.66	0.29
	22-45	Bg	0.20	3.80	25.2	29.9	0.79	0.35
	45-69	2Ahgb	0.31	2.06	24.6	29.6	0.75	0.51
	69-106	2Blg1	0.41	2.27	23.2	29.6	1.22	0.54
P1000	0-10	Alp	0.21	6.36	26.9	30.1	0.37	0.39
	10-16	Al(d)p1	0.19	5.82	26.4	29.8	0.50	0.29
	16-21	Aldp2	0.20	3.77	26.2	29.8	0.72	0.33
	21-40	2Ahgb	0.34	2.19	25.2	29.5	0.69	0.55
	40-55	2Bg	0.23	2.91	26.5	29.7	0.66	0.43
	55-80	2Bl1	0.23	3.42	25.6	29.6	0.52	0.41
	80-93	3Ahlb	0.33	2.04	25.5	29.4	0.57	0.53
	93-100	3Bl	0.29	2.07	24.9	30.0	0.52	0.54
P2000	0-15	Alp	0.21	4.53	27.7	30.0	0.79	0.41
	15-20	Ar(d)p	0.18	3.96	29.0	30.3	0.80	0.43
	20-27	Bdg	0.21	2.37	22.9	30.3	0.75	0.50
	27-35	2AhgB	0.30	1.83	23.0	30.1	0.83	0.54
	35-50	2Bg1	0.30	2.13	22.8	29.8	0.70	0.52
	50-70	2Bg2	0.28	1.78	21.7	29.8	0.80	0.47
	70-100	2Blg	0.26	1.73	22.1	30.6	0.85	0.64
P50	0-7	Alp	0.13	7.95	27.9	30.0	0.57	0.30
	7-14	Arp	0.15	8.03	28.6	30.0	0.68	0.32
	14-23	Ardp	0.11	8.84	27.9	29.8	0.63	0.23
	23-38	Bwg1	0.25	2.75	21.7	30.7	1.35	0.59

Site	Depth (cm)	Horizon FAO[a]	P_{aq}[b]	CPI[c]	ACL^{d}_{all}	$ACL^{d}_{long\ (26-35)}$	C_{23}/C_{25}	$(C_{25}+C_{33})/(C_{29}+C_{31})$
NP50	0-9	Ap	0.11	8.46	28.6	29.9	0.72	0.19
	9-17	Abw	0.10	6.72	28.3	29.7	0.80	0.15
	17-24	Bw	0.19	3.10	24.6	30.3	0.65	0.44
	24-45	Bcg	0.18	3.68	21.8	30.1	0.87	0.33
	45-70	CBg	0.19	2.81	22.1	30.3	0.79	0.41
	70-100	CBg	0.20	2.73	22.3	30.1	0.58	0.52
NP100	0-14	Ap1	0.06	9.03	28.5	29.8	0.62	0.14
	14-25	Ap2	0.06	10.47	26.0	29.7	0.72	0.14
	25-30	Bw	0.10	5.11	22.7	29.8	0.86	0.18
	30-38	BCwg1	0.18	3.30	21.5	30.1	0.78	0.34
	38-70	BCwg2	0.17	2.95	21.7	30.1	0.69	0.33
	70-100	BCwIg	0.46	1.68	22.0	29.1	0.70	0.70
NP300	0-11	Ah	0.07	8.50	28.4	30.1	0.44	0.19
	11-22	Bw1	0.09	7.66	28.7	30.0	0.64	0.20
	22-32	Bw2	0.16	3.72	22.5	30.2	0.73	0.35
	32-50	Bwg1	0.23	2.46	23.7	30.5	0.76	0.57
	50-70	Bwg2	0.22	2.14	21.8	30.4	0.77	0.53
	70-100	BwIg	0.30	3.18	22.9	30.1	0.95	0.55
NP500	0-12	Ap1	0.09	7.82	26.8	29.8	0.51	0.21
	12-17	Ap2	0.11	7.81	26.9	29.5	0.55	0.16
	17-23	BCw	0.13	5.83	23.9	29.9	0.73	0.23
	23-45	C1	0.29	3.43	24.5	29.6	0.62	0.44
	45-70	C2	0.29	3.24	23.8	29.7	0.77	0.41
	70-100	C3	0.28	3.23	23.3	29.5	0.60	0.38
NP700	0-12	Ap1	0.06	9.78	28.4	29.8	0.54	0.16
	12-17	Ap2	0.10	7.79	28.3	30.0	0.56	0.23
	17-23	Bw1	0.14	4.96	28.2	30.3	0.64	0.34
	23-45	Bw2	0.23	2.68	21.4	30.5	0.90	0.55
	45-70	BwI1	0.23	2.32	21.7	30.4	1.00	0.53
	70-100	BwI2	0.24	2.54	21.8	30.2	0.90	0.45
TW			0.16	5.32	27.1	29.7	0.62	0.40
Marsh			0.18	4.73	25.7	30.2	0.62	0.63
FW			0.26	3.32	27.3	30.6	0.96	0.71

[a] Guidelines for soil profile description, FAO (2006).
[b] $P_{aq} = \sum C_{23}+C_{25}/\sum C_{23}+C_{25}+C_{29}+C_{31}$
[c] $CPI = [(\sum C_{21}+C_{23}+C_{25}+C_{27}+C_{29}+C_{31} / \sum C_{20}+C_{22}+C_{24}+C_{26}+C_{28}+C_{30})$
$^{+}(\sum C_{21}+C_{23}+C_{25}+C_{27}+C_{29}+C_{31}/\sum C_{22}+C_{24}+C_{26}+C_{28}+C_{30}+C_{32})]/2$
[d] ACL $[\sum(a \times A) \times n^{-1}]$; a is the amount of compound with **A** carbons and **n** number of alkanes

Supplementary data associated with this article can additionally be found in the online version at http://dx.doi.org/10.1016/j.orggeochem.2015.03.016.

4. Distribution of tetraether lipids in agricultural soils – differentiation between paddy and upland management

C. Mueller-Niggemann[1], S. R. Utami[2], A. Marxen[3], K. Mangelsdorf[4], T. Bauersachs[1], L., Schwark[1, 5]

[1]Institute of Geosciences, Christian-Albrechts-University of Kiel University, Kiel, Germany
[2]Soil Science, Faculty of Agriculture, University of Brawijaya, Malang, Indonesia
[3]Department of Soil Physics, Helmholtz Centre for Environmental Research UFZ, Halle (Saale), Germany
[4]Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, Section 4.3 Org. Geochem., Potsdam, Germany
[5]WA-OIGC, Curtin University, Perth, Australia

Published in *Biogeosciences Discuss.* **12** (2015) 16709-16754.
doi:1010.5194/bgd-12-16709-2015

Abstract. Insufficient knowledge of the composition and variation of isoprenoid and branched glycerol dialkyl glycerol tetraethers (GDGTs) in agricultural soils exists, despite of the potential effect of different management types (e.g. soil/water and redox conditions, cultivated plants) on GDGT distribution. Here, we determined the influence of different soil management types on the GDGT composition in paddy (flooded) and adjacent upland (non-flooded) soils, and if available also forest, bushland and marsh soils. To compare the local effects on GDGT distribution patterns, we collected comparable soil samples in various locations from tropical (Indonesia, Vietnam and Philippines) and subtropical (China and Italy) sites. We found that differences in the distribution of isoprenoid GDGTs (iGDGTs) as well as of branched GDGTs (brGDGTs) are predominantly controlled by management type and only secondarily by climatic exposition. In general upland soil had higher crenarchaeol contents than paddy soil, which on the contrary was more enriched in GDGT-0. The GDGT-

0/crenarchaeol ratio was 3-27 times higher in paddy soil and indicates the enhanced presence of methanogenic archaea, which were additionally linked to the number of rice cultivation cycles per year (higher number of cycles was coupled with an increase in the ratio). The TEX_{86} values were 1.3 times higher in upland, bushland and forest soils than in paddy soils. In all soils brGDGT predominated over iGDGTs, with the relative abundance of brGDGTs increasing from subtropical to tropical soils. Higher BIT values in paddy soils compared to upland soils together with higher BIT values in soil from subtropical climates indicate effects on the amounts of brGDGT through differences in management as well as climatic zones. In acidic soil CBT values correlated well with soil pH. In neutral to alkaline soils, however, no apparent correlation but an offset between paddy and upland managed soils was detected, which may suggest that soil moisture may exert an additional control on the CBT in these soils. Lower MBT' values and calculated temperatures (T_{MC}) in paddy soils compared to upland soils may indicate a management (e.g. enhanced soil moisture through flooding practises) induced effect on mean annual soil temperature (MST).

4.1 Introduction

Glycerol dialkyl glycerol tetraethers (GDGTs) are characteristic cell membrane lipids of archaea (Pearson and Ingalls, 2013; Schouten et al., 2013 and references therein) and bacteria (Weijers et al., 2006a; Sinninghe Damsté et al., 2011). The GDGT core structures differ in both domains, with isoprenoid alkyl chains being specific for archaea and branched alkyl chains for bacteria (for structures see Appendix). Both types of tetraether lipids have a high potential to be preserved in the sediment record (Schouten et al., 2013) and have been reported in abundance from terrestrial and marine environments, e.g. in the water column and sediments of oceans and lakes (Hopmans et al., 2000, 2004; Schouten et al., 2012; Tierney and Russel, 2009; Zink et al., 2010; Naeher et al., 2014), in ponds (Tierney et al., 2012; Loomis et al., 2014; Huguet et al., 2015), in hot springs (Pearson et al., 2004; Peterse et al., 2009a; Pitcher et al., 2009), in peat bogs (Sinninghe Damsté et al., 2000; Weijers et al., 2006a, 2010), in grassland soils (Weijers et al., 2007, 2010; Naeher et al., 2014), in forest soils (Hopmans et al., 2004; Weijers et al., 2007, 2010), in permafrost soils (Peterse et al., 2009b; Bischoff et al., 2014), in loess soils (Huguet et al., 2012), in Podzols (Huguet et al.,

2010), in garden and agricultural soils (Leininger et al., 2006; Weijers et al., 2010; Sinninghe Damsté et al., 2012) as well as in paddy soils (Bannert et al., 2011b; Ayari et al., 2013).

It is well known that archaea are involved in biogeochemically important processes, including methanogenesis, anaerobic methane oxidation (AMO) and aerobic ammonia oxidation (Kuypers et al., 2001; Pancost et al., 2001; Leininger et al., 2006; Pearson and Ingalls, 2013). Distributions of isoprenoid GDGTs (iGDGTs) were initially used to characterize archaeal communities in marine environments with two major groups of archaea being distinguished: *Crenarchaeota, Thaumarchaeota* and *Euryarchaeota* (see Pearson and Ingalls, 2013 and references therein). The archaeal phylum comprising the ammonia-oxidizing *Thaumarchaeota* has been identified more recently (Brochier-Armanet et al., 2008; Spang et al., 2010). Members of this phylum are currently the only known biological sources of crenarchaeol and in addition they contain varying amounts of tetraether lipids with 0 to 4 cyclopentane rings (Sinninghe Damsté et al., 2012; Schouten et al. 2013; Pearson and Ingalls, 2013). Tetraether lipids of methanogenic archaea generally contain GDGT-0 (Koga et al., 1998; Koga and Morii, 2005; Pearson and Ingalls, 2013, Schouten et al., 2013), although in some instances iGDGTs with cyclopentyl moieties have been reported (De Rosa 1986; Bauersachs et al., 2015). iGDGTs with cyclopentane rings were also reported from methanotrophic archaea of the ANME-1 cluster, *Thaumarchaeota* as well as extremophilic *Euryarchaeota* and *Crenarchaeota* (Blumenberg et al., 2004; Pearson and Ingalls, 2013, Schouten et al., 2013 and references therein). The cell membrane of mesophilic archaea consists, among others, of iGDGT structures usually containing 1 to 4 cyclopentyl moieties (GDGT-1 to GDGT-4) with members of the *Thaumarchaeota* also possessing crenarchaeol, a GDGT structure that contains four cyclopentane ring systems and an additional cyclohexane ring moiety (Sinninghe Damsté et al., 2002).

High abundances of branched GDGT (brGDGTs) have previously been reported from soils worldwide (Weijers et al., 2007, 2010; Peterse et al., 2009a; Huguet et al., 2010, 2012). Information on the biological sources of these components, however, is still very limited (Hopmans et al., 2004; Weijers et al., 2007, 2010; Peterse et al., 2009b, 2009c; Tierney and Russell, 2009; Huguet et al., 2010, 2012; Tierney et al., 2012). Molecular investigations in

peat bogs demonstrated that brGDGTs occurred in highest concentrations in the catotelm, the bottom layer of peats (Weijers et al., 2006a, 2010), which suppose an anaerobic and acid tolerant bacterial species as origin, e.g. belonging to *Acidobacteria* the most abundant bacteria in this environment (Weijers et al., 2006a, 2009, 2010). This is supported by the presence of the tetra-methylated brGDGT that was recently identified in two cultured acidobacterial strains (Sinninghe Damsté et al., 2011). In addition, the ether-bound 5-methyl *iso*-diabolic acid was detected in four mesophilic species of the subdivision 4 of the *Acidobacteria* as a potential breakdown product of penta-methylated brGDGT (Sinninghe Damsté et al., 2014). It has consequently been suggested that bacteria producing these lipids are obligate anaerobes and follow a heterotrophic mode of life (Oppermann et al., 2010; Weijers et al., 2006a, 2010). The presence of brGDGTs in oxic soils does not exclude that aerobically living bacteria produce these lipids, but anaerobic bacteria residing in anoxic microhabitats are also possible sources (Schouten et al., 2013). The distribution of brGDGTs in soils is related to growth temperature (mean annual air and soil temperature) and soil pH (Schouten et al., 2002; Weijers et al., 2007, 2009; Peterse et al., 2009a, 2012). Indices which denote the degree of methylation and cyclization of brGDGTs, the MBT and the CBT indices, have previously been employed to reconstruct mean annual air temperatures (MAT) using a global soil calibration (Weijers et al., 2009). More recently, Peterse et al., (2012) defined the MBT', which represents the ratio of tetra-methylated brGDGT (GDGT-Ia, Ib and Ic) *vs.* the seven most abundant brGDGTs (GDGT-Ia, Ib, Ic, IIa, IIb, IIc and IIIa).

However, factors other than temperature and pH also seem to affect the distribution of brGDGTs in natural ecosystems. For example, the relative broad scatter of calculated MAT in arid soils (Peterse et al., 2012) as well as values deviating from the trend in the highest elevations of a transect sampled on Mt. Kilimanjaro (Sinninghe Damsté et al., 2008) are interpreted to indicate an influence of water content and vegetation type on the brGDGT pool. In addition, several authors noted that changes in the distribution of brGDGT are strongly correlated with MAT on local scales as, for example, in altitudinal transects of Mt. Rungwe and Mt. Gongga (Peterse et al., 2009c; Coffinet et al., 2014). In agricultural soils from the same area, the type of soil management and the vegetation cover can differ, leading to

variable soil water contents and soil temperatures (Liu et al., 2014; Awe et al., 2015), which affect the local microbial community. In addition, soil microbes respond to environmental stresses induced by e.g. starvation, oxygen limitation or acidification. The latter results in the predominance of brGDGTs without cyclopentyl moieties in the bacterial cell membrane and explain the dependency of soil pH and CBT (Weijers et al., 2007).

In addition to the pH, the redox potential (Eh) is an important factor that affects the diversity and abundance of soil microorganisms. The Eh expresses the activity of electrons (measured in volts), which influence microbial metabolic reactions in soils. As individual microorganisms are adapted to specific Eh conditions, an increase in e.g. soil moisture is accompanied by a decrease in Eh because of the consumption of oxygen by microbes (Husson, 2013). Agricultural management may control redoximorphic conditions. In contrast to upland soil, i.e. without water flooding and associated crop plants, including corn/maize, wheat, barley, rape, cassava, sugar cane, cotton, banana and other vegetables, rice paddy soil management with repetitive puddling of the surface soil as well as frequent flooding and alternating draining practices leads to a reduced Eh in the surface layer (Kögel-Knabner et al., 2010; Kölbl et al., 2014). Prevailing anoxic conditions are assumed to restrict the decomposition rate of organic matter (Lal, 2002; Sahrawat, 2005), leading to high activities of methanogenic archaea (Liesack et al., 2000) and in combination with the application of mineral fertilizer to high denitrification rates producing nitrous oxide (Xiong et al., 2007). In contrast, oxic conditions are associated with high Eh, as in upland soil and in paddy soil after draining ammonia oxidation occurs. The latter is either performed by ammonia-oxidizing archaea (AOA) or bacteria (AOB) (Leininger et al., 2006) depending on the soil pH, with AOA being more active in acidic soils and AOB in alkaline soils (Jiang et al., 2015).

Here, we investigated the environmental controls that affected the tetraether lipid composition in soils of different management systems, which developed in subtropical (Italy, SW-China) as well as in tropical (Indonesia, Philippines, Vietnam) climates. Additionally to the management type, including differences in cropping style (upland crop plants *vs.* wetland rice), the intensity of the management and the duration of utilization were distinctive criteria in the investigation of effects on microbial lipids in upland, paddy and forest soils. Only

limited information on the distribution of tetraether lipids in paddy soils is currently available (Bannert et al., 2011b; Ayari et al., 2013), although an area of 157 million ha, contributing 18% area to the ten major crops worldwide, is covered by rice paddy management (FAO, 2003). To the best of our knowledge, this is the first study, which compares non-flooded and flooded agroecosystems with respect to their GDGT composition. The variation in GDGT distribution patterns between soils with different agricultural usage will provide additional information on the sources and properties of GDGTs in terrestrial ecosystems on local and global scale.

4.2 Material and methods

Sampling

From 2008 to 2014, a total of 170 Indonesian, Vietnamese, Philippine, Chinese and Italian soils with different land-use systems were collected, including 119 paddy, 37 upland, 9 forest, 2 bushland and 3 marsh samples from the topsoil horizon (0-30 cm depth). The study sites are located in tropical as well as in subtropical climate zones (Fig. 4.1, Table 4.1) and agricultural soils were subject to different management techniques. Detailed soil characteristics and geographical positions of the sampling sites are given in Table S4.1 (Supplementary material). Topsoils were sampled with a soil auger as described by Klotzbücher et al. (2014).

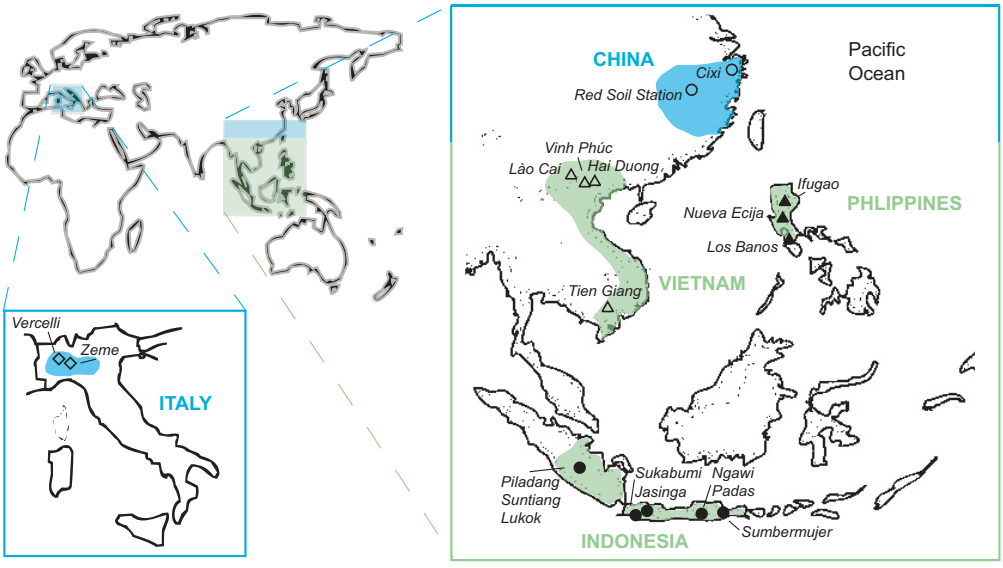


Fig. 4.1. Map of sampling locations. Blue coloured area denotes subtropical sampling locations and green denotes tropical sampling locations.

In addition, successive land reclamation in the Chinese location Cixi via dyke construction on marine tidal flats over the last > 1000 yr (Feng and Bao, 2005) led to differently aged soils, which allow studying a 2000 yr chronosequence. Based on the time of dyke construction and information from the Edit Committee of Chorography of Cixi County (1992), differently aged marsh soils (10-35 yr) and agricultural soils under continuous non-irrigated upland use (50-700 yr) as well as wetland rice cultivation (50-2000 yr) were selected and sampled. The local cropping system on paddy fields is paddy-upland rotation, with one wetland rice season and one inter-crop (vegetables, wheat or cereals) season per year (Cheng et al., 2009). Paddy and upland topsoils were sampled with a soil auger. Three composite samples, composed of 7 sub-samples, being representative for the complete field were investigated at each location.

Bulk geochemistry

All soils were lyophilized, sieved to a size < 2 mm and ground to a fine powder using agate pestle and mortar prior to analyses. Soil pH was measured in a suspension of the soil in 0.01 M $CaCl_2$, using a 1:2.5 (w/v) soil/liquid ratio. The pH was determined with a pH meter Model FG2-438 (Mettler-Toledo AG, Switzerland) at ambient temperature and atmospheric pressure. The total carbon (TC) and total nitrogen (TN) contents were measured on a CNS elemental analyser Vario EL III (Elementar Analysensysteme GmbH, Germany). The total inorganic carbon (TIC) content was determined using the Vario EL III elemental analyser coupled to SoliTIC module. The soil organic carbon (SOC) was calculated as the difference between TC and TIC.

GDGT preparation and HPLC-MS analysis

Core lipids of iGDGTs and brGDGTs were obtained by automated solvent extraction using an ASE 200 (Dionex, USA) at a temperature of 75°C and a pressure 5.0 x 10^6 Pa. Each sample was extracted for 20 min using a solvent mixture of dichloromethane (DCM)/MeOH (93:7, v/v). The total lipid extracts were separated over an aluminium oxide column into apolar and polar fractions using *n*-hexane/DCM (9:1, v/v) and DCM/MeOH (1:1, v/v) as respective eluents. The polar fractions were dried under a gentle stream of N_2, re-dissolved in *n*-

hexane/2-propanol (99:1, v/v) and filtered through a 0.45 µm polytetrafluoroethylene (PTFE) filter prior to analysis.

All samples were analysed by high performance liquid chromatography coupled to atmospheric pressure positive ionisation mass spectrometry (HPLC/APCI-MS) using an Alliance 2690 HPLC (Waters, UK) and a Quattro LC triple quadrupole mass spectrometer (Micromass, UK) and following the analytical protocol described by Hopmans et al. (2000) and Schouten et al. (2007). Briefly, 10 µl of the filtered polar fractions were injected on an analytical Prevail Cyano column (2.1 x 150 mm, 3 µm particle size, Grace, USA), maintained at a temperature of 30 °C with a constant flow rate of 0.2 ml/min. Tetraether lipids were eluted isocratically with 99% *n*-hexane and 1% 2-propanol for 5 min, followed by a linear gradient to 1.8% 2-propanol in 36 min and subsequently to 10% 2-propanol in 5 min, after which the system was held isocratic for 5 min. The column was re-equilibrated with 99% *n*-hexane and 1% 2-propanol for 12 min before the next injection. The MS was operated as outlined in Heyng et al. (2015) with isoprenoid and branched GDGTs being detected in selected ion recording (SIR) mode of their protonated molecules $[M+H]^+$.

Calculation of GDGT indices

Acronyms in the below equations refer to the relative abundance of GDGTs displayed in the Appendix A. The relationship between the commonly less occurring cyclopentane ring containing iGDGTs (GDGT-1 to GDGT-3 *vs*. the crenarchaeol regioisomer) was considered with using the TEX_{86} (tetraether index of tetraethers consisting of 86 carbons). The TEX_{86} was calculated according to Schouten et al. (2002):

TEX_{86} = (GDGT-2 + GDGT-3 + Cren regioisomer)/(GDGT-1 + GDGT-2 + GDGT-3 + Cren regioisomer) (1)

The Cyclization ratio of Branched Tetraethers (CBT) was calculated using the relative abundance of tetra- and penta-methylated brGDGT according to Weijers et al. (2007):

CBT = -log ((Ib + IIb)/ (Ia + IIa)) (2)

The Methylation index of Branched Tetraethers (MBT') index was calculated as described by Peterse et al. (2012):

$$MBT' = (Ia + Ib + Ic)/ (Ia + Ib + Ic + IIa + IIb + IIc + IIIa) \quad (3)$$

The MBT' and CBT derived MAT (T_{MC}) was calculated after Peterse et al. (2012):

$$T_{MC} = 0.81 - 5.67 \times CBT + 31.0 \times MBT' \quad (4)$$

The Branched and Isoprenoid Tetraether (BIT) index was determined as given in Hopmans et al. (2004):

$$BIT = (Ia + IIa + IIIa)/ (Ia + IIa + IIIa + Cren) \quad (5)$$

Statistical analysis

Statistical analysis was conducted using the PASW Statistics 18 software. Principal component analysis (PCA) was performed on relative abundances of iGDGTs, brGDGTs and the different GDGT-based indices, to explore and characterize the variability within the GDGT distribution in differently managed soils. To identify relationships between variables, a correlation analysis was performed. Results were given as r for Pearson's correlation regression coefficient together with the *p*-value (two-tailed test), which denotes a significance if p is < 0.001.

4.3 Results

SOC (Table 4.1) varied from 0.4 to 5.0% with highest contents present in paddy soils from the Philippine Ifugao (5.0%) and Laguna (4.0%), the Indonesian Sukabumi (4.4%) and the Vietnamese Tien Giang (4.4%) sites. The forest and bushland soils had a mean SOC of 2.7±0.9% (n = 11), which was higher than in most upland soils (1.6±0.9%, n = 37). The pH ranged from 3.7 in Tien Giang (Vietnam) to 8.2 in Cixi (China; Table 4.1). In general, no differences in pH values were noticed for soils with paddy (5.3±1.0, n = 119) or upland (5.3±1.1, n = 37) management. Forest and bushland soils had the lowest mean pH of 4.5±0.5 (n = 11).

Table 4.1. List of sampling areas, environmental characteristics and minimum as well as maximum of GDGT proportions (expressed as a percentage of total GDGTs or as indices).

Country	Sampling area	Soil type	Dataset code	N	MAT (°C)	MAP (mm)	SOC (%)		pH		iGDGTs (%)		brGDGTs (%)		GDGT-0/cren		$Tex_{86'}$		CBT		MBT'		T_{MC} (°C)
							Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max	mean
Italy	Zeme	Upland	IT-NP	1	12.5	954	0.73		4.1		25.1		74.9		0.42		0.66		1.41		0.52		8,8
		Paddy	IT-P	1	12.5	954	1.15		4.9		9.6		90.4		2.67		0.44		0.90		0.51		11,6
	Vercelli	Paddy	IT-P	4	12.1	923			6.1	7.0	5.5	11.5	88.5	94.5	0.37	1.53	0.54	0.71	0.14	0.65	0.33	0.49	11,6
China	Cixi	Marsh	C-Marsh	3	16.6	1266	0.43	0.63	8.0	8.0	12.4	29.8	70.2	87.6	0.22	0.57	0.64	0.72	-0.03	0.38	0.47	0.50	14,7
		Upland	C-NP	5	16.6	1266	0.72	1.10	6.0	8.2	15.2	35.0	65.0	84.8	0.14	0.37	0.62	0.72	-0.02	0.19	0.53	0.63	18,2
		Paddy	C-P	21	16.6	1266	0.92	2.88	5.2	7.5	7.7	22.5	77.5	92.3	0.29	5.77	0.30	0.68	0.26	0.67	0.49	0.70	16,8
	Red Soil Station	Upland	C-NP	3	18.5	1731	0.70	0.85	4.1	5.1	15.5	16.0	84.0	84.5	0.32	0.48	0.76	0.78	1.56	2.13	0.72	0.77	13,3
		Paddy	C-P	5	18.5	1731	2.04	2.75	4.2	4.5	6.6	11.4	88.6	93.4	2.07	3.51	0.49	0.68	0.99	1.21	0.69	0.76	17,1
Indonesia	Jasinga	Upland	JAV-NP	3	26.9	3252	2.08	3.22	3.8	5.6	5.6	9.1	90.9	94.4	0.20	0.89	0.72	0.84	0.64	1.86	0.92	0.96	22,0
		Paddy	JAV-P	4	26.9	3252	1.97	2.30	4.2	4.4	0.9	2.0	98.0	99.1	2.01	2.26	0.61	0.68	1.60	1.83	0.91	0.92	19,3
	Ngawi	Upland	JAV-NP	3	27.0	2034	1.46	1.74	4.7	5.4	6.9	14.2	85.8	93.1	0.12	0.16	0.72	0.74	0.84	1.15	0.92	0.94	24,0
		Paddy	JAV-P	3	27.0	2034	1.40	1.81	6.4	7.2	6.8	9.5	90.5	93.2	0.58	1.20	0.68	0.71	0.34	0.65	0.72	0.80	21,8
	Padas	Paddy	JAV-P	1	26.7	2162	1.73		6.8		15.3		84.7		0.40		0.70		0.42		0.83		24,1
	Simo village	Paddy	JAV-P	3	26.9	2100	1.52	1.86	6.9	7.5	15.4	23.2	76.8	84.6	0.38	1.24	0.71	0.75	0.29	0.38	0.67	0.82	21,8
	Sukabumi	Upland	JAV-NP	3	23.5	2806	3.50	4.34	4.4	4.8	13.6	22.9	77.1	86.4	0.36	1.28	0.66	0.72	0.90	1.48	0.88	0.90	21,3
		Paddy	JAV-P	3	23.5	2806	4.02	4.41	5.1	5.3	5.5	6.1	93.9	94.5	0.38	0.45	0.68	0.71	1.16	1.24	0.77	0.80	18,4
	Sumbermujer	Paddy	JAV-P	1	17.8	2693	2.49		5.2		11.5		88.5		2.73		0.42		0.82		0.79		20,6
		Bamboo	JAV-Bamb	1	17.8	2693	3.57		5.2		3.1		96.9		1.80		0.63		1.10		0.95		23,9
	Sumatra	Paddy	SUM-P	4	21.8	2170	1.39	2.54	4.7	5.4	6.5	10.2	89.8	93.5	0.49	5.78	0.46	0.71	0.94	1.34	0.75	0.82	19,1
Philippines	Ifugao	Forest	PH-For	3	21.4	2376	2.38	3.22	4.8	5.2	1.8	3.5	96.5	98.2	0.32	1.05	0.59	0.69	0.74	0.88	0.80	0.87	22,3
		Upland	PH-NP	5	21.4	2376	1.21	2.09	4.4	5.6	2.7	7.3	92.7	97.3	0.39	2.02	0.59	0.70	0.78	1.27	0.81	0.90	22,1
		Paddy	PH-P	10	21.4	2376	1.16	5.04	4.3	5.5	3.6	17.6	82.4	96.4	3.67	121.6	0.45	0.58	0.70	1.23	0.63	0.80	18,1
	Laguna	Upland	PH-NP	5	27.1	2064	1.77	2.17	5.1	5.7	4.0	10.0	90.0	96.0	0.14	2.48	0.68	0.85	0.56	1.39	0.87	0.94	23,8
		Paddy	PH-P	10	27.1	2064	1.59	4.01	4.7	6.2	7.8	13.9	86.1	92.2	0.19	5.65	0.50	0.86	0.70	1.08	0.77	0.89	21,2
	Nueva Ecija	Upland	PH-NP	4	27.1	1821	0.54	1.30	4.6	6.5	6.7	25.7	74.3	93.3	0.17	0.92	0.74	0.83	0.51	1.33	0.85	0.91	23,0
		Paddy	PH-P	10	27.1	1821	0.83	1.95	4.3	6.2	5.7	14.4	85.6	94.3	0.15	9.66	0.48	0.81	0.52	1.65	0.73	0.86	19,2
Vietnam	Hai Duong	Upland	VN-NP	2	24.1	1608	0.79	1.17	4.9	7.4	7.7	10.4	89.6	92.3	0.40	1.66	0.59	0.76	-0.04	0.91	0.71	0.73	20,6
		Paddy	VN-P	8	24.1	1608	1.13	1.68	4.8	5.7	4.6	9.0	91.0	95.4	1.42	5.63	0.45	0.59	0.45	0.81	0.65	0.72	18,3
	Lào Cai	Bamboo	VN-Bamb	1	16.2	2223	2.97		4.2		2.3		97.7		0.95		0.66		1.26		0.89		21,2
		Bushland	VN-Bush	2	16.2	2223	2.56	3.32	4.1	4.4	4.1	4.4	95.6	95.9	1.31	3.08	0.65	0.73	1.36	1.61	0.90	0.90	20,3
		Forest	VN-For	2	16.2	2223	2.77	3.88	4.1	4.1	3.0	3.6	96.4	97.0	0.83	1.10	0.63	0.72	1.23	1.60	0.87	0.89	20,1
		Paddy	VN-P	10	16.2	2223	0.83	2.48	4.3	5.2	4.8	10.7	89.3	95.2	0.79	20.73	0.35	0.62	0.80	1.44	0.59	0.86	15,7
	Tien Giang	Paddy	VN-P	13	27.4	1450	2.06	4.43	3.7	4.8	7.6	10.9	89.1	92.4	0.72	17.39	0.54	0.61	0.99	1.14	0.79	0.85	20,4
	Vinh Phúc	Bamboo	VN-Bamb	1	23.6	1687	0.69		4.3		4.4		95.6		0.66		0.75		1.83		0.95		19,8
		Forest	VN-For	1	23.6	1687	1.30		3.8		8.1		91.9		0.55		0.79		2.00		0.86		16,1
		Upland	VN-NP	3	23.6	1687	0.58	1.64	4.0	6.1	5.0	18.8	81.2	95.0	0.57	1.30	0.75	0.77	0.88	1.77	0.87	0.93	20,7
		Paddy	VN-P	8	23.6	1687	1.12	2.41	4.3	4.8	9.1	16.1	83.9	90.9	0.88	8.19	0.50	0.70	0.88	1.60	0.75	0.85	18,4

Both iGDGT and brGDGT were detected in variable abundances in all soils. The brGDGT/iGDGT ratio was > 80 in Indonesian paddy soils (Jasinga), 20-80 in forest and bushland soils, and as low as 1.9 in the remaining soils (Supplementary material, Fig. S4.1). The lowest proportion of brGDGT was noted in Italian upland soils, in very young Chinese marsh soils (< 30 yr) and upland soils. A specific feature of soil from the Chinese Cixi area is its development from tidal wetland sediment. The GDGT signature of these soils was distinct from the one in other soils investigated in this study and represents a mixed signature of the parent substrate (tidal wetland sediments) and the recent soil organic matter (SOM).

4.4 Discussion

Distribution of isoprenoid GDGTs in soils

iGDGTs constitute between 0.9 and 25.7% (and in soils of Cixi 35%) of all GDGTs (Table 4.1), indicating substantial contributions of archaeal lipids to most investigated soils. Forest and bushland soils had lowest relative mean abundances of iGDGTs (5.8±2.6%), followed by tropical paddy (9.3±4.0%) and upland soils (9.8±6.0%). The proportion of iGDGTs was highest in Chinese and Italian upland soils (21.1±8.0%) compared to their adjacent paddy soils and all other remaining soils (13.3±5.0%). The fact that the iGDGT content was lower in tropical compared to subtropical soils suggests that the composition of the microbial consortia varies on regional to global scales. In addition, the differentiation between upland and paddy soils with higher amounts of iGDGTs in the former may indicate management (regulating the water regime, nutrient availability, oxygen availability and/or redox conditions) induced variations of GDGT containing microorganism. In general, at locations with the same climate and substrate, different management types best explain different GDGT distribution. Regardless of whether paddy, upland or forest management, all differ in their microbial lipid pattern that may be influenced by differing inputs of plant organic matter, differing fertilization practises and redox conditions. The latter is controlled by flooding and draining practises on paddy soils, which seem to favour growth and input of brGDGT containing bacteria compared to the adjacently located upland soils.

The distribution of iGDGTs in soils may provide detailed insights into the archaeal community structure and the biological processes that they mediate (Koga et al., 1998; Pancost et al., 2001; Blumenberg et al., 2004; Koga and Morii, 2005). The most abundant iGDGTs in our sample set are GDGT-0 and crenarchaeol. The latter is considered a highly specific biological marker for ammonia-oxidizing *Thaumarchaeota* (Leininger et al., 2006; Pitcher et al., 2010; Sinninghe Damsté et al., 2012; Pearson and Ingalls, 2013). Molecular investigations on cultivated *Thaumarchaeota* revealed separation between group I.1a *Thaumarchaeota* (aquatic) and group I.1b *Thaumarchaeota* (terrestrial/soil) based on the relative abundance of the crenarchaeol regioisomer. Abundances of the crenarchaeol regioisomer < 5% are indicative for group I.1a and > 10-20% for group I.1b *Thaumarchaeota* (Sinninghe Damsté et al., 2012). The same authors demonstrated that in soils group I.1a *Thaumarchaeota* and group I.1b *Thaumarchaeota* produce higher abundances of the crenarchaeol regioisomer than in marine or lacustrine environments (Sinninghe Damsté et al., 2012). Crenarchaeol and its regioisomer are present in all analysed soil samples, which is in agreement with a previous study (Weijers et al., 2006b). The amount of crenarchaeol is generally higher in upland soils (46.4±12.9%, n = 37) compared to adjacent paddy soils (22.5±14.5%, n = 119; Fig. 4.2a), possibly suggesting management induced differences in the archaeal community structure. The abundance of the crenarchaeol regioisomer varies from 3 to 21% to that of crenarchaeol (mean value of 9±4%, n = 170), and shows no differences between soils and/or management types (Fig. S4.2).

Angel et al. (2012) observed that methanogenic archaea are ubiquitous in soils and being active only in anoxic, highly reducing environments, e.g. under flooded conditions. One distinct feature of paddy soil management *vs*. management of all other soils is the periodic flooding and draining of soils, which leads to highly variable redox conditions throughout the time course of a year (Kögel-Knabner et al., 2010; Kölbl et al., 2014). Paddy soils are known for high methanogenic activity and as significant sources of atmospheric CH_4 (Conrad, 2007; Thauer et al., 2008; Serano-Silva et al., 2014) without any changes in the methanogenic community structure between floodings (Krüger et al., 2005; Watanabe et al., 2006, 2009). In

turn, this suggests that the overall lipid pool in paddies does not change significantly after draining the fields for harvesting.

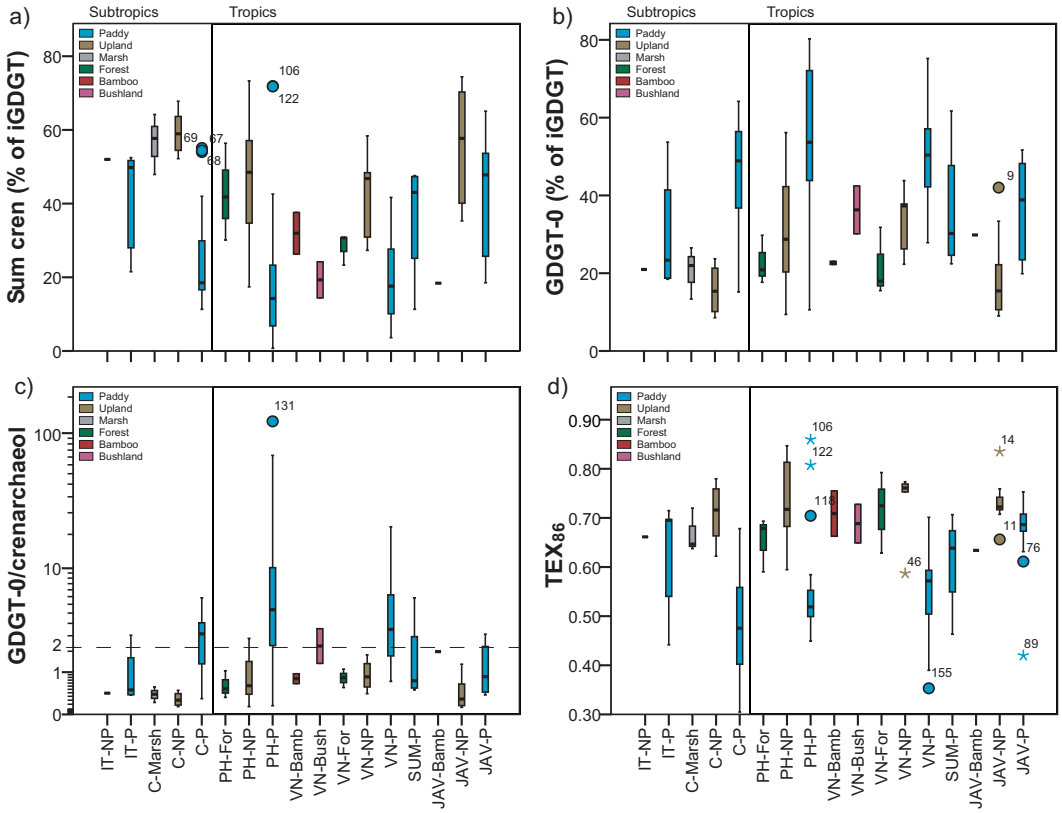


Fig. 4.2. Box-plot diagrams of **(a)** crenarchaeol, **(b)** GDGT-0, **(c)** GDGT-0/crenarchaeol ratio and **(d)** TEX_{86} in upland (NP, brown), paddy (P, blue), marsh (grey), forest (For), bamboo cultivated (Bamb, red) and bushland (Bush, violet) soils. Abbreviations refer to different sampling locations: Italy (IT), China (C), Philippines (PH), Vietnam (VN), Sumatra (SUM) and Java (JAV). The vertical line separates subtropical from tropical locations. Numbers in all plots indicate samples listed in Table S4.1. The dotted line in (c) marks the GDGT-0/crenarchaeol value of 2 that is the boundary to higher proportions of methanogens, which reveal values > 2. Note the logarithmic scale for GDGT-0/crenarchaeol ratios.

The ratio of GDGT-0/crenarchaeol, initially proposed for lake environments, may be used to indicate the dominance of methanogenic archaea (Blaga et al., 2009) or of *Thaumarchaeota* in a given sedimentary environment. The latter are members performing the first and rate-limiting step in nitrification: the aerobic oxidation of ammonia (Stahl and de la Torre, 2012; Stieglmeier et al., 2014). In various studies, it was shown that a GDGT-0/crenarchaeol ratio

> 2 is diagnostic for methanogens (Blaga et al., 2009; Naeher et al., 2014). In the analysed soils, the GDGT-0/crenarchaeol ratio ranged from 0.1 to 121.6, with highest ratios observed in Philippine and Vietnamese paddy soils (Fig. 4.2c, Table 4.1). In oxic upland and forest soils the mean GDGT-0/crenarchaeol ratio was ≤1, which indicates that methanogenic archaea are only a minor component of the microbial community at these sites. In addition, a few paddy soils (e.g. sites in Chinese Cixi and in Italy) had GDGT-0/crenarchaeol ratios comparable to those observed in upland soils, which can be explained by the management form including higher intensities of crop-rotation with upland crops under non-flooded conditions on these fields. However, if soils from the same region are compared, the ratio was generally 3-27 times higher in soils which are under paddy management compared to adjacent upland soils, indicating increased abundances and activity of methanogens in flooded soils.

The TEX_{86} values determined ranged from 0.3 to 0.9 (Fig. 4.2d, Table 4.1) without an apparent geographical trend. However, within a region TEX_{86} values were on average 1.3 times higher in upland, bushland and forest soils compared to the adjacent paddy soils. Highest values (upland/paddy-TEX_{86} = 1.5) were observed in the subtropical locations of Cixi and Italy (Table 4.1). None or only minor differences in TEX_{86} values were noted in the Jasinga and Ngawi soils of Indonesia. Because of the relation between the TEX_{86} and temperature, one explanation for the difference could be that the periodic water layer on paddy soils may protect the soil surface from excessive heating and therefore results in lower mean annual soil temperatures (MST) in both soil types. Previous studies of altitudinal mountain transects support this suggestion, as the soil TEX_{86} was negatively correlated with elevation and therefore with decreasing temperatures e.g. in the Qinghai-Tibetan Plateau ($r = -0.81$, $r^2 = 0.65$, $p < 0.01$; Liu et al. 2013) and Tanzania ($r = -0.71$, $r^2 = 0.50$, $p < 0.0001$; Coffinet et al., 2014).

In the soils investigated here, the relative proportion of GDGT-3 and the crenarchaeol regioisomer together with GDGT-1 mainly affected the TEX_{86}. Low TEX_{86} values, as observed in paddy soils, are the result of high relative abundances of GDGT-1 and low proportions of GDGT-3. This suggests that paddy soil characteristics such as alternating redox conditions and higher water content control the presence of GDGT-1. High contents of

cyclopentyl moieties in archaeal membrane lipids were associated with anaerobic methanotrophic (ANME) archaea, which synthesize significant quantities of GDGT-1, GDGT-2 and GDGT-3 (Pancost et al., 2001; Blumenberg et al. 2004). Interestingly, two divergent trends in direction of increased TEX_{86} values were observed for GDGT-2 (Fig. 4.3a), with an increase of the GDGT-2 content to a TEX_{86} value of 0.70 and a subsequent decrease if values exceed this threshold (Fig. 4.3a). This change may again indicate that the archaeal community differs in dry upland/forest soils and flooded soils.

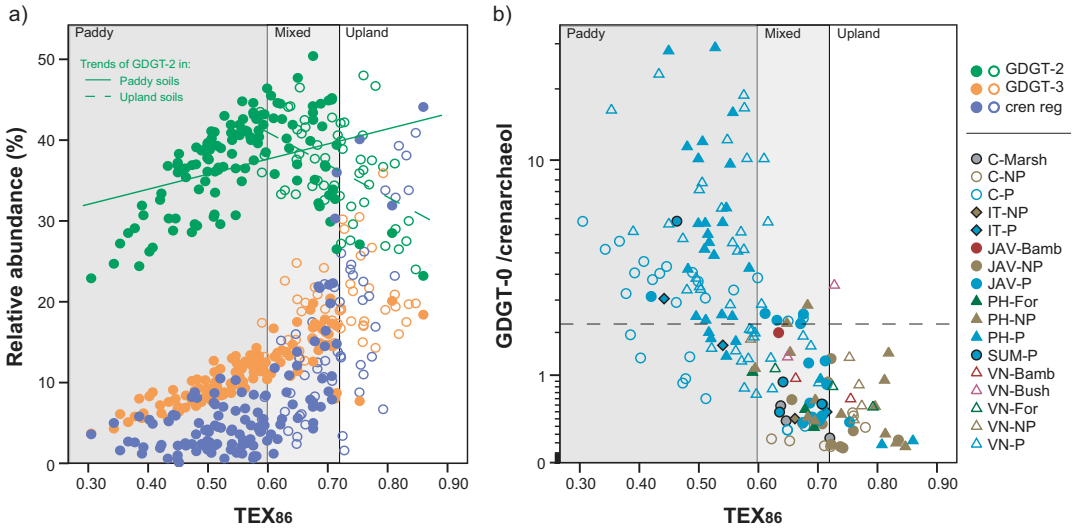


Fig. 4.3. Cross-plots showing **(a)** the relative abundance (% of the sum of GDGT-1, -2, -3 and crenarchaeol regioisomer) versus TEX_{86} and **(b)** the relationship between the most abundant iGDGTs (GDGT-0 and crenarchaeol) and lower concentrated iGDGTs (GDGT-1, -2, -3, and crenarchaeol regioisomer) as TEX_{86}. GDGT-0/crenarchaeol > 2 and TEX_{86} < 0.6 are diagnostic for methanogens. Two outliers from the Ifugao site (Philippines) with GDGT-0/crenarchaeol ratio > 69 were excluded from the figure. Note the logarithmic scale for GDGT-0/crenarchaeol ratios. The filled circles in (a) denote paddy soils and the non-filled circles denote upland, marsh, forest, bamboo and bushland soils.

Fig. 4.3b shows that there is only a weak relationship between the relative abundance of GDGT-0 and TEX_{86} (logarithmic r = -0.67, r^2 = 0.45, $p < 0.0001$). However, both the TEX_{86} and the GDGT-0/crenarchaeol ratio show clear differences in soils under paddy (grey background in Fig. 4.3b) and upland management for adjacent sites suggesting that they may be used to determine anoxic or oxic conditions in soils. In general, paddy soils plotted within a field characterized by GDGT-0/crenarchaeol ratios > 2 and TEX_{86} values < 0.6 (Fig. 4.3b),

possibly denoting a diagnostic area for the abundance of methanogenic archaea. The GDGT-0/crenarchaeol ratio also differs between the various paddy soils, with exceptional high ratios in the Philippine Ifugao and Vietnamese Lào Cai soil (Table S4.1). At these sites, longer flooding periods (> 5 month per year) compared to Chinese and Indonesian soils are the likely explanation for the high ratios.

Distribution of branched GDGTs in soils

In the soils investigated here, the relative proportion of brGDGTs within the total GDGT pool was high and varied from 65.0 to 99.1% (Table 4.1). Forest soils generally contained the highest abundances of brGDGTs (> 92%), while they were significantly lower in upland and paddy soils (Fig. 4.4a). Pearson's correlation analysis indicated that the SOC content was not related to the relative abundance of brGDGT ($r = 0.22$, $r^2 = 0.05$, $p < 0.01$).

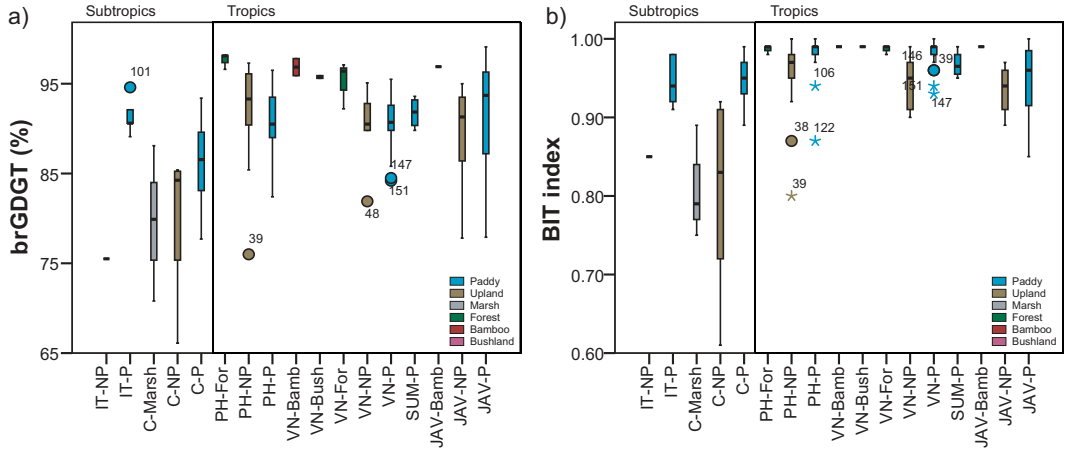


Fig. 4.4. Box-plot diagrams of **(a)** relative proportion of brGDGT in the total GDGT pool and **(b)** the BIT index in soil. Note different symbols (circle or asterisk) for outliers that are more than 1.5 (or 3) box lengths from one hinge of the box. Abbreviations and subdivisions as in Fig. 4.2.

In general, the tetra-methylated GDGT-Ia was the most abundant brGDGT in acidic soil and was the only brGDGT to increase in relative abundance with decreasing pH ($r = -0.75$, $r^2 = 0.56$, $p < 0.001$; Fig. 4.5). All other brGDGTs increased in relative abundance with pH ($p < 0.001$; Table S4.2), with the highest correlations observed for GDGT Ib ($r = 0.83$, $r^2 = 0.69$),

GDGT IIb (r = 0.79, r^2 = 0.62) and GDGT IIIb (r = 0.71, r^2 = 0.50). Our results thus suggest that especially the monocyclization of brGDGT is strongly controlled by pH (r = 0.86, r^2 = 0.74, $p < 0.001$) with alkaline conditions favouring the synthesis of brGDGT with one cyclopentane moiety (Fig. 4.5). Similar observations have previously been made in a set of globally distributed upland soils (Weijers et al., 2007; Peterse et al., 2012).

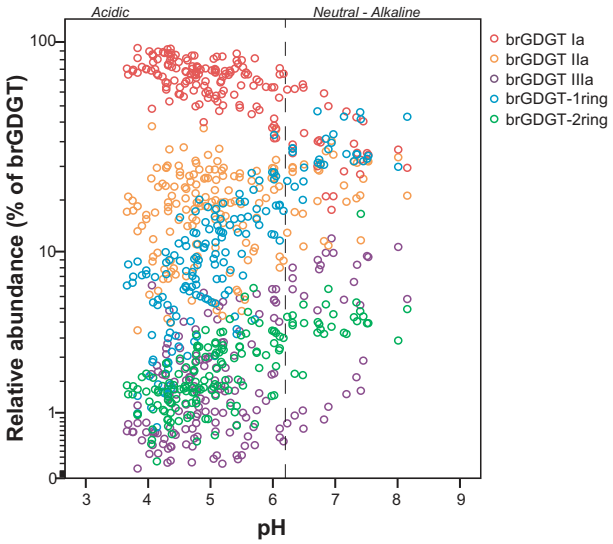


Fig. 4.5. Relative abundance of brGDGT plotted versus measured soil pH. Note logarithmic scale for relative abundance. Dotted line separates acidic from neutral/alkaline soil.

Weijers et al. (2007) explained the lower number of cyclopentyl moieties in brGDGT as a protection mechanism of bacterial cell membranes within acidic soils. The decrease in the amount of cyclopentyl moieties in brGDGT is associated with a decrease in membrane permeability, which regulates the internal pH of bacteria under acidic conditions (Weijers et al., 2007). In soils investigated here, the CBT ratio varied between -0.04 to 2.13 (Table 4.1) and showed a negative correlation with increasing soil pH (r = -0.81, r^2 = 0.65, p < 0.001; Fig. 4.6a). In neutral to alkaline soils (with pH values > 6.5) CBT values stayed rather constant with an offset observed between paddy soils (mean 0.34) and upland soils (mean -0.01; Fig. 4.6a). Wang et al. (2014) also found no apparent correlation between pH and CBT in alkaline soils in a study of arid and subhumid Chinese soils. However, a predominant dependency of CBT with soil water content and the mean annual precipitation (MAP) was observed (Wang et al., 2014). In our study, soil moisture could be one potential factor for the varying CBT values in paddy and upland soil, especially under alkaline conditions (Fig. 4.6a).

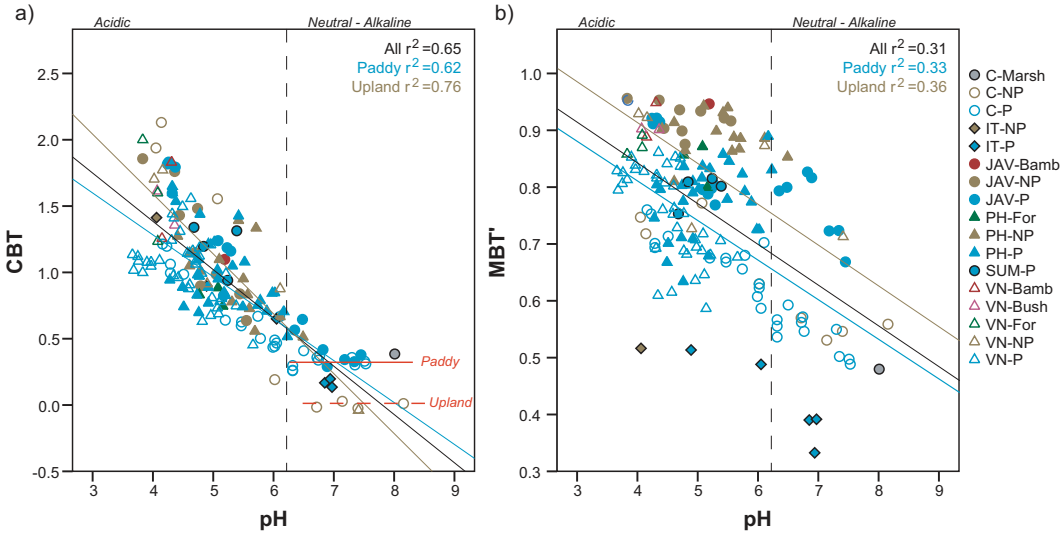


Fig. 4.6. Plot of **(a)** the cyclization ratio of branched tetraethers (CBT) versus soil pH and of **(b)** the revised methylation index of branched tetraethers (MBT') versus soil pH. Dotted line separates acidic from neutral/alkaline soil. Regressions line of all soils is coloured in black, the line of upland, marsh, forest, bamboo and bushland soils is brown and the line for paddy soils is blue. Abbreviations as in Fig. 4.2. Red lines in (a) show the offset between paddy and upland soil, which have > 6.2 pH values.

The degree of methylation of brGDGTs (MBT') has previously been shown to correlate with MAT and pH (Weijers et al., 2007; Peterse et al., 2012). Our results demonstrate that the MBT' generally shows low values in paddy soils compared to the adjacently located upland soils, except for the Chinese soils of Cixi (Table 4.1). The difference in MBT' between soils from the same sampling area denotes a lower influence of MAT on the MBT' than on the pH, which was weakly related to the MBT' ($r = -0.55$, $r^2 = 0.31$, $p < 0.001$; Fig. 4.6b). The MBT' was mainly controlled by the relative abundance of GDGT-Ia and GDGT-IIa, both of which were strongly related to MAP (Peterse et al., 2012). As the latter is largely similar at adjacent sites, we consider the paddy soil specific management techniques, including periodically flooding of soils, as responsible for the low GDGT-Ia and high GDGT-IIa content in paddy soils compared to upland soils (Table S4.1). This indicates that moisture is an important environmental variable affecting the distribution of brGDGT in soil. Moisture is also known to affect soil temperature, in particular in surface soils. Indeed, calculated T_{MC} values were generally lower in paddy soils compared to the adjacent upland soils (Table 4.1), indicating that temperature denotes more the mean annual soil temperature.

A recently developed method separates the structural isomers of brGDGTs with their methyl groups located in positions 5 and 6 (De Jonge et al., 2013). De Jonge et al. (2014) showed that the new CBT_{5ME}, calculated without 6-methyl brGDGTs, to correlate stronger with soil pH than the regular CBT, which includes both isomers, the 5- and 6-methyl brGDGTs. In addition, they found no correlation between pH and the newly developed MBT'_{5ME}, which is calculated without the 6-methyl isomer but a stronger correlation of MBT'_{5ME} with MAT. De Jonge et al. (2014) thus demonstrated that co-elution of GDGTs can affect estimation of pH values. Conventional methods as applied in this study, use a Prevail cyano column upon HPLC-MS analysis, which does not separate these isomers. Therefore, it is possible that our CBT-based pH reconstruction revealed some scatter (Fig. 4.6a) due to the presence of unresolved 5- and 6-methyl brGDGTs. The overall covariation of CBT and pH, however, was unaffected by this co-elution.

Influence of management systems on GDGT distribution

The BIT index quantifies the relationship between acyclic brGDGTs and crenarchaeol and has been used previously to determine the input of terrestrially derived organic matter to marine and lake environments (Hopmans et al., 2004; Weijers et al., 2007). The interpretation of BIT values in soil is not that straight forward as crenarchaeol originates from terrestrial *Thaumarchaeota* with less well constrained crenarchaeol abundances. Wang et al. (2013) observed a positive correlation between increasing soil water content and BIT values in Chinese marsh soils. In our sample set, the BIT index was slightly higher in paddy soils than in the adjacent upland soils (Fig. 4.4b). Furthermore, higher values were observed generally in paddy soil from tropic (1.02-1.04 fold) compared to subtropic (1.07-1.11 fold) locations. In contrast to the general trend, we found highest BIT values (1.27 fold) in the subtropical paddy soil of the Chinese Cixi location. In this area the BIT values in marsh and upland soil (0.61-0.89) were comparatively low, indicating that the latter have a mixed lipid composition with crenarchaeol originating predominantly from the residual parent substrate (tidal wetland sediment) and in smaller quantities also from the current microbial soil community. Comparable results were observed in a study of the plant wax lipids, which confirm the mixed lipid composition in these soils (Mueller-Niggemann and Schwark, 2015). Despite the higher

contribution of crenarchaeol to the marsh soils, our results show that brGDGT producing bacteria clearly dominate over *Thaumarchaeota* in all of the investigated soil types. Interestingly, crenarchaeol producing *Thaumarchaeota* seem to be more abundant in upland soils compared to forest and periodical flooded paddy soils (Fig. 4.4b). This is the opposite to results of an 152 day experimental study, with a higher production rate of crenarchaeol in soils that were incubated with different types of water (river, ocean or distilled water) to simulate the development of an aquatic environment under aerobic conditions (Peterse et al., 2015). Low redox conditions as assumed for paddy soils may thus lead to an enrichment of brGDGTs either by higher production or increased preservation of brGDGTs compared to crenarchaeol in wetland soils.

PCA was performed to obtain information on the major factors that control the variability of the distribution of iGDGTs and brGDGTs. Results of this analysis indicate that crenarchaeol exerts a major iGDGT in upland soils (Fig. 4.7a). The component loading score of GDGT-0 is opposite to crenarchaeol and has the highest negative score in PC1. In general, soils can be sorted into two groups on the basis of their scores on the first component. Paddy soils load negatively and all other soils load positively on PC1. Paddy soils that plot in the quadrant of upland soils are characterized by a higher intensity of crop-rotation with upland crops on the fields. The iGDGT composition of periodically flooded paddy soils is mainly controlled by GDGT-0 and that of non-paddy upland soils by crenarchaeol derived from *Thaumarchaeota*. In flooded rice paddy soils, oxygen availability determines the development of microbial consortia adapted to more anoxic conditions such as GDGT-0 synthesizing methanogenic archaea (Koga et al., 1998; Koga and Morii, 2005). The variance on PC2 is mainly associated with the relative abundance of GDGT-2 and separating forest and bushland soils from all other soils. The larger scatter of paddy soils on PC2 is explained by the number of rice cultivation cycles per year, which apparently influence the GDGT-2 contents significantly (Fig. 4.7b). Methanogenic archaea were found to be phylogenetically related to ANME archaea (Krüger et al., 2003; Shima et al., 2012). ANME archaea are a well known source of iGDGTs (including GDGT-2) in natural environments (Pancost et al., 2001; Blumenberg et al. 2004). Both, the interaction of methanogenic and methanotrophic archaea as well as the fact

that ANME are an abundant source of GDGT-2, could explain the relationship between higher numbers of rice cultivation cycles, which induce increased methanogenesis through abundant redox cycling, and the presence of GDGT-2. MAT and MAP had no obvious influence on discrimination of agricultural soil via iGDGT distribution (Fig. S4.3).

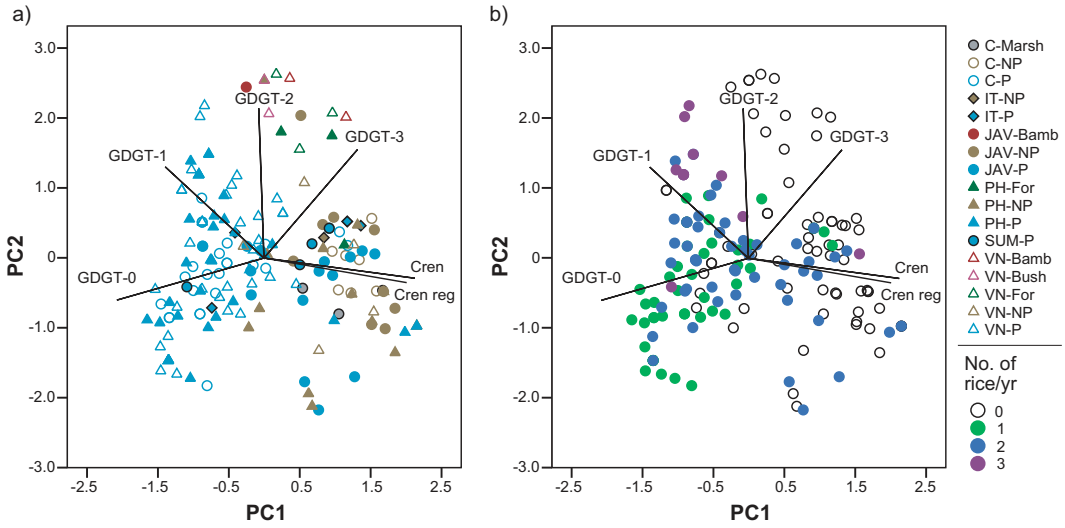


Fig. 4.7. Principal component analysis (PCA) based on standardized relative abundances of six iGDGTs in 170 investigated soils. The first principal component (PC1) accounted for 53.9% of the total variance and the second (PC2) for 29.9%. **(a)** Symbols and colours denote different management forms. Abbreviations as in Fig. 4.2. **(b)** The sample site symbols are indicative of the number of rice cultivation cycles per year.

PCA analysis on the relative abundances of brGDGT shows an opposite relation of GDGT-Ia to all other brGDGTs, with the highest component loading score on PC1 for GDGT-Ia (Fig. 4.8). The cyclopentane containing GDGT-IIb and -IIIb plot negatively on PC1. Higher contents of GDGT-Ia in upland soils compared to adjacent paddy soils (Table S4.1) confirm that tetra-methylated brGDGTs may be useful in separating different agricultural soils. GDGT-IIa has the lowest loading score on PC1 but the highest on PC2. Upland soils load separately from paddy soils along the PC2 with variation of relative abundance of the cyclic GDGT-Ib and GDGT-Ic playing the most important role. In contrast, paddy soils are mainly influenced by the abundance of GDGT-IIa and GDGT-IIIa, which both show only a low correlation with pH (Table S4.2). We rather, assume their dependency on soil moisture. The first PC, explaining 69.11% of the variance, indicates a separation between locations, with a

strong negative score in subtropical Italian and Chinese soils and more positive scores in soils originating from the tropics (Fig. 4.8a). The MAP (Fig. 4.8b) and MAT (Fig. S4.4) gradients of sampling locations on PC1, confirms a relation of climatic parameters to the variation of acyclic brGDGTs.

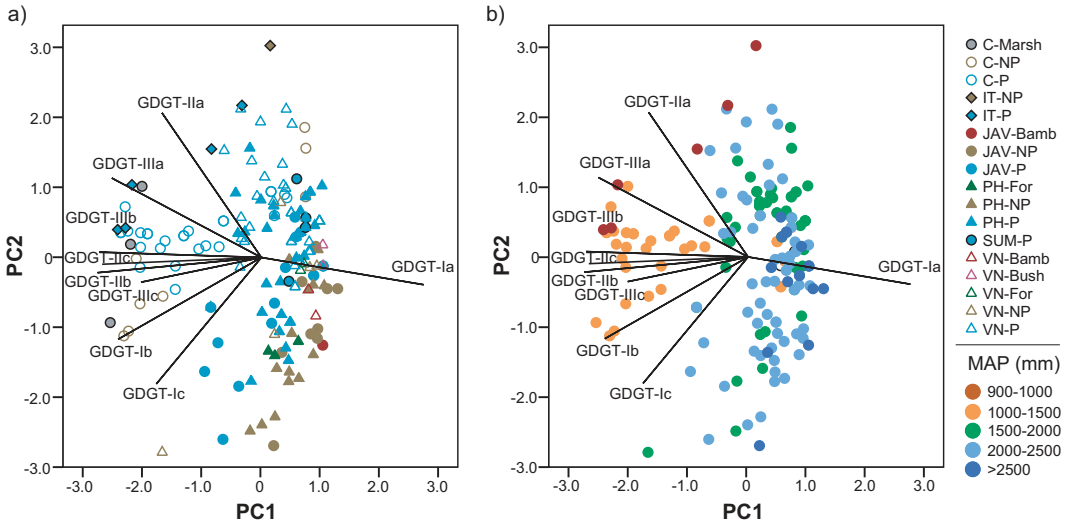


Fig. 4.8. Principal component analysis (PCA) based on standardized relative abundances of nine brGDGTs in 170 investigated soils. The first principal component (PC1) accounts for 69.1% of the variance and the second (PC2) for 14.3%. **(a)** Symbols and colours denote different management forms. Abbreviations as in Fig. 4.2. **(b)** The sample site symbols are indicative of the mean annual precipitation.

PCA analysis on environmental parameters as well as on indices of bacterial and archaeal GDGTs indicated that separation of paddy and upland soil is mainly controlled by the intensity of methanogenesis (Fig. 4.9a). The GDGT-0/crenarchaeol ratio and the BIT index had the highest positive loading score on PC2. The SOC and TN loaded in the same quadrant as the BIT index, suggesting that a positive correlation between the amount of organic matter and acyclic brGDGT, especially in paddy soils, prevailed. Alternating anoxic conditions in paddy soils are known to favour the preservation and therefore the accumulation of organic matter (Lal et al., 2002), which could lead to an increase of heterotrophic and brGDGT producing bacteria. In general, the CBT loaded opposite of the soil pH on PC1, indicating

their negative relation to each other. The internal separation of paddy soils via the number of rice cultivation cycles is evident by high loading scores of the CBT and MBT' (Fig. 4.9b).

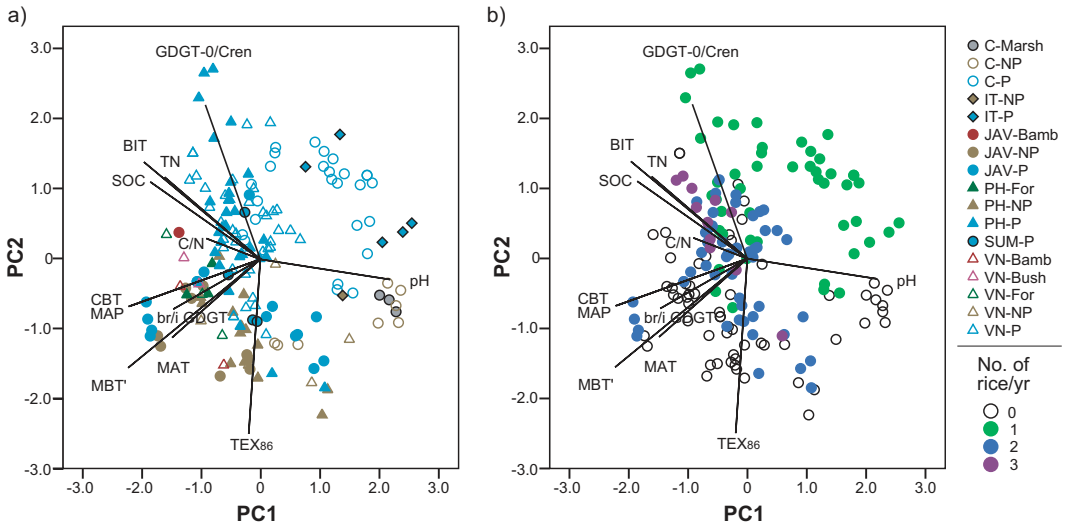


Fig. 4.9. Principal component analysis (PCA) based on commonly used indices and ratios for the 170 investigated soils. The first principal component (PC1) accounts for 33.5% of the variance and the second (PC2) for 21.4%. **(a)** Symbols and colours denote different management forms. Abbreviations as in Fig. 4.2. **(b)** The sample site symbols are indicative of the number of rice cultivation cycles per year.

Apparently, the increase of the MBT' is linked with the number of rice cycles, and therefore with lowering of penta- and hexa-methylated brGDGT during increasing redox cycles. Similar loading scores as well as similar directions of climatic parameters, such as MAP and MAT, and of CBT and MBT' also indicated a linkage to each other. In addition to methanogenesis, differences in MAT and soil water content seemed to be secondary factors controlling the distribution of brGDGT in soils, which also allowed a separation between upland and paddy management. It should be considered though that MAT is not identical to MST as the latter was also affected by e.g. the albedo and soil management, which can be different in the adjacent soils (Liu et al., 2014; Awe et al., 2015 and references therein). The reflection coefficient of the surface differs in agricultural soils as a consequence of management practises, which influence the soil bulk density (via tillage), the plant cover (function of the crop leaf area index) and the soil water content. For example, Awe et al.

(2015) found differences in soil temperature as consequence of management practises with lower temperatures in soils under chiselling and conventional tillage compared to no-tillage.

Effects of long-term management on GDGT distributions

Changes in GDGT distribution within two Cixi chronosequences with different cropping systems, one under continuous non-flooded upland and the other under paddy management, indicated specific adaption processes during the long-term usage at each site. Marsh soils were the first soils to develop after the construction of dykes on tidal wetland sediments and therefore represent the starting point of the subsequent soil development. We observed high BIT values (~0.77) already in the surface horizon of the marsh soils, indicating the rapid adaption of the microbial community to more terrestrial conditions. A plot of the brGDGT/iGDGT ratio over time provides evidence for a dominance of brGDGT over iGDGT in all soils, with values of this ratio varying between 2 and 6 in upland soils (Fig. 4.10a). In contrast to paddy soils, which had a fourfold increase of the ratio after 2000 yr rice cultivation, this suggests an influence of long-term processes on the proportion of archaeal and bacterial soil microorganism. These processes may include desalinization, decalcification through leaching as shown in changes of soil pH values (Fig. S4.5a), fertilization activities, organic matter input and accumulation (Fig. S4.5b). Paddy soil management is known to strongly affect the accumulation of organic matter (Wu, 2011; Mueller-Niggemann et al., 2012; Kölbl et al., 2014) as the periodically anaerobic conditions result in a slower degradation of organic matter (Lal et al., 2002). Kölbl et al. (2014) investigated the response of redox dynamics to changing water conditions over a one year time period in 100, 700 and 2000 yr old paddy soils. They noted a change of the redox potential towards anoxic conditions, already after 5 days of flooding. After stabilization, the redox potential was in the same range in all soils (-170 to -200 mV), independent of the duration of paddy management. In upland soils, permanent oxic conditions were persistent throughout the time period investigated. Results of Kölbl et al. (2014) demonstrate that the rapid establishment of anoxic conditions and the long-term usage of paddy soils may lead to an increase of organic carbon concentrations over time.

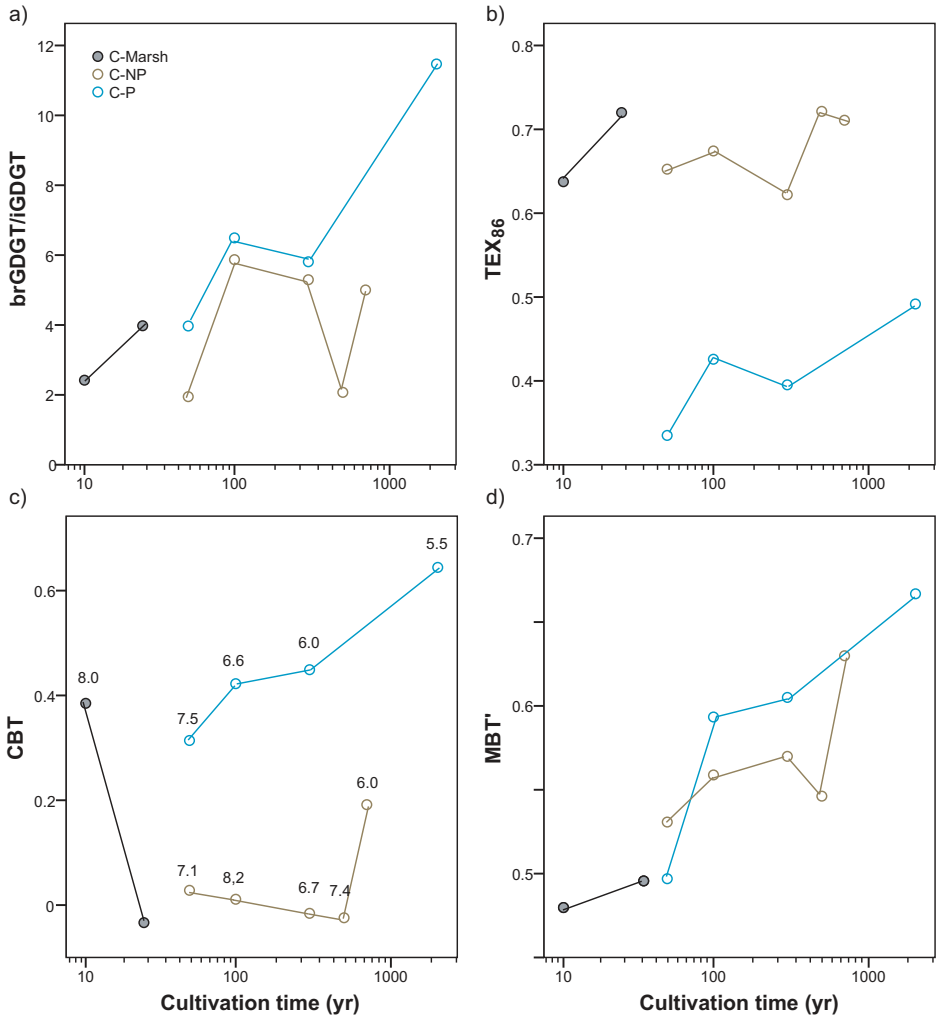


Fig. 4.10. Time plots of various GDGT ratios and indices in soils of the Chinese Cixi region: **(a)** ratio of branched *vs.* isoprenoid GDGTs, **(b)** the TEX_{86}, **(c)** the CBT and **(d)** MBT'. Note logarithmic scale for the cultivation time. Numbers in plot (c) reflect soil pH values.

Within the upland soil chronosequence, the TEX_{86} does not change significantly over the 700 yr cultivation time and averages 0.7 (Fig. 4.10b). In paddy soils, on the contrary, the TEX_{86} decreased from the initial marsh soil value of 0.7 to values of 0.3 within only 50 yr of paddy management. Rotation between paddy- and upland-type of cultivation resulted in a high TEX_{86} value of 0.5 in the 2000 yr-old paddy soils (Fig. 4.10b). Our results thus suggest that management systems significantly affect the microbial soil community. Long-time paddy management also led to the successive increase of ammonia-oxidizing *Thaumarchaeota* based

on high relative abundances of crenarchaeol, indicating either a recovering process of water-stressed soil *Thaumarchaeota* or the enrichment of fossil crenarchaeol. The latter is potentially explainable by the management type used in the Cixi area, with one wetland rice season and one dry inter-crop season per year that influence the presence of aerobic and anaerobic microbes in these paddy soils. In particular, the periodically anaerobic conditions may result in a slower degradation of organic matter (Lal et al., 2002). GDGTs may originate from a mixed source of microbial membrane lipids that were recently deposited (during the oxic as well as in the anoxic period) additionally to the previously preserved ones. Thus, higher proportions of crenarchaeol e.g. as marker for terrestrial ammonia oxidizers, being active during the oxic inter-crop period, were detected but in lower amounts as commonly observed in upland soils (Table S4.1). At the same time, the proportion of methanogenic archaea, which was estimated by using GDGT-0/crenarchaeol ratio, decreased during the long-term paddy management from 5.0 in the 50 yr to 2.8 in the 2000 yr old paddy soil.

The pH values ranged between 8.0 in marsh soil and 5.5 in the 2000 yr paddy soil. The paddy management (including flooding practises) thus leads to enhanced decalcification of soils compared to the non-flooded upland management. However, most soils have an alkaline or neutral pH with exceptions of the 700 yr upland soil and the 2000 yr paddy soils, which all had pH values < 6.5 (Fig. S4.5a). It has previously been demonstrated that the CBT is negatively correlated with increasing pH values (Weijers et al., 2007; Peterse et al., 2012). In the alkaline soils of the Cixi chronosequences a negative correlation was also observed, which was higher for paddy soils ($r = -0.94$, $r^2 = 0.88$, $n = 4$, $p < 0.001$) than for upland soils ($r = -0.69$, $r^2 = 0.47$, $n = 5$, $p < 0.001$). Interestingly, an offset of CBT values between paddy and upland soils with no apparent changes during cultivation time was noted (Fig. 4.10c). In addition, the CBT was higher in the younger of both marsh soils, probably because of the greater soil water content in the ~10 yr old compared to the ~35 yr old marsh soil as result of the progressive dewatering during marsh soil pedogenesis. The observation for the CBT values supports the idea that soil moisture in addition to pH controls the degree of cyclization of brGDGTs under alkaline conditions; possibly as a reaction to water stress or oxygen deprivation on microorganisms. The increase of CBT values in acidic soils (Fig. 4.10c) also

suggests that low soil pH results in the increased synthesis of brGDGTs with no cyclopentyl moieties.

Except for the youngest paddy soils (50 yr), the MBT' was slightly lower in Cixi upland soils compared to their corresponding paddy soils with identical cultivation time (Fig. 4.10d). This is in contrast to the observations that paddy soils in general showed a lower MBT' compared to the adjacent upland soils (Fig. 4.6b). This may indicate that soil bacteria living under contrasting pH regimes adapt the composition of their membrane lipids in a different fashion, even if the agricultural management is comparable.

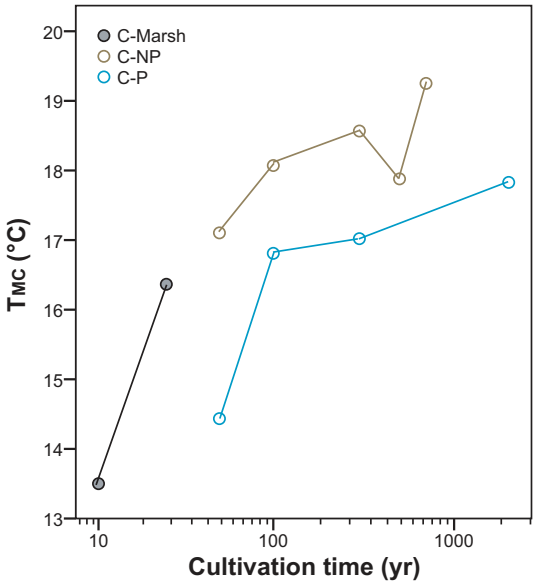


Fig. 4.11. Time plot of MBT'-CBT derived temperatures (T_{MC}) in soils of the Chinese Cixi. Note logarithmic scale for cultivation time.

The CBT and MBT' are both considered to be strongly related to MAT (Weijers et al., 2007; Peterse et al., 2012), which is largely similar for paddy and upland soils from the same sampling region. However, the calculated T_{MC} was different in adjacent paddy and upland soils (Table S4.1) and gradually increased during long-term management in both chronosequences (Fig. 4.11) from 14.4 °C to 17.8 °C in paddy soils and from 17.1 °C to 19.3 °C in upland soils, respectively. In general, temperatures were approximately 1.4°C higher in upland soils compared to soils under paddy management with the same cultivation time. This

implies that the management type affects the MST, which in turn controls the membrane lipid composition of brGDGT producing bacteria.

4.5 Conclusions

Our results show that archaeal and bacterial GDGTs were ubiquitously distributed in paddy, upland, forest, bushland and marsh soils of tropical and subtropical climate regimes. Independent of the soil usage, the brGDGTs predominated over iGDGTs in all soils, but had lower relative proportions in soils located in the subtropics compared to soils in tropical latitudes. This implies that warm and humid environments favour the growth of bacteria that produce brGDGT.

Agricultural management was a major factor that controlled the distribution of the archaeal community in soils. Biomarker for methanogens were enhanced in subaquatic paddy soils compared to predominantly thaumarchaeal ammonia oxidation in dry upland soils. In addition, the number of -or a long-term duration of- rice cultivation cycles per year significantly affected the composition of iGDGT with an increase of the GDGT-0/crenarchaeol ratio in soils with a higher number of cultivation cycles.

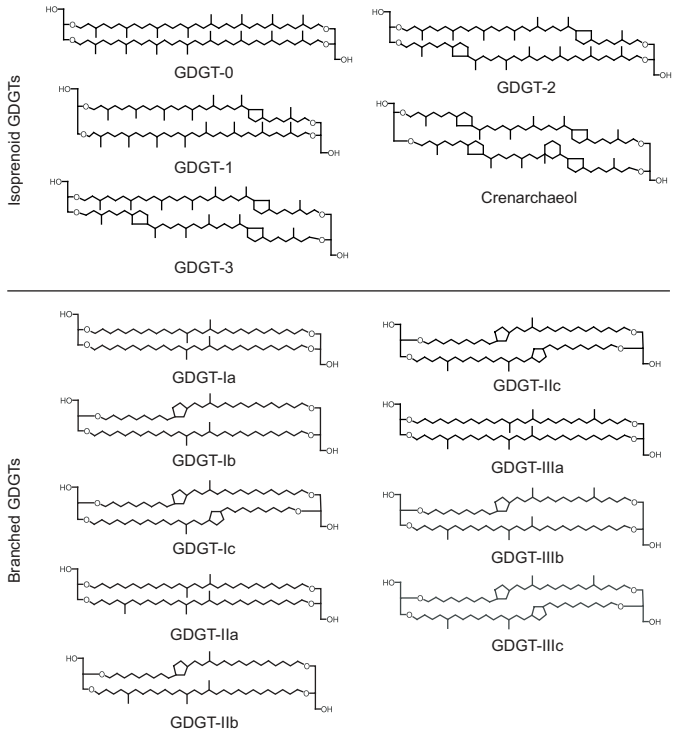
CBT values were correlated with soil pH and were controlled by a predominance of acyclic brGDGT in acidic soils. In alkaline soils, CBT values were rather invariant but the offset between soils under periodical flooding (paddy soils) and soils under non-flooded upland management suggests that parameters other than pH affected the distribution of brGDGTs as well (e.g. soil moisture that in addition to soil pH and MAT exerts a control on the degree of cyclization of brGDGTs). MBT' values differed in adjacent paddy and upland soils, confirming that other factors than MAT and MAP affect the degree of methylation of brGDGT on a regional scale. brGDGT-based temperatures (T_{MC}) were higher in soils under upland management than under paddy management and these differences in T_{MC} suggest that the specific management influenced the soil moisture, which in turn affects MST. The results of the Cixi chronosequence covering 2000 yr soil development confirm that the SOC, the pH value and the soil moisture controlled the distribution of brGDGT during long-term paddy soils usage.

4.6 Acknowledgements

We thank the German Research Foundation (DFG) for financial support (Schw554/20). Asian and European partners of Research Initiative FOR 995 as well as of the LEGATO project are thanked for field work collaboration. The anonymous reviewers are thanked for constructive comments.

4.7 Appendix A.

Chemical structures of branched GDGTs (brGDGTs) and isoprenoid GDGTs (iGDGTs) investigated in this study.



4.8 Appendix B. Supplementary data

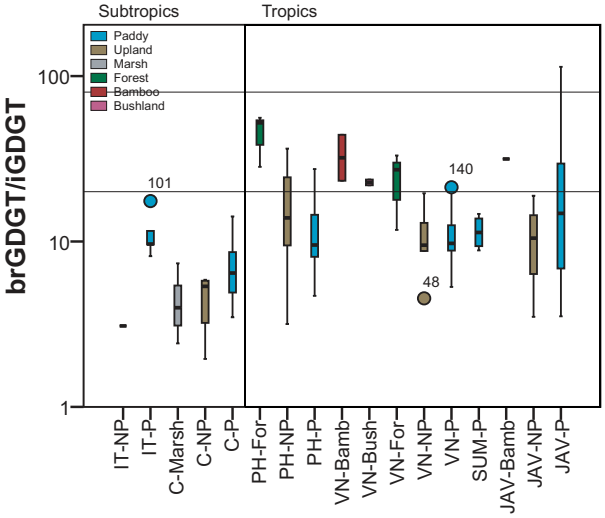


Fig. S 4.1. Box-plot diagrams of brGDGT/iGDGT ratio in soil. Note different symbols (circles or asterisk) for outliers that are more than 1.5 (or 3) box lengths from one hinge of the box. Abbreviations refer to different sampling locations: Italy (IT), China (C), Philippines (PH), Vietnam (VN), Sumatra (SUM) and Java (JAV). The vertical line separates subtropical from tropical locations. Numbers in all plots indicate samples listed in Table S4.1.

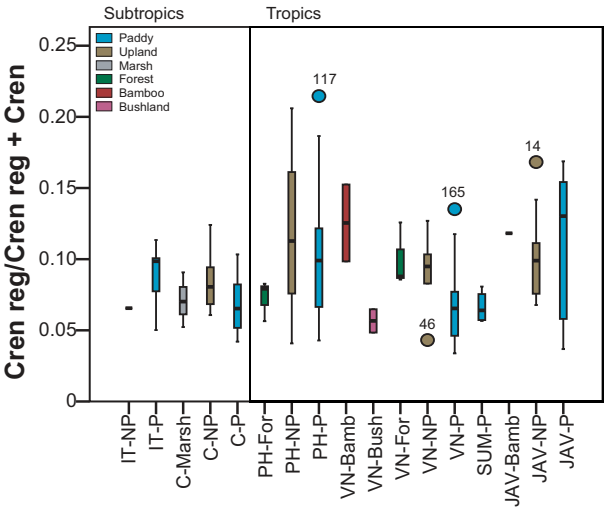


Fig. S 4.2. Box-plot diagrams of cren reg/cren reg + cren ratio in soils. Note different symbols (circles or asterisk) for outliers that are more than 1.5 (or 3) box lengths from one hinge of the box. Abbreviations refer to different sampling locations: Italy (IT), China (C), Philippines (PH), Vietnam (VN), Sumatra (SUM) and Java (JAV). The vertical line separates subtropical from tropical locations. Numbers in all plots indicate samples listed in Table S4.1.

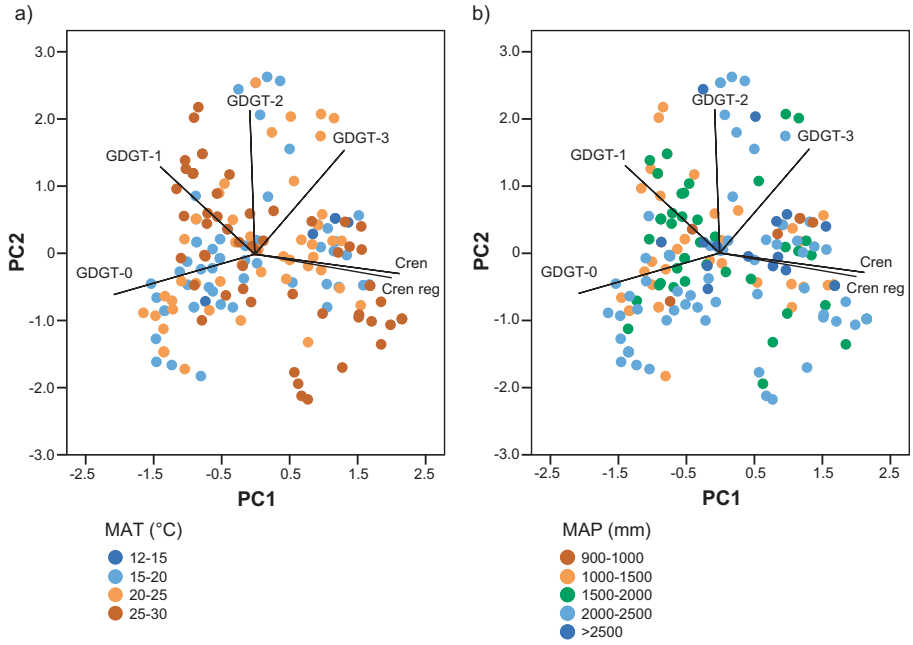


Fig. S 4.3. Principal component analysis (PCA) based on standardized relative abundances of six iGDGTs in 170 investigated soils. The first principal component (PC1) accounted for 53.9% of the total variance and the second (P2) for 29.9%. **(a)** The sample site symbols are indicative of the mean annual air temperature (MAT). **(b)** The sample site symbols are indicative of the mean annual precipitation (MAP).

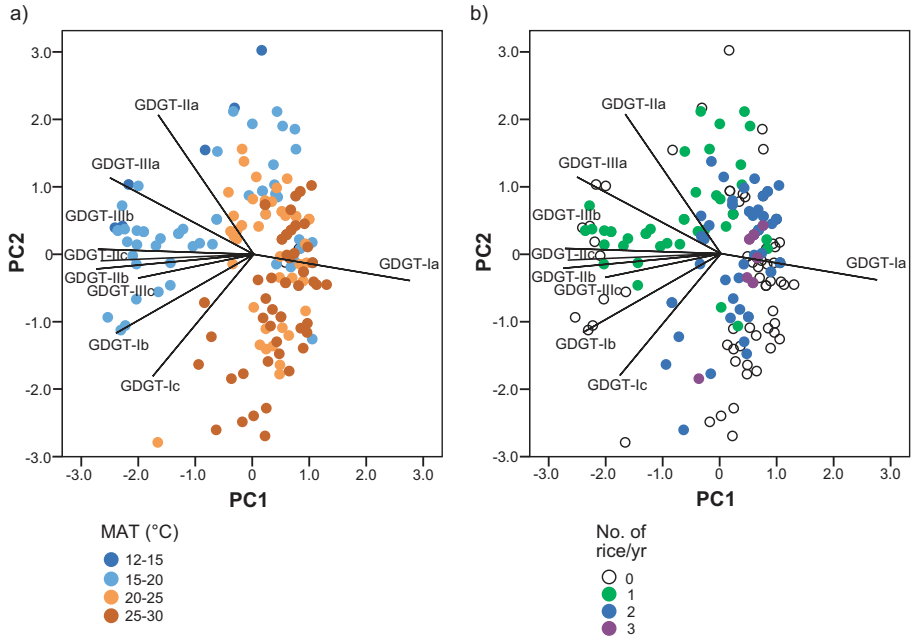


Fig. S 4.4. Principal component analysis (PCA) based on standardized relative abundances of nine brGDGTs in 170 investigated soils. The first principal component (PC1) accounts for 69.1% of the variance and the second (PC2) for 14.3%. **(a)** The sample site symbols are indicative of the mean annual air temperature (MAT). **(b)** The sample site symbols are indicative of the number of rice cultivation cycles per year.

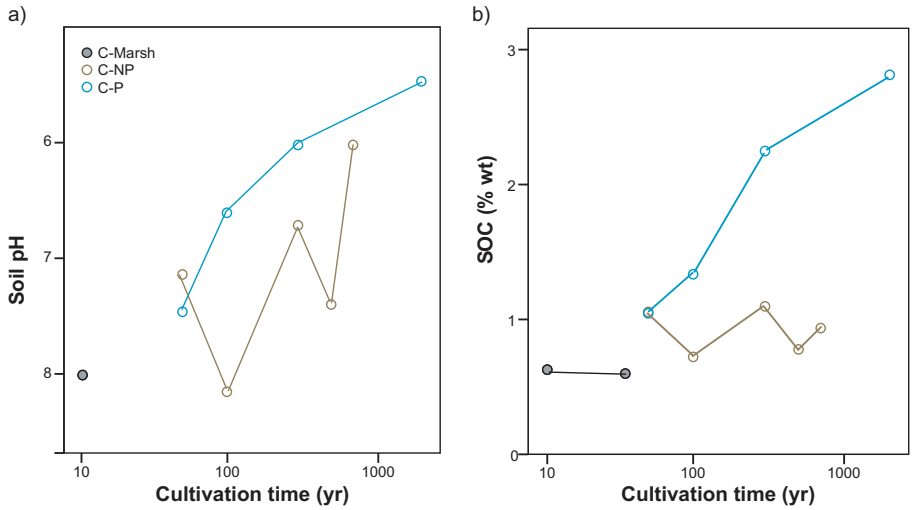


Fig. S 4.5. Time plots of **(a)** soil pH and **(b)** soil organic carbon (SOC) content in Chinese soils of Cixi region.

Table S 4.1. Detailed list of 170 soil samples including information about location, soil type, land management, bulk parameter and relative abundances of iGDGT and brGDGT.

ID	Dataset code	Country	Sampling Area	Grad N	Grad E	Soil type	Vegetation	No. of rice/ year	No. of crops/ year	MAT (°C)	MAP (mm)	Altitude (m.a.s.l)	SOC (%)	soil pH	iGDGT (% GDGT)	brGDGT (% GDGT)	brGDGT/ iGDGT	CBT	MBT'	T_{MC} (°C)
51	IT-NP	Italy	Zeme	45°11.555	8°40.104	upland	maize			12.5	954	80	0.73	4.1	25.1	74.9	2.98	1.41	0.52	8.8
4	C-NP	China	Cixi	30°13.152	121°21.382´	upland	water melon, cotton, vegetables			16.6	1266	5	1.06	7.1	35.0	65.0	1.86	0.03	0.53	17.1
5	C-NP	China	Cixi	30°11.884	121°21.196´	upland	melons			16.6	1266	5	0.72	8.2	15.2	84.8	5.56	0.01	0.56	18.1
6	C-NP	China	Cixi	30°06.932´	121°30.646´	upland	grassland			16.6	1266	5	1.10	6.7	16.5	83.5	5.06	-0.02	0.57	18.6
7	C-NP	China	Cixi	30°11.680´	121°05.078´	upland	soybeans with fruit trees			16.6	1266	5	0.78	7.4	33.5	66.5	1.98	-0.02	0.55	17.9
8	C-NP	China	Cixi	30°10.967´	121°08.706´	upland	cotton-rape			16.6	1266	5	0.94	6.0	17.6	82.4	4.68	0.19	0.63	19.3
9	JAV-NP	Indonesia	Sukabumi	S 06°52.029	E 106°56.725	upland	cassava/ maize rotation			23.5	2806	981	4.34	4.4	22.9	77.1	3.37	1.43	0.90	20.7
10	JAV-NP	Indonesia	Sukabumi	S 06°52.034	E 106°56.725	upland	cassava/ maize rotation			23.5	2806	969	4.33	4.7	16.6	83.4	5.02	1.48	0.90	20.3
11	JAV-NP	Indonesia	Sukabumi	S 06°52.046	E 106°56.714	upland	cassava/ maize/ chilli rotation			23.5	2806	972	3.50	4.8	13.6	86.4	6.38	0.90	0.88	22.9
12	JAV-NP	Indonesia	Jasinga	S 06°32.158	E 106°31.030	upland	cassava/ maize/ chilli rotation			26.9	3252		3.22	3.8	9.1	90.9	10.01	1.86	0.96	19.9
13	JAV-NP	Indonesia	Jasinga	S 06°32.212	E 106°31.033	upland	cassava (8month rotation), banana, tuber, vegetables			26.9	3252	236	2.97	4.4	5.7	94.3	16.66	1.76	0.95	20.4
14	JAV-NP	Indonesia	Jasinga	S 06°32.239	E 106°31.027	upland	cassava (8month rotation), banana, tuber, vegetables			26.9	3252	238	2.08	5.6	5.6	94.4	16.78	0.64	0.92	25.6
15	JAV-NP	Indonesia	Ngawi	S 07°26.691	E 111°36.667	upland	sugar cane			27	2034	80	1.46	4.7	14.2	85.8	6.03	1.15	0.94	23.3
16	JAV-NP	Indonesia	Ngawi	S 07°26.719	E 111°36.672	upland	sugar cane			27	2034	78	1.48	5.1	6.9	93.1	13.41	0.99	0.93	24.2
17	JAV-NP	Indonesia	Ngawi	S 07°26.640	E 111°36.530	upland	sugar cane			27	2034	80	1.74	5.4	8.5	91.5	10.71	0.84	0.92	24.6
18	C-NP	China	Red Soil Station	N 28°14.035	E 116°53.784	upland	rape			18.5	1731	44	0.83	4.1	15.5	84.5	5.47	2.13	0.72	11.0
19	C-NP	China	Red Soil Station	N 28°14.029	E 116°53.830	upland	rape			18.5	1731	44	0.70	5.1	15.5	84.5	5.45	1.56	0.77	15.9
20	C-NP	China	Red Soil Station	N 28°14.029	E 116°53.830	upland	rape			18.5	1731	44	0.85	4.1	16.0	84.0	5.24	1.94	0.75	13.0
32	PH-NP	Phillipines	Laguna Los Banos	14,22921	121,33537	upland	fruit trees			27.1	2064	10	1.77	5.5	4.9	95.1	19.46	0.95	0.94	24.5
33	PH-NP	Phillipines	Laguna Los Banos	14,11583	121,41298	upland	squash field			27.1	2064	287	1.90	5.1	10.0	90.0	9.05	1.10	0.94	23.8
34	PH-NP	Phillipines	Laguna Los Banos	14,21248	121,33754	upland	fruit trees			27.1	2064	14	2.09	5.4	4.0	96.0	23.77	1.39	0.93	21.8
35	PH-NP	Phillipines	Laguna Los Banos	14,13992	121,39901	upland	coconut plantation			27.1	2064	192	2.17	5.7	8.1	91.9	11.30	0.56	0.87	24.5
36	PH-NP	Phillipines	Laguna Los Banos	14,18689	121,36572	upland	coconut plantation			27.1	2064	32	1.93	5.6	8.4	91.6	10.89	0.73	0.89	24.2
37	PH-NP	Phillipines	Nueva Ecija	15,67066	120,84277	upland	rice/vegetable			27.1	1821	46	1.30	5.7	6.7	93.3	13.96	1.33	0.89	20.7
38	PH-NP	Phillipines	Nueva Ecija	15,59845	120,91125	upland	onion			27.1	1821	45	0.88	6.5	15.3	84.7	5.54	0.51	0.85	24.3
39	PH-NP	Phillipines	Nueva Ecija	15,57902	120,9112	upland	bitter gourd			27.1	1821	40	0.54	6.1	25.7	74.3	2.89	0.67	0.89	24.5
40	PH-NP	Phillipines	Nueva Ecija	15,61836	120,94529	upland	patola (Cucurbitaceae)			27.1	1821	52	1.14	4.6	15.1	84.9	5.64	1.15	0.91	22.6
41	PH-NP	Phillipines	Ifugao	16,90742	121,07405	upland	abandoned rice field			21.4	2376	947	1.21	5.3	2.9	97.1	33.93	0.78	0.90	24.3
42	PH-NP	Phillipines	Ifugao	16,9101	121,12833	upland	abandoned rice			21.4	2376	912	1.71	5.6	2.8	97.2	34.53	0.83	0.86	22.9

ID	Dataset code	Country	Sampling Area	Grad N	Grad E	Soil type	Vegetation	No. of rice/ year	No. of crops/ year	MAT (°C)	MAP (mm)	Altitude (m.a.s.l.)	SOC (%)	soil pH	iGDGT (% GDGT)	brGDGT (% GDGT)	brGDGT/ iGDGT	CBT	MBT'	T_{MC} (°C)
43	PH-NP	Phillipines	Ifugao	16,85995	121,09858	upland	abandoned rice field			21.4	2376	871	1.94	4.4	2.7	97.3	36.01	1.27	0.90	21.6
44	PH-NP	Phillipines	Ifugao	16,8586	121,10159	upland	abandoned rice field			21.4	2376	937	2.09	4.8	4.8	95.2	19.64	1.01	0.86	21.9
45	PH-NP	Phillipines	Ifugao	16,90534	121,06454	upland	abandoned rice field			21.4	2376	1074	1.58	4.6	7.3	92.7	12.62	1.05	0.81	20.0
46	VN-NP	Vietnam	Hai Duong	21,03488	106,3505	upland	tomato			24.1	1608		1.17	4.9	10.4	89.6	8.66	0.91	0.73	18.2
47	VN-NP	Vietnam	Hai Duong	21,03488	106,3505	upland	banana			24.1	1608		0.79	7.4	7.7	92.3	11.99	-0.04	0.71	23.1
48	VN-NP	Vietnam	Vinh Phúc	21,34784	105,70521	upland	vegetable field			23.6	1687	20	0.58	6.1	18.8	81.2	4.33	0.88	0.87	22.9
49	VN-NP	Vietnam	Vinh Phúc	21,37265	105,72034	upland	litchi/tea			23.6	1687	44	1.64	4.0	5.0	95.0	18.94	1.70	0.93	20.0
50	VN-NP	Vietnam	Vinh Phúc	21,31602	105,73824	upland	peanut field			23.6	1687	20	0.84	4.2	9.9	90.1	9.12	1.77	0.92	19.4
1	C-Marsh	China	Cixi	30,349196	121,195581	marsh				16.6	1266	3	0.63	8.0	29.8	70.2	2.36	0.38	0.48	13.5
2	C-Marsh	China	Cixi	30,26714	121,197167	marsh				16.6	1266	3	0.43	nd	12.4	87.6	7.08	0.20	0.47	14.2
3	C-Marsh	China	Cixi	30,315612	121,196849	marsh				16.6	1266	3	0.60	nd	21.1	78.9	3.75	-0.03	0.50	16.4
23	JAV-Bamb	Indonesia	Sumbermujer	E 08°08.285	113°00.836	bamboo	bamboo forest			17.8	2693	742	3.57	5.2	3.1	96.9	30.85	1.10	0.95	23.9
27	VN-Bamb	Vietnam	Vinh Phúc	21,34715	105,70399	bamboo	bamboo forest			23.6	1687	19	0.69	4.3	4.4	95.6	21.95	1.83	0.95	19.8
30	VN-Bamb	Vietnam	Lào Cai	22,30122	103,88925	bamboo	bamboo forest			16.2	2223	1010	2.97	4.2	2.3	97.7	43.02	1.26	0.89	21.2
21	VN-Bush	Vietnam	Lào Cai	22,41038	103,90191	bushland	bushland			16.2	2223	735	2.56	4.4	4.1	95.9	23.51	1.36	0.90	21.1
22	VN-Bush	Vietnam	Lào Cai	22,39546	103,84473	bushland	bushland			16.2	2223	1302	3.32	4.1	4.4	95.6	21.53	1.61	0.90	19.6
24	PH-For	Phillipines	Ifugao	16,8586	121,10159	forest	forest			21.4	2376	939	2.66	4.8	3.5	96.5	27.28	0.83	0.86	22.6
25	PH-For	Phillipines	Ifugao	16,90845	121,07412	forest	forest			21.4	2376	994	3.22	5.1	1.8	98.2	54.19	0.88	0.87	22.8
26	PH-For	Phillipines	Ifugao	16,92825	121,05477	forest	forest			21.4	2376	1299	2.38	5.2	1.9	98.1	50.91	0.74	0.80	21.4
28	VN-For	Vietnam	Vinh Phúc	21,31624	105,74211	forest	forest			23.6	1687	24	1.30	3.8	8.1	91.9	11.29	2.00	0.86	16.1
29	VN-For	Vietnam	Lào Cai	22,29854	103,9134	forest	forest			16.2	2223	1049	2.77	4.1	3.0	97.0	32.43	1.23	0.87	20.8
31	VN-For	Vietnam	Lào Cai	22,31435	103,85794	forest	forest			16.2	2223	1254	3.88	4.1	3.6	96.4	26.43	1.60	0.89	19.4
52	C-P	China	Cixi	30°11.031	121°21.366	paddy		1		16.6	1266	5	0.92	7.5	20.4	79.6	3.91	0.31	0.49	14.2
53	C-P	China	Cixi	30°11.031	121°21.366	paddy		1		16.6	1266	5	1.11	7.5	20.7	79.3	3.84	0.33	0.50	14.4
54	C-P	China	Cixi	30°11.031	121°21.366	paddy		1		16.6	1266	5	1.11	7.4	19.9	80.1	4.02	0.30	0.50	14.7
55	C-P	China	Cixi	30°09.827	121°20.97	paddy		1		16.6	1266	5	1.68	6.0	13.6	86.4	6.37	0.49	0.63	17.5
56	C-P	China	Cixi	30°09.827	121°20.97	paddy		1		16.6	1266	5	1.02	7.3	15.1	84.9	5.60	0.36	0.55	15.8
57	C-P	China	Cixi	30°09.827	121°20.97	paddy		1		16.6	1266	5	1.31	6.5	12.2	87.8	7.22	0.41	0.59	16.9
58	C-P	China	Cixi	30°06.437	121°30.28	paddy	rice-water melon	1		16.6	1266	5	2.25	6.0	13.6	86.4	6.35	0.44	0.62	17.6
59	C-P	China	Cixi	30°06.437	121°30.28	paddy	rice-water melon	1		16.6	1266	5	2.40	6.0	15.7	84.3	5.37	0.44	0.61	17.2
60	C-P	China	Cixi	30°06.437	121°30.28	paddy	rice-water melon	1		16.6	1266	5	2.10	6.1	15.0	85.0	5.66	0.47	0.59	16.4
61	C-P	China	Cixi	30°12.417	121°07.74	paddy	rice-rape	1		16.6	1266	5	1.31	6.3	17.3	82.7	4.79	0.26	0.57	16.9
62	C-P	China	Cixi	30°12.417	121°07.74	paddy	rice-rape	1		16.6	1266	5	1.50	6.3	17.0	83.0	4.87	0.26	0.54	16.0
63	C-P	China	Cixi	30°12.417	121°07.74	paddy	rice-rape	1		16.6	1266	5	1.30	6.3	16.3	83.7	5.12	0.30	0.56	16.3
64	C-P	China	Cixi	30°10.408	121°09.180	paddy		1		16.6	1266	5	2.05	6.8	21.8	78.2	3.60	0.37	0.55	15.6
65	C-P	China	Cixi	30°10.408	121°09.180	paddy		1		16.6	1266	5	2.10	6.7	21.6	78.4	3.63	0.36	0.56	16.2
66	C-P	China	Cixi	30°10.408	121°09.180	paddy		1		16.6	1266	5	1.93	6.7	22.5	77.5	3.44	0.34	0.57	16.6

ID	Dataset code	Country	Sampling Area	Grad N	Grad E	Soil type	Vegetation	No. of rice/ year	No. of crops/ year	MAT (°C)	MAP (mm)	Altitude (m.a.s.l.)	SOC (%)	Soil pH	iGDGT (% GDGT)	brGDGT (% GDGT)	brGDGT/ iGDGT	CBT	MBT'	T_{MC} (°C)
67	C-P	China	Cixi	30°09.763	121°06.957	paddy		1		16.6	1266	5	1.09	5.5	10.6	89.4	8.42	0.60	0.67	18.3
68	C-P	China	Cixi	30°09.763	121°06.957	paddy		1		16.6	1266	5	1.19	6.1	10.8	89.2	8.25	0.67	0.70	18.8
69	C-P	China	Cixi	30°09.763	121°06.957	paddy		1		16.6	1266	5	0.95	5.8	10.9	89.1	8.17	0.50	0.68	19.0
70	C-P	China	Cixi	30°05.455	121°26.738	paddy		1		16.6	1266	5	2.83	5.7	8.8	91.2	10.43	0.67	0.66	17.3
71	C-P	China	Cixi	30°05.455	121°26.738	paddy		1		16.6	1266	5	2.73	5.2	7.7	92.3	11.98	0.64	0.68	18.1
72	C-P	China	Cixi	30°05.455	121°26.738	paddy		1		16.6	1266	5	2.88	5.5	7.9	92.1	11.71	0.63	0.67	18.0
73	JAV-P	Indonesia	Sukabumi	S 06°52.802	E 106°56.457	paddy	rice(Ciherang)-rice-Bok Choi	2	1	23.5	2806	870	4.37	5.2	5.5	94.5	17.02	1.19	0.80	18.8
74	JAV-P	Indonesia	Sukabumi	S 06°52.800	E 106°56.463	paddy	rice(Ciherang)-rice-Bok Choi	2	1	23.5	2806	871	4.41	5.3	6.1	93.9	15.26	1.16	0.77	18.1
75	JAV-P	Indonesia	Sukabumi	S 06°52.796	E 106°56.465	paddy	rice(Ciherang)-rice-Bok Choi	2	1	23.5	2806	871	4.02	5.1	5.6	94.4	16.75	1.24	0.80	18.5
76	JAV-P	Indonesia	Jasinga	S 06°32.205	E 106°31.062	paddy	rice(Ciherang)-rice-Maize	2	1	26.9	3252	248	1.97	4.3	1.4	98.6	71.57	1.83	0.91	18.7
77	JAV-P	Indonesia	Jasinga	S 06°32.209	E 106°31.062	paddy	rice(Ciherang)-rice-Maize	2	1	26.9	3252	240	2.06	4.2	0.9	99.1	109.48	1.82	0.92	19.1
78	JAV-P	Indonesia	Jasinga	S 06°32.218	E 106°31.059	paddy	rice(Ciherang)-rice-Maize	2	1	26.9	3252	240	2.30	4.3	2.0	98.0	48.79	1.60	0.92	20.3
79	JAV-P	Indonesia	Jasinga	S 06°32.218	E 106°31.059	paddy	rice(Ciherang)-rice-Maize	2	1	26.9	3252	238	2.15	4.4	0.9	99.1	112.04	1.79	0.91	19.0
80	JAV-P	Indonesia	Ngawi	S 07°26.878	E 111°36.599	paddy	rice(Ciherang)-rice-tobacco	2	1	27	2034	78	1.41	7.2	9.5	90.5	9.57	0.34	0.72	21.3
81	JAV-P	Indonesia	Ngawi	S 07°26.819	E 111°36.515	paddy	rice(Ciherang)-rice-tobacco	2	1	27	2034	79	1.81	6.4	6.8	93.2	13.63	0.57	0.79	22.2
82	JAV-P	Indonesia	Ngawi	S 07°27.095	E 111°36.576	paddy	rice(Ciherang)-rice-tobacco	2	1	27	2034	77	1.40	6.5	9.2	90.8	9.82	0.65	0.80	21.9
83	C-P	China	Red Soil Station	N 28°14.020	E 116°53.866	paddy	rice-peanut-clover			18.5	1731	44	2.09	4.3	6.6	93.4	14.05	1.00	0.69	16.7
84	C-P	China	Red Soil Station	N 28°14.020	E 116°53.866	paddy	rice-peanut-clover			18.5	1731	44	2.09	4.5	7.4	92.6	12.52	0.99	0.71	17.3
85	C-P	China	Red Soil Station	N 28°14.020	E 116°53.866	paddy	rice-peanut-clover			18.5	1731	44	2.75	4.2	11.0	89.0	8.08	1.20	0.75	17.4
86	C-P	China	Red Soil Station	N 28°14.034	E 116°53.865	paddy	rice-peanut-clover			18.5	1731	44	2.50	4.2	10.8	89.2	8.27	1.21	0.76	17.5
87	C-P	China	Red Soil Station	N 28°14.048	E 116°53.860	paddy	rice-peanut-clover			18.5	1731	44	2.04	4.3	11.4	88.6	7.76	1.06	0.70	16.5
88	JAV-P	Indonesia	Padas village	S 07°26.154	111°31.699	paddy	3xrice	3		26.7	2162	76	1.73	6.8	15.3	84.7	5.52	0.42	0.83	24.1
89	JAV-P	Indonesia	Sumbermujer	S 08°08.775	113°01.220	paddy	2x rice 2xfallow	2	2	17.8	2693	685	2.49	5.2	11.5	88.5	7.72	0.82	0.79	20.6
90	JAV-P	Indonesia	Simo village	S 07°29.957	111°36.129	paddy	2xrice 1x legumes	2	1	26.9	2100	85	1.52	6.9	15.4	84.6	5.50	0.29	0.82	24.5
91	JAV-P	Indonesia	Simo village	S 07°07.947	111°36.418	paddy	rice,bean,corn	2	1	26.9	2100	85	1.69	7.3	22.1	77.9	3.53	0.33	0.72	21.4
92	JAV-P	Indonesia	Simo village	S 07°07.947	111°36.418	paddy	rice,bean,corn	2	1	26.9	2100	85	1.86	7.5	23.2	76.8	3.31	0.38	0.67	19.4
93	SUM-P	Indonesia	Piladang	S 00°15.569	100°35.116	paddy	3xrice	3		21.8	2170	560	2.10	5.2	10.2	89.8	8.78	0.94	0.82	20.8
94	SUM-P	Indonesia	Suntiang	S 00°21.618	100°23.709	paddy	2xrice 1x cabbage	2	1	21.8	2170	1085	1.50	4.7	6.5	93.5	14.35	1.34	0.75	16.6
95	SUM-P	Indonesia	Suntiang	S 00°21.618	100°23.709	paddy	2xrice 1x cabbage	2	1	21.8	2170	1085	1.39	5.4	7.4	92.6	12.49	1.31	0.80	18.2

ID	Dataset code	Country	Sampling Area	Grad N	Grad E	Soil type	Vegetation	No. of rice/ year	No. of crops/ year	MAT (°C)	MAP (mm)	Altitude (m.a.s.l.)	SOC (%)	soil pH	iGDGT (% GDGT)	brGDGT (% GDGT)	brGDGT/ iGDGT	CBT	MBT'	T_{MC} (°C)
96	SUM-P	Indonesia	Lukok	S 00°22.025	100°24.625	paddy	1xrice, 2x legumes	2	1	21.8	2170	1178	2.54	4.8	9.5	90.5	9.57	1.20	0.81	19.1
97	IT-P	Italy	Zeme	45°11.536	8°40.078	paddy		1		12.5	954	79	1.15	4.9	9.6	90.4	9.43	0.90	0.51	11.6
98	IT-P	Italy	Vercelli	45°19.450	08°22.432	paddy		1		12.1	923		nd	6.9	11.5	88.5	7.73	0.17	0.39	12.0
99	IT-P	Italy	Vercelli	45°19.450	08°22.432	paddy		1		12.1	923		nd	6.9	8.3	91.7	11.02	0.20	0.33	10.0
100	IT-P	Italy	Vercelli	45°19.450	08°22.432	paddy		1		12.1	923		nd	7.0	9.9	90.1	9.07	0.14	0.39	12.2
101	IT-P	Italy	Vercelli	45°19.450	08°22.432	paddy		1		12.1	923		nd	6.1	5.5	94.5	17.30	0.65	0.49	12.3
102	VN-P	Vietnam	Hai Duong	10°26.495	106°3.524	paddy				27.4	1450		nd	4.0	10.9	89.1	8.16	1.08	0.83	20.5
103	VN-P	Vietnam	Hai Duong	10°26.495	106°3.524	paddy				27.4	1450		nd	4.1	10.0	90.0	9.04	1.09	0.83	20.3
104	VN-P	Vietnam	Hai Duong	10°23.384	106°3.946	paddy				27.4	1450		nd	4.7	7.6	92.4	12.12	1.09	0.85	21.0
105	VN-P	Vietnam	Hai Duong	10°23.384	106°3.946	paddy				27.4	1450		nd	4.7	7.8	92.2	11.85	1.06	0.84	20.7
106	PH-P	Phillipines	Laguna Los Banos	14,22857	121,33542	paddy		2		27.1	2064	9	1.59	6.2	7.8	92.2	11.85	0.70	0.89	24.4
107	PH-P	Phillipines	Laguna Los Banos	14,22862	121,3398	paddy		2		27.1	2064	5	4.01	5.8	9.6	90.4	9.38	0.79	0.79	20.9
108	PH-P	Phillipines	Laguna Los Banos	14,11495	121,41266	paddy		2		27.1	2064	288	2.27	4.9	7.9	92.1	11.58	0.89	0.79	20.3
109	PH-P	Phillipines	Laguna Los Banos	14,1154	121,4107	paddy		2		27.1	2064	275	2.90	5.1	9.8	90.2	9.21	0.95	0.84	21.4
110	PH-P	Phillipines	Laguna Los Banos	14,21577	121,33724	paddy		2		27.1	2064	8	2.29	5.5	10.5	89.5	8.55	1.08	0.85	20.9
111	PH-P	Phillipines	Laguna Los Banos	14,2174	121,3351	paddy		2		27.1	2064	7	3.73	5.9	13.9	86.1	6.17	0.71	0.77	20.8
112	PH-P	Phillipines	Laguna Los Banos	14,13942	121,39849	paddy		2		27.1	2064	190	2.06	4.8	10.3	89.7	8.68	0.97	0.84	21.2
113	PH-P	Phillipines	Laguna Los Banos	14,13856	121,4007	paddy		2		27.1	2064	187	2.84	5.2	8.9	91.1	10.25	0.83	0.83	21.9
114	PH-P	Phillipines	Laguna Los Banos	14,18674	121,3655	paddy		2		27.1	2064	31	2.36	4.7	10.9	89.1	8.15	0.98	0.80	20.1
115	PH-P	Phillipines	Laguna Los Banos	14,18912	121,36439	paddy		2		27.1	2064	18	2.80	5.1	11.1	88.9	7.97	1.02	0.82	20.3
116	PH-P	Phillipines	Nueva Ecija	15,67389	120,84252	paddy		2		27.1	1821	47	1.26	5.3	8.0	92.0	11.53	1.05	0.80	19.7
117	PH-P	Phillipines	Nueva Ecija	15,6724	120,84193	paddy		2		27.1	1821	45	1.70	4.9	9.8	90.2	9.19	1.21	0.81	19.0
118	PH-P	Phillipines	Nueva Ecija	15,6666	120,87578	paddy		2		27.1	1821	55	1.29	5.4	6.2	93.8	15.23	1.43	0.86	19.3
119	PH-P	Phillipines	Nueva Ecija	15,66395	120,87556	paddy		2		27.1	1821	52	1.95	4.8	5.7	94.3	16.54	1.44	0.80	17.6
120	PH-P	Phillipines	Nueva Ecija	15,66869	120,91821	paddy		2		27.1	1821	55	1.67	4.3	6.1	93.9	15.51	1.65	0.80	16.4
121	PH-P	Phillipines	Nueva Ecija	15,6726	120,92022	paddy		2		27.1	1821	58	1.27	4.4	6.2	93.8	15.06	1.53	0.80	16.8
122	PH-P	Phillipines	Nueva Ecija	15,59983	120,91127	paddy		2		27.1	1821	45	0.83	6.2	14.4	85.6	5.93	0.52	0.83	23.6
123	PH-P	Phillipines	Nueva Ecija	15,60254	120,91089	paddy		2		27.1	1821	44	1.26	6.0	6.1	93.9	15.40	0.84	0.73	18.5
124	PH-P	Phillipines	Nueva Ecija	15,61412	120,94629	paddy		1		27.1	1821	50	1.54	5.7	6.6	93.4	14.23	0.75	0.82	22.1
125	PH-P	Phillipines	Nueva Ecija	15,61707	120,94716	paddy		1		27.1	1821	52	1.64	6.0	6.0	94.0	15.69	0.85	0.73	18.5
126	PH-P	Phillipines	Ifugao	16,90493	121,07573	paddy		1		21.4	2376	898	4.05	4.7	11.0	89.0	8.10	0.83	0.71	18.2
127	PH-P	Phillipines	Ifugao	16,91033	121,12625	paddy		1		21.4	2376	851	1.56	4.8	7.1	92.9	13.02	1.10	0.78	18.6
128	PH-P	Phillipines	Ifugao	16,85875	121,10124	paddy		1		21.4	2376	933	2.14	4.3	9.8	90.2	9.20	0.94	0.75	18.6
129	PH-P	Phillipines	Ifugao	16,90534	121,06454	paddy		1		21.4	2376	1075	2.07	4.5	14.1	85.9	6.11	0.89	0.67	16.5
130	PH-P	Phillipines	Ifugao	16,9332	121,13409	paddy		1		21.4	2376	901	1.16	5.0	3.6	96.4	27.13	1.23	0.80	18.6
131	PH-P	Phillipines	Ifugao	16,86122	121,09905	paddy		1		21.4	2376	867	3.96	4.5	15.7	84.3	5.38	0.74	0.70	18.3
132	PH-P	Phillipines	Ifugao	16,91008	121,12891	paddy		1		21.4	2376	933	2.73	5.5	6.5	93.5	14.31	0.74	0.78	20.7
133	PH-P	Phillipines	Ifugao	16,92061	121,05901	paddy		1		21.4	2376	1089	3.87	4.8	17.6	82.4	4.69	0.70	0.63	16.5
134	PH-P	Phillipines	Ifugao	16,92063	121,05906	paddy		1		21.4	2376	1155	2.21	5.2	12.1	87.9	7.26	0.85	0.68	17.0
135	PH-P	Phillipines	Ifugao	16,93278	121,1375	paddy		1		21.4	2376	782	5.04	4.9	13.2	86.8	6.56	0.77	0.71	18.4

ID	Dataset code	Country	Sampling Area	Grad N	Grad E	Soil type	Vegetation	No. of rice/ year	No. of crops/ year	MAT (°C)	MAP (mm)	Altitude (m.a.s.l.)	SOC (%)	soil pH	iGDGT (% GDGT)	brGDGT (% GDGT)	brGDGT/ iGDGT	CBT	MBT'	T_{MC} (°C)
136	VN-P	Vietnam	Hai Duong	21,03488	106,3505	paddy		2		24.1	1608	4	1.58	5.0	7.1	92.9	13.12	0.81	0.69	17.7
137	VN-P	Vietnam	Hai Duong	21,03296	106,35387	paddy		2		24.1	1608	3	1.13	5.4	8.3	91.7	10.98	0.75	0.71	18.5
138	VN-P	Vietnam	Hai Duong	20,98858	106,41315	paddy		2		24.1	1608	4	1.52	5.1	6.6	93.4	14.12	0.73	0.69	18.0
139	VN-P	Vietnam	Hai Duong	20,99121	106,40808	paddy		2		24.1	1608	4	1.26	5.7	9.0	91.0	10.15	0.45	0.68	19.2
140	VN-P	Vietnam	Hai Duong	20,96091	106,44228	paddy		2		24.1	1608	2	1.46	5.1	4.6	95.4	20.90	0.69	0.72	19.2
141	VN-P	Vietnam	Hai Duong	20,96094	106,44429	paddy		2		24.1	1608	2	1.68	5.0	6.1	93.9	15.30	0.68	0.71	18.9
142	VN-P	Vietnam	Hai Duong	20,94533	106,36392	paddy		2		24.1	1608	2	1.36	5.1	7.9	92.1	11.72	0.78	0.65	16.4
143	VN-P	Vietnam	Hai Duong	20,94395	106,36802	paddy		2		24.1	1608	4	1.30	4.8	7.5	92.5	12.36	0.63	0.68	18.4
144	VN-P	Vietnam	Vinh Phúc	21,3487	105,70621	paddy		2		23.6	1687	23	1.45	4.4	10.3	89.7	8.72	0.96	0.76	18.9
145	VN-P	Vietnam	Vinh Phúc	21,3513	105,70691	paddy		2		23.6	1687	26	1.71	4.6	13.9	86.1	6.19	0.91	0.76	19.3
146	VN-P	Vietnam	Vinh Phúc	21,37207	105,71867	paddy		2		23.6	1687	40	1.12	4.3	13.1	86.9	6.61	1.60	0.82	17.0
147	VN-P	Vietnam	Vinh Phúc	21,37282	105,72001	paddy		2		23.6	1687	43	1.37	4.5	16.1	83.9	5.21	1.50	0.84	18.4
148	VN-P	Vietnam	Vinh Phúc	21,31653	105,7387	paddy		2		23.6	1687	20	1.41	4.5	9.1	90.9	10.04	1.31	0.81	18.4
149	VN-P	Vietnam	Vinh Phúc	21,31931	105,7411	paddy		2		23.6	1687	21	2.41	4.6	14.3	85.7	6.01	0.88	0.75	19.2
150	VN-P	Vietnam	Vinh Phúc	21,31455	105,73573	paddy		2		23.6	1687	22	2.21	4.8	9.4	90.6	9.64	1.12	0.75	17.8
151	VN-P	Vietnam	Vinh Phúc	21,31782	105,73461	paddy		2		23.6	1687	23	1.31	4.6	15.8	84.2	5.33	1.56	0.85	18.4
152	VN-P	Vietnam	Lào Cai	22,41071	103,90204	paddy		1		16.2	2223	729	1.65	4.3	9.5	90.5	9.56	1.44	0.82	18.1
153	VN-P	Vietnam	Lào Cai	22,40992	103,90203	paddy		1		16.2	2223	748	0.83	4.4	8.2	91.8	11.22	1.41	0.86	19.5
154	VN-P	Vietnam	Lào Cai	22,29836	103,91269	paddy		1		16.2	2223	1043	1.23	5.2	6.6	93.4	14.06	0.99	0.68	16.3
155	VN-P	Vietnam	Lào Cai	22,29643	103,91056	paddy		1		16.2	2223	979	1.61	5.1	10.7	89.3	8.32	0.80	0.59	14.4
156	VN-P	Vietnam	Lào Cai	22,30199	103,88959	paddy		1		16.2	2223	998	2.48	4.3	6.3	93.7	14.96	1.24	0.69	15.2
157	VN-P	Vietnam	Lào Cai	22,30416	103,88726	paddy		1		16.2	2223	989	1.86	4.3	9.7	90.3	9.26	0.98	0.61	14.2
158	VN-P	Vietnam	Lào Cai	22,31438	103,858	paddy		1		16.2	2223	1252	2.02	5.0	9.8	90.2	9.19	1.21	0.64	13.7
159	VN-P	Vietnam	Lào Cai	22,31789	103,85866	paddy		1		16.2	2223	1269	1.77	4.8	4.9	95.1	19.27	1.18	0.69	15.4
160	VN-P	Vietnam	Lào Cai	22,39314	103,8437	paddy		1		16.2	2223	1284	1.84	4.7	4.8	95.2	19.92	1.16	0.73	16.8
161	VN-P	Vietnam	Lào Cai	22,39481	103,84429	paddy		1		16.2	2223	1289	1.61	4.6	5.2	94.8	18.31	1.14	0.62	13.4
162	VN-P	Vietnam	Tien Giang	10,44085	106,05783	paddy		3		27.4	1450	3	4.43	4.0	10.0	90.0	9.00	1.05	0.84	20.9
163	VN-P	Vietnam	Tien Giang	10,44382	106,05849	paddy		3		27.4	1450	2	4.01	3.7	9.8	90.2	9.16	1.13	0.83	20.1
164	VN-P	Vietnam	Tien Giang	10,40293	106,10302	paddy		3		27.4	1450	3	2.28	3.8	9.8	90.2	9.21	1.08	0.83	20.2
165	VN-P	Vietnam	Tien Giang	10,40512	106,10195	paddy		3		27.4	1450	2	2.24	3.7	10.0	90.0	8.98	1.02	0.79	19.7
166	VN-P	Vietnam	Tien Giang	10,37051	106,12832	paddy		3		27.4	1450	1	3.82	3.8	10.9	89.1	8.18	1.00	0.81	20.4
167	VN-P	Vietnam	Tien Giang	10,37998	106,11175	paddy		3		27.4	1450	2	2.46	4.6	8.0	92.0	11.45	1.14	0.80	19.2
168	VN-P	Vietnam	Tien Giang	10,37927	106,11247	paddy		3		27.4	1450	2	2.06	4.8	7.8	92.2	11.79	0.99	0.80	20.1
169	VN-P	Vietnam	Tien Giang	10,40726	106,1079	paddy		3		27.4	1450	2	3.20	3.8	10.2	89.8	8.77	1.12	0.81	19.5
170	VN-P	Vietnam	Tien Giang	10,40724	106,10898	paddy		3		27.4	1450	3	2.43	3.9	10.7	89.3	8.31	1.08	0.85	21.1

ID	Dataset code	Country	Sampling Area	Soil type	isoprenoid GDGTs (% of iGDGT)						branched GDGTs (% of brGDGT)								
					GDGT -0	GDGT -1	GDGT -2	GDGT -3	Cren	Cren reg	GDGT -IIIa	GDGT -IIa	GDGT -Ia	GDG -Ib	GDGT -Ic	GDGT -IIIb	GDGT -IIIc	GDGT -IIb	GDGT -IIc
51	IT-NP	Italy	Zeme	upland	20.3	10.6	13.3	4.0	48.6	3.4	6.7	40.4	48.8	2.4	0.4	0.1	0.0	1.1	0.1
4	C-NP	China	Cixi	upland	10.4	12.0	13.2	4.8	55.2	4.5	5.6	21.2	24.7	24.8	2.4	1.9	0.3	18.1	1.0
5	C-NP	China	Cixi	upland	18.8	9.7	10.2	4.6	51.4	5.2	5.7	18.9	25.7	25.8	3.5	1.4	0.3	17.7	1.2
6	C-NP	China	Cixi	upland	12.1	11.2	9.9	4.1	58.3	4.5	4.9	17.8	26.1	27.1	2.4	2.0	0.3	18.4	1.0
7	C-NP	China	Cixi	upland	8.2	8.8	13.3	5.3	60.3	4.2	5.2	18.9	23.9	26.1	3.5	1.7	0.2	19.2	1.1
8	C-NP	China	Cixi	upland	8.8	8.7	10.3	4.5	61.2	6.5	4.4	20.2	35.5	24.5	2.2	1.0	0.2	11.2	0.7
9	JAV-NP	Indonesia	Sukabumi	upland	40.5	7.8	11.1	5.4	31.6	3.7	0.6	7.6	87.2	2.4	0.7	0.0	0.1	1.1	0.4
10	JAV-NP	Indonesia	Sukabumi	upland	17.3	10.1	14.7	6.1	48.3	3.5	0.5	8.3	86.5	2.2	0.9	0.1	0.1	0.9	0.4
11	JAV-NP	Indonesia	Sukabumi	upland	21.6	15.5	18.4	8.5	33.3	2.7	0.6	9.5	76.0	8.8	2.5	0.1	0.1	2.0	0.4
12	JAV-NP	Indonesia	Jasinga	upland	31.9	9.2	13.2	5.6	35.7	4.5	0.1	3.8	93.8	1.2	0.6	0.0	0.1	0.2	0.3
13	JAV-NP	Indonesia	Jasinga	upland	14.2	8.7	13.2	6.2	49.5	8.2	0.2	4.2	93.1	1.5	0.5	0.0	0.1	0.2	0.2
14	JAV-NP	Indonesia	Jasinga	upland	10.9	5.5	10.0	6.6	55.7	11.3	0.4	5.6	70.5	15.5	5.4	0.1	0.1	2.1	0.2
15	JAV-NP	Indonesia	Ngawi	upland	8.5	5.7	6.6	4.7	69.2	5.2	0.2	5.5	86.5	6.0	1.0	0.1	0.1	0.6	0.1
16	JAV-NP	Indonesia	Ngawi	upland	9.9	7.4	7.3	5.1	63.4	7.0	0.2	5.4	83.5	8.3	1.5	0.0	0.0	0.9	0.1
17	JAV-NP	Indonesia	Ngawi	upland	8.9	6.6	6.9	5.6	66.0	6.0	0.2	6.2	79.3	11.0	1.8	0.1	0.1	1.4	0.0
18	C-NP	China	Red Soil Station	upland	17.2	6.5	13.8	4.2	53.2	5.0	2.3	25.7	71.1	0.6	0.1	0.0	0.1	0.2	0.0
19	C-NP	China	Red Soil Station	upland	22.2	7.7	13.0	4.8	45.8	6.5	1.8	20.1	74.9	1.9	0.3	0.0	0.1	0.7	0.2
20	C-NP	China	Red Soil Station	upland	22.0	7.0	13.9	4.9	49.1	3.2	2.0	22.8	73.6	0.9	0.2	0.0	0.0	0.2	0.2
32	PH-NP	Phillipines	Laguna Los Banos	upland	8.2	4.8	7.5	6.2	60.4	12.8	0.3	4.9	82.4	9.0	2.4	0.0	0.2	0.8	0.1
33	PH-NP	Phillipines	Laguna Los Banos	upland	19.4	11.3	12.5	7.5	44.8	4.5	0.3	4.9	85.4	6.8	2.0	0.0	0.1	0.4	0.1
34	PH-NP	Phillipines	Laguna Los Banos	upland	54.6	7.4	9.2	4.0	22.0	2.7	0.3	6.3	88.4	3.6	1.0	0.0	0.1	0.2	0.1
35	PH-NP	Phillipines	Laguna Los Banos	upland	26.0	5.6	7.4	5.2	46.3	9.3	0.7	10.0	62.6	17.8	6.2	0.0	0.1	2.4	0.2
36	PH-NP	Phillipines	Laguna Los Banos	upland	45.5	3.9	5.8	3.5	32.9	8.5	0.6	9.0	70.7	13.5	4.5	0.0	0.1	1.4	0.2
37	PH-NP	Phillipines	Nueva Ecija	upland	22.2	7.9	9.6	5.3	47.6	7.4	0.4	10.5	84.0	4.0	0.5	0.0	0.1	0.4	0.2
38	PH-NP	Phillipines	Nueva Ecija	upland	10.1	5.1	8.3	5.3	59.8	11.5	0.7	10.6	61.3	19.1	4.8	0.0	0.1	3.0	0.4
39	PH-NP	Phillipines	Nueva Ecija	upland	14.9	5.0	7.8	5.2	58.3	8.8	0.5	8.3	70.6	14.7	3.1	0.0	0.1	2.3	0.3
40	PH-NP	Phillipines	Nueva Ecija	upland	40.8	2.8	5.5	3.1	44.2	3.5	0.4	7.5	83.9	5.9	1.4	0.0	0.1	0.6	0.1
41	PH-NP	Phillipines	Ifugao	upland	19.8	9.1	9.7	4.4	50.5	6.6	0.6	7.8	73.6	12.0	4.2	0.0	0.0	1.5	0.2
42	PH-NP	Phillipines	Ifugao	upland	27.7	8.9	11.0	5.6	43.1	3.7	0.9	11.1	72.1	10.7	3.4	0.0	0.1	1.5	0.2
43	PH-NP	Phillipines	Ifugao	upland	32.8	18.0	19.2	12.7	16.2	1.2	0.6	8.2	84.5	4.3	1.4	0.0	0.0	0.7	0.2
44	PH-NP	Phillipines	Ifugao	upland	35.7	13.1	12.3	4.2	32.1	2.6	1.0	10.9	76.5	7.3	2.6	0.0	0.1	1.3	0.3
45	PH-NP	Phillipines	Ifugao	upland	44.8	8.0	10.8	3.0	32.1	1.4	1.2	16.1	72.5	6.5	1.9	0.0	0.1	1.4	0.3
46	VN-NP	Vietnam	Hai Duong	upland	43.3	12.6	13.3	3.5	26.1	1.2	2.3	22.8	62.6	8.5	1.5	0.1	0.1	1.9	0.3
47	VN-NP	Vietnam	Hai Duong	upland	20.6	6.8	9.2	5.0	51.0	7.4	1.5	11.4	28.0	29.6	13.2	0.5	0.2	13.6	1.9
48	VN-NP	Vietnam	Vinh Phúc	upland	36.1	5.0	7.4	4.7	42.4	4.4	0.7	9.4	75.6	9.2	2.4	0.1	0.2	2.1	0.4
49	VN-NP	Vietnam	Vinh Phúc	upland	36.1	8.9	13.0	11.0	27.7	3.2	0.3	6.0	90.5	1.5	0.7	0.0	0.1	0.4	0.4
50	VN-NP	Vietnam	Vinh Phúc	upland	25.2	6.9	11.4	8.1	44.3	4.0	0.3	6.3	90.3	1.1	0.7	0.1	0.1	0.6	0.6
1	C-Marsh	China	Cixi	marsh	25.9	10.4	12.7	3.1	45.4	2.5	10.6	28.8	31.4	13.8	2.1	1.2	0.2	11.0	0.9
2	C-Marsh	China	Cixi	marsh	21.1	8.9	8.4	3.9	53.7	4.1	6.9	28.4	24.8	18.0	3.5	1.2	0.1	15.8	1.5
3	C-Marsh	China	Cixi	marsh	12.6	8.1	9.6	5.5	58.4	5.8	7.4	20.9	18.0	23.9	6.7	1.4	0.6	18.2	3.0
23	JAV-Bamb	Indonesia	Sumbermujer	bamboo	29.2	20.0	24.1	8.4	16.2	2.2	0.2	4.6	85.6	6.7	2.2	0.1	0.0	0.5	0.1

ID	Dataset code	Country	Sampling Area	Soil type	isoprenoid GDGTs (% of iGDGT)						branched GDGTs (% of brGDGT)								
					GDGT -0	GDGT -1	GDGT -2	GDGT -3	Cren	Cren reg	GDGT -IIIa	GDGT -IIa	GDGT -Ia	GDG -Ib	GDGT -Ic	GDGT -IIIb	GDGT -IIIc	GDGT -IIb	GDGT -IIc
27	VN-Bamb	Vietnam	Vinh Phúc	bamboo	21.0	11.6	17.8	12.1	31.9	5.7	0.2	4.4	92.5	1.2	1.1	0.0	0.1	0.2	0.3
30	VN-Bamb	Vietnam	Lào Cai	bamboo	22.5	18.2	22.4	10.7	23.7	2.6	0.6	9.9	83.1	4.6	1.0	0.1	0.1	0.6	0.1
21	VN-Bush	Vietnam	Lào Cai	bushland	42.1	12.0	18.6	12.8	13.7	0.7	0.4	9.0	85.4	3.7	0.9	0.0	0.0	0.4	0.1
22	VN-Bush	Vietnam	Lào Cai	bushland	29.7	16.7	19.7	9.7	22.6	1.6	0.4	8.9	87.6	2.1	0.5	0.1	0.1	0.3	0.1
24	PH-For	Phillipines	Ifugao	forest	20.1	13.3	15.7	9.0	38.4	3.5	0.9	10.9	72.2	10.2	3.1	0.1	0.1	2.1	0.4
25	PH-For	Phillipines	Ifugao	forest	17.1	9.1	10.2	7.2	53.2	3.2	0.8	10.5	74.4	9.6	3.1	0.0	0.0	1.5	0.1
26	PH-For	Phillipines	Ifugao	forest	29.0	17.7	15.6	7.5	27.8	2.4	1.6	15.3	64.9	11.6	3.3	0.1	0.1	2.9	0.2
28	VN-For	Vietnam	Vinh Phúc	forest	14.9	12.1	21.3	20.9	26.9	3.9	1.2	12.6	84.2	0.8	0.8	0.0	0.1	0.2	0.2
29	VN-For	Vietnam	Lào Cai	forest	17.6	16.8	23.8	18.4	21.3	2.1	0.6	11.5	80.5	4.7	1.5	0.0	0.2	0.7	0.2
31	VN-For	Vietnam	Lào Cai	forest	31.0	15.2	13.9	9.3	28.1	2.6	0.6	9.4	86.4	1.9	0.5	0.1	0.4	0.5	0.2
52	C-P	China	Cixi	paddy	59.1	18.1	6.8	1.2	13.4	1.4	9.4	27.5	29.6	16.0	2.4	1.3	0.3	11.9	1.4
53	C-P	China	Cixi	paddy	59.7	17.6	7.9	1.5	12.5	0.8	9.5	27.8	30.2	16.2	2.6	1.3	0.4	10.9	1.2
54	C-P	China	Cixi	paddy	58.3	22.0	7.3	1.2	10.1	1.2	8.6	27.1	30.3	16.6	2.5	1.1	0.4	12.0	1.3
55	C-P	China	Cixi	paddy	36.9	18.0	11.9	3.3	28.1	1.8	5.6	24.2	43.9	15.8	2.7	0.6	0.3	6.3	0.8
56	C-P	China	Cixi	paddy	41.0	21.9	8.7	2.9	23.4	2.1	7.0	27.5	33.4	17.7	3.1	0.9	0.5	9.0	1.0
57	C-P	China	Cixi	paddy	36.0	21.6	9.5	3.0	27.9	2.0	5.9	26.7	37.5	18.1	3.0	0.7	0.3	6.9	0.9
58	C-P	China	Cixi	paddy	51.5	18.9	11.0	2.1	15.8	0.8	5.6	25.5	40.7	18.3	2.8	0.5	0.2	5.6	0.7
59	C-P	China	Cixi	paddy	53.2	19.2	9.7	2.0	15.2	0.7	6.4	25.6	39.7	17.7	2.9	0.5	0.2	6.2	0.7
60	C-P	China	Cixi	paddy	48.2	21.6	10.1	2.0	17.1	0.9	5.4	28.0	39.5	16.1	2.8	0.4	0.1	6.9	0.7
61	C-P	China	Cixi	paddy	26.1	16.8	11.6	3.5	39.5	2.6	7.3	24.5	31.8	20.9	3.2	0.9	0.3	9.9	1.0
62	C-P	China	Cixi	paddy	31.0	17.7	9.7	3.5	35.3	2.8	8.3	25.6	30.1	19.9	2.9	1.1	0.3	10.8	1.0
63	C-P	China	Cixi	paddy	30.9	19.0	10.8	3.3	32.7	3.2	6.4	25.9	33.2	18.9	2.9	0.8	0.3	10.8	0.9
64	C-P	China	Cixi	paddy	60.8	14.0	6.4	2.1	15.5	1.2	8.5	27.1	33.9	17.5	2.6	0.8	0.3	8.3	0.9
65	C-P	China	Cixi	paddy	56.4	15.5	8.3	2.1	16.2	1.5	9.1	25.2	34.3	18.2	2.9	1.1	0.3	7.9	0.9
66	C-P	China	Cixi	paddy	63.4	10.8	6.0	1.6	17.0	1.2	8.1	24.9	34.7	19.0	2.9	0.8	0.2	8.3	1.0
67	C-P	China	Cixi	paddy	14.4	12.9	13.6	5.1	48.9	5.1	3.7	23.1	50.8	14.0	2.0	0.6	0.3	4.7	0.7
68	C-P	China	Cixi	paddy	15.1	11.9	11.9	6.0	51.0	4.1	4.2	20.9	55.0	12.7	1.9	0.5	0.4	3.7	0.6
69	C-P	China	Cixi	paddy	16.4	11.1	12.1	6.0	49.1	5.3	4.7	20.6	48.2	16.2	2.9	0.7	0.4	5.5	0.9
70	C-P	China	Cixi	paddy	48.7	17.4	11.8	2.3	19.0	0.9	6.0	24.8	50.0	13.1	2.2	0.2	0.2	2.9	0.5
71	C-P	China	Cixi	paddy	47.4	19.0	14.3	3.4	14.8	1.1	5.5	23.5	50.6	14.1	2.5	0.3	0.2	2.9	0.5
72	C-P	China	Cixi	paddy	44.6	19.0	16.0	3.0	16.6	0.9	5.7	24.2	49.4	14.7	2.4	0.3	0.1	2.7	0.5
73	JAV-P	Indonesia	Sukabumi	paddy	22.1	8.3	11.2	4.5	51.0	2.8	1.5	17.6	73.3	5.0	1.4	0.0	0.1	0.9	0.2
74	JAV-P	Indonesia	Sukabumi	paddy	23.1	7.5	11.5	4.4	51.5	2.0	1.8	20.0	70.3	5.1	1.4	0.1	0.0	1.1	0.1
75	JAV-P	Indonesia	Sukabumi	paddy	19.4	9.4	12.8	4.4	51.8	2.2	1.4	18.0	73.9	4.4	1.2	0.1	0.0	0.9	0.2
76	JAV-P	Indonesia	Jasinga	paddy	50.0	10.9	10.0	3.7	22.1	3.3	0.5	8.1	88.9	1.3	0.9	0.0	0.1	0.1	0.1
77	JAV-P	Indonesia	Jasinga	paddy	46.3	10.1	12.3	5.9	23.0	2.5	0.4	7.2	89.8	1.3	0.9	0.0	0.1	0.2	0.1
78	JAV-P	Indonesia	Jasinga	paddy	49.0	10.2	11.1	3.8	23.4	2.5	0.4	7.3	88.6	2.2	1.3	0.0	0.0	0.2	0.1
79	JAV-P	Indonesia	Jasinga	paddy	49.1	9.4	12.4	5.8	21.9	1.5	0.4	7.9	88.9	1.4	1.1	0.0	0.0	0.2	0.1
80	JAV-P	Indonesia	Ngawi	paddy	23.4	10.6	12.4	5.3	40.1	8.2	1.4	17.6	47.4	21.6	3.0	0.3	0.1	8.0	0.6
81	JAV-P	Indonesia	Ngawi	paddy	24.1	11.3	11.5	5.4	40.0	7.8	0.9	15.7	60.5	17.0	1.7	0.1	0.0	3.7	0.2
82	JAV-P	Indonesia	Ngawi	paddy	39.0	9.0	9.2	4.0	32.4	6.3	1.0	15.4	63.5	14.5	1.7	0.1	0.0	3.3	0.2
83	C-P	China	Red Soil Station	paddy	50.1	18.2	13.8	2.7	14.2	0.9	3.0	25.8	60.0	7.3	1.9	0.1	0.0	1.4	0.4
84	C-P	China	Red Soil Station	paddy	52.3	13.2	12.8	3.0	17.8	0.9	2.5	24.5	61.5	7.7	2.1	0.1	0.0	1.2	0.4
85	C-P	China	Red Soil Station	paddy	49.7	8.8	13.6	3.4	23.3	1.2	2.0	21.6	69.0	4.9	1.3	0.0	0.0	0.9	0.3
86	C-P	China	Red Soil Station	paddy	47.1	10.6	14.4	3.5	22.7	1.7	2.1	20.8	70.0	4.8	1.1	0.1	0.1	0.8	0.3
87	C-P	China	Red Soil Station	paddy	56.0	10.9	12.6	2.9	16.8	0.8	2.7	25.7	61.6	6.5	1.9	0.1	0.0	1.1	0.4
88	JAV-P	Indonesia	Padas village	paddy	18.3	10.6	11.8	5.2	46.1	7.9	0.9	11.3	57.7	21.7	3.1	0.2	0.1	4.7	0.4
89	JAV-P	Indonesia	Sumbermujer	paddy	48.6	19.5	11.0	2.4	17.8	0.7	1.7	17.5	65.9	10.7	2.2	0.1	0.0	1.8	0.2
90	JAV-P	Indonesia	Simo village	paddy	21.1	5.7	6.3	1.8	55.8	9.3	1.1	10.7	50.2	25.4	5.7	0.2	0.1	5.8	0.6

ID	Dataset code	Country	Sampling Area	Soil type	isoprenoid GDGTs (% of iGDGT)						branched GDGTs (% of brGDGT)								
					GDGT -0	GDGT -1	GDGT -2	GDGT -3	Cren	Cren reg	GDGT -IIIa	GDGT -IIa	GDGT -Ia	GDG -Ib	GDGT -Ic	GDGT -IIIb	GDGT -IIIc	GDGT -IIb	GDGT -IIc
91	JAV-P	Indonesia	Simo village	paddy	35.7	6.5	6.0	2.0	41.6	8.2	1.9	18.3	44.5	22.9	4.8	0.2	0.1	6.7	0.6
92	JAV-P	Indonesia	Simo village	paddy	44.3	5.8	6.6	1.6	35.7	6.1	2.5	22.8	42.6	20.1	3.9	0.3	0.1	7.2	0.6
93	SUM-P	Indonesia	Piladang	paddy	61.3	15.1	10.0	2.4	10.6	0.6	2.2	15.0	71.0	8.8	1.7	0.0	0.0	1.1	0.1
94	SUM-P	Indonesia	Suntiang	paddy	32.9	10.9	12.7	4.5	36.7	2.3	1.9	22.0	70.8	3.6	0.8	0.0	0.0	0.6	0.1
95	SUM-P	Indonesia	Suntiang	paddy	25.9	8.9	13.6	4.4	43.8	3.3	1.5	17.7	75.2	4.0	0.9	0.0	0.0	0.5	0.1
96	SUM-P	Indonesia	Lukok	paddy	21.5	12.7	13.6	4.7	43.7	3.8	1.3	16.8	74.5	5.0	1.4	0.0	0.0	0.9	0.1
97	IT-P	Italy	Zeme	paddy	52.8	15.3	7.9	2.5	19.8	1.7	6.1	39.6	42.3	7.8	1.2	0.1	0.0	2.5	0.4
98	IT-P	Italy	Vercelli	paddy	17.7	10.9	13.2	6.4	46.5	5.2	9.9	30.8	18.7	16.0	3.6	1.6	0.2	17.7	1.5
99	IT-P	Italy	Vercelli	paddy	22.2	9.4	13.0	5.6	44.9	4.9	11.6	34.5	16.1	14.0	2.5	1.6	0.2	18.1	1.3
100	IT-P	Italy	Vercelli	paddy	17.3	11.0	13.3	5.9	46.6	6.0	9.8	29.7	18.8	16.7	3.0	1.6	0.2	18.8	1.4
101	IT-P	Italy	Vercelli	paddy	40.8	15.0	13.0	3.1	26.6	1.4	6.4	38.9	35.5	11.5	1.7	0.2	0.0	5.2	0.6
102	VN-P	Vietnam	Hai Duong	paddy	60.4	15.1	15.4	4.8	3.9	0.4	0.6	15.2	75.0	6.7	1.2	0.1	0.1	0.8	0.2
103	VN-P	Vietnam	Hai Duong	paddy	58.3	16.3	16.7	5.1	3.4	0.3	0.6	15.6	74.8	6.6	1.2	0.1	0.2	0.7	0.2
104	VN-P	Vietnam	Hai Duong	paddy	27.9	15.8	14.1	4.3	34.7	3.2	0.4	13.5	77.1	6.7	1.2	0.0	0.1	0.8	0.2
105	VN-P	Vietnam	Hai Duong	paddy	27.0	14.3	13.7	4.4	37.6	3.0	0.7	14.7	75.1	7.0	1.4	0.1	0.0	0.7	0.2
106	PH-P	Phillipines	Laguna Los Banos	paddy	11.1	4.3	7.1	5.6	58.5	13.4	0.5	9.0	71.2	14.3	3.3	0.1	0.0	1.5	0.1
107	PH-P	Phillipines	Laguna Los Banos	paddy	53.0	17.3	13.1	3.2	11.2	2.1	1.5	17.4	65.4	11.8	2.0	0.1	0.1	1.7	0.2
108	PH-P	Phillipines	Laguna Los Banos	paddy	48.0	14.0	12.5	2.7	21.5	1.2	1.0	18.6	65.6	9.8	3.5	0.1	0.0	1.2	0.2
109	PH-P	Phillipines	Laguna Los Banos	paddy	41.1	14.6	13.8	3.1	25.5	1.8	0.6	14.5	72.0	8.7	2.9	0.0	0.0	1.0	0.2
110	PH-P	Phillipines	Laguna Los Banos	paddy	36.4	16.4	11.1	4.9	27.5	3.8	1.0	13.5	76.3	6.7	1.5	0.0	0.1	0.8	0.1
111	PH-P	Phillipines	Laguna Los Banos	paddy	53.6	15.9	12.4	2.9	12.8	2.4	1.3	18.4	61.8	13.1	2.4	0.1	0.0	2.7	0.3
112	PH-P	Phillipines	Laguna Los Banos	paddy	40.4	17.0	10.5	4.2	24.1	3.8	0.8	14.3	72.8	8.2	2.5	0.1	0.1	1.1	0.2
113	PH-P	Phillipines	Laguna Los Banos	paddy	48.7	12.9	10.2	3.4	22.3	2.6	1.0	13.9	69.0	10.8	3.4	0.1	0.0	1.6	0.3
114	PH-P	Phillipines	Laguna Los Banos	paddy	61.7	12.0	7.5	3.1	13.9	1.8	1.6	16.3	70.8	7.6	1.8	0.1	0.1	1.5	0.2
115	PH-P	Phillipines	Laguna Los Banos	paddy	58.5	15.6	11.3	2.9	10.4	1.3	1.0	16.1	72.6	7.4	1.5	0.1	0.1	1.1	0.2
116	PH-P	Phillipines	Nueva Ecija	paddy	49.3	19.2	11.0	3.9	13.6	3.0	1.2	17.4	72.2	7.0	0.9	0.0	0.1	1.0	0.1
117	PH-P	Phillipines	Nueva Ecija	paddy	58.0	15.6	9.8	3.6	10.2	2.8	1.3	16.8	75.4	4.9	0.6	0.0	0.1	0.8	0.1
118	PH-P	Phillipines	Nueva Ecija	paddy	34.7	8.4	8.5	5.8	37.0	5.6	0.8	13.0	81.9	3.2	0.6	0.0	0.1	0.3	0.1
119	PH-P	Phillipines	Nueva Ecija	paddy	49.8	19.1	15.9	5.4	8.7	1.0	1.5	17.5	77.0	3.0	0.3	0.0	0.1	0.5	0.1
120	PH-P	Phillipines	Nueva Ecija	paddy	47.4	21.0	17.1	4.2	9.3	1.1	1.6	17.5	77.9	1.9	0.5	0.0	0.1	0.3	0.1
121	PH-P	Phillipines	Nueva Ecija	paddy	52.8	18.7	17.5	4.8	5.5	0.7	1.2	18.8	76.5	2.5	0.6	0.0	0.0	0.3	0.0
122	PH-P	Phillipines	Nueva Ecija	paddy	9.7	5.3	7.8	5.5	63.0	8.7	0.8	12.4	60.0	18.7	4.1	0.1	0.1	3.4	0.3
123	PH-P	Phillipines	Nueva Ecija	paddy	43.1	17.8	13.2	3.7	20.5	1.8	2.0	22.9	61.5	9.7	1.2	0.1	0.0	2.4	0.2
124	PH-P	Phillipines	Nueva Ecija	paddy	38.8	19.2	14.5	4.2	21.5	1.8	1.1	14.7	66.9	12.8	2.5	0.1	0.0	1.7	0.2
125	PH-P	Phillipines	Nueva Ecija	paddy	43.7	18.4	12.7	3.8	20.0	1.5	2.0	22.9	61.6	9.6	1.2	0.1	0.0	2.4	0.2
126	PH-P	Phillipines	Ifugao	paddy	72.6	11.6	10.6	2.2	2.8	0.2	2.1	23.9	59.4	9.8	1.8	0.1	0.0	2.5	0.4
127	PH-P	Phillipines	Ifugao	paddy	71.9	10.3	10.5	2.2	4.8	0.3	1.6	19.4	70.0	5.9	1.4	0.1	0.1	1.2	0.3
128	PH-P	Phillipines	Ifugao	paddy	70.9	11.8	8.9	1.7	6.3	0.4	2.0	21.0	65.0	7.8	1.6	0.1	0.1	2.1	0.3
129	PH-P	Phillipines	Ifugao	paddy	79.1	9.8	6.6	1.2	3.1	0.2	3.5	26.3	57.2	7.8	1.6	0.2	0.1	3.0	0.4
130	PH-P	Phillipines	Ifugao	paddy	59.4	10.1	10.1	3.1	16.2	1.1	1.6	17.6	73.8	4.5	1.3	0.1	0.1	0.9	0.2

ID	Dataset code	Country	Sampling Area	Soil type	isoprenoid GDGTs (% of iGDGT)						branched GDGTs (% of brGDGT)								
					GDGT -0	GDGT -1	GDGT -2	GDGT -3	Cren	Cren reg	GDGT -IIIa	GDGT -IIa	GDGT -Ia	GDG -Ib	GDGT -Ic	GDGT -IIIb	GDGT -IIIc	GDGT -IIb	GDGT -IIc
131	PH-P	Phillipines	Ifugao	paddy	74.9	13.4	9.4	1.6	0.6	0.1	3.0	23.6	57.1	11.6	1.3	0.1	0.1	3.0	0.2
132	PH-P	Phillipines	Ifugao	paddy	71.7	7.9	6.4	2.0	11.0	1.1	1.8	17.2	62.8	11.8	3.1	0.1	0.1	2.7	0.4
133	PH-P	Phillipines	Ifugao	paddy	80.2	10.3	7.0	1.3	1.2	0.1	3.8	27.8	50.1	11.2	1.9	0.2	0.1	4.5	0.4
134	PH-P	Phillipines	Ifugao	paddy	74.7	9.0	7.1	1.5	7.4	0.3	3.2	24.9	57.2	8.4	2.2	0.2	0.1	3.3	0.7
135	PH-P	Phillipines	Ifugao	paddy	74.2	9.6	7.9	1.6	6.3	0.4	2.5	23.0	57.8	10.7	2.2	0.2	0.1	3.1	0.4
136	VN-P	Vietnam	Hai Duong	paddy	50.1	17.6	13.5	2.9	15.1	0.8	2.6	25.2	57.3	10.4	1.6	0.1	0.1	2.4	0.3
137	VN-P	Vietnam	Hai Duong	paddy	59.7	16.3	10.7	1.9	10.6	0.7	2.7	23.4	56.8	11.5	2.1	0.1	0.1	2.8	0.4
138	VN-P	Vietnam	Hai Duong	paddy	51.6	15.7	11.8	2.7	17.0	1.2	2.7	25.2	54.5	11.9	2.3	0.2	0.0	2.9	0.4
139	VN-P	Vietnam	Hai Duong	paddy	39.5	14.4	12.0	4.2	27.7	2.3	3.1	22.2	46.1	17.8	3.6	0.2	0.1	6.2	0.6
140	VN-P	Vietnam	Hai Duong	paddy	36.6	18.5	15.1	3.3	24.8	1.8	2.2	22.5	56.2	13.1	2.3	0.1	0.1	3.1	0.4
141	VN-P	Vietnam	Hai Duong	paddy	47.7	18.7	13.4	2.6	16.3	1.3	2.6	22.9	54.9	13.2	2.5	0.2	0.0	3.3	0.4
142	VN-P	Vietnam	Hai Duong	paddy	54.6	18.0	12.7	2.7	11.1	0.9	4.4	27.6	52.8	10.2	1.4	0.1	0.0	3.1	0.2
143	VN-P	Vietnam	Hai Duong	paddy	42.2	14.4	15.1	4.1	22.7	1.4	2.5	25.2	51.2	14.3	2.6	0.2	0.0	3.6	0.4
144	VN-P	Vietnam	Vinh Phúc	paddy	60.5	11.0	11.6	2.7	13.7	0.6	1.8	20.9	66.2	8.4	1.4	0.0	0.0	1.2	0.1
145	VN-P	Vietnam	Vinh Phúc	paddy	69.3	11.0	9.1	1.8	8.5	0.3	1.8	20.4	65.6	9.3	1.3	0.0	0.0	1.4	0.2
146	VN-P	Vietnam	Vinh Phúc	paddy	45.8	9.0	12.1	4.7	26.5	2.0	1.4	16.6	78.7	2.1	0.7	0.0	0.1	0.3	0.1
147	VN-P	Vietnam	Vinh Phúc	paddy	44.6	8.1	11.6	4.5	29.3	1.9	1.0	14.4	80.7	2.7	0.7	0.0	0.1	0.3	0.1
148	VN-P	Vietnam	Vinh Phúc	paddy	42.4	13.8	14.7	3.9	23.6	1.6	1.4	17.3	75.5	4.1	1.1	0.0	0.0	0.4	0.2
149	VN-P	Vietnam	Vinh Phúc	paddy	69.2	10.9	8.7	1.8	9.0	0.4	2.2	20.7	64.0	9.6	1.4	0.2	0.0	1.6	0.2
150	VN-P	Vietnam	Vinh Phúc	paddy	52.1	10.7	12.3	3.0	20.8	1.0	2.5	21.4	68.1	6.1	1.1	0.0	0.0	0.7	0.2
151	VN-P	Vietnam	Vinh Phúc	paddy	34.8	7.7	11.6	4.2	39.4	2.3	1.0	13.4	82.1	2.5	0.7	0.0	0.1	0.2	0.1
152	VN-P	Vietnam	Lào Cai	paddy	41.0	10.3	12.1	3.4	31.8	1.3	1.3	15.9	78.1	2.9	1.0	0.0	0.1	0.5	0.2
153	VN-P	Vietnam	Lào Cai	paddy	28.0	13.9	15.7	5.2	35.2	2.1	0.7	12.6	81.8	3.2	1.0	0.0	0.1	0.5	0.1
154	VN-P	Vietnam	Lào Cai	paddy	47.9	11.0	10.7	3.6	25.6	1.2	3.2	25.7	59.4	6.4	2.2	0.1	0.1	2.3	0.7
155	VN-P	Vietnam	Lào Cai	paddy	75.0	13.0	5.5	1.3	4.9	0.4	5.2	31.6	47.7	8.8	2.0	0.2	0.1	3.7	0.7
156	VN-P	Vietnam	Lào Cai	paddy	54.2	13.6	10.7	2.4	18.4	0.7	3.2	26.3	64.4	4.1	0.6	0.1	0.0	1.1	0.2
157	VN-P	Vietnam	Lào Cai	paddy	73.6	13.0	8.4	1.3	3.5	0.2	4.9	31.4	52.9	6.7	1.3	0.1	0.0	2.2	0.4
158	VN-P	Vietnam	Lào Cai	paddy	72.6	5.7	6.6	2.0	12.7	0.5	4.7	29.6	59.2	3.8	0.5	0.1	0.1	1.7	0.3
159	VN-P	Vietnam	Lào Cai	paddy	38.0	13.4	12.5	3.8	30.8	1.5	2.7	27.0	63.2	4.6	1.0	0.0	0.0	1.3	0.2
160	VN-P	Vietnam	Lào Cai	paddy	53.5	11.2	10.2	2.7	21.4	0.9	1.9	23.5	66.5	4.9	1.4	0.0	0.0	1.3	0.4
161	VN-P	Vietnam	Lào Cai	paddy	58.7	18.4	9.6	1.7	11.2	0.4	4.2	32.1	55.7	4.7	1.1	0.1	0.0	1.6	0.5
162	VN-P	Vietnam	Tien Giang	paddy	55.6	16.3	16.6	5.6	5.5	0.4	0.8	14.5	74.6	7.4	1.8	0.0	0.1	0.6	0.2
163	VN-P	Vietnam	Tien Giang	paddy	51.3	17.0	18.6	7.3	5.1	0.6	0.6	15.8	75.2	6.1	1.4	0.0	0.1	0.6	0.2
164	VN-P	Vietnam	Tien Giang	paddy	48.0	17.8	16.9	5.2	11.1	1.1	0.6	16.0	74.3	6.8	1.4	0.0	0.1	0.7	0.2
165	VN-P	Vietnam	Tien Giang	paddy	50.0	20.7	18.7	5.7	4.2	0.7	0.8	18.7	70.5	7.7	1.1	0.0	0.2	0.9	0.2
166	VN-P	Vietnam	Tien Giang	paddy	53.7	17.6	16.0	3.8	8.2	0.7	0.8	16.8	72.0	8.0	1.3	0.0	0.1	0.9	0.2
167	VN-P	Vietnam	Tien Giang	paddy	37.2	16.7	17.2	4.8	21.8	2.3	0.7	18.3	73.2	6.0	1.0	0.0	0.1	0.7	0.1
168	VN-P	Vietnam	Tien Giang	paddy	35.5	15.3	14.6	4.5	27.7	2.4	1.1	17.4	70.9	8.0	1.4	0.1	0.0	0.9	0.2
169	VN-P	Vietnam	Tien Giang	paddy	47.5	18.8	17.5	5.1	10.1	1.0	0.6	17.9	73.2	6.2	1.1	0.0	0.2	0.7	0.2
170	VN-P	Vietnam	Tien Giang	paddy	49.7	17.5	16.5	5.5	9.5	1.3	0.5	13.4	76.6	6.9	1.6	0.0	0.1	0.6	0.2

Table S 4.1. Correlation analyses of individual GDGTs and soil pH.

		Pearson correlation soil pH
GDGT-IIIa	r	0.588**
	P	0.000
GDGT-IIa	r	0.251**
	p	0.001
GDGT-Ia	r	-0.746**
	p	0.000
GDGT-Ib	r	0.831**
	p	0.000
GDGT-Ic	r	0.595**
	p	0.000
GDGT-IIIb	r	0.707**
	p	0.000
GDGT-IIIc	r	0.533**
	p	0.000
GDGT-IIb	r	0.788**
	p	0.000
GDGT-IIc	r	0.717**
	p	0.000
GDGT-0	r	-0.250**
	p	0.001
GDGT-1	r	-0.073
	p	0.345
GDGT-2	r	-,454**
	p	0.000
GDGT-3	r	-0.270**
	p	0.000
Crenarchaeol	r	0.364**
	p	0.000
Crenarchaeol reg	r	0.393**
	p	0.000
brGDGT/iGDGT	r	-0.258**
	p	,001
BIT index	r	-0.572**
	p	0.000
CBT	r	-0.808**
	p	0.000
MBT'	r	-0.553**
	p	0.000
Tex86'	r	-0.049
	p	0.526
Methane-Index	r	-0.416**
	p	0.000

** Correlation is significant at the 0,01 level (2-tailed).
* Correlation is significant at the 0,05 level (2-tailed).

n = 168

Supplementary data associated with this article can additionally be found in the online version at doi:10.5194/bgd-12-1-2015-supplement.

5. Comparison of lipid biomarker and gene abundance characterizing the archaeal ammonia-oxidizing community in flooded soils

A. Bannert[1], C. Mueller-Niggemann[2], K. Kleineidam[3], L. Wissing[4], Z.H. Cao[5], L. Schwark[2], M. Schloter[3]

[1]Soil Ecology, Technische Universität München, Neuherberg, Germany

[2]Institute of Geosciences, Christian-Albrechts-University of Kiel, Kiel, Germany

[3]Institute of Soil Ecology, Department of Terrestrial Ecogenetics, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany

[4]Soil Science, Technische Universität München, Freising-Weihenstephan, Germany

[5]Institute of Soil Science, Chinese Academy of Sciences, Nanjing, China

Published as "Short Communication" in *Biol. Fertil. Soils* **47** (2011) 839-843. doi:10.1007/s00374-011-0552-6

Abstract. In the last years, archaea have been identified as key players in global N cycling, especially in nitrification. Ammonia-oxidizing archaea (AOA) are postulated to belong to the new phylum Thaumarchaeota for which the lipid crenarchaeol should be specific. The ratios between two independent markers for AOA, the ammonia monooxygenase gene and crenarchaeol have been studied in different aerated soils, but so far not in flooded soils. This study investigated ammonia-oxidizing archaea in four paddy soils and a tidal wetland. Ratios were significantly higher in the paddy soils compared to the tidal wetland and in general higher as in upland soils, leading to the assumption that archaeal ammonia oxidizers different from crenarchaeol-containing Thaumarchaeota may play an important role in paddy soils.

5.1 Introduction

Archaea represent a considerable fraction of microorganisms in terrestrial ecosystems. They play an important role in the global nutrient cycles mainly of C and N (Gattinger et al. 2004;

Leininger et al. 2006). In the N cycle, archaea have been found to be particularly involved in nitrification (Schleper et al. 2005; Venter et al. 2004). This process results in the formation of nitrate, which is a substrate for denitrification that leads to N losses from soil (Wrage et al. 2001).

As recently postulated, ammonia-oxidizing archaea belong to the new archaeal phylum Thaumarchaeota (Spang et al. 2010) for which the glycerol dialkyl glycerol tetraether lipid crenarchaeol is a good biomarker. Leininger et al. (2006) could prove for many different upland soils a relative constant ratio between gene copy numbers of archaeal ammonium monooxygenase genes (*amoA*) and crenarchaeol being in the range of 15–60 ($\times 10^7$ gene copies g^{-1} dw / μg g^{-1} dw), which supports assumption of crenarchaeol as biomarker for AOA (Spang et al. 2010). However, so far no data exist on the ratio of *amoA* gene copy numbers and crenarchaeol from flooded soils, e.g., natural wetlands or paddy soils; hence, it is unclear if archaeal nitrification is also exclusively performed by microbes of the phylum Thaumarchaeota in these particular soils. As a reference parameter estimating the total archaeal abundance, caldarchaeol was assessed as an overall lipid marker being present in all major archaeal phyla (De Rosa and Gambacorta 1988).

In this study, we analysed the ratio of archaeal *amoA* genes to crenarchaeol as well as the amount of caldarchaeol in a natural tidal wetland and four agriculturally used paddy soils in the southeast of China. The abundance of *amoA* genes was determined by real-time PCR while crenarchaeol and caldarchaeol, respectively, were measured via pressurized liquid extraction.

5.2 Material and methods

Site description and soil sampling

The study sites are located in Cixi, Zhejiang Province, China, in a subtropical monsoon area, with a mean annual temperature of 16.3°C and precipitation of 1325 mm (Zhang et al. 2004). We sampled five flooded soils, one of them being a natural tidal wetland (TW) and four cultivated paddy soils (P50, P100, P300, and P2000). The coordinates of the sampled sites are: TW: 30°19′ N, 121°09′ E; P50: 30°11′ N, 121°22′ E; P100: 30°09′ N, 121°21′ E; P300:

30°06′ N, 121°31′ E; and P2000: 30°05′ N, 121°27′E. All five sites under investigation are located within 40 km. The paddy soils which had been obtained by land reclamation from the tidal wetland are used for rice cultivation for 50, 100, 300, and 2000 years, respectively, and differ in pH value, total organic C and total N content (Table 5.1). The duration of rice cultivation at the respective sites was estimated according to Cheng et al. (2009).

All soils were sampled in July 2009 at the beginning of the vegetation period. As all paddy fields are located in the same region and the agricultural management is centrally controlled in China since 1949 by instructions of the technical service bureau, a comparable management has been performed for all sites. Five independent field replicates were taken at each site with a soil auger from 0–20 cm depth. Soil aliquots for DNA extraction were shock-frozen in liquid N directly after sampling and stored at −80°C.

DNA was extracted with the FastDNA Spin Kit for soil (MP Biomedicals, USA), according to the protocol of the manufacturer. Quality and quantity of the DNA extracts were checked with a spectrophotometer (Nanodrop, PeqLab, Germany).

Quantitative real-time PCR of archaeal as well as bacterial *amoA* genes was carried out according to Töwe et al. (2010). Dilution series of the different DNA extracts were tested in a pre-experiment with all soils to avoid inhibition of PCR, e.g., by co-extracted humic substances. DNA extract dilution of 1:128 turned out to be best suited (data not shown). PCR efficiencies, calculated from the formula $Eff=[10^{(-1/slope)}-1]\times100\%$, were 94.1–98.1% for archaeal *amoA* genes and 83.1–83.5% for bacterial ones.

Glycerol dialkyl glycerol tetraether lipids were recovered from lyophilized soil via pressurized liquid extraction (DIONEX ASE 200) using a mixture of dichloromethane/ methanol (3:1; v/v) at 100°C and 7×10^6 Pa. Extracts were cleaned by Al_2O_3-solid phase extraction and filtered through 0.45-mm polytetrafluoroethylene filters. Glycerol dialkyl glycerol tetraether fractions were analysed by liquid chromatography-atmospheric pressure chemical ionization mass spectrometry on a cyanopropyl column and protonated molecular ions were recorded in selected ion monitoring as described previously (Reigstad et al. 2008).

Data were subjected to analysis of variance using the statistic program SPSS 13.0. Normal distribution of the variables was checked by Kolmogorov–Smirnov test and boxplot analysis, and homogeneity of variances by Levene test.

5.3 Results and discussion

Soil pH values significantly decreased with cultivation time (from 8.1 in TW to 7.3 in P2000), most probably as consequence of a continuing decalcification due to flooding (Zou et al. 2011). In contrast, total organic C contents significantly increased with cultivation time from 0.58% in TW up to 3.1% in P2000. The same trend with significant differences between all soils was found for total N concentrations (Table 5.1). This may be caused by the agricultural management and again especially by the flooding of the paddy fields during rice growth, because under waterlogged conditions, soil organic matter decomposition proceeds at slower rates than in well-drained, aerobic soils (Neue et al. 1997).

Based on the amount of caldarchaeol, the highest archaeal biomass values were found in the paddy soils which have a long history of rice cultivation (P300 and P2000). In contrast in the tidal wetland the lowest amounts of caldarchaeol were detected (Fig. 5.1a). As the total amount of extracted DNA increased significantly from 140 μg g^{-1} dw in TW to 1100 μg g^{-1} dw in P2000 (Table 5.1), the relative abundance of archaea based on ng of DNA decreased from TW to P2000 (data not shown).

The highest total and relative amounts of crenarchaeol were observed in TW (25.7 ng g^{-1} dw), whereas values in the paddy soils ranged between 8.1 and 22.1 ng g^{-1} dw. Compared to the amounts of crenarchaeol measured by Leininger et al. (2006) for upland soils, lower values of crenarchaeol were assessed in flooded soils. A shift in the ratio of caldarchaeol vs. crenarchaeol from 1.6 in TW up to 7.7 in the paddy soils might reflect an increasing contribution of other archaea like methanogens in the soils under rice cultivation (data not shown).

Table 5.1. Characterization of the five examined soils (tidal wetland 50, 100, 300, and 2000 years cultivated paddy soils) by different parameters: soil texture, pH value ($CaCl_2$), total organic C, and total N, nitrate and ammonium concentrations, microbial biomass C, microbial biomass N, and DNA content.

Soil parameters	**TW**	**P50**	**P100**	**P300**	**P2000**
Soil texture (% sand, silt, clay)	7.4 80.4 12.2	0.4 83.6 16.0	2.0 81.2 16.8	3.4 81.2 15.4	4.0 85.1 10.9
pH	8.1a (0.13)	7.6b (0.08)	7.6b (0.13)	7.5b (0.08)	7.3c (0.10)
TOC (%)	0.58 a (0.17)	1.7 b (0.14)	1.7 b (0.16)	2.5 c (0.16)	3.1 d (0.11)
TN (%)	0.060 a (0.012)	0.17 b (0.014)	0.19 c (0.015)	0.27 d (0.020)	0.36 e (0.019)
Nitrate (µg N g^{-1} dw)	2.1 a (0.70)	12 ab (1.2)	8.3 ab (2.9)	16 b (6.5)	2.2 a (1.7)
Ammonium (µg N g^{-1} dw)	0.42 a (0.12)	6.0 a (3.5)	25 a (19)	27 a (30)	22 a (11)
C_{mic} (µg g^{-1} dw)	150 a (58)	720 b (120)	1000 b (330)	1800 b (780)	5100 c (1300)
N_{mic} (µg g^{-1} dw)	39 a (6.8)	28 a (18)	110 ab (55)	92 ab (36)	150 b (30)
DNA-Content (µg g^{-1} dw)	140 a (35)	760 bc (24)	630 b (120)	810 bc (220)	1100 c (140)

Standard deviations are given in parentheses (n = 5). Significant differences are indicated by different letters TW tidal wetland, TOC total organic C, TN total N, C_{mic} microbial biomass C, N_{mic} microbial biomass N

With exception of the TW, the absolute gene copy numbers of *amoA* AOA determined in the four paddy sites (related to g^{-1} dw) were higher compared to values measured by Leininger et al. (2006). Lowest archaeal *amoA* copy numbers g^{-1} dw were measured in the TW compared to the four paddy soils (Fig. 5.1b). A reason could be the significantly higher pH value in TW because decreasing gene copy numbers of archaeal *amoA* genes with increasing soil pH values were described in several studies (Erguder et al. 2009; Gubry-Rangin et al. 2010; Nicol et al. 2008). Furthermore, *amoA* gene copy numbers of the ammonia-oxidizing archaea followed the increasing microbial (C_{mic}, DNA) and archaeal (caldarchaeol) biomass, respectively. Relative abundances (normalized on total amount of extracted DNA) showed no significant difference between all sites and ranged between 1.7×10^3 and 5.5×10^3 copies ng^{-1}

DNA (data not shown). One reason may be that archaeal ammonia oxidizers are able to adapt to changing environmental conditions as suggested by Leininger et al. (2006).

Concerning ammonia-oxidizing bacteria, absolute and relative abundances were lower in the paddy soils than in TW (relative amounts between 9.9×10^{1} and 5.1×10^{2} copies ng^{-1} DNA) and in general at least one order of magnitude lower than AOA expect in TW (data not shown). A dominance of ammonia-oxidizing archaea over their bacterial counterparts in paddy soils has been shown previously (Chen et al. 2008).

Calculating the ratios of archaeal *amoA* copy numbers to caldarchaeol ($\times10^{7}$ copies g^{-1} dw / $\mu g\ g^{-1}$ dw) showed no significant difference between all paddy soils (Fig. 5.1c). This indicates that the share of ammonia-oxidizing archaea on the archaeal community remained relatively constant with cultivation time. In contrast, a trend of higher ratios in the paddy soils compared to TW could be found, which was only significant for P50, supporting the hypothesis of archaeal ammonia oxidizers adapting well to the conditions in a paddy soil.

Ratios of archaeal *amoA* copy numbers to crenarchaeol ($\times10^{7}$ copies g^{-1} dw / $\mu g\ g^{-1}$ dw) were (a) higher in the paddy soils (between 610 and 2200) compared to TW (210) and (b) in general significantly higher compared to values observed by Leininger et al. (2006) which ranged between 15 and 60. This may lead to the assumption that archaeal ammonia oxidizers different from crenarchaeol-containing Thaumarchaeota could play an important role in flooded soils. However, the multidisciplinary approach in assessing microbial processes in soils and sediments by molecular genetics and lipid analysis requires a specification of the compatibility of the methods. DNA extraction and subsequent molecular analysis rather reflect the composition of the recent microbial community at this time point of sampling (snapshot), whereas extraction and analysis of core glycerol dialkyl glycerol tetraether lipids provides a time-integrated (decades to millennia) view of the preservable microbial input into soils and sediments (Kuypers et al. 2001). Comparing the data of this study with data measured in aerated soils, the reduced turnover rates in the paddy soils, which were shown by constantly increasing total organic C and total N values with cultivation time, should be taken into account that may also lead to higher amounts of extracellular DNA and lipids in soil (Lindahl 1993; Poinar et al. 1996; Willerslev et al. 2004; Pietramellara et al. 2009; Harvey

and Macko 1997). However, based on the observation that *amoA* AOB gene copy numbers decreased from TW to P2000, while *amoA* AOA showed the opposite tendency higher enrichment rates of extracellular DNA in paddy soils compared to tidal wetlands in our study are unlikely.

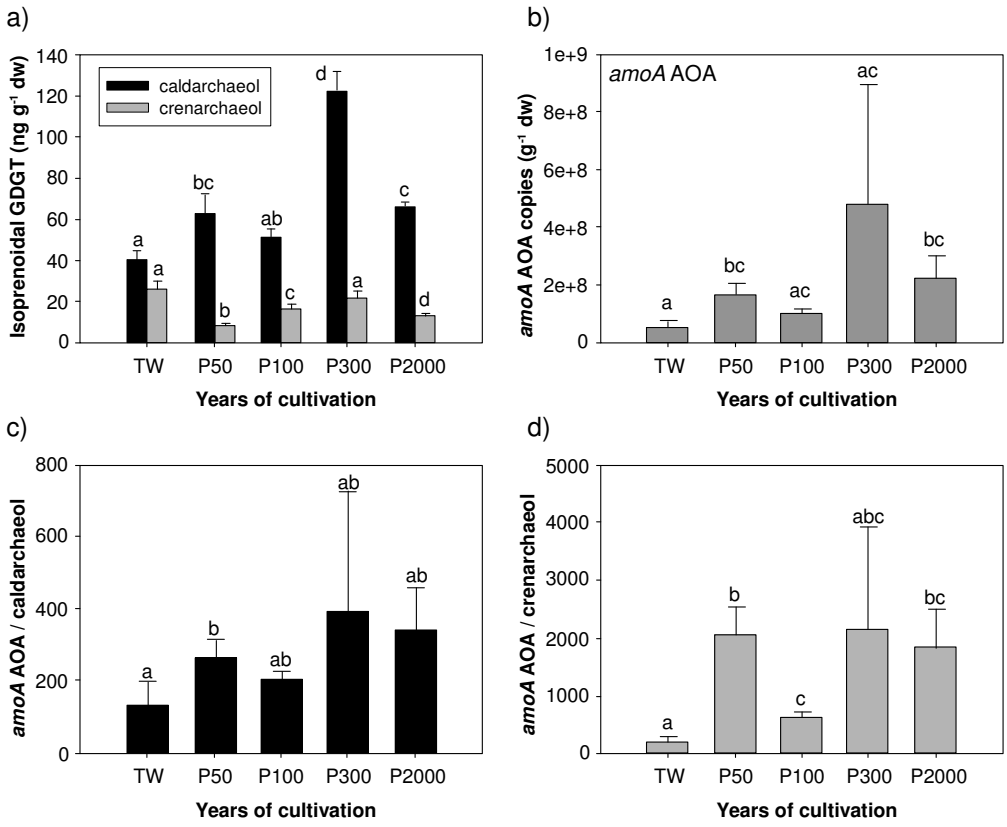


Fig. 5.1. Total copy numbers of *amoA* AOA genes **(a)** and values of isoprenoidal glycerol dialkyl glycerol tetraether lipids (caldarchaeol and crenarachaeol) **(b)** as well as ratios of total *amoA* AOA copy numbers to amounts of caldarchaeol **(c)** and crenarchaeol **(d)**, respectively ($\times 10^7$ copies g^{-1} dw / ng g^{-1} dw), in the tidal wetland (TW), the 50, 100, 300, and the 2000 years cultivated paddy soils (n = 5, error bars represent standard deviations). Significant differences are indicated by different letters.

Thus, the question remains if the higher ratios of archaeal *amoA* copy numbers to crenarchaeol in the paddy soils compared to TW are (a) due to a different community structure with more ammonia-oxidizing archaea containing no crenarchaeol in the paddy soils or (b) due to a better conservation of fossil lipids in TW and a lower input in the paddy soils.

To get a better insight which organisms are involved in archaeal ammonia oxidation in flooded soils, further molecular studies are necessary, e.g., by using metagenomic tools.

5.4 Acknowledgements

We thank Gudrun Hufnagel for excellent technical support in measuring ammonium and nitrate concentrations. Dr. Kai Mangelsorf, GFZ Potsdam is thanked for access to LC/MS/MS instrumentation. Many thanks also to Adrian Ho and Peter Frenzel for their help during soil sampling. We also thank the reviewers of the manuscript and the editor of Biol Fertil Soils for their valuable input. Financial support was provided by the German Research Foundation DFG. This paper represents a contribution to the DFG FOR 995 Biogeochemistry of paddy soil evolution.

6. Source determination and depth translocation of combustion residues in Chinese agricultural soils

C. Mueller-Niggemann[1], E. Lehndorff[2], W. Amelung[2], L. Schwark[1,3]

[1]Institute of Geosciences, Christian-Albrechts-University of Kiel, Kiel, Germany

[2]Institute of Crop Science and Resource Conservation, Soil Sciences, Bonn University, Bonn, Germany

[3]WA-OIGC, Curtin University, Perth, Australia

Submitted to Organic Geochemistry.

Abstract. Rice straw burning in agroecosystems delivers BC to soil/sediment but its contribution relative to other BC sources is unspecified. We differentiated for the first time combustion residue derived from autochthonous and pyrogenic agricultural input from allochthonous and biogenic riverine input by way of complementary analysis of the abundance and composition of BC and PAHs in natural substrates. Both combustion markers in four upland and six flooded paddy soils having different cultivation time (0-700 and 50-2000 yr) were analysed. Macromolecular BC comprised the quantitatively dominant fraction of combustion residues, whereas monomolecular PAHs occurred in trace amounts. Summed PAH concentration in paddy exceeded that in upland topsoil. The lowest BC and PAH concentration occurred in subsoils ≤ 700 yr old, independent of paddy or upland use, but BC content increased in paddy subsoil ≥ 700 yr old. Exceptional BC enrichment and diagnostic composition patterns of the derived benzene polycarboxylic acid products allowed dark buried horizons (700 and 1000 yr sites) to be identified as former topsoil. Relative distributions of 5-ring PAHs differentiated pyrogenic combustion products from allochthonous biogenic residues in substrate, topsoil and subsoil. Perylene was enriched but benzofluoranthenes and benzopyrenes depleted in tidal wetland substrate, dominantly supplied by soil/sediment erosion in the Yangtze River catchment. An identical pattern was observed in regular subsoil. Topsoil and buried horizons yielded a distinctive distribution of pyrogenic rice ash PAHs.

Although PAHs occurred in trace amounts, their source discrimination power exceeded that of BC. The former are recommended for source identification and the latter for quantifying combustion residue input to soil/sediment. Parallel application of both methodological approaches is advocated for investigations targeting combustion processes in ecosystems.

6.1 Introduction

Widespread post-harvest practices with agricultural soils are in-field burning of crop residues or adding ashes from domestic burning to a field to increase soil pH and release nutrients for better soil fertility (González-Pérez et al., 2004; Yang et al., 2006). However, open burning contributes to air pollution through the dispersal of particulate matter and emission of harmful air pollutants (Gadde et al., 2009, Ni et al., 2014). In general, biomass combustion occurs as a continuum of various reactions including, amongst others, hydrolysis, oxidation, dehydration and pyrolysis of organic matter (OM) through the effect of high temperatures that result in combustible, volatile, tarry substances and highly reactive carbonaceous char (Simoneit, 2002 and reference therein). Subsequently, large quantities of highly condensed OM are produced and finally accumulate in soil. The usually incomplete combustion of biomass residues leads to the formation of ubiquitous highly refractory polycyclic aromatic hydrocarbons (PAHs; Simoneit, 2002; Lima et al., 2005) and so-called black carbon (BC; Goldberg, 1985). The latter is characterized by stacked sheet-like macromolecular structures including cross-linked clusters of condensed (fused) aromatic rings (Schmidt and Noack 2000). As macromolecular BC is not amenable to gas chromatography (GC), it has to be broken down into monomolecular benzene polycarboxylic acids (BPCAs) using acid digestion (Brodowski et al., 2005; Wiedemeier et al., 2014). The number of carboxyl groups in the BPCAs indicates the degree of aromatic cross-linking in the original BC. The relative proportions of B3CAs (∑hemimellitic, trimellitic and trimesic acids), B4CAs (∑pyromellitic, mellophanic and prehnitic acids), B5CA (benzene pentacarboxylic acid) and B6CA (mellitic acid) can guide in the identification of the charring temperature of combusted matter, i.e. with a high contents of B6CA after OM combustion at high temperature (Brodowski et al., 2007; Schneider et al., 2010; Wolf et al., 2013; Wiedemeier et al., 2015).

PAHs are composed of two or more fused aromatic rings that have a planar structure. They can have strongly adverse effects on human health as they are carcinogenic and mutagenic. In general, the highest amounts of PAH are produced upon combustion at lower temperatures that favour the release or formation of alkylated PAH analogues. In contrast, high temperatures initiate the production of highly condensed PAHs, in particular with 5 rings (Wilcke, 2000 and reference therein). In soil, PAHs undergo various transformation processes, which include sorption and desorption, e.g. to surface-active soil OM (SOM), BC or the clay fraction. In addition, the availability of PAHs in deeper soil horizons is often associated with the translocation of dissolved OM (DOM) and/or of soil particles. Within topsoil horizons, volatilization of low molecular weight (LMW) PAHs may be important for the loss of PAHs (Wilcke, 2000 and references therein). In addition, aerobic and anaerobic microbial degradation is possible, although the latter is a slow process (Haritash and Kaushik, 2009) and the former only occurs under dry conditions or in flooded soils near the oxic/anoxic interface of rice aerenchymatous tissue (He et al., 2015 and references therein).

Besides of the dominant combustion source, some biological PAH sources may exist. Such PAHs are thought to form from biogenic precursor compounds, whose structural information is mainly preserved. For instance, retene may be a degradation product of abietic acid, a conifer resin component. Perylene may originate from terrestrial perylene quinone pigments or be derived from microbial (fungal) in-situ generation, in particular from wood decay, and has often been found in marine and freshwater sediments (Wakeham et al. 1980; Venkatesan, 1988; Grice et al., 2009).

We investigated upland and paddy soils from a 2000 yr old Chinese chronosequence (Zhejiang Province) to elucidate the incorporation and behaviour of different combustion residues in agriculturally managed soil. Previous biogeochemical studies of the same chronosequence have focussed on soil pedogenesis and heterogeneity via investigations of soil organic carbon (SOC), minerogenic content, texture, plant lipid input and microbial composition (Cheng et al., 2009; Bannert et al., 2011a,b; Chen et al., 2011; Roth et al., 2011; Wissing et al., 2011; Mueller-Niggemann et al., 2012, 2015; Kölbl et al., 2014). The

contribution of BC to the SOM and the accumulation rate in some of these paddy soils were described by Lehndorff et al. (2014).

We analysed the molecular composition of BC via BPCAs and that of PAHs in natural systems for the first time and thereby attempted to achieve source discrimination of combustion residues and spatially resolve their accumulation within topsoil, subsoil and buried soil horizons. Long term effects of combustion residue influx to soil were examined via the comparison of soils under 50, 100, 300, 700, 1000 and 2000 yr of paddy or upland management.

6.2 Material and methods

Sampling

The study sites were in a Chinese land reclamation area in Hangzhou Bay near Cixi (30° 10'N, 121° 140'E), northeast of Zhejiang Province. The bay receives Yangtze River (YR) runoff and tidal sediment transport from the East China Sea. The climate is subtropical monsoonal, with mean annual temperature 16.3 °C and mean annual precipitation 1325 mm (Cheng et al., 2009). More information about the sampling area is given by Mueller-Niggemann et al. (2012, 2015) and Kölbl et al. (2014). Briefly, in the region of Cixi successive land reclamation via dyke construction on marine tidal wetland sediments (TWs) since >1000 yr (Feng and Bao, 2005) has led to the development of differently soils of different age (50-2000 yr). Based on the time of dyke construction and information from the Edit Committee of Chorography of Cixi County (1992), different agricultural soils under continuous non-irrigated upland use (50, 100, 300 and 700 yr; denoted as NP50, NP100, NP300, NP700) as well as wetland rice cultivation (50, 100, 300, 700, 1000 and 2000 yr; (P50 to P2000, respectively).

In June 2008 one TW, one limnic freshwater sediment (FW), four upland (NP50, NP100, NP300 and NP700) and six paddy (P50, P100, P300, P700, P1000 and P2000) soil profiles were excavated and composite samples from soil horizons to a depth >100 cm were taken. Within the subsoils of the P700, P1000 and P2000 sites several buried topsoil layers were

identified, on the basis of on their dark colour and high SOC content (Wissing et al., 2011). A detailed soil description for all sampling sites according to FAO (2006) is given by Kölbl et al. (2014).

In parallel with soil samples, combustion residues were collected from burning heaps of rice straw on the field. These authentic combustion residues served as reference material for elucidation of the source end member properties and compositions of PAHs and BC. This material, though designated below as rice ash, represents a continuum of completely ashed and only marginally charred rice straw that still exhibited cellular plant structure. This type of incompletely burned rice straw is typical for field combustion conditions (Wiesenberg et al., 2009) and may add a substantial amount of refractory OM to the top soil. Samples were air-dried, sieved to < 2mm, and homogenized via grinding prior to chemical analysis.

Organic carbon analysis

Total carbon (TC) content was measured with an elemental analyser Vario EL III (Elementar Analysensysteme GmbH, Germany). Total inorganic carbon (TIC) content was determined using the Vario EL III elemental analyser coupled to a SoliTIC module. SOC was calculated by subtracting the TIC from the TC.

Polycyclic aromatic hydrocarbons

PAHs were obtained from automated solvent extraction using an ASE 200 (Dionex, USA) at75°C and 5.0 x 106 Pa. Each sample (10-12 g soil or 1 g ash) was extracted for 20 min using dichloromethane (DCM)/MeOH, 93:7, v/v). The extract was dried and separated into neutral and acid fractions via solid phase extraction (SPE) using silica gel impregnated with KOH in MeOH (10%, w/w). The fractions were obtained by successive elution with DCM and DCM/formic acid (99:1 v/v). Neutral components were further separated into three fractions using small scale chromatography (Pasteur pipette filled with silica gel) and a solvent polarity gradient. Aliphatic hydrocarbons, aromatic hydrocarbons and polar compounds were eluted with n-hexane, n-hexane/DCM (7:3, v/v) and MeOH, respectively. For quantification of PAHs, known amounts of perdeuterated d10-anthracene, d10-pyrene and d12-benzo[a]pyrene

were added as internal standards prior to separation. The aromatic fractions were analysed using GC–mass spectrometry (GC–MS) employing an Agilent 7890A GC instrument equipped with a split/splitless injector and a ZB-5HT fused silica column (30m × 0.25mm i.d., 0.25 µm film thickness; Phenomenex, USA). The oven programme was: 70 °C (3 min) to 140 °C at 10 °C/min and then to 340 °C (held 13 min) at 3 °C/min. He was the carrier gas at a constant 1 ml/min. The GC instrument was coupled to an Agilent 5975B mass spectrometer operated in scan mode (m/z 100–350) with electron ionization (EI) at 70 eV. Assignment of individual compounds was based on mass spectra or mass chromatograms and comparison with retention times of standards. Quantification was conducted by way of normalization of peak areas vs. internal standard.

Benzenepolycarboxylic acids

The analysis of BC in soil was performed according to Brodowski et al. (2005). First, potentially interfering metals in the samples (equivalent to ca. 5 mg carbon) were eliminated via digestion with trifluoroacetic acid (105 °C, 4 h). The residue was oxidized to BPCAs using 65% HNO3 (170 °C, 8 h). BPCAs were purified using a cation exchange column (Dowex 50 W X 8, 200–400 mesh, Fluka, Steinheim, Germany). After derivatization with pyridine and N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA), the trimethylsilyl derivatives of BPCAs were analysed using GC–flame ionization detector (GC–FID) with an Agilent 6890 GC instrument equipped with a split/splitless injector and an Equity-5 column (30m × 0.25mm i.d., 0.25 µm film thickness; Supelco, Steinheim, Germany). For quantification of BPCAs, known amounts of biphenylene dicarboxylic acid and citric acid were added, with the latter added immediately before the cation exchange step. The summed BPCAs were used to calculate BC using a minimum conversion factor of 2.27 (Glaser et al., 1998; Brodowski et al., 2005).

6.3 Results and discussion

The total SOC concentration ranged from 0.13 to 3.62%. In general it was higher in topsoils than in subsoil, with the highest values in paddy soils. Within the deeper soil layers of the P700, P1000 and P2000 paddy soil profiles, buried horizons occurred that had higher SOC content than the adjacent subsoils (Table 6.1).

PAHs and BCs were present in all soil layers (Fig. 6.1; Table 6.1). The quantitatively dominant fraction of these condensed, highly aromatic substances in the SOM was the macromolecular BC, with abundance ranging between 16 and 512 mg/g SOC. The lowest BC contents were in the youngest soils (< 700 yr), with no significant difference between paddy and upland usage. Paddy soils ≥ 700 yr showed a BC increase with depth and possessed remarkably high proportions of BC in dark buried horizons found in the P700 site and P1000 site. These were considered either former topsoil horizons or sedimentary layers due to inundation of soil by marine or limnic flooding.

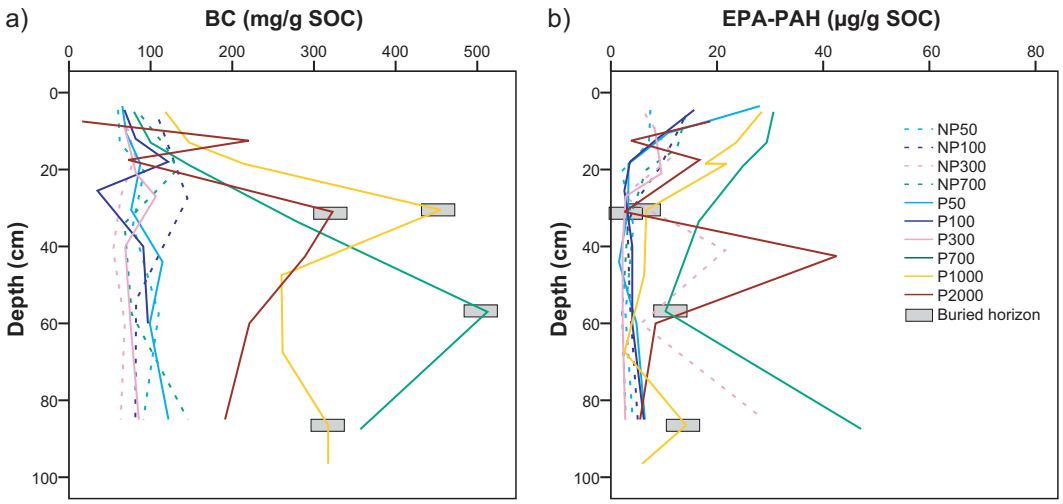


Fig. 6.1. Soil profiles with **(a)** BC content and **(b)** EPA-PAH content normalized to SOC. Grey squares denote dark coloured buried topsoil horizons. Dotted lines indicate upland soil and filled lines paddy soil. Abbreviations refer to different sampling sites: upland (NP), paddy (P), number indicates the cultivation time of soil. EPA-PAHs in **(b)** include phenanthrene, anthracene, chrysene, fluoranthene, benz[a]anthracene, pyrene, benzo[b]fluoranthene, benzo[k]fluoranthene, benzo[a]pyrene, indenopyrene and benzo[g,h,i]perylene.

PAHs occurred in only trace amount (e.g., EPA-PAH range between 1.5 and 47.0 µg/g SOC). In contrast to BC, the proportion of PAHs in SOM in paddy soils ≥ 700 yr was slightly higher than in upland topsoil (Fig. 6.1). In the deeper soil horizons, higher PAH concentration was noted in the first to second soil horizon below the buried topsoil horizons, suggesting translocation of PAHs after long time paddy management.

PAHs, despite low concentration, could be used for differentiation of combustion effects in substrates, topsoils and subsoils due to higher compositional variability. In particular, the dark buried horizons at the P700 site exhibited a composition of BC and PAHs similar to that of upland topsoil (Fig. 6.2a and b). Based on these signatures, buried horizons were interpreted as former topsoil horizons and not as sedimentary layers.

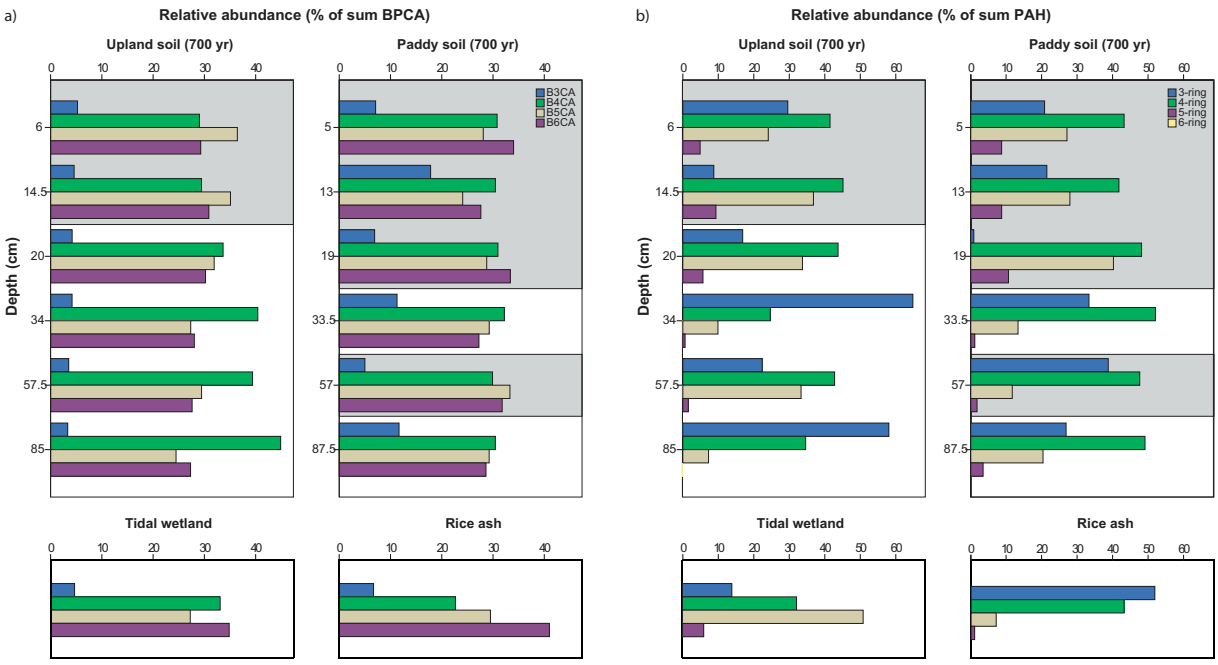


Fig. 6.2. (a) BPCA pattern and **(b)** PAH pattern from 700 yr upland and paddy soil, tidal wetland substrate and rice ash. B3CAs, ∑tricarboxylic acids; B4CAs, ∑tetracarboxylic acids; B5CA, pentacarboxylic acid; B6CA, mellitic acid. Grey shaded areas denote topsoil horizons.

The BPCA pattern (Fig. 6.2), especially the B6CA/B4CA ratio, indicated a higher proportion of rice straw ash (1.8) in paddy topsoil (1.1) (Lehndorff et al., 2014). The relative proportion of 3-, 4-, 5- and 6-ring PAHs did not agree with that of BC. Potential causes for the deviation are LMW (3- and 4-ring) PAHs prone to volatilization or translocation with DOM (Wilcke, 2000 and reference therein) as well as microbial degradation (Haritash and Kaushik, 2009).

Table 6.1. Concentrations of SOC and EPA-PAHs including (phenanthrene, anthracene, chrysene, fluoranthene, benz[a]anthracene, pyrene, benzo[b]fluoranthene, benzo[k]fluoranthene, benzo[a]pyrene, indenopyrene and benzo[g,h,i]perylene) and content of BC in different soil horizons, substrate and rice ash. PAH %) indicates proportion of EPA-PAHs to sum of all PAHs. Age denotes duration of agricultural soil cultivation.

Horizon	Age	N	SOC		EPA-PAH						BC			
	(yr)		(%)		(ng/g dw)		(µg/g SOC)		(% of PAH)		(mg/g dw)		(mg/g SOC)	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Upland topsoils	50	2	0.93	0.25	65	21	6.9	0.2	80.0	1.4	0.60	0.07	61	2
	100	2	0.74	0.06	86	17	11.6	3.2	80.5	0.7	0.97	.	110	.
	300	2	1.17	0.53	67	16	5.7	3.7	82.5	3.5	0.58	0.18	74	10
	700	2	0.77	0.28	100	45	12.9	1.1	81.5	3.5	0.77	0.15	104	22
Paddy topsoils	50	3	1.79	0.68	205	231	11.4	12.7	73.3	11.5	0.92	0.39	74	12
	100	3	1.77	0.62	153	131	8.6	6.1	83.0	1.7	1.18	0.32	90	28
	300	2	2.51	0.72	175	45	7.0	0.9	83.5	0.7	1.76	0.09	75	10
	700	3	2.45	0.87	562	289	22.9	3.0	81.7	3.8	1.84	0.49	109	35
	1000	3	1.12	0.35	193	116	17.2	4.4	85.0	0.8	1.32	0.27	173	48
	2000	2	3.64	0.91	534	204	14.7	1.4	86.0	0.0	1.15	0.88	44	40
Upland subsoil	50	4	0.42	0.07	13	3	3.3	1.1	60.0	4.5	0.48	0.07	93	13
	100	4	0.49	0.17	21	12	4.2	1.0	62.5	9.0	0.67	0.35	103	37
	300	4	0.35	0.12	47	41	14.6	12.7	70.8	10.6	0.31	0.03	62	6
	700	4	0.34	0.04	13	9	3.7	2.0	67.5	8.1	0.40	0.09	106	39
Paddy subsoil	50	4	0.33	0.02	14	8	4.0	2.1	55.5	9.3	0.53	0.13	103	20
	100	4	0.40	0.14	16	3	4.2	1.5	62.8	6.0	0.47	0.02	74	34
	300	4	0.36	0.06	9	2	2.5	0.3	65.5	6.2	0.42	0.10	84	16
	1000	1	0.27	.	17	.	6.3	.	70.0	.	0.76	.	260	.
	2000	2	0.14	0.01	10	4	7.0	2.1	64.0	11.3	0.33	0.06	206	21
Buried horizon	700	3	0.65	0.04	124	110	18.9	16.2	84.0	1.7	2.75	0.56	463	43
	1000	3	1.07	0.52	98	85	8.7	4.9	82.7	2.3	4.01	1.53	363	80
	2000	1	0.34	.	9	.	2.6	.	72.0	.	1.24	.	323	.
P subsoils adjacent to buried	700	4	0.27	0.08	111	67	42.5	25.6	82.0	2.4	1.00	0.18	336	38
	1000	3	0.34	0.03	16	6	4.7	1.8	64.3	18.8	1.09	0.02	280	32
	2000	3	0.32	0.14	42	53	16.1	22.9	84.7	5.8	0.91	0.20	266	40
Marsh soil		1	0.63	.	8	.	1.2	.	85.0	.				
TW		1	0.61	.	84	.	13.7	.	54.0	.	0.65	.	60	.
Rice ash		1	7.41	.	405	.	5.5	.	76.0	.	44.15	.	55	.

In the following, from the wide suite of PAH components, the 5-ring PAH analogues were preferentially used for source differentiation as these components are less volatile, less water soluble and less prone to microbial degradation than LMW counterparts. The TW substrate was highly enriched in perylene but depleted in benzofluoranthenes and benzopyrenes (Fig. 6.3). This signature contrasted sharply with that in the rice ash, whereas the freshwater site exhibited a mixed signature (Fig. 6.3). Perylene is a common soil PAH of anaerobic microbial rather than combustion origin (Venkatesan, 1988; Wilcke, 2007; Grice et al., 2009). It has been reported to occur in Yangtze estuarine sediments in concentration and composition relative to other PAHs which are somewhat similar to our observations (Bouloubassi et al., 2001; Hu et al., 2014). Other tributaries of the East China Sea near Shanghai, e.g. the Huangpu and Suzhou rivers, drain a much smaller part of the hinterland and also exhibit an elevated concentration of perylene indicative of soil erosion, but not as pronounced as for the Yangtze estuary (Liu et al., 2008). Consequently, the high proportions of perylene vs. those of other 5-ring PAH can be utilized as a source indicator for allochthonous PAH input, preferably via the Yangtze River. Regular subsoil had a PAH pattern identical to that of marine substrate, whereas topsoil and buried horizons yielded a distinctive autochthonous rice ash combustion PAH pattern (Fig. 6.3).

Further differentiation of paddy vs. upland soil based on BC distribution was difficult to achieve as in-field combustion of crop residues was practised in both, though in upland soil the preferential cultivation of vegetables led to less combustion activity than for cereal cropping. The concentrations of BC and EPA-PAHs in general gave good agreement and BC concentration in paddy topsoil outweighed that in upland topsoil (Fig. 6.4c). The highest BC concentration was encountered in buried horizons and exceeded the values for paddy topsoils (Fig. 6.4c). The lowest BC concentration values were in subsoil, with no significant difference between paddy and upland use. Subsoil adjacent to buried horizon but not classified as such, exhibited BC concentration intermediate between buried topsoil horizons and regular subsoil horizons (Fig. 6.4c). This potentially indicates either that colloidal BC was translocated within the soil profile, or that this BC was ancient in origin, already deposited with the Yangtze sediment.

Variation in PAH composition was more complex and systematic than for BC, potentially due to more variable input sources, i.e. from local combustion, longer distance transport in particulate and gaseous phase, soil erosion in the hinterland and higher structural variability and associated physicochemical properties than for BC. The concentration of PAHs allowed the same differentiation as for BC (Fig. 6.4c) but in addition PAHs offered more discriminative power than BC composition (Fig. 6.4). The degree of BC condensation was more uniform than the relative proportions of benzopyrenes (BaPY/BaPY+BePY). The latter ratio grouped topsoils, subsoils and buried horizons (Fig. 6.4d). End member marine sediment and rice ash fell into appropriate categories using this binary discrimination diagram. Improved separation of pedogenic and pyrogenic PAH distribution was achieved by discrimination diagrams using the relative abundance of 6-ring PAHs indenopyrene vs. benzoperylene and 4-ring PAH benz[a]anthracene (Fig. 6.4a). Subsoil except for buried horizons and directly adjacent horizons lacked benz[a]anthracene and exhibited high variability in relative indenopyrene abundance. Topsoils showed invariant relative indenopyrene proportions but differentiated paddy vs. upland soils according to relative benz[a]anthracene proportion (Fig. 6.4a).

The benzofluoranthenes (BF) vs. perylene ratio was consistently high for topsoil, buried horizons and adjacent subsoil horizons (Fig. 6.4b). This reinforced an origin of buried horizons from plant ash input that showed identical properties but also indicated that PAHs had been redistributed from buried horizons into adjacent layers, as supported by the benz[a]anthracene distribution (Fig. 6.4a). Due to the low water solubility of high MW PAHs, the most probable mechanism for translocation is by transport in colloidal solution.

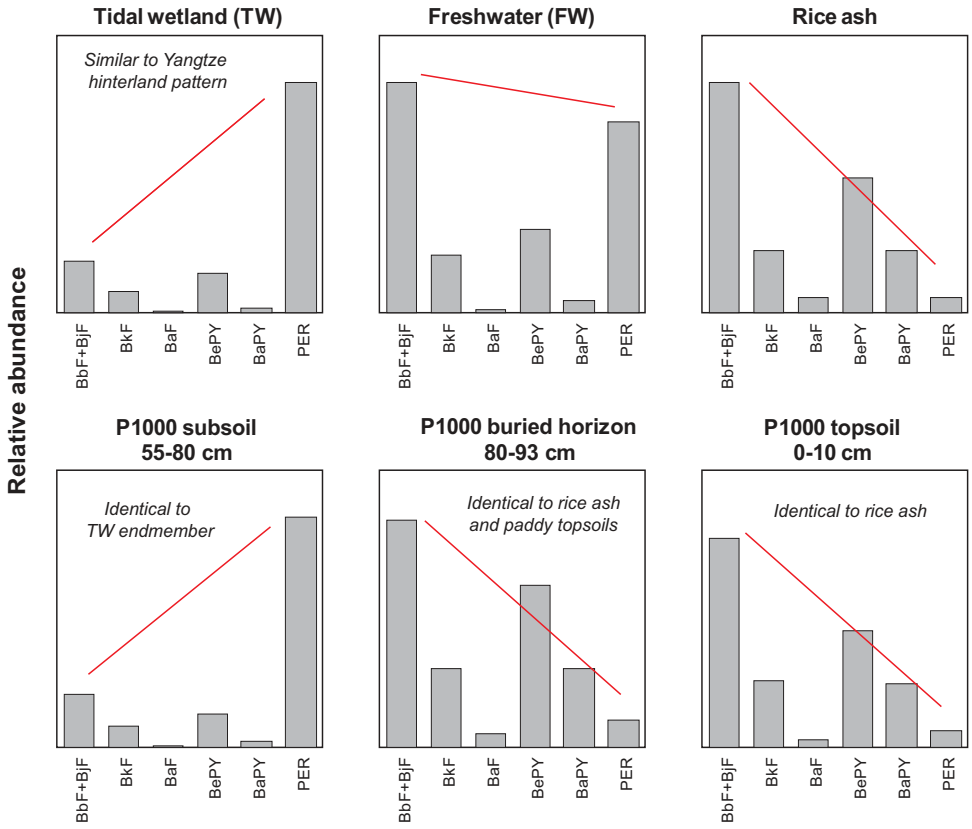


Fig. 6.3. Relative distribution of 5-ring PAH in substrate end members and soils. BxF, benzo[x]fluoranthenes; BxP, benzo[x]pyrenes; PER, perylene.

The ratio of benzo[a]pyrene vs. benzo[e]pyrene separated paddy from upland soil (Fig. 6.4b and d), which can be explained by (i) different combustion intensity in the two management types (more frequent combustion of rice straw; potentially less combustion or other combustion processes in upland cultivation of vegetables or (ii) less degradation of the more labile benzo[a]pyrene in paddy soils. Both explanations are consistent with higher total PAH concentration in paddy vs. upland soils (Fig. 6.4c).

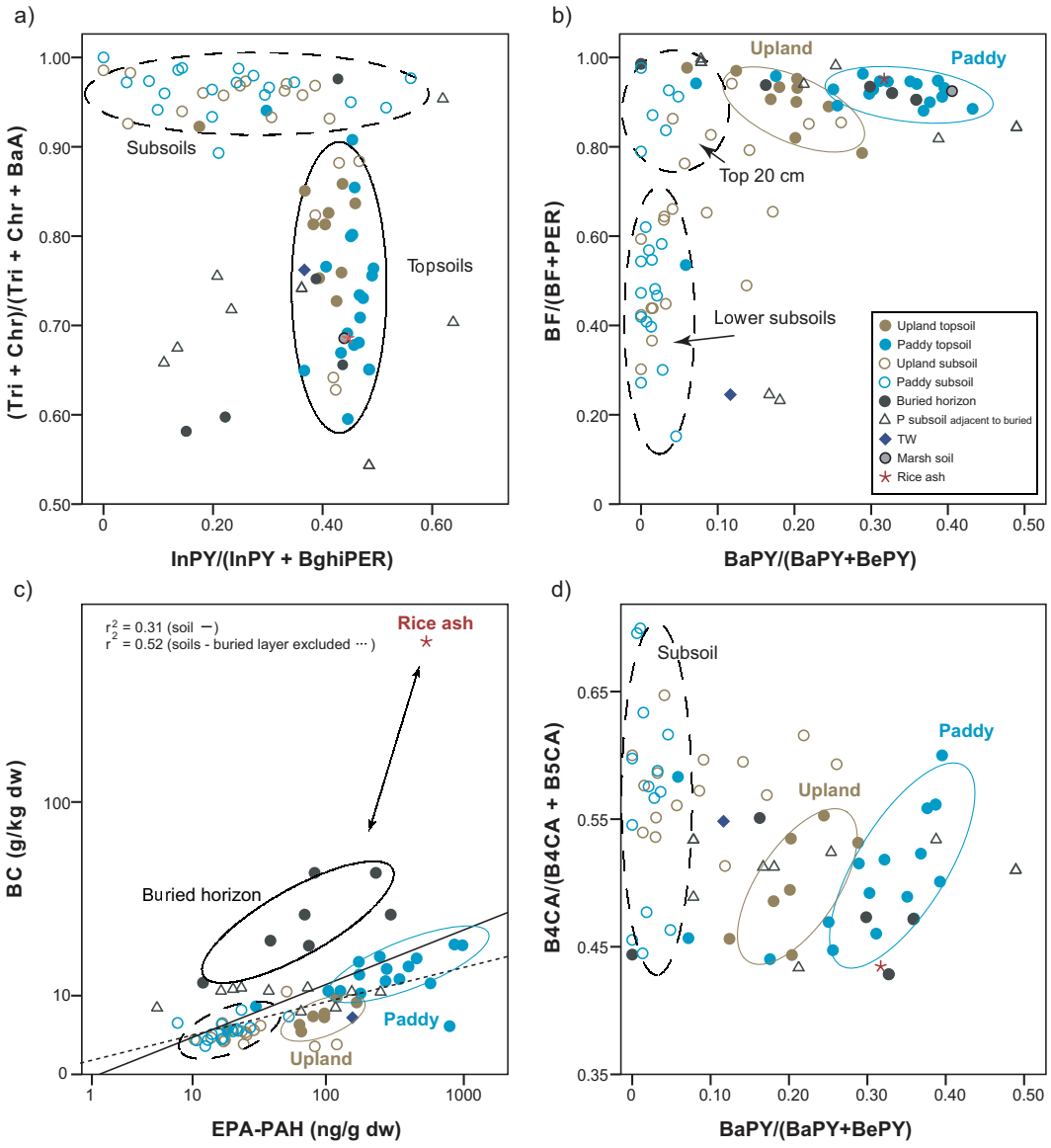


Fig. 6.4. Discrimination diagrams based on BC and PAH composition showing the relative distribution of 4- to 6-ring PAH in substrate end members and soils. BxF, benzo[x]fluoanthenes; BxP, benzo[x]pyrenes; PER, perylene; Tri+Chr+BaA, triphenylene + chrysene + benz[a]anthracene; InPY+BghiPER, indenopyrene + benzo[ghi]perylene; BxCA, = benzene with x carboxyl groups => from black carbon of different degree of condensation. EPA-PAHs include phenanthrene, anthracene, chrysene, fluoranthene, benz[a]anthracene, pyrene, benzo[b]fluoranthene, benzo[k]fluoranthene, benzo[a]pyrene, indenopyrene and benzo[g,h,i]perylene.

6.4 Conclusions

PAH concentration and compositional variation allowed specific source allocation for the soil substrates. As deduced from the n-alkane distribution of subsoils, described by Mueller-Niggemann et al. (2015), the marine tidal wetland substrate contained almost exclusively OM of terrigenous origin, derived from soil and sediment erosion in the catchment area of the Yangtze River. Perylene abundance indicated that sediment input from small local rivers was insignificant. Buried horizons were shown to be former cultivated topsoil horizons that received substantial input of in-field plant combustion residues. Colloidal transport of PAHs and eventually BC might have translocated some these components from buried horizons into adjacent strata above and below.

In general, results obtained from BC and PAH analyses were conformal but PAHs possessed a higher differentiation potential due to higher structural diversity for this compound group. Based on BC and PAH distribution it was evident that combustion residues play a major role in paddy soil development.

6.5 Acknowledgements

We thank the German Research Foundation (DFG) for financial support (Schw554/20). Asian and European partners of Research Initiative FOR 995 are thanked for field work collaboration.

7. Summary

The objective of this thesis was to identify and characterize plant and microbial lipid input and its alteration during paddy soil development. For this initially, a set of sedimentary substrates and two chronosequence of soils developed (in land reclamation areas of Zhejiang Province, China) under paddy and upland (non-paddy) conditions were investigated for their lipid composition. In order to assess the degree of topsoil heterogeneity a wide suite of parameters were determined, including those that were highly variable over short time intervals like growth seasons, e.g. soil moisture or ammonium, nitrate, and phosphate content. These parameters were contrasted with those representing time-integrated values including soil colour, magnetic susceptibility, TIC, TOC, TN and TS content, ^{13}C and ^{15}N isotope signatures, total lipid and aliphatic hydrocarbon yields. All properties or parameters underwent rigorous statistical analyses (e.g. Kruskal-Wallis test, cluster analysis) to determine intra- vs. inter-site variability in biogeochemical soil properties. Results showed that differences between sites were higher than variability within the topsoils (10 cm depth) of individual sites. Except for two paddy fields (P500, P1000) that have not been under constant paddy use, the oldest soils ≥ 300 to 2000 yr separated strongly from the younger soils ≤ 100 yr. Tight clustering indicated full evolution of paddy soil characteristic in the ≥ 300 yr paddy soils.

The type of predominant plant input into soils and the proportion of rice plant derived organic matter was investigated using firstly a chemotaxonomic study of the wax alkanes composition of rice plants and various crop plants (rice, maize, sorghum, rape, mustard, beans and cotton) followed by the comparison of these lipids in soils. Results showed that the *n*-alkanes in agricultural topsoil reflected the input of plant organic matter remaining in the field after harvesting. The soil *n*-alkane distributions in paddy and upland topsoil differed, with alkane distribution patterns indicating a predominant rice root input in paddy soils and an input from other crop plants in upland soil. Agricultural usage altered the lipid composition of topsoil already after 50 yr to reflect the *n*-alkane patterns of crops. In addition, their relative proportion was found to increase with rice cultivation time. Lipid compositional changes from topsoil to subsoil indicated different plant origins as well as advanced pedogenesis in tidal

wetland sediment after land reclamation and associated development into lacustrine/limnic sediment, marsh soil and finally arable soil. Subsoil did not exhibit notable differences in lipid composition between sites, suggesting good preservation of substrate lipid signatures.

Glycerol dialkyl glycerol tetraethers (GDGTs) are characteristic cell membrane lipids of archaea and bacteria and have been reported in abundance from terrestrial and marine environments. Comparison of paddy and upland soils indicate effects of different management types (e.g. soil/water and redox conditions, cultivated plants) on the distribution of GDGTs rather than original soil type on which soils developed. Bacterial tetraether lipids as branched GDGTs (brGDGTs) predominate in all soils, but were higher in paddies and showed a fourfold increase after 2000 yr rice cultivation suggesting an influence of long-term processes onto the proportions of soil microorganisms. Furthermore, the lipid signature of paddy and upland soils with highly variable substrate composition, in contrast to the uniform tidal wetland substrate in the Cixi area, and locations in different climate zones were investigated. Results confirmed that differences in microbial tetraether lipid composition were predominantly controlled by the management type and only secondarily by the climatic conditions. Thus, for instance, the higher intensity of rice cultivation on tropically located paddy soils, which had been exposed to longer durations of water flooding and anoxic conditions prevailing therefrom, had an impact on the microbial consortia. Bacterial indices indicated a strong influence of soil moisture, soil temperature and soil air regime. The cyclization ratio of brGDGTs (CBT) was generally negatively correlated to increasing soil pH. Interestingly, in neutral to alkaline soils (with pH values > 6.5) CBT values stayed rather constant with an offset observed between paddy soils and upland soils suggesting soil moisture could be one further potential factor for the varying CBT values in paddy and upland soil, especially under alkaline conditions. The degree of methylation of brGDGTs (MBT') indicated a dual influence of pH and mean annual air temperature (MAT) on methylation, as had been shown previously. However, results of this study demonstrate that the MBT' usually showed low values in paddy soil compared to the adjacently located upland soil. These differences in MBT' confirm that other factors than MAT and MAP affect the degree of methylation of brGDGT on a regional scale. In addition, higher brGDGT-based temperatures

(T_{MC}) in soils under upland management than under paddy management approve the influence of specific paddy management practises that control the soil moisture and in turn affect the mean soil temperature (MST) as a regulator of bacterial tetraether lipid composition. Archaeal lipid signatures indicated an increased presence of methanogenic archaea in paddy soils that had an abundant proportion of GDGT-0. In upland soil higher contents of crenarchaeol, considered to be a highly specific biological maker for ammonia-oxidizing *Thaumarchaeota,* prevailed. Despite enhanced concentrations of molecular markers for methanogens in subaquatic paddy soils, long-time paddy management led to the successive increase of the relative abundances of crenarchaeol.

Microbial processes in paddy soils were highly variable due to alternating changes in environmental conditions. Therefore, a combined approach using microbial lipid analysis and molecular genetics was used to characterize the archaeal community involved in nitrification. Nitrogen cycling in alternate dry and flooded paddy soils was rapid due to intensive fertilization and substantial nitrogen loss via leaching. The oxidation of ammonia to nitrite is the first step in the nitrification and is carried out by bacteria and archaea, with the latter shown to be predominant in paddy soils. The key enzyme for this process is the ammonium monooxygenase, whose subunit A is encoded by the functional gene *amoA* in ammonia-oxidizing archaea (AOA). Genetic methods target those microbes that were present at the time of sampling, so that consequently the results of genomics reveal a snapshot of the microbial activity at high temporal and functional resolution. In contrast, lipidomic analysis yields a time-integrated picture of microbial structure in soils. The comparison of gene copy numbers with the abundances of archaeal tetraether lipids revealed diagnostic covariance. as The sum of isoprenoid GDGTs (iGDGTs) that have been identified in all *Euryarchaeota, Crenarchaeota* and *Thaumarchaeota* lineages (reviewed in Schouten et al., 2013), versus crenarchaeol that is exclusively biosynthesized by *Thaumarchaeota*, revealed a covariance with *amoA* AOA that was more significant for the summed iGDGTs than for crenarchaeol. Nevertheless, independently from each other both approaches evidenced that periods of different length of paddy cultivation (50 to 2000 yr) did not lead to systematic increase in archaeal activity. Rather, high abundance of *amoA* gene copies in soil could be attributed to

temporary nitrogen shortage at this site, which would favour archaeal ammonium oxidation. In conjunction with the parallel high abundance of crenarchaeol in the time-integrative soil lipids it was deduced that nitrogen shortage at this site was not due to current management practice but must have prevailed over extended periods of time. In the paddy soils investigated, the ratio of *amoA* AOA gene abundances to crenarchaeol was partially inconsistent, which was contrary to observations in aerated agricultural soils. Especially the exceptional low ratio in the 100 yr old paddy soil indicated that the higher content of crenarchaeol was not supported by complementary present day gene abundances, which was explained by a higher proportion of "fossil" crenarchaeol at this site. Furthermore, it may have to be considered that other thaumarchaeotal groups are capable of synthesizing crenarchaeol in flooded soils. These thaumarchaeota potentially lacking the functional gene *amoA* may not rely on ammonia oxidation as recently observed for Group 1.1c *Thaumarchaeota* (Weber et al., 2015). Alternatively, the presence of ammonia oxidizer not belonging to the phylum *Thaumarchaeota* (recently found the exclusive group of archaea to produce crenarchaeol) or environmental effects (soil temperature, soil moisture, nutrient content) influencing phenotypic variation like the tetraether lipid composition cannot be ruled out and require further research. Results also indicate the presence of different archaeal ammonia oxidizers in paddy soils compared to tidal wetland sediments.

Agricultural soils are strongly influenced by human activities that primarily attempt to improve soil fertility and to maximize crop yields. Consequences are strong effects on soil properties (e.g. pH, porosity and moisture) and soil processes (e.g. mobilization and relocation of minerals), next to the composition of soil organic matter. The latter denotes mainly the sum of organic carbon-containing substances that originate from plant and animal remains (in various stages of decomposition) as well as from microorganism. Furthermore, soils under agricultural management may also receive anthropogenic SOM as from e.g., lubricant and fossil fuel contaminations or combustion residues. Especially in paddy soils the combination of increased insertion of ashes (burnt either *in situ* or in heaps scattered on the field) from combustion of rice straw into soil, mediated by the redox conditions prevailing, promote the presence of substantial amounts of black carbon (BC) and polycyclic aromatic hydrocarbons

(PAHs). Results indicate a relation between the concentrations of the predominating BC and the PAHs in paddy and upland soils, with highest amounts in topsoils under paddy management. Distribution patterns of BPCAs, the break down products of the complex BC structure, showed no relationship to those of single PAHs. The latter reveal diagnostic compositional variation, allowing source allocation and discrimination.

References

Albro, P.W. (1976) Bacterial waxes. In: Kolattukudy, P.E. (Ed.), *Chemistry and Biochemistry of Natural Waxes*. Elsevier, Amsterdam, pp. 419–450.

Amblès, A., Parlanti, E., Jambu, P., Mayoungou, P., Jacquesy, J.-C. (1994) *n*-Alkane oxidation in soil. Formation of internal monoalkenes. *Geoderma 64*, 111–124.

Amelung, W., Brodowski, S., Sandhage-Hofmann, A., Bol, R. (2008) Combining biomarker with stable isotope analyses for assessing the transformation and turnover of soil organic matter. In: Sparks, D.L. (Ed.), *Advances in Agronomy* 100. Elsevier Inc., pp. 155–250.

Amundson, R. (2001) The carbon budget in soils. *Annual Review of Earth and Planetary Sciences* 29, 535–562.

Angel, R., Claus, P., Conrad, R. (2012) Methanogenic archaea are globally ubiquitous in aerated soils and become active under wet anoxic conditions. *ISME J.* 6, 847–862.

Ariosa, Y., Quesada, A., Aburto, J., Carrasco, D., Carreres, R., Leganés, F., Valiente, E.F. (2004) Epiphytic cyanobacteria on *Chara vulgaris* are the main contributors to N_2 fixation in rice fields. *Appl. Enrivon. Microbiol. 70*, 5391–5397.

Atahan, P., Grice, K., Dodson, J. (2007) Agriculture and environmental change at Qingpu, Yangtze delta region, China: a biomarker, stable isotope and palynological approach. *The Holocene 17*, 507–515.

Awe, G.O., Reichert, J.M., Wendroth, O.O. (2015) Temporal variability and covariance structures of soil temperature in a sugarcane field under different management practices in southern Brazil. *Soil Tillage Res.* 150, 93–106.

Ayari, A., Yang, H., Xie, S. (2013) Flooding impact on the distribution of microbial tetraether lipids in paddy rice soil in China. *Front. Earth Sci.* 7, 384–394.

Bai, Q., Gattinger, A., Zelles, L. (2000) Characterization of microbial consortia in paddy rice soil by phospholipid analysis. *Microb. Ecol. 39*, 273–281.

Baldock, J.A., Skjemstad, J.O. (2000) Role of the soil matrix and minerals in protecting natural organic materials against biological attack. *Organic Geochemistry* 31, 697–710.

Balesdent, J., Mariotti, A. (1996) Measurement of soil organic matter turnover using ^{13}C-natural abundance. In: Boutton, T.W., Yamasaki, S. (Eds.), *Mass Spectrometry of Soils*. Marcel Dekker, Inc., New York, pp. 83–111.

Bannert, A., Kleineidam, K., Wissing, L., Mueller-Niggemann, C., Vogelsang, V., Welzl, G., Cao, Z.H., Schloter, M. (2011a) Changes in diversity and functional gene abundances of microbial communities involved in nitrogen fixation, nitrification, and denitrification in a tidal wetland versus paddy soils cultivated for different time periods. *Appl. Enrivon. Microbiol. 77*, 6109–6116.

Bannert, A., Mueller-Niggemann, C., Kleineidam, K., Wissing, L., Cao, Z.H., Schwark, L., Schloter, M. (2011b) Comparison of lipid biomarker and gene abundance characterizing the archaeal ammonia-oxidizing community in flooded soils. *Biol. Fertil. Soils 47*, 839–843.

Bastow, T., van Aarssen, B.G.K., Lang, D. (2007) Rapid smallscale separation of saturate, aromatic and polar components in petroleum. *Org. Geochem. 38*, 1235–1250.

Bauersachs, T., Weidenbach, K., Schmitz, R.A., Schwark, L. (2015) Distribution of glycerol ether lipids in halophilic, methanogenic and hyperthermophilic archaea. *Organic Geochemistry* 83-84, 101–108.

Bianchi, A., Bianchi, G. (1990) Surface lipid composition of C_3 and C_4 plants. *Biochem: Syst. Ecol. 18*, 533–537.

Bianchi, G. (1995) Plant waxes. In: Hamilton R.J. (Ed.) *Waxes: Chemistry, Molecular Biology and Functions.* The Oily Press, Dundee, pp. 175–222.

Bianchi, G., Lupotto, E., Russo, S. (1979) Composition of epicuticular wax of rice, Oryza sativa. *Experientia 35*, 1417–1417.

Bischoff, J., Mangelsdorf, K., Schwamborn, G., Wagner, D. (2014) Impact of lake-level and climate changes on microbial communities in a terrestrial permafrost sequence of the El'gygytgyn crater, Far East Russian Arctic. *Permafrost and Periglac. Process.* 25, 107–116.

Blaga, C.I., Reichart, G.J., Heiri, O., Sinninghe Damsté, J.S. (2009) Tetraether membrane lipid distributions in water-column particulate matter and sediments: a study of 47 European lakes along a north-south transect. *J. Paleolimnol.* 41, 523–540.

Blume, H.P., Brümmer, G.W., Horn, R., Kandeler, E., Kögel-Knabner, I., Kretschmar, R., Stahr, K., Wilke, B.M. (2010) *Scheffer/Schachtschabel: Lehrbuch der Bodenkunde*. 16th ed., Spektrum Akademischer Verlag, Heidelberg.

Blumenberg, M., Seifert, R., Reitner, J., Pape, T., Michaelis, W. (2004) Membrane lipid patterns typify distinct anaerobic methanotrophic consortia. *Proceedings of the National Academy of Sciences of the United States of America* 101, 11111–11116.

Blumer, M. (1957) Removal of elemental sulfur from hydrocarbon fractions. *Anal. Chem. 29*, 1039–1041.

Borggaard, O.K. (1997) Composition, properties and development of Nordic soils. In: Saether, O.M., de Caritat, P. (Ed.) *Geochemical processes, weathering and groundwater recharge in catchments.* A.A. Balkema, Rotterdam, pp. 21-76.

Bouloubassi, I., Fillaux, J., Saliot, A. (2001) Hydrocarbons in surface sediments from the Changjiang (Yangtze River) Estuary, East China Sea. *Mar. Pollut. Bull. 42*, 1335–1346.

Brady, N.C., Weil, R.R. (2002) *The nature and properties of soils.* 13th ed., Prentice-Hall, NJ., pp 960.

Bräuer, T., Grootes, P.M., Nadeau, M.-J., Andersen, N. (2013) Downward carbon transport in a 2000-year rice paddy soil chronosequence traced by radiocarbon measurements. *Nuclear Instruments and Methods in Physics Research B 294*, 584–587.

Bray E.E., Evans E.D. (1961) Distribution of *n*-paraffins as a clue to recognition of source beds. *Geochim. Cosmochim. Acta 22*, 2-15.

Brochier-Armanet, C., Boussau, B., Gribaldo, S., Forterre, P. (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245–252.

Brodowski, S., Amelung, W., Haumaier, L., Zech, W. (2007) Black carbon contribution to stable humus in German arable soils. *Geoderma, 139*, 220–228.

Brodowski, S., Rodionov, A., Haumaier, L., Glaser, B., Amelung, W. (2005) Revised black carbon assessment using benzene polycarboxylic acids. *Org. Geochem. 36*, 1299–1310.

Bush, R.T., McInerney, F.A. (2013) Leaf wax *n*-alkane distributions in and across modern plants: Implications for paleoecology and chemotaxonomy. *Geochim. Cosmochim. Acta 117,* 161–179.

Cameron, K.D., Teece, M.A., Bevilacqua, E., Smart, L.B. (2002) Diversity of cuticular wax among Salix species and Populus species hybrids. *Phytochemistry 60*, 715–725.

Cao, Z.H., Ding, J.L., Hu, Z.Y., Knicker, H., Kögel-Knabner, I., Yang, L.Z., Yin, R., Lin, X.G., Dong, Y.H. (2006) Ancient paddy soils from the Neolithic age in China's Yangtze River Delta. *Naturwissenschaften 93*, 232–236.

Cao, Z.H., Ding, J.L., Hu, Z.Y., Knicker, H., Kögel-Knabner, I., Yang, L.Z., Yin, R., Lin, X. G., Dong, Y.H. (2006) Ancient paddy soils from the Neolithic age in China's Yangtze River Delta. *Die Naturwissenschaften* 93, 232–236.

Carr, A.S., Boom, A., Grimes, H.L., Chase, B.M., Meadows, M.E., Harris, A. (2014) Leaf wax *n*-alkane distributions in arid zone South African flora: Environmental controls, chemotaxonomy and palaeoecological implications. *Org. Geochem. 67*, 72–84.

Chantigny, M.H., Prévost, D., Angers, D.A., Vézina, L.-P., Chalifour, F.P. (1996) Microbial biomass and N transformations in two soils cropped with annual and perennial species. *Biol. Fert. Soils 21*, 239–244.

Chen, L.-M., Zhang, G.-L., Effland, W.R. (2011) Soil characteristic response times and pedogenic thresholds during the 1000-year evolution of a paddy soil chronosequence. *Soil Sci. Soc. Am. J. 75*, 1807–1820.

Chen, S., Huang, Y., and Zou, J. (2008) Relationship between nitrous oxide emission and winter wheat production. Biol Fertil Soils, 44, 985–989.

Cheng, W., Sakai, H., Matsushima, M., Yagi, K., Hasegawa T. (2010) Response of the floating aquatic fern Azolla filiculoides to elevated CO2, temperature, and phosphorus levels. *Hydrobiologia* 656, 5–14.

Cheng, Y.-Q., Yang, L.-Z., Cao, Z.-H., Ci, E., Yin, S. (2009) Chronosequential changes of selected pedogenic properties in paddy soils as compared with non-paddy soils. *Geoderma 151*, 31–41.

Chikaraishi, Y., Naraoka, H. (2003) Compound-specific δD–$\delta^{13}C$ analyses of *n*-alkanes extracted from terrestrial and aquatic plants. *Phytochemistry 63*, 361–371.

Coffinet, S., Huguet, A., Williamson, D., Fosse, C., Derenne, S. (2014) Potential of GDGTs as a temperature proxy along an altitudinal transect at Mount Rungwe (Tanzania). *Org. Geochem.* 68, 82–89.

Collister, J.W., Rieley, G., Stern, B., Eglinton, G., Fry, B. (1994) Compound-specific $\delta^{13}C$ analyses of leaf lipids from plants with differing carbon dioxide metabolisms. *Org. Geochem. 21*, 619–627.

Colombo, J.C., Pelletier, E., Brochu, C., Khalil, M., Catoggio, J.A. (1989) Determination of hydrocarbon sources using *n*-alkane and polyaromatic hydrocarbon distribution indexes. Case study: Rio de la Plata Estuary, Argentina. *Environ. Sci. Technol. 23*, 888–894.

Comeau, Y. (2008) Microbial metabolism. In: Henze, M., van Loodfrecht, M.C.M., Ekama, G.A., Brdjanovic, D. (ed.) *Biological wastewater treatment: principles, modelling and design*. Cambridge University Press, Cambridge, pp. 9-32.

Conrad, R. (1992) The global methane cycle: recent advances in understanding the microbial processes involved. *Environ. Microbiol. Reports 1*, 285–292.

Conrad, R. (2007) Microbial ecology of methanogens and methanotrophs. *Adv. Agron.* 96, 1–63.

Conte, M.H., Weber, J.C., Carlson, P.J., Flanagan, L.B. (2003) Molecular and carbon isotopic composition of leaf wax in vegetation and aerosols in a northern prairie ecosystem. *Oecologia 135*, 67–77.

Cranwell, P.A. (1982) Lipids of aquatic sediments and sedimenting particulates. *Prog. Lipid Res.21*, 271–308.

Davidson, E.A., Hart, S.C., Firestone, M.K. (1992) Internal cycling of nitrate in soils of a mature coniferous forest. *Ecology 73*, 1148–1156.

De Jonge, C., Hopmans, E. C., Stadnitskaia, A., Rijpstra, W. I. C., Hofland, R., Tegelaar, E., Sinninghe Damsté, J. S. (2013) Identification of novel penta- and hexamethylated branched glycerol dialkyl glycerol tetraethers in peat using HPLC-MS^2, GC-MS and GC-SMB-MS. *Org. Geochem.* 54, 78–82.

De Jonge, C., Hopmans, E. C., Zell, C. I., Kim, J. H., Schouten, S., Sinninghe Damsté, J. S. (2014) Occurrence and abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate reconstruction. *Geochim. Cosmochim. Acta* 141, 97–112.

De Rosa, M., and Gambacorta, A. (1988) The lipids of Archaebacteriaea. Prog Lipid Res, 27, 153–175.

De Rosa, M., Gambacorta, A., Lanzotti, V., Trincone, A., Harris, J. E., Grant, W. D. (1986) A range of ether core lipids from the methanogenic archaebacterium Methanosarcina barkeri. *Biochim. Biophys. Acta* 875, 487–492.

Diefendorf, A.F., Freeman, K.H., Wing, S.L., Graham, H. V. (2011) Production of *n*-alkyl lipids in living plants and implications for the geologic past. *Geochim. Cosmochim. Acta 75*, 7472–7485.

Dinel, H., Schnitzer, M., Mehuys, G.R. (1990) Soil lipids: origin, nature, content, decomposition, and effect on soil physical properties. In: Bollag, J.M., Stotzky, G. (Eds.), *Soil Biochemistry, Vol 6*. Marcel Dekker, New York, pp 397–423.

Dove, H. (1992) Using the *n*-alkanes of plant cuticular wax to estimate the species composition of herbage mixtures. *Aust. J. Agric. Res. 43*, 1711–1724.

Dove, H., Mayes, R.W., Freer, M. (1996) Effects of species, plant part, and plant age on the *n*-alkane concentrations in the cuticular wax of pasture plants. *Aust. J. Agric. Res. 47*, 1333–1347.

Dupuis, E.M., Whalen, J.K. (2007) Soil properties related to the spatial pattern of microbial biomass and respiration in agroecosystems. *Canadian J. Soil Sci. 87*, 479–484.

Eckmeier, E., Wiesenberg, G.L.B. (2009) Short-chain *n*-alkanes ($C_{16–20}$) in ancient soil are useful molecular markers for prehistoric biomass burning. *J. Archaeol. Sci. 36*, 1590–1596.

Edit Committee of Chorography of Cixi County (1992) *Chorography of Cixi County, Zhejiang*, Peoples Express House, Hangzhou.

Eglinton, G., Gonzalez, A.G., Hamilton, R.J., Raphael, R.A. (1962) Hydrocarbon constituents of the wax coatings of plant leaves: a taxonomic survey. *Phytochemistry 1*, 89–102.

Eglinton, G., Hamilton, R.J. (1967) Leaf epicuticular waxes. *Science 156*, 1322–1335.

Erguder, T.H., Boon, N., Wittebolle, L., Marzorati, M., Verstraete, W. (2009) Environmental factors shaping the ecological niches of ammoniaoxidizing archaea. FEMS Microbiol Rev, 33, 855–869.

Fahy, E., Subramaniam, S., Brown, H.A., Glass, C.K., Merrill, A.H., Murphy, R.C., Raetz, C.R.H., Russell, D.W., Seyama, Y., Shaw, W., Shimizu, T., Spener, F., van Meer, G., van Nieuwenhze, M.S., White, S.H., Witztum, J.L., Dennis, E.A (2005) A comprehensive classification system for lipids. *Journal of lipid research* 46, 839–61.

FAO (2003) World agriculture: towards 2015/2030. An FAO perspective. ed. J. Bruinsma, Earthscan Publications Ltd, London.

FAO (2006) *Guidelines for soil description, 4th edition*. Publishing Management Service, Information Division. FAO, Rome, 97 pp.

FAO statistical division (2015) http://faostat3.fao.org/browse/Q/*/E

Feng, L.-H., Bao, Y.-X. (2005) Impact of human activity on the estuary of the Qiantang River and the reclamation of tidal flats and river regulation. *Environ. Geol. 49*, 76–81.

Fernando, C.H. (1993) Rice field ecology and fish culture - an overview. *Hydrobiologia* 259, 91–113.

Ficken, K.J., Li, B., Swain, D.L., Eglinton, G. (2000) An *n*-alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes. *Org. Geochem. 31*, 745–749.

Ficken, K.J., Street-Perrott, F.A., Perrott, R.A., Swain, D.L., Olago, D.O., Eglinton, G. (1998) Glacial/interglacial variations in carbon cycling revealed by molecular and isotope stratigraphy of Lake Nkunga, Mt. Kenya, East Africa. *Org. Geochem. 29*, 1701–1719.

Frysinger, G.S., Gaines, R.B., Xu, L., Reddy, C.M. (2003) Resolving the unresolved complex mixture in petroleum-contaminated sediments. *Environ. Sci. Technol. 37*, 1653–1662.

Fuller, D.Q., Allaby, R.G., Stevens C. (2010) Domestication as innovation: the entanglement of techniques, technology and chance in the domestication of cereal crops. *World Archaeology* 42, 13–18.

Fuller, D.Q., Qin, L. (2009) Water management and labour in the origins and dispersal of Asian rice. *World Archaeology* 41, 88–111.

Gadde, B., Bonnet, S., Menke, C., Garivait, S. (2009) Air pollutant emissions from rice straw open field burning in India, Thailand and the Philippines. *Environ. Pollut. 157*, 1554–1558.

Gattinger, A., Bausenwein, U., Bruns, C. (2004) Microbial biomass and activity in composts of different composition and age. J. Plant. Nutr. Soil Sci. Z Pflanzenernährung Bodenkunde, 167, 556–561.

Gaunt, J.L., Neue, H.-U., Cassman, K.G., Olk, D.C., Arah, J.R.M., Witt, C., Ottow ,J.C.G., Grant, I.,F. (1995) Microbial biomass and organic matter turnover in wetland rice soils. *Biology and Fertility of Soils* 19, 333–342.

Gerhardt, E.C.M., Rodrigues, T.E., Müller-Santos, M., Pedrosa, F.O., Souza, E.M., Forchhammer, K., Huergo, L.F. (2015) The Bacterial signal transduction protein GlnB regulates the committed step in fatty acid biosynthesis by acting as a dissociable regulatory subunit of acetyl-CoA carboxylase. *Molecular Microbiology* 95, 1025–1035.

Goldberg, E.D. (1985) *Black carbon in the environment.* Wiley, New York, 198 pp.

Gong, Z., Chen, H., Yuan, D., Zhao, Y., Wu, Y., Zhang, G. (2007) The temporal and spatial distribution of ancient rice in China and its implications. *Chinese Science Bulletin* 52, 1071–1079.

González-Pérez, J.A, González-Vila, F.J., Almendros, G., Knicker, H. (2004) The effect of fire on soil organic matter-a review. *Environ. Int. 30*, 855–870.

Gratton, W.S., Nkongolo, K.K., Spiers, G.A. (2000) Heavy metal accumulation in soil and Jack Pine (Pinus banksiana) needles in Sudbury, Ontario, Canada. *B. Environ. Contam. Tox. 64,* 550–557.

Grice, K., Lu, H., Atahan, P., Asif, M., Hallmann, C., Greenwood, P., Maslen, E., Tulipani, S., Williford, K., Dodson, J. (2009) New insights into the origin of perylene in geological samples. *Geochim. Cosmochim. Acta 73*, 6531–6543.

Gubry-Rangin, C., Nicol, G.W., Prosser, J.I. (2010) Archaea rather than bacteria control nitrification in two agricultural acidic soils. *FEMS Microbiol Ecol 74*, 566–574.

Gülz, P.-G. (1994) Epicuticular leaf waxes in the evolution of the plant kingdom. *J. Plant Physiol. 143*, 453–464.

Haritash, A.K., Kaushik, C.P. (2009) Biodegradation aspects of polycyclic aromatic hydrocarbons (PAHs): a review. *J. Hazard. Mater.* 169, 1–15.

Harvey, H.R., Macko, S.A. (1997) Kinetics of phytoplankton decay during simulated sedimentation: changes in lipids under oxic and anoxic conditions. *Org. Geochem. 27*, 129–140.

Hays, M.D., Fine, P.M., Geron, C.D., Kleeman, M.J., Gullett, B.K. (2005) Open burning of agricultural biomass: Physical and chemical properties of particle-phase emissions. *Atmos. Environ. 39*, 6747–6764.

He, Y., Xia, W., Li, X., Lin, J., Wu, J., Xu, J. (2015) Dissipation of phenanthrene and pyrene at the aerobic–anaerobic soil interface: differentiation induced by the rhizosphere of PAH-tolerant and PAH-sensitive rice (Oryza sativa L.) cultivars. *Environ. Sci. Pollut. Res.* 22, 3908–3919.

Heffer, P. (2013) *Assessment of fertiliser use by crop at the global level 2010-2010/11*., International Fertilizer Industry Association, Paris.

Herbin, G.A., Robins, P.A. (1969) Patterns of variation and development in leaf wax alkanes. *Phytochemistry 8*, 1985–1998.

Heyng, A.M., Mayr, C., Lücke, A., Moschen, R., Wissel, H., Striewski, B., Bauersachs, T. (2015) Middle and Late Holocene paleotemperatures reconstructed from oxygen isotopes and GDGTs of sediments from Lake Pupuke, New Zealand. *Quat. Int.* 374, 3-14.

Hopmans, E.C., Schouten, S., Pancost, R.D., van der Meer, M.T.J., Sinninghe Damsté, J.S. (2000) Analysis of intact tetraether lipids in archaeal cell material and sediments by high performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* 14, 585–589.

Hopmans, E.C., Weijers, J.W.H., Schefuß, E., Herfort, L., Sinninghe Damsté, J.S., Schouten, S. (2004) A novel proxy for terrestrial organic matter in sediments based on branched and isoprenoid tetraether lipids. *Earth Planet. Sci. Lett.* 224, 107–116.

Hu, L., Ren, W., Tang, J., Li, N., Zhang, J., Chen, X. (2013) The productivity of traditional rice-fish co-culture can be increased without increasing nitrogen loss to the environment. *Agriculture, Ecosystems and Environment* 177, 28–34.

Hu, L., Shi, X., Lin, T., Guo, Z., Ma, D., Yang, Z. (2014) Perylene in surface sediments from the estuarine-inner shelf of the East China Sea: A potential indicator to assess the sediment footprint of large river influence. *Continental Shelf Research 90*, 142 – 150.

Hua, X.L., Zhu, H. (2000) *Ningbo Yearbook*. Publishing House of China, Beijing (in Chinese).

Huang, L.-M., Thompson, A., Zhang, G.-L., Chen, L.-M., Han, G.-Z., Gong, Z.-T. (2015) The use of chronosequences in studies of paddy soil evolution: A review. *Geoderma 237-238*, 199–210.

Huang, Y., Bol, R., Harkness, D.D., Ineson, P., Eglinton, G. (1996) Post-glacial variations in distributions, ^{13}C and ^{14}C contents of aliphatic hydrocarbons and bulk organic matter in three types of British acid upland soils. *Org. Geochem.* 24, 273–287.

Huguet, A., Fosse, C., Metzger, P., Fritsch, E., Derenne, S. (2010) Occurrence and distribution of extractable glycerol dialkyl glycerol tetraethers in podzols. *Org. Geochem.* 41, 291–301.

Huguet, A., Grossi, V., Belmahdi, I., Fosse, C., Derenne, S. (2015) Archaeal and bacterial tetraether lipids in tropical ponds with contrasted salinity (Guadeloupe, French West Indies): Implications for tetraether-based environmental proxies. *Org. Geochem.* 83-84, 158-169.

Huguet, A., Wiesenberg, G.L.B., Gocke, M., Fosse, C., Derenne, S. (2012) Branched tetraether membrane lipids associated with rhizoliths in loess: Rhizomicrobial overprinting of initial biomarker record. *Org. Geochem.* 43, 12–19.

Husson, O. (2013) Redox potential (Eh) and pH as drivers of soil/plant/microorganism systems: a transdisciplinary overview pointing to integrative opportunities for agronomy. *Plant Soil* 362, 389–417.

Hwang, S., Hanaki, K. (2000) Effects of oxygen concentration and moisture content of refuse on nitrification, denitrification and nitrous oxide production. *Bioresource technology* 71, 159–165.

Iost, S., Landgraf, D., Makeschin, F. (2007) Chemical soil properties of reclaimed marsh soil from Zhejiang Province P.R. China. *Geoderma 142*, 245–250.

IPPC (2007) Fourth Assessment Report: Climate Change 2007, available at: http://www.ipcc.ch/publications_and_data/publications_and_data_reports.shtml (last access: March 2011).

IRRI (2010) available at: http://www.knowledgebank.irri.org (last access: September 2011).

Islam, M.A., Du, H., Ning, J., Ye, H., Xiong, L. (2009) Characterization of glossy1-homologous genes in rice involved in leaf wax accumulation and drought resistance. *Plant Mol. Biol. 70*, 443–456.

Itzstein-Davey, F., Atahan, P., Dodson, J., Taylor, D., Zheng, H. (2007) Environmental and cultural changes during the terminal Neolithic: Qingpu, Yangtze delta, eastern China. *The Holocene 17*, 875–887.

IUSS Working Group WRB (2007) World Reference Base for Soil Resources 2006, first update 2007. *World Soil Resources Reports No. 103*. FAO, Rome.

Jaffé, R., Mead, R., Hernandez, M.E., Peralba, M.C., DiGuida, O.A. (2001) Origin and transport of sedimentary organic matter in two subtropical estuaries: a comparative, biomarker-based study. *Org. Geochem. 32*, 507–526.

Jansen, B., Nierop, K.G.J., Hageman, J. A., Cleef, A.M., Verstraten, J.M. (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.

Jeng, W.-L., Huh, C.-A. (2004) Lipids in suspended matter and sediments from the East China Sea Shelf. *Org. Geochem. 35*, 647–660.

Jiang, X., Hou, X., Zhou, X., Xin, X., Wright, A., Jia, Z. (2015) pH regulates key players of nitrification in paddy soils. *Soil Biol. Biochem.* 81, 9–16.

Jilan, S., Kangshan, W. (1989) Changjiang river plume and suspended sediment transport in Hangzhou Bay. *Cont. Shelf Res. 9*, 93–111.

Jones, C.M., Stres, B., Rosenquist, M., Hallin, S. (2008) Phylogenetic analysis of nitrite, nitric oxide, and nitrous oxide respiratory enzymes reveal a complex evolutionary history for denitrification. *Mol. Biol. Evol.* 25, 1955–1966.

Jörgensen, R., Brookes, P. (1991) Die Bestimmung der mikrobiellen Biomasse des Bodens mit der Fumigations-Extraktions-Methode, *VDLUFA-Schriftenreihe*, 666–671.

Kahmen, A., Schefuß, E., Sachse, D. (2013) Leaf water deuterium enrichment shapes leaf wax *n*-alkane δD values of angiosperm plants I: experimental evidence and mechanistic insights. *Geochim. Cosmochim. Acta 111*, 39–49.

Kalbitz, K., Kaiser, K., Fiedler, S., Kölbl, A., Amelung, W., Bräuer, T., Cao, Z., Don, A., Grootes, P., Jahn, R., Schwark, L., Vogelsang, V., Wissing, L., Kögel-Knabner, I. (2013) The carbon count of 2000 years of rice cultivation. *Glob. Chang. Biol. 19*, 1107–1113.

Kalbitz, K., Schmerwitz, J., Schwesig, D., Matzner, E. (2003) Biodegradation of soil-derived dissolved organic matter as related to its properties. *Geoderma 113*, 273–291.

Kao, S.J., Lin, F.J., Liu, K.K. (2003) Organic carbon and nitrogen contents and their isotopic compositions in surficial sediments from the East China Sea shelf and the southern Okinawa Trough. *Deep Sea Research Part II: Topical Studies in Oceanography 50*, 1203–1217.

Katsaounos, C.Z., Giokas, D.L., Vlessidis, A.G., Karayannis, M.I. (2007) Identification of longitudinal and temporal patterns of phosphorus fractionation in river sediments by non-parametric statistics and pattern recognition techniques. *Desalination 213*, 311–333.

Klotzbücher, T., Marxen, A., Vetterlein, D., Schneiker, J., Türke, M., van Sinh, N., Manh, N. H., van Chien, H., Marquez, L., Villareal, S., Bustamante, J.V., Jahn, R. (2014) Plant-available silicon in paddy soils as a key factor for sustainable rice production in Southeast Asia. *Basic and Applied Ecology*. doi:10.1016/j.baae.2014.08.002.

Knicker, H., Hilscher, A., de la Rosa, J.M., González-Pérez, J.A., González-Vila, F.J. (2013) Modification of biomarkers in pyrogenic organic matter during the initial phase of charcoal biodegradation in soils. *Geoderma 197-198*, 43–50.

Koga, Y., Morii, H. (2005) Recent advances in structural research on ether lipids from archaea including comparative and physiological aspects. *Bioscience, Biotechnology, and Biochemistry* 69, 2019–2034.

Koga, Y., Morii, H., Akagawa-Matsushita, M., Ohga, M. (1998) Correlation of polar lipid composition with 16S rRNA phylogeny in methanogens. Further analysis of lipid component parts. *Bioscience, Biotechnology, and Biochemistry* 62, 230–236.

Kögel-Knabner, I. (2002) The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. *Soil Biology and Biochemistry* 34, 139–162.

Kögel-Knabner, I., Amelung, W., Cao, Z. H., Fiedler, S., Frenzel, P., Jahn, R., Kalbitz, K., Kölbl, A., Schloter, M. (2010) Biogeochemistry of paddy soils. *Geoderma* 157, 1–14.

Kolattukudy, P., Croteau, R., Buckner, J. (1976) Biochemistry of plant waxes. In: Kolattukudy, P.E. (Ed.), *Chemistry and Biochemistry of Natural Waxes*. Elsevier, Amsterdam, pp. 289–347.

Kölbl, A., Schad, P., Jahn, R., Amelung, W., Bannert, A., Cao, Z.H., Fiedler, S., Kalbitz, K., Lehndorff, E., Müller-Niggemann, C., Schloter, M., Schwark, L., Vogelsang, V., Wissing, L., Kögel-Knabner, I. (2014) Accelerated soil formation due to paddy management on marshlands (Zhejiang Province, China). *Geoderma 228-229*, 67–89.

Kool, D.M., Dolfing, J., Wrage, N., Van Groenigen, J.W. (2011) Nitrifier denitrification as a distinct and significant source of nitrous oxide from soil. *Soil Biology and Biochemistry* 43, 174–178.

Krüger, M., Frenzel, P., Kemnitz, D., Conrad, R. (2005) Activity, structure and dynamics of the methanogenic archaeal community in a flooded Italian rice field. *FEMS Microbiol. Ecol.* 51, 323–331.

Krüger, M., Meyerdierks, A., Glöckner, F.O., Amann, R., Widdel, F., Kube, M., Reinhardt, R., Kahnt, J., Böcher, R., Thauer, R.K., Shima, S. (2003) A conspicuous nickel protein in microbial mats that oxidize methane anaerobically. *Nature* 426, 878–881.

Kuhn, T.K., Krull, E.S., Bowater, A., Grice, K., Gleixner, G. (2010) The occurrence of short chain *n*-alkanes with an even over odd predominance in higher plants and soils. *Org. Geochem. 41,* 88–95.

Kuypers,, M.M.M., Blokker,, P., Erbacher,, J., Kinkel,, H., Pancost,, R.D., Schouten,, S., Sinninghe Damsté,, J.S. (2001) Massive expansion of marine archaea during a mid-cretaceous oceanic anoxic event. *Science* 293, 92–94.

Lal, R. (2002) Soil carbon sequestration in China through agricultural intensification, and restoration of degraded and desertified ecosystems. *L. Degrad. Dev. 13*, 469–478.

Lehndorff, E., Roth, P.J., Cao, Z.H., Amelung, W. (2014) Black carbon accrual during 2000 years of paddy-rice and non-paddy cropping in the Yangtze River Delta, China. *Glob. Chang. Biol. 20,* 1968–1978.

Lehndorff, E., Schwark, L. (2008) Accumulation histories of major and trace elements on pine needles as function of air quality. *Atmos. Environ. 42*, 833–845.

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G.W., Prosser, J. I., Schuster, S.C., Schleper, C. (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442, 806–809.

Lennartz, B., Horn, R., Duttmann, R., Gerke, H.H., Tippkötter, R., Eickhorst, T., Janssen, I., Janssen, M., Rüth, B., Sander, T., Shi, X., Sumfleth, K., Taubner, H., Zhang, B. (2009) Ecological safe management of terraced rice paddy landscapes. *Soil Till. Res. 102*, 179–192.

Li, L., Zhang, X., Zhang, P., Zheng, J., Pan, G. (2007) Variation of organic carbon and nitrogen in aggregate size fractions of a paddy soil under fertilisation practices from Tai Lake Region, China. *J. Sci. Food Agr. 87*, 1052–1058.

Li, X., Mander, Ü., Ma, Z., Jia, Y. (2009) Water quality problems and potential for wetlands as treatment systems in the Yangtze River Delta, China. *Wetlands 29*, 1125–1132.

Lichtfouse, É., Chenu, C., Baudin, F., Leblond, C., Da Silva, M., Behar, F., Derenne, S., Largeau, C., Wehrung, P., Albrecht, P. (1998) A novel pathway of soil organic matter formation by selective preservation of resistant straight-chain biopolymers: Chemical and isotope evidence. *Org. Geochem. 28*, 411–415.

Lichtfouse, É., Elbisser, B., Balesdent, J., Mariotti, A., Bardoux, G. (1994) Isotope and molecular evidence for direct input of maize leaf wax *n*-alkanes into crop soils. *Org. Geochem. 22*, 349–351.

Liesack, W., Schnell, S., Revsbech, N. P. (2000) Microbiology of flooded rice paddies. *FEMS Microbiol. Rev.* 24, 625–645.

Lima, A.L.C., Farrington, J.W., Reddy, C. (2005) Combustion-derived polycyclic aromatic hydrocarbons in the environment – a review. *Environ. Forensics 6*, 109–131.

Lindahl ,T. (1993) Instability and decay of the primary structure of DNA. *Nature 362*, 709–715.

Liu, W., Huang, Y., An, Z., Clemens, S.C., Li, L., Prell, W.L., Ning, Y. (2005) Summer monsoon intensity controls C_4/C_3 plant abundance during the last 35 ka in the Chinese Loess Plateau: Carbon isotope evidence from bulk organic matter and individual leaf waxes. *Palaeogeogr. Palaeoclimatol. Palaeoecol. 220*, 243–254.

Liu, W., Wang, H., Zhang, C.L., Liu, Z., He, Y. (2013) Distribution of glycerol dialkyl glycerol tetraether lipids along an altitudinal transect on Mt. Xiangpi, NE Qinghai-Tibetan Plateau, China. *Org. Geochem.* 57, 76–83.

Liu, Y., Chen, L., Jianfu, Z., Qinghui, H., Zhiliang, Z., Hongwen, G. (2008) Distribution and sources of polycyclic aromatic hydrocarbons in surface sediments of rivers and an estuary in Shanghai, China. *Environ. Pollut. 154*, 298–305.

Liu, Y., Wang, J., Liu, D., Li, Z., Zhang, G., Tao, Y., Xie, J., Pan, J., Chen, F. (2014) Straw mulching reduces the harmful effects of extreme hydrological and temperature conditions in citrus orchards. *PLoS One* 9, e87094.

Lockey, K.H. (1988) Lipids of the insect cuticle: origin, composition and function. *Comp. Biochem. Physiol. Part B Comp. Biochem. 89,* 595–645.

Lockey, K.H. (1991) Insect hydrocarbon classes: Implications for chemotaxonomy. *Insect Biochemistry 21*, 91–97.

Lockheart, M.J., van Bergen, P.F., Evershed, R.P. (1997) Variations in the stable carbon isotope compositions of individual lipids from the leaves of modern angiosperms: implications for the study of higher land plant-derived sedimentary organic matter. *Org. Geochem. 26,* 137–153.

Loomis, S.E., Russell, J.M., Heureux, A.M., D'Andrea, W.J., Sinninghe Damsté, J.S. (2014) Seasonal variability of branched glycerol dialkyl glycerol tetraethers (brGDGTs) in a temperate lake system. *Geochim. Cosmochim. Acta* 144, 173–187.

Luce, M.St., Whalen, J.K., Ziadi, N., Zebarth, B.J. (2011) Nitrogen dynamics and indices to predict soil nitrogen supply in humid temperate soils. *Advances in Agronomy* 112. Elsevier Inc., pp. 55-102.

Marschner, B., Brodowski, S., Dreves, A., Gleixner, G., Gude, A., Grootes, P.M., Hamer, U., Heim, A., Jandl, G., Ji, R., Kaiser, K., Kalbitz, K., Kramer, C., Leinweber, P., Rethemeyer, J., Schäffer, A., Schmidt, M.W.I., Schwark, L., Wiesenberg, G.L.B. (2008) How relevant is recalcitrance for the stabilization of organic matter in soils? *J. Plant Nutr. Soil Sci. 171*, 91–110.

Meyers, P.A. (1997) Organic geochemical proxies of paleoceanographic, paleolimnologic, and paleoclimatic processes. *Org. Geochem. 27*, 213–250.

Mueller-Niggemann, C., Bannert, A., Schloter, M., Lehndorff, E., Schwark, L. (2012) Intra- versus inter-site macroscale variation in biogeochemical properties along a paddy soil chronosequence. *Biogeosciences 9*, 1237–1251.

Mueller-Niggemann, C., Schwark, L. (2015) Chemotaxonomy and diagenesis of aliphatic hydrocarbons in rice plants and soils from land reclamation areas in the Zhejiang Province, China. *Org. Geochem.* 83-84, 215-226.

Mullins, C.E. (1977) Magnetic-susceptibility of soils and its significance in soil science-review, *Int. J. Soil Sci, 28*, 223–246.

Myrold, D.D., Tiedje, J. M. (1986) Simultaneous estimation of several nitrogen cycle rates using 15N: Theory and application. *Soil Biol. Biochem. 18*, 559–568.

Naeher, S., Peterse, F., Smittenberg, R.H., Niemann, H., Zigah, P.K., Schubert, C.J. (2014) Sources of glycerol dialkyl glycerol tetraethers (GDGTs) in catchment soils, water column and sediments of Lake Rotsee (Switzerland) - Implications for the application of GDGT-based proxies for lakes. *Org. Geochem.* 66, 164–173.

Nanzyo, M., Nakamaru, Y., Yamasaki, S.I., Samonte, H.P. (1999) Effect of reducing conditions on the weathering of Fe^{3+}-rich biotite in the new lahar deposit from Mt. Pinatubo, Philippines. *Soil Sci.* 164, 206–214.

Nelson, D.L., Cox, M.M. (2005*) Lehninger Principles of biochemistry*. 4th ed., W.H. Freeman, Gordonsville, V., pp. 1100.

Neue, H.U., Gaunt, J.L., Wang, Z.P., Becker Heidmann, P., Quijano, C. (1997) Carbon in tropical wetlands. *Geoderma 79*, 163–185.

Neue, H.U., Gaunt, J.L., Wang, Z.P., Becker-Heidmann, P., Quijano, C. (1997) Carbon in tropical wetlands. *Geoderma* 79, 163-185.

Ni, M., Huang, J., Lu, S., Li, X., Yan, J., Cen, K. (2014) A review on black carbon emissions, worldwide and in China. *Chemosphere 107*, 83–93.

Nicol, G.W., Leininger, S., Schleper, C., Prosser, J.I. (2008) The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Environ. Microbiol. 10*, 2966–2978.

Nieder R., Benbi D.K. (2008) *Carbon and nitrogen in the terrestrial environment*. Springer Netherlands.

Ohlrogge, J., Browse, J. (1995) Lipid biosynthesis. *The Plant cell* 7, 957–970.

Oppermann, B.I., Michaelis, W., Blumenberg, M., Frerichs, J., Schulz, H.M., Schippers, A., Beaubien, S.E., Krüger, M. (2010) Soil microbial community changes as a result of long-term exposure to a natural CO_2 vent. *Geochim. Cosmochim. Acta* 74, 2697–2716.

Ourisson, G., Rohmer, M., Poralla, K. (1987) Prokaryotic hopanoids and other polyterpenoid sterol surrogates. *Ann. Rev. Microbiol. 41*, 301–333.

Pan, G., Li, L., Wu, L., Zhang, X. (2003) Storage and sequestration potential of topsoil organic carbon in China's paddy soils. *Glob. Chang. Biol. 10*, 79–92.

Pancost, R.D., Hopmans, E.C., Sinninghe Damsté, J.S. (2001) Archaeal lipids in mediterranean cold seeps: Molecular proxies for anaerobic methane oxidation. *Geochim. Cosmochim. Acta* 65, 1611–1627.

Pearson, A., Huang, Z., Ingalls, A.E., Romanek, C.S., Wiegel, J., Freeman, K.H., Smittenberg, R.H., Zhang, C.L. (2004) Nonmarine crenarchaeol in Nevada hot springs. *Appl. Environ. Microbiol.* 70, 5229–5237.

Pearson, A., Ingalls, A.E. (2013) Assessing the Use of Archaeal Lipids as Marine Environmental Proxies. *Annu. Rev. Earth Planet. Sci.* 41, 359–384.

Peters, K.E., Walters, C.C., Moldowan, J.M. (2005) *The Biomarker Guide*. Biomarkers and Isotopes in Petroleum Exploration and Earth History. Second Edition. Cambridge University Press.

Peterse, F., Kim, J.H., Schouten, S., Kristensen, D.K., Koç, N., Sinninghe Damsté, J.S. (2009b) Constraints on the application of the MBT/CBT palaeothermometer at high latitude environments (Svalbard, Norway). *Org. Geochem.* 40, 692–699.

Peterse, F., Moy, C.M., Eglinton, T.I. (2015) A laboratory experiment on the behaviour of soil-derived core and intact polar GDGTs in aquatic environments. *Biogeosciences* 12, 933–943.

Peterse, F., Schouten, S., van der Meer, J., van der Meer, M.T.J., Sinninghe Damsté, J.S. (2009a) Distribution of branched tetraether lipids in geothermally heated soils: Implications for the MBT/CBT temperature proxy. *Org. Geochem.* 40, 201–205.

Peterse, F., van der Meer, J., Schouten, S., Weijers, J.W.H., Fierer, N., Jackson, R.B., Kim, J.H., Sinninghe Damsté, J.S. (2012) Revised calibration of the MBT-CBT paleotemperature proxy based on branched tetraether membrane lipids in surface soils. *Geochim. Cosmochim. Acta* 96, 215–229.

Peterse, F., van der Meer, M.T.J., Schouten, S., Jia, G., Ossebaar, J., Blokker, J., Sinninghe Damsté, J.S. (2009c) Assessment of soil *n*-alkane δD and branched tetraether membrane lipid distributions as tools for paleoelevation reconstruction. *Biogeosciences* 6, 2799–2807.

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M., Guerri, G., Nannipieri, P. (2009) Extracellular DNA in soil and sediment: fate and ecological relevance, *Biol Fertil Soils 45*, 219–235.

Pitcher, A., Rychlik, N., Hopmans, E.C., Spieck, E., Rijpstra, W.I.C., Ossebaar, J., Schouten, S., Wagner, M., Sinninghe Damsté, J.S. (2010) Crenarchaeol dominates the membrane lipids of Candidatus Nitrososphaera gargensis, a thermophilic group I.1b Archaeon. *ISME J.* 4, 542–552.

Pitcher, A., Schouten, S., Sinninghe Damsté, J.S. (2009) In situ production of crenarchaeol in two California hot springs. *Appl. Environ. Microbiol.* 75, 4443–4451.

Poinar, H.N., Hoss, M., Bada, J.L., Paabo, S. (1996) Amino acid racemization and the preservation of ancient DNA, *Science 272*, 864–866.

Ponnamperuma, F.N. (1972) The chemistry of submerged soils. *Adv. Agron.* 24, 29–96.

Poynter, J.G., Farrimond, P., Brassell, S.C., Eglinton, G. (1989) Aeolian-derived higher-plant lipids in the marine sedimentary record: links with paleoclimate. In: Leinen, M., Sarnthein, M. (Eds.), *Palaeoclimatology and Palaeometeorology: Modern and Past Patterns of Global Atmosphere Transport*. Kluwer, pp. 435–462.

Press, F., Siever, R, (1995) *Allgemeine Geologie: Eine Einführung*. Spektrum Akademischer Verlag, Heidelberg.

Qiu, Y.J., Saliot, A. (1991) Non-aromatic hydrocarbons in "dissolved" phase (< 0-7 µm) and their fractionation between "dissolved" and particulate phases in the Changjiang (Yangtse River) Estuary. *Mar. Environ. Res. 31*, 287–308.

Quénéa, K., Largeau, C., Derenne, S., Spaccini, R., Bardoux, G., Mariotti, A. (200). Molecular and isotopic study of lipids in particle size fractions of a sandy cultivated soil (Cestas cultivation sequence, southwest France): Sources, degradation, and comparison with Cestas forest soil. *Org. Geochem. 37*, 20–44.

Reigstad, L.J., Richter, A., Daims, H., Urich, T., Schwark, L., Schleper, C. (2008) Nitrification in terrestrial hot springs of Iceland and Kamchatka. *FEMS Microbiol. Eco., 64*, 167–174.

Ringrose-Voase, A.J., Kirby, J.M., Djoyowasito, G., Sanidad, W.B., Serrano, C., Lando, T.M. (2000) Changes to the physical properties of soils puddled for rice during drying. *Soil Till. Res. 56*, 83–104.

Rohmer, M., Bouvier-Nave, P., Ourisson, G. (1984) Distribution of hopanoid triterpenes in prokaryotes. *J. Gen. Microbiol. 130*, 1137–1150.

Rommerskirchen, F., Plader, A., Eglinton, G., Chikaraishi, Y., Rullkötter, J. (2006) Chemotaxonomic significance of distribution and stable carbon isotopic composition of long-chain alkanes and alkan-1-ols in C_4 grass waxes. *Org. Geochem. 37*, 1303–1332.

Roth, P.J., Lehndorff, E., Cao, Z.H., Zhuang, S., Bannert, A., Wissing, L., Schloter, M., Kögel-Knabner, I., Amelung, W. (2011) Accumulation of nitrogen and microbial residues during 2000 years of rice paddy and non-paddy soil development in the Yangtze River Delta, China. *Glob. Chang. Biol. 17*, 3405–3417.

Rüth, B., Lennartz, B. (2008) Spatial Variability of Soil Properties and Rice Yield Along Two Catenas in Southeast China. *Pedosphere 18*, 409–420.

Sachse, D., Billault, I., Bowen, G.J., Chikaraishi, Y., Dawson, T.E., Feakins, S.J., Freeman, K.H., Magill, C.R., McInerney, F.A., van der Meer, M.T.J., Polissar, P., Robins, R.J., Sachs, J.P., Schmidt, H.-L., Sessions, A.L., White, J.W.C., West, J.B., Kahmen, A. (2012) Molecular paleohydrology: interpreting the hydrogen-isotopic composition of lipid biomarkers from photosynthesizing organisms. *Annu. Rev. Earth Planet. Sci. 40*, 221–249.

Sachse, D., Radke, J., Gaupp, R., Schwark, L., Lüniger, G., Gleixner, G. (2004) Reconstruction of palaeohydrological conditions in a lagoon during the 2nd Zechstein cycle through simultaneous use of δD values of individual *n*-alkanes and $\delta^{18}O$ and $\delta^{13}C$ values of carbonates. *Int. J. Earth Sci. 93*, 554–564.

Sahrawat, K.L. (2004) Organic matter accumulation in submerged soils. *Adv. Agron*. 81, 169–201.

Sahrawat, K.L. (2005) Fertility and organic matter in submerged rice soils. *Curr. Sci. 88*, 735–739.

Samuels, L., Kunst, L., Jetter, R. (2008) Sealing plant surfaces: cuticular wax formation by epidermal cells. *Annu. Rev. Plant Biol. 59*, 683–707.

Sander, T., Gerke, H.H. (2007) Preferential flow patterns in paddy fields using a dye tracer. *Vadose Zone J. 6*, 105–115.

Sasaki, Y., Nagano, Y. (2004) Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding. *Bioscience, biotechnology, and biochemistry* 68, 1175–1184.

Schaetzl, R.J., Thompson, M.L. (2015) *Soils: genesis and geomorphology*. 2nd ed.. Cambridge Univ. Press, Cambridg, pp. 778.

Scharpenseel, H.W., Pfeiffer, E,-M,, Becker-Heidmann, P. (1996) Organic carbon storage in tropical hydromorphic soils. In: Carter, M.R., Stewart, B.A. (Eds.). *Structure and organic matter storage in agricultural soils.* Adv. Soil Sci.. Lewis Publishers, Boca Raton, pp. 361-392.

Schefuß, E., Schouten, S., Jansen, J.H.F., Sinninghe Damsté, J.S., 2003. African vegetation controlled by tropical sea surface temperatures in the mid-Pleistocene period. Nature 422, 418–421.

Schleper, C., Jurgens, G., Jonuscheit, M. (2005) Genomic studies of uncultivated archaea. *Nat. Rev. Microbiol. 3*, 479–488.

Schmidt, M.W.I., Noack, A.G. (2000) Black carbon in soils and sediments: analysis, distribution, implications, and current challenges. *Global Biogeochem. Cycles 14*, 777–793.

Schmitter, P., Dercon, G., Hilger, T., Thi Le Ha, T., Huu Thanh, N., Lam, N., Duc Vien, T., Cadisch, G. (2010) Sediment induced soil spatial variation in paddy fields of Northwest Vietnam. *Geoderma 155*, 298–307.

Schneider, M.P.W., Hilf, M., Vogt, U.F., Schmidt, M.W.I. (2010) The benzene polycarboxylic acid (BPCA) pattern of wood pyrolyzed between 200 °C and 1000 °C. *Org. Geochem. 41*, 1082–1088.

Schouten, S., Hopmans, E.C., Schefuß, E., Sinninghe Damsté, J.S. (2002) Distributional variations in marine crenarchaeotal membrane lipids: A new tool for reconstructing ancient sea water temperatures? *Earth Planet. Sci. Lett.* 204, 265–274.

Schouten, S., Hopmans, E.C., Sinninghe Damsté, J.S. (2013) The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. *Org. Geochem.* 54, 19–61.

Schouten, S., Huguet, C., Hopmans, E.C., Kienhuis, M.V.M., Sinninghe Damsté, J.S. (2007) Analytical methodology for TEX_{86} paleothermometry by high-performance liquid chromatography/atmospheric pressure chemical ionization-mass spectrometry. *Anal. Chem.* 79, 2940–2944.

Schouten, S., Rijpstra, W.I.C., Durisch-Kaiser, E., Schubert, C.J., Sinninghe Damsté, J.S. (2012) Distribution of glycerol dialkyl glycerol tetraether lipids in the water column of Lake Tanganyika. *Org. Geochem.* 53, 34–37.

Schwark, L. (2010) Hydrocarbons in the Pedosphere. In: Timmis, K.N. (Ed.), *Handbook of Hydrocarbon and Lipid Microbiology*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 280–295.

Schwark, L., Zink, K., Lechterbeck, J. (2002) Reconstruction of postglacial to early Holocene vegetation history in terrestrial Central Europe via cuticular lipid biomarkers and pollen records from lake sediments. *Geology 30*, 463–466.

Serrano-Silva, N., Sarria-Guzmán, Y., Dendooven, L., Luna-Guido, M. (2014) Methanogenesis and methanotrophy in soil: A review. *Pedosphere* 24, 291–307.

Shi, Z., Li, Y., Wang, R.C., Makeschine, F. (2005) Assessment of temporal and spatial variability of soil salinity in a coastal saline field. *Environ. Geol. 48*, 171–178.

Shima, S., Krueger, M., Weinert, T., Demmer, U., Kahnt, J., Thauer, R.K., Ermler, U. (2012) Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* 481, 98–101.

Shrestha, R.K., Ladha, J.K. (1998) Nitrate in groundwater and integration of nitrogen-catch crop in rice-sweet pepper cropping system. *Soil Sci. Soc. Am. J. 62*, 1610–1619.

Simoneit, B.R.T. (2002) Biomass burning - a review of organic tracers for smoke from incomplete combustion. *App. Geochem. 17*, 129-162.

Simpson, M.J., Simpson, A.J. (2012) The chemical ecology of soil organic matter molecular constituents. *Journal of chemical ecology* 38, 768–784.

Sinninghe Damsté J.S., Rijpstra W.I.C., Hopmans E.C., Weijers J.W.H., Foesel B.U., Overmann J., Dedysh S.N. (2011) 13,16-Dimethyl octacosanedioic acid (iso-Diabolic Acid), a common membrane-spanning lipid of Acidobacteria subdivisions 1 and 3. *Appl. Environ. Microbiol.* 77, 4147–4154.

Sinninghe Damsté, J.S., Hopmans, E.C., Pancost, R.D., Schouten, S., Geenevasen, J.A.J. (2000) Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments. *Chem. Commun.* 17, 1683–1684.

Sinninghe Damsté, J.S., Ossebaar, J., Schouten, S., Verschuren, D. (2008) Altitudinal shifts in the branched tetraether lipid distribution in soil from Mt. Kilimanjaro (Tanzania): Implications for the MBT/CBT continental palaeothermometer. *Org. Geochem.* 39, 1072–1076.

Sinninghe Damsté, J.S., Rijpstra, W.I.C., Hopmans, E.C., Foesel, B.U., Wüst, P.K., Overmann, J., Tank, M., Bryant, D.A., Dunfield, P.F., Houghton, K., Stott, M.B. (2014) Ether- and ester-bound iso-diabolic acid and other lipids in members of Acidobacteria subdivision 4. *Applied and Environmental Microbiology* 80, 5207–5218.

Sinninghe Damsté, J.S., Rijpstra, W.I.C., Hopmans, E.C., Jung, M.Y., Kim, J.G., Rhee, S.K., Stieglmeier, M., Schleper, C. (2012) Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids of group I.1a and I.1b Thaumarchaeota in soil. *Appl. Environ. Microbiol.* 78, 6866–6874.

Sinninghe Damsté, J.S., Schouten, S., Hopmans, E.C., van Duien, A.C.T., Geenevasen, J.A. J. (2002) Crenarchaeol: the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota. *Journal of Lipid Research* 43, 1641–1651.

Soil Survey Staff (1999) *Soil taxonomy. A basic system of soil classification for making and interpreting soil surveys.* 2nd ed. USDA Agric. Handbook 436. Washington, DC.

Spang, A., Hatzenpichler, R., Brochier-Armanet, C., Rattei, T., Tischler, P., Spieck, E., Streit, W., Stahl, D.A, Wagner, M., Schleper, C. (2010) Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends Microbiol.* 18, 331–340.

Stahl, D.A., de la Torre, J.R. (2012) Physiology and diversity of ammonia-oxidizing archaea. *Annu. Rev. Microbiol.* 66, 83–101.

Stark, J.M., Hart, S.C. (1997) High rates of nitrification and nitrate turnover in undisturbed coniferous forests. *Nature 385*, 61–64.

Stieglmeier, M., Klingl, A., Alves, R.J.E., Rittmann, S.K.M.R., Melcher, M., Leisch, N., Schleper, C. (2014) Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* 64, 2738–2752.

Street, J.H., Anderson, R.S., Rosenbauer, R.J., Paytan, A. (2013) *n*-Alkane evidence for the onset of wetter conditions in the Sierra Nevada, California (USA) at the mid-late Holocene transition, ~3.0 ka. *Quaternary Research 79*, 14–23.

Targulian, V.O., Krasilnikov, P.V. (2007) Soil system and pedogenic processes: Self-organization, time scales, and environmental significance. *Catena 71*, 373–381.

Tatsuya, I., Goto, K., Iida, M., Nonami, K., Inoue, H., Umeda, M. (2004) Geostatistical analysis of yield, soil properties and crop management practices in paddy rice fields. *Plant Prod. Sci. 7*, 230–239.

Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W., Hedderich, R. (2008) Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* 6, 579–591.

Tierney, J. E., Schouten, S., Pitcher, A., Hopmans, E. C., Sinninghe Damsté, J.S. (2012) Core and intact polar glycerol dialkyl glycerol tetraethers (GDGTs) in Sand Pond, Warwick, Rhode Island (USA): Insights into the origin of lacustrine GDGTs. *Geochim. Cosmochim. Acta* 77, 561–581.

Tierney, J.E., Russell, J. M. (2009) Distributions of branched GDGTs in a tropical lake system: Implications for lacustrine application of the MBT/CBT paleoproxy. *Org. Geochem.* 40, 1032–1036.

Töwe, S., Albert A., Kleineidam, K., Brankatschk, R., Dümig, A., Welzl, G., Munch, J., Zeyer, J., Schloter, M. (2010) Abundance of microbes involved in nitrogen transformation in the rhizosphere of Leucanthemopsis alpina (L.) heywood grown in soils from different sites of the Damma Glacier Forefield. *Microb. Ecol. 60*, 762–770.

Tsuruta, T., Yamaguchi, M., Abe, S.I., Iguchi, K. (2011) Effect of fish in rice-fish culture on the rice yield. *Fisheries Science* 77, 95–106.

Valentine, D. L. (2007) Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* 5, 316–323.

Valiela, I. (1995) Marine ecolocial processes, 2nd ed., Springer Science & Business Media New York, pp. 686.

Van den Berg, G.A., Loch, J.P.G. (2000) Decalcification of soils subject to periodic waterlogging. *European Journal of Soil Science* 51, 27–33.

Venkatesan, M. I. (1988) Occurrence and possible sources of perylene in marine sediments-a review. *Mar. Chem. 25*, 1–27.

Venkatesan, M.I., Kaplan, I.R. (1982) Distribution and transport of hydrocarbons in surface sediments of the Alaskan Outer Continental Shelf. *Geochim. Cosmochim. Acta 46*, 2135–2149.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D.Y., Paulsen, I., Nelson, K.E., Nelson, W., Fouts, D.E., Levy, S., Knap, A.H., Lomas, M.W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y.H., Smith, H.O. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science 304*, 66–74.

Viscarra Rossel, R.A., Minasny, B., Roudier, P., McBratney, A.B. (2006) Colour space models for soil science. *Geoderma 133*, 320–337.

Vogts, A., Moossen, H., Rommerskirchen, F., Rullkötter, J. (2009) Distribution patterns and stable carbon isotopic composition of alkanes and alkan-1-ols from plant waxes of African rain forest and savanna C_3 species. *Org. Geochem. 40*, 1037–1054.

Wakeham, S.G., Schaffner, C., Giger, W. (1980) Polycyclic aromatic hydrocarbons in recent lake sediments-II: compounds derived from biogenic precursors during early diagenesis. *Geochim. Cosmochim. Acta 44*, 415–429.

Walker, L.R., Wardle, D.A., Bardgett, R.D., Clarkson, B.D. (2010) The use of chronosequences in studies of ecological succession and soil development. *J. Ecol. 98*, 725–736.

Wang, H., Liu, W., Zhang, C.L. (2014) Dependence of the cyclization of branched tetraethers on soil moisture in the Chinese Loess Plateau and the adjacent areas: implications for palaeorainfall. *Biogeosciences* 11, 6755–6768.

Wang, H., Liu, W., Zhang, C.L., Liu, Z., He, Y. (2013) Branched and isoprenoid tetraether (BIT) index traces water content along two marsh-soil transects surrounding Lake Qinghai: Implications for paleo-humidity variation. *Org. Geochem.* 59, 75–81.

Wang, H., Saito, Y., Zhang, Y., Bi, N., Sun, X., Yang, Z. (2011) Recent changes of sediment flux to the western Pacific Ocean from major rivers in East and Southeast Asia. *Earth-Science Rev. 108,* 80–100.

Wang, H., Yang, Z., Wang, Y., Saito, Y., Liu, J.P. (2008a) Reconstruction of sediment flux from the Changjiang (Yangtze River) to the sea since the 1860s. *J. Hydrol. 349*, 318–332.

Wang, L., Wu, J.P., Liu, Y.X., Huang, H.G., Fang, Q.F. (2009) Spatial Variability of Micronutrients in Rice Grain and Paddy Soil. *Pedospher, 19*, 748–755.

Wang, W.M., Ding, J.L., Shu, J.W., Chen, W. (2010) Exploration of early rice farming in China. *Quaternary International* 227, 22–28.

Wang, X.-C., Sun, M.-Y., Li, A.-C. (2008b) Contrasting chemical and isotopic compositions of organic matter in Changjiang (Yangtze River) Estuarine and East China Sea Shelf Sediments. *J. Oceanogr. 64*, 311–321.

Watanabe, I. (1982) Azolla–Anabaena symbiosis–its physiology and use in tropical agriculture. In: Dommerguesand, Y.R., Diem, H.S. (Ed.) Microbiology of tropical soils and plant productivity, *Developments in Plant and Soil Sciences* 5, pp. 169-185.

Watanabe, T., Kimura, M., Asakawa, S. (2006) Community structure of methanogenic archaea in paddy field soil under double cropping (rice-wheat). *Soil Biol. Biochem.* 38, 1264–1274.

Watanabe, T., Kimura, M., Asakawa, S. (2009) Distinct members of a stable methanogenic archaeal community transcribe *mcrA* genes under flooded and drained conditions in Japanese paddy field soil. *Soil Biol. Biochem.* 41, 276–285.

Weber, E.B., Lehtovirta-Morley, L.E., Prosser, J.I., Gubry-Rangin, C. (2015) Ammonia oxidation is not required for growth of Group 1.1c soil Thaumarchaeota. *FEMS Microbiol. Ecol.* 91, fiv001. doi: http://dx.doi.org/10.1093/femsec/fiv001

Wei, Y-C., Bai, Y-L., Jin, J-Y., Zhang, F., Zhang, L-P., Liu, X.Q. (2009) Spatial variability of soil chemical properties in the reclaiming marine foreland to Yellow Sea of China, *Agr. Sci. Chin. 8*, 1103–1111.

Weijers, J.W.H., Panoto, E., van Bleijswijk, J., Schouten, S., Rijpstra, W.I.C., Balk, M., Stams, A.J.M., Sinninghe Damsté, J.S. (2009) Constraints on the biological source(s) of the orphan branched tetraether membrane lipids. *Geomicrobiol. J.* 26, 402–414.

Weijers, J.W.H., Schouten, S., Hopmans, E.C., Geenevasen, J.A.J., David, O.R.P., Coleman, J.M., Pancost, R.D., Sinninghe Damsté, J.S. (2006a) Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits. *Environ. Microbiol.* 8, 648–657.

Weijers, J.W.H., Schouten, S., Spaargaren, O.C., Sinninghe Damsté, J.S. (2006b) Occurrence and distribution of tetraether membrane lipids in soils: Implications for the use of the TEX_{86} proxy and the BIT index. *Org. Geochem.* 37, 1680–1693.

Weijers, J.W.H., Schouten, S., van den Donker, J.C., Hopmans, E.C., Sinninghe Damsté, J.S. (2007) Environmental controls on bacterial tetraether membrane lipid distribution in soils. *Geochim. Cosmochim. Acta* 71, 703–713.

Weijers, J.W.H., Wiesenberg, G.L.B., Bol, R., Hopmans, E.C., Pancost, R.D. (2010) Carbon isotopic composition of branched tetraether membrane lipids in soils suggest a rapid turnover and a heterotrophic life style of their source organism(s). *Biogeosciences* 7, 2959–2973.

Wiedemeier, D.B., Abiven, S., Hockaday, W.C., Keiluweit, M., Kleber, M., Masiello, C.A., McBeath, A.V., Nico, P.S., Pyle, L.A., Schneider, M.P.W., Smernik, R.J., Wiesenberg, G.L.B., Schmidt, M.W.I. (2015) Aromaticity and degree of aromatic condensation of char. *Org. Geochem. 78*, 135–143.

Wiedemeier, D.B., Brodowski, S., Wiesenberg, G.L.B. (2014) Pyrogenic molecular markers: Linking PAH with BPCA analysis. *Chemosphere 119*, 432–437.

Wiesenberg, G.L.B., Lehndorff, E., Schwark, L. (2009) Thermal degradation of rye and maize straw: Lipid pattern changes as a function of temperature. *Org. Geochem. 40*, 167–174.

Wiesenberg, G.L.B., Schwark, L. (2006) Carboxylic acid distribution patterns of temperate C_3 and C_4 crops. *Org. Geochem. 37*, 1973–1982.

Wiesenberg, G.L.B., Schwark, L., Schmidt, M.W.I. (2006) Extractable lipids and colour in particle-size fractions and bulk arable soils. *Euro. Jo. Soil Sci. 57*, 634–643.

Wiesenberg, G.L.B., Schwarzbauer, J., Schmidt, M.W.I., Schwark, L. (2004) Source and turnover of organic matter in agricultural soils derived from *n*-alkane/*n*-carboxylic acid compositions and C-isotope signatures. *Org. Geochem. 35*, 1371–1393.

Wiesenberg, G.L.B., Schwarzbauer, J., Schmidt, M.W.I., Schwark, L. (2008) Plant and soil lipid modifications under elevated atmospheric CO_2 conditions: II. Stable carbon isotopic values ($\delta^{13}C$) and turnover. *Org. Geochem. 39*, 103–117.

Wilcke, W. (2000) Polycyclic aromatic hydrocarbons (PAHs) in soil - a review. *J. Plant Nutr. Soil Sci. 163*, 229–248.

Wilcke, W. (2007) Global patterns of polycyclic aromatic hydrocarbons (PAHs) in soil. *Geoderma 141*, 157–166.

Willerslev, E., Hansen, A.J., Poinar, H.N. (2004) Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends Ecol. Evol. 19*, 141–147.

Wissing, L., Kölbl, A., Vogelsang, V., Fu, J.-R., Cao, Z.-H., Kögel-Knabner, I. (2011) Organic carbon accumulation in a 2000-year chronosequence of paddy soil evolution. *CATENA 87*, 376–385.

Wolf, M., Lehndorff, E., Wiesenberg, G. L. B., Stockhausen, M., Schwark, L., Amelung, W. (2013) Towards reconstruction of past fire regimes from geochemical analysis of charcoal. *Org. Geochem. 55*, 11–21.

Wrage, N., Velthof, G.L., van Beusichem, M.L., Oenema, O. (2001) Role of nitrifier denitrification in the production of nitrous oxide. *Soil Biol. Biochem. 33*, 1723–1732.

Wu, J. (2011) Carbon accumulation in paddy ecosystems in subtropical China: evidence from landscape studies. *Eur. J. Soil Sci. 62*, 29–34.

Xie, D., Wang, Z., Gao, S., De Vriend, H.J. (2009) Modeling the tidal channel morphodynamics in a macro-tidal embayment, Hangzhou Bay, China. *Cont. Shelf Res. 29*, 1757–1767.

Xie, S., Lai, X., Yi, Y., Gu, Y., Liu, Y., Wang, X., Liu, G., Liang, B. (2003) Molecular fossils in a Pleistocene river terrace in southern China related to paleoclimate variation. *Org. Geochem. 34*, 789–797.

Xie, S., Nott, C.J., Avsejs, L.A., Volders, F., Maddy, D., Chambers, F.M., Gledhill, A., Carter, J.F., Evershed, R.P. (2000) Palaeoclimate records in compound-specific δD values of a lipid biomarker in ombrotrophic peat. *Org. Geochem. 31*, 1053–1057.

Xiong, Z.Q., Xing, G.X., Zhu, Z.L. (2007) Nitrous oxide and methane emissions as affected by water, soil and nitrogen. *Pedosphere* 17, 146–155.

Yanai, J., Lee, C.K., Kaho, T., Iida, M., Matsui, T., Umeda, M., Kosaki, T. (2001) Geostatistical analysis of soil chemical properties and rice yield in a paddy field and application to the analysis of yield determining factors, *Soil Sci. Plant Nutr. 47*, 291–301.

Yang, H.-H., Tsai, C.-H., Chao, M.-R., Su, Y.-L., Chien, S.-M. (2006) Source identification and size distribution of atmospheric polycyclic aromatic hydrocarbons during rice straw burning period. *Atmos. Environ. 40*, 1266–1274.

Yerima, B.P.K., van Ranst, E. (2005) *Introduction to soil science: soils of the tropics.* Trafford Publishing, Victoria, BC.

Yu, Q., Yu, Y., You, A., Song, L., Cao, Y. (2003) Contemporary evolution process of the Hangzhou Bay, *International Conference on Estuaries and Coasts*, November 9–11, Hangzhou, China.

Zegouagh, Y., Derenne, S., Largeau, C., Saliot, A. (2000) A geochemical investigation of carboxylic acids released via sequential treatments of two surficial sediments from the Changjiang delta and East China Sea. *Org. Geochem. 31*, 375–388.

Zhang, J., Wu, Y., Jennerjahn, T.C., Ittekkot, V., He, Q. (2007) Distribution of organic matter in the Changjiang (Yangtze River) Estuary and their stable carbon and nitrogen isotopic ratios: Implications for source discrimination and sedimentary dynamics. *Mar. Chem. 106*, 111–126.

Zhang, M., He, Z., Zhao, A., Zhang, H., Endale, D.M., Schomberg, H.H. (2011) Water extractable soil organic carbon and nitrogen affected by tillage and manure application., *Soil Sci. 176*, 307–312.

Zhang, M., Lu, H., Zhao, X.J., Li, R.A. (2004) A comparative study of soil fertility change of upland soil in Cixi Country. *Chinese J. Soil Sci. 35*, 91–93 (in Chinese).

Zhao, Y., Xu, X., Darilek, J. L., Huang, B., Sun, W., Shi, X. (2009) Spatial variability assessment of soil nutrients in an intense agricultural area, a case study of Rugao County in Yangtze River Delta Region, China. *Environ. Geol. 57*, 1089–1102.

Zhou, P., Song, G., Pan, G., Li, L., Zang, X. (2009) Role of chemical protection by binding to oxyhydrates in SOC sequestration in three typical paddy soils under long-term agro-ecosystem experiments from South China. *Geoderma 153*, 52–60.

Zhou, W., Xie, S., Meyers, P.A., Zheng, Y. (2005) Reconstruction of late glacial and Holocene climate evolution in southern China from geolipids and pollen in the Dingnan peat sequence. *Org. Geochem. 36*, 1272–1284.

Zhu, J.G., Han, Y., Liu, G., Zhang, Y.L., Shao, S.H. (2000) Nitrogen in percolation water in paddy fields with a rice/wheat rotation. *Nutr. Cycl. Agroecosys. 57*, 75–82.

Zink, K.-G., Vandergoes, M.J., Mangelsdorf, K., Dieffenbacher-Krall, A.C., Schwark, L. (2010) Application of bacterial glycerol dialkyl glycerol tetraethers (GDGTs) to develop modern and past temperature estimates from New Zealand lakes. *Org. Geochem.* 41, 1060–1066.

Zohary, D., Hopf, M., Weiss, E. (2012) *Domestication of plants in the old world.* 4th ed., Oxford University Press, Oxford, pp. 264.

Zou, P., Fu, J., Cao, Z.H. (2011) Chronosequence of paddy soils and phosphorus sorption–desorption properties. *J. Soils Sediments 11*, 249–259.

Acknowledgments

I would like to express my gratitude to those numerous people who have supported and encouraged me during the period of my life that I dedicated to this work.

I am deeply grateful to my supervisor Prof. Dr. Lorenz Schwark for making this doctoral thesis possible for me and sharing his expertise throughout the entire time.

Furthermore, I would like to thank the German Research Foundation (DFG) for financial support of the Research Initiative FOR 995

A special thanks goes to all my motivators, friends and helping hands at the University of Cologne (foremost Mareike Wolf, Nicole Mantke, Sina Grewe and Bianca Stapper) and at the University of Kiel (foremost Kirsten Hartmann, Petra Fiedler, Inge Dold, Eyke Kirchhoff., Nicole Häuser, Jan Kraume-Flügel, Marieke Sieverding, Vanessa Grothe and Robert Kräft).

I would like to acknowledge the various collaborators in China (Prof. Zhi-Hong Cao, Dr. Jin Zhang) and in Indonesia (Prof. Sri Rahayu Utami) for organization and preparation of sampling locations and the help during field work.

I would like to thank Andrea Bannert, Angelika Kölbl, Philipp Roth, Eva Lehndorff, Tino Bräuer, Livia Urbanski, Alexander Hanke, Vanessa Vogelsang, Adrian Ho, Pauline Winkler, Andreas Hofmann, Miriam Houtermans and Erwin Prastowo from the Chinese and Indonesian paddy soil groups for serious scientific discussions during the PhD meetings, for sharing their data, for the great time and for the enjoyable dinners.

I would like to express my gratitude to my colleagues Martin Stockhausen, Wolfgang Rübsam, Dr. Thorsten Bauersachs and Ann-Sophie Jonas for help both on a scientific as well as a personal level.

I am very beholden to my family for their love, endurance and belief in my skills. I would to dedicate this work to them.

Personal contribution to multiple-author publications

The present Ph.D. thesis comprises four publications (Chapter 2, 3, 4 and 5) and one manuscript submitted for publication (Chapter 6), which were prepared in cooperation with various co-authors. My contribution to these manuscripts was:

Chapter 2: All soil/sediment samples were recovered in June 2008 by Prof. L. Schwark. A. Bannert and Prof. M. Schloter provided data of Nmic, Cmic, NO^{3-}, NH^{4+}, DON and DOC contents. E. Lehndorff provided data of bulk isotope. Bulk analyses and lipid extractions were performed by C. Müller-Niggemann. Statistical analyses and calculations were performed by C. Müller-Niggemann. Data interpretation, synthesis of results and preparation of the manuscript was done by C. Müller-Niggemann under supervision of Prof. L. Schwark.

Chapter 3: Soil/sediment and plant samples were recovered in June 2008 by Prof. L. Schwark. All biogeochemical analyses were performed by C. Müller-Niggemann. GC-MS analyses, identifications and evaluation of substances were performed by C. Müller-Niggemann. Data interpretation, synthesis of results and preparation of the manuscript was done by C. Müller-Niggemann under supervision of Prof. L. Schwark.

Chapter 4: Soil samples were sampled upon various field campaigns ranging from 2008 to 2014 by Prof. L. Schwark (China, Sumatra), C. Müller-Niggemann (China, Java), Dr. A. Marxen (Philippines, Vietnam), Dr. C. Lüke (Italy, Vietnam), Dr. A. Kölbl (Italy). All biogeochemical analyses were performed by C. Müller-Niggemann. GC-MS analyses were performed by C. Müller-Niggemann. APCI LC-MS analyses were performed by K. Mangelsdorf and T. Bauersachs. Identification and evaluation of discussed substances were performed by C. Müller-Niggemann. Statistical analyses and calculations were performed by C. Müller-Niggemann. Data interpretation, synthesis of results and preparation of the manuscript was done by C. Müller-Niggemann supported by Prof. L. Schwark.

Chapter 5: The manuscript describes a twofold approach with the microbiology conducted by the Helmholtz group and the lipidomics carried out by the Kiel Group. All soil samples were recovered in July 2009. C. Müller-Niggemann was involved in field work and performed the entire lipid analyses and data interpretation. C. Müller-Niggemann was responsible for material and methods part for lipid analyses contributed the lipids sections of the discussion part.

Chapter 6: All soil/sediment and rice ash samples were recovered in June 2008 by Prof. L. Schwark. All biogeochemical analyses were performed by C. Müller-Niggemann. GC-MS analyses, identification and evaluation of substances were performed by C. Müller-Niggemann. The BC data were provided by Dr. E. Lehndorff for evaluation. Data interpretation, synthesis of results and preparation of the manuscript was done by C. Müller-Niggemann under supervision of Prof. L. Schwark.

List of publications

Research article

Mueller-Niggemann, C., Utami, S.R., Marxen, A., Mangelsdorf, K., Bauersachs, T., Schwark, L. (2015) Distribution of tetraether lipids in agricultural soils – differentiation between paddy and upland management, *Biogeosciences Discuss.* 12, 16709-16754, doi:1010.5194/bgd-12-16709-2015.

Mueller-Niggemann, C., Schwark, L. (2015) Chemotaxonomy and diagenesis of aliphatic hydrocarbons in rice plants and soils from land reclamation areas in the Zhejiang Province, China, *Org. Geochem. 83-84*, 215-226, doi:10.1016/j.orggeochem.2015.03.016.

Kölbl, A., Mueller-Niggemann, C., Schwark, L., Cao, Z.H., Kögel-Knabner, I. (2015) Spatial distribution of soil organic matter in two fields on tidal flat sediments (Zhejiang Province, China) differing in duration of paddy management, *J. Plant Nutr. Soil Sci. 178*, 649-657, doi:10.1002/jpln.201400119.

Kölbl, A., Schad, P., Jahn, R., Amelung, W., Bannert, A., Cao, Z.H., Fiedler, S., Kalbitz, K., Lehndorff, E., Müller-Niggemann, C., Schloter, M., Schwark, L., Vogelsang, V., Wissing, L., Kögel-Knabner, I. (2014) Accelerated soil formation due to paddy management on marshlands (Zhejiang Province, China), *Geoderma 228-229*, 67-89, doi:10.1016/j.geoderma.2013.09.005.

Mueller-Niggemann, C., Bannert, A., Schloter, M., Lehndorff, E., Schwark, L. (2012) Intra- versus inter-site macroscale variation in biogeochemical properties along a paddy soil chronosequence, *Biogeosciences 9*, 1237-1251, doi:10.5194/bg-9-1237-2012.

Bannert, A., Kleineidam, K., Wissing, L., Mueller-Niggemann, C., Vogelsang, V., Welzl, G., Cao, Z.-H., Schloter, M. (2011) Changes in diversity and functional gene abundances of microbial communities involved in nitrogen fixation, nitrification and denitrification comparing a tidal wetland to paddy soils cultivated for different time periods, *Appl. Environ. Microbiol. 77*, 6109-6116, doi:10.1128/aem.01751-10.

Bannert, A., Mueller-Niggemann, C., Kleineidam, K., Wissing, L., Cao, Z.-H., Schwark, L., Schloter, M. (2011) Comparison of lipid biomarker and gene abundance characterizing the archaeal ammonia-oxidizing community in flooded soils, *Biol. Fertil. Soils 47*, 839-843, doi:10.1007/s00374-011-0552-6.

Zhang, J., Mueller-Niggemann, C., Wang, M., Cao, Z.H., Luo, X., Wong, M., Chen, W. (2013) Change of PAHs with evolution of paddy soils from prehistoric to present over the last six millennia in the Yangtze River Delta region, China, submitted to *Sci. of the Total Environ. 449, 328-335.* doi: 10.1016/j.scitotenv.2013.01.084.

Poster presentations at scientific conferences

Müller-Niggemann, C., Winkler, P., Kaiser, K., Bauersachs, T., Schwark, L.: Redox-mediated changes in soil microbial community composition after a one year paddy soil formation experiment, *IMOG*, Prague, Czech Republic 2015. *(Best poster award)*

Müller-Niggemann, C., Kölbl, A., Utami, S.R., Schwark, L.: GDGT lipid distribution in rice paddy and upland soil profiles, *IMOG*, Prague, Czech Republic 2015.

Mueller-Niggemann, C., Roth, P., Lehndorff, E., Schwark, L.: Source determination and depth translocation of PAH in Chinese paddy and nonpaddy soils, *International workshop - Biogeochemistry of submerged agro-ecosystems: Properties, processes, cycles and function*, Freising, Germany 2014.

Mueller-Niggemann, C., Mangelsdorf, K., Schwark, L.: Changes in microbial lipids during paddy soil development on marine tidal sediments, *International workshop - Biogeochemistry of submerged agro-ecosystems: Properties, processes, cycles and function*, Freising, Germany 2014.

Kölbl, A., Kaiser, K., Urbanski, L., Schad, P., Geier, P., Vogelsang, V., Jahn, R., Roth, P., Lehndorff, E., Amelung, W., Utami, S.R., Cao, Z.H., Bannert, A., Schloter, M., Kalbitz, K., Müller-Niggemann, C., Schwark, L., Kögel-Knabner, I.: Changes in properties of different soils exposed to paddy management, *WCSS*, Jeju, South Korea 2014. *(Best poster award)*

Mueller-Niggemann, C., Utami, S.R., Marxen, A., Jahn, R., Lueke, C., Frenzel, P. Bauersachs, T., Schwark, L.: Substrate control on microbial lipids in agricultural soils, *IMOG*, Costa Adeje, Spain 2013.

Mueller-Niggemann, C., Roth, P., Lehndorff, E., Schwark, L.: Source determination and depth translocation of PAH in Chinese paddy and non-paddy soils, *IMOG*, Interlaken, Switzerland 2011.

Mueller-Niggemann, C., Zhang, J., Cao, Z.H., Schwark, L.: Composition of aliphatic hydrocarbons in prehistoric rice paddy soils in China, *IMOG*, Interlaken, Switzerland 2011.

Mueller-Niggemann C., Cao Z.H., Schwark L.: Transformation of marine sediment to paddy soil: Primary marine, lacustrine, and land plant lipids, *EGU*, Vienna, Austria 2010.

Mueller-Niggemann, C., Schwark, L.: A 2000 year chronosequence study of marine sediment to paddy soil transformation II: microbial lipids, *IMOG*, Bremen, Germany 2009.

Mueller-Niggemann, C., Schwark, L.: A 2000 year chronosequence study of marine sediment to paddy soil transformation I: primary marine, lacustrine, and land plant lipids, *IMOG*, Bremen, Germany 2009.

Oral presentations

Kögel-Knabner, I., Kölbl, A., Urbanski, L., Schad, P., Kaiser, K., Geier, P., Vogelsang, V., Jahn, R., Roth, P., Lehndorff, E., Amelung, W., Utami, S.R., Cao, Z.H., Bannert, A., Schloter, M., Kalbitz, K., Müller-Niggemann, C., Schwark, L.: Central aims of FOR995: Paddy soil evolution and its dependence on different soil types, *International workshop - Biogeochemistry of submerged agro-ecosystems: Properties, processes, cycles and function*, Freising, Germany 2014.

Mueller-Niggemann, C., Bannert, A., Schloter, M., Mangelsdorf, K., Schwark, L.: Denitrifying bacteria as potential sources of isoalkane-GDGT, *IMOG*, Interlaken, Switzerland 2011.

Mueller-Niggemann C., Bannert A., Schloter M., Cao Z.H., Schwark L.: Microbial lipids in Paddy Soils of the Yangtze Area, *EGU*, Vienna, Austria 2010.

Kölbl, A., Mueller-Niggemann, C., Schwark, L., Cao, Z.H., Fu, J., Kögel-Knabner, I.: Spatial distribution of SOM parameters during paddy soil evolution, *EGU*, Vienna, Austria 2010.

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig, ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe. Die Arbeit entstand unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft. Ferner erkläre ich, dass ich diese Dissertation weder in ihrer Gesamtheit noch in Teilen einer anderen wissenschaftlichen Hochschule im Rahmen eines Prüfungsverfahrens vorgelegt habe.

Kiel, den 17.11.2015 gez. Cornelia Müller-Niggemann