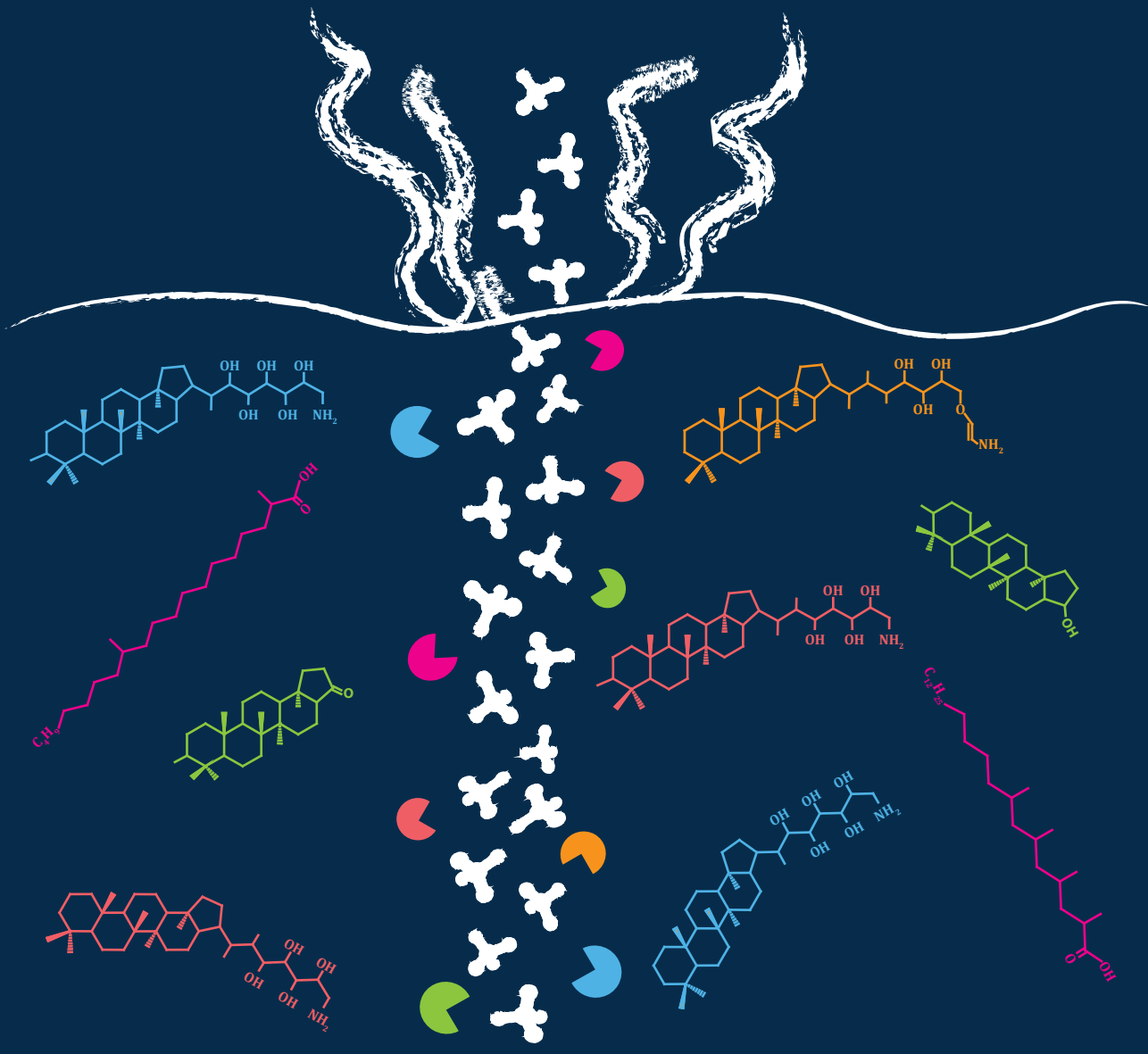
Novel lipid biomarkers for detecting microbial oxidation of methane in the environment

Nadine T. Smit



Novel lipid biomarkers for detecting microbial oxidation of methane in the environment

Nadine Talea Smit

Printed by Ipskamp Printing, Enschede
Cover design: Nadine T. Smit
ISBN 978-94-6421-424-6

Novel lipid biomarkers for detecting microbial oxidation of methane in the environment

Nieuwe biomarker lipiden voor detectie van microbiële oxidatie van methaan in het milieu
(met een samenvatting in het Nederlands)

Neuartige Lipid-Biomarker zum Nachweis von mikrobieller Methanoxidation in der Umwelt
(mit einer Zusammenfassung in Deutsch)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 27 augustus 2021 des ochtends te 10.15 uur

door

Nadine Talea Smit

geboren op 7 juni 1990
te Meerbusch, Duitsland

Promotoren:

Prof. dr. ir. S. Schouten

Prof. dr. ir. J.S. Sinninghe Damsté

Copromotor:

Dr. D. Rush

This work has been funded by the Netherlands Earth System Science Center (NESSC) through a gravitation grant (NWO 024.002.001) from the Dutch Ministry for Education, Culture and Science.

“If we knew what we were doing, it would not be called research, would it?”

Albert Einstein

Table of Contents

Summary

Methane is a powerful greenhouse gas which has impacted Earth's climate over geological timescales. Anthropogenic emissions have led to a significant increase in its atmospheric concentration in the last two centuries. Various natural and anthropogenic sources are known to emit methane into the atmosphere, e.g. fossil fuels, wetlands, thawing permafrost, marine sediments and hydrocarbon seeps. However, much of this emitted methane is subject to microbial oxidation processes which act as a large sink. Anaerobic methanotrophic archaea (ANMEs) are mainly responsible for the anaerobic oxidation of methane (AOM), while the aerobic oxidation of methane is conducted by methane-oxidizing bacteria (MOBs). In addition to DNA-based techniques, lipid biomarkers offer an additional tool to investigate microbial communities involved in methane oxidation in present-day and past environments. However, to this date, only a few unambiguous biomarkers for methanotrophs like specific fatty acids or bacteriohopanepolyols (BHPs) are known to trace microbes involved in the aerobic oxidation of methane. Therefore, this thesis aims to expand the current toolbox of characteristic biomarkers used to trace aerobic oxidation of methane and other hydrocarbon gasses, especially in terrestrial environments.

In the first part of this thesis, lipid biomarkers for nitrite/nitrate-dependent methane oxidation were investigated in enrichment cultures of the intra-aerobic methanotroph '*Candidatus* Methylomirabilis oxyfera' ('*Ca*. M. oxyfera') and the ANME-2d archaeon '*Candidatus* Methanoperedens'. '*Ca*. M. oxyfera' directly synthesizes four unique demethylated hopanoids (22,29,30-trisnorhopan-21-ol, 3-methyl-22,29,30-trisnor hopan-21-one and 3-methyl- 22,29,30-trisnorhopan-21-ol), of which only 22,29,30-trisnorhopan-21-one had been identified previously. A new multiple reaction monitoring method was developed and used to successfully detect these trisnorhopanoids in a peatland. '*Ca*. Methanoperedens' produced archaeol and hydroxyarchaeol as well as isoprenoidal GDGTs and unusually high amounts of hydroxy-GDGTs. The novel demethylated hopanoids and high abundances of hydroxy-GDGTs may thus be used in future studies to trace nitrite/nitrate-dependent methane oxidation in various present and past environments.

Mud volcanoes (MVs) release high amounts of methane into the atmosphere and, therefore, three prominent terrestrial Sicilian mud volcanoes were investigated for their lipid biomarker inventories to study microbial communities involved in aerobic and anaerobic methane oxidation. The freshly emitted mud breccias

showed high abundances of petroleum-derived hydrocarbons but low amounts of microbially derived lipid biomarkers such as fatty acids and hopanoic acids, which were not depleted in ^{13}C. These results suggest that these lipids might originate from bacteria other than methane-oxidizing bacteria. Additional analysis of isoprenoidal GDGTs present in one MV show a distribution indicative for anaerobic oxidation of methane by ANME archaea. The possible presence of AOM in the Maccalube di Aragona MV but not in the other two MVs implies that microbial processes other than methane oxidation may play an important role in these terrestrial MVs in contrast to marine MV systems and warrants further investigations in the future.

The biomarker potential of known and novel bacteriohopanepolyols (BHPs) and their derivatives to trace aerobic methane oxidizing bacteria was investigated in this thesis. First, a new UHPLC-HRMS^2 method to detect underivatized BHPs was applied to bacterial cultures and a soil near a terrestrial methane seep (Fuoco di Censo everlasting fire in Sicily, Italy). This revealed mass spectra with distinct fragmentation patterns of common BHPs and their methylated and unsaturated counterparts as well as aminoBHPs and various novel composite BHPs (e.g. ethenolamineBHPs). Application of this methodology to two terrestrial methane seep transects (Censo and Bissana seep) showed high relative abundances of aminoBHPs (i.e. aminopentol) and the novel ethenolamineBHPs at and close to the seeps, which are likely synthesized by Type I MOBs (e.g. *Methylococcaceae*) and potentially other Proteobacteria species. The Censo seep showed high relative abundances of a novel late-eluting aminotriol isomer, which was also identified in the verrucomicrobial strain, *Methylacidimicrobium cyclopophantes* 3B, suggesting it as a new biomarker for Verrucomicrobia. A new index (AminoBHP-index) was developed based on selected BHPs, which shows high values close to the seeps (≥ 0.4) and drops to < 0.2 at distances >3 m from the active seeps. This novel AminoBHP-index offers a new biomarker proxy to reconstruct present and past aerobic oxidation of methane in the terrestrial realm.

Soils from a natural everlasting fire named Fuoco di Censo in Sicily, Italy, were also studied. They showed high abundances of novel mycobacteria as revealed by 16S rRNA gene sequencing and the presence of mycocerosic acids (MAs), multi-methyl branched fatty acids, at and close to the seep. These novel soil mycobacteria are phylogenetically closely related to the *Mycobacterium simiae* complex but more remotely to well-studied pathogenic mycobacteria such as *M. tuberculosis*. A range of new MAs were identified including unusual ones containing mid-chain methyl branches at positions C-12 and C-16. These MAs

were depleted in ^{13}C, suggesting a direct or indirect utilization of ^{13}C depleted gases such as methane and ethane. Confirmation of the use of these hydrocarbon gasses as a carbon source came from SIP incubations using ^{13}C-labeled methane and ethane, which revealed ^{13}C label incorporation of ethane, but not of methane, into the MAs. These results imply that mycobacteria at the Censo seep oxidize ethane but not methane, in agreement with previous studies. Thus, these novel MAs in tandem with ^{13}C depleted isotopic signatures offer new unique biomarkers to trace higher gaseous hydrocarbon oxidation and gas-consuming soil mycobacteria in modern and past environments.

In summary, the results presented in this thesis have substantially expanded the lipid biomarker toolbox for organic geochemists to detect recent and past methane and ethane oxidation. Together with more established biomarker proxies, these newly developed proxies such as mycocerosic acids and the AminoBHP-index may be tested and applied further to trace past methane and higher gaseous hydrocarbon cycles in present and past environments in the future.

Samenvatting

Methaan is een krachtig broeikasgas met een grote invloed op het Aardse klimaat over geologische tijdschalen. Antropogene emissies hebben de afgelopen twee eeuwen geleid tot een aanzienlijke toename van de atmosferische concentratie van methaan. Van verschillende natuurlijke en antropogene bronnen is bekend dat ze methaan in de atmosfeer uitstoten, zoals moerassen, ontdooiende permafrost, mariene sedimenten en het natuurlijk weglekken van fossiele brandstoffen. Veel van dit uitgestoten methaan is echter onderhevig aan microbiële oxidatieprocessen die als een belangrijke bezinkput fungeren. Anaërobe methanotrofe archaea (ANME's) zijn voornamelijk verantwoordelijk voor de anaërobe oxidatie van methaan (AOM), terwijl de aërobe oxidatie van methaan wordt uitgevoerd door methaan-oxiderende bacteriën (MOB's). Naast op DNA gebaseerde technieken bieden lipide biomarkers een extra hulpmiddel om microbiële gemeenschappen op te sporen die actief zijn (geweest) in de methaanoxidatie in zowel huidige als vroegere milieus. Tot op heden zijn er echter slechts enkele eenduidige biomarkers voor methanotrofen bekend om microben te traceren die betrokken zijn bij de aerobe oxidatie van methaan, zoals specifieke vetzuren of bacteriohopaanpolyolen (BHP's). Daarom is het doel van dit proefschrift om de huidige verzameling van karakteristieke biomarkers die kunnen worden gebruikt om aerobe oxidatie van methaan en andere koolwaterstofgassen te traceren, uit te breiden, vooral voor terrestrische milieus.

In het eerste deel van dit proefschrift werden lipide biomarkers voor nitriet/nitraat-afhankelijke methaanoxidatie onderzocht in verrijkingsculturen van de intra-aërobe methanotrofe bacterie '*Candidatus* Methylomirabilis oxyfera' ('*Ca*. M. oxyfera') en de ANME-2d archaeon ' *Candidatus* Methanoperedens'. *'Ca.* M. oxyfera' synthetiseert vier unieke gedemethyleerde hopanoïden (22,29,30-trisnorhopan-21-ol, 3-methyl-22,29,30-trisnorhopan-21-on en 3-methyl- 22,29,30- trisnorhopan-21-ol), waarvan slechts 22,29,30-trisnorhopan-21-on eerder was geïdentificeerd. Om deze trisnorhopanoïden succesvol te detecteren in een veengebied, is een nieuwe massapectrometrische methode die gebruik maakt van meervoudige reactie-monitoring ontwikkeld en toegepast. *'Ca.* Methanoperedens' produceerde naast archaeol en hydroxyarchaeol ook isoprenoïde GDGT's en hydroxy-GDGT's, die laatsten in ongebruikelijk grote hoeveelheden. Deze specifieke biomarkerdistributie kan in toekomstige studies worden gebruikt om nitriet/nitraat-afhankelijke methaanoxidatie in zowel huidige als vroegere milieus te traceren.

Moddervulkanen (MV's) vormen een belangrijke bron voor atmosferisch methaan. Daarom werden drie prominente terrestrische Siciliaanse MV's onderzocht op lipide biomarkers om microbiële gemeenschappen te bestuderen die betrokken zijn bij zowel aerobe als anaërobe methaanoxidatie. De vers uitgespuwde modderbreccia's bevatte grote hoeveelheden koolwaterstoffen afkomstig van aardolie, maar slechts kleine hoeveelheden microbiële lipide biomarkers zoals vetzuren en hopaanzuren, die niet verarmd waren in ^{13}C. Deze resultaten suggereren dat deze lipiden afkomstig kunnen zijn van andere bacteriën dan MOB's. Aanvullende analyse van isoprenoïdale GDGT's aanwezig in de Maccalube di Aragona MV liet een distributie zien die indicatief is voor anaërobe oxidatie van methaan door ANME's. De mogelijke aanwezigheid van AOM in de Maccalube di Aragona MV maar niet in de andere twee MV's impliceert dat andere microbiële processen dan methaanoxidatie een belangrijke rol kunnen spelen in deze terrestrische MV's, in tegenstelling tot mariene MV-systemen. Dit rechtvaardigt verder toekomstig onderzoek.

In het tweede deel van dit proefschrift werd het biomarker potentieel van bekende en nieuw-geïdentificeerde bacteriohopaanpolyolen (BHP's) om aerobe MOB's op te sporen onderzocht. Een nieuwe UHPLC-HRMS^2-methode om niet-gederivatiseerde BHP's te detecteren werd gebruikt in de analyse van culturen van bacteriën en van een bodem in de buurt van een terrestrische methaanbron (de 'Fuoco di Censo', een eeuwig vuur op Sicilië, Italië). De verkregen massaspectra lieten verschillende fragmentatiepatronen zien die het mogelijk maakte bekende BHP's en hun gemethyleerde en onverzadigde tegenhangers, amino-BHP's en verschillende nieuwe samengestelde BHP's (bijv. ethenolamine-BHP's) te identificeren. Toepassing van deze analysetechniek op bodems rondom twee terrestrische methaanbronnen (Censo en Bissana seep) toonde grote hoeveelheden van aminoBHP's (voornamelijk aminopentol) en de nieuwe ethenolamine-BHP's in bodems nabij de methaanbronnen aan. Deze BHP's worden waarschijnlijk gesynthetiseerd door Type I MOB's (bijvoorbeeld *Methylococcaceae*) en mogelijk ook door andere proteobacteria. De Censo-methaanbron bevatte ook grote hoeveelheden van een nieuwe, laat-eluerende, aminotriol-isomeer. Deze BHP werd ook aangetroffen in een bacterie behorende tot de Verrucomicrobia, i.e. *Methylacidimicrobium cyclopophantes 3B*. Mogelijk is dit een nieuwe biomarker voor methaanoxiderende Verrucomicrobia. Een nieuwe index (de AminoBHP-index) werd ontwikkeld op basis van geselecteerde specifieke BHP's. Deze index laat hoge waarden laat zien dichtbij de methaanbron (>0,4) en daalt tot < 0,2 op afstanden van >3 m van de bron. Deze nieuwe AminoBHP-index biedt een nieuwe biomarker-proxy om

aerobe oxidatie van methaan in huidige en vroegere milieus in het geologisch verleden te reconstrueren.

Ook werden in bodems van een natuurlijk eeuwigdurend vuur, genaamd Fuoco di Censo op Sicilië, Italië, grote hoeveelheden van nieuwe mycobacteriën gevonden. Deze mycobacteriën werden aangetoond in bodems vlakbij de methaanbron middels 16S rRNA-gen sequencing alsmede door de aanwezigheid van mycocerosinezuren (MA's), multi-methyl vertakte vetzuren, karakteristieke biomarker lipiden voor mycobacteriën. Deze nieuwe bodembacteriën zijn fylogenetisch nauw verwant aan het *Mycobacterium simiae*-complex, maar meer op afstand van goed-bestudeerde, pathogene mycobacteriën zoals *M. tuberculosis*. Een aantal nieuwe MA's werd geïdentificeerd. Deze MA's zijn ongebruikelijke omdat zij methylgroepen in het midden van de koolstofketen, op posities C-12 en C-16, bevatten. Deze MA's bevatte relatief weinig ^{13}C, hetgeen wijst op een direct of indirect gebruik van ^{13}C-verarmde gassen zoals methaan en ethaan. Het gebruik van koolwaterstofgassen als koolstofbron werd bevestigd door middel van incubatiestudies met ^{13}C-gelabeld methaan en ethaan. Hieruit bleek dat ^{13}C-label van ethaan wel werd opgenomen in de MA's, maar niet dat van methaan. De implicatie van deze bevindingen is dat mycobacteriën in de Censo bodems ethaan maar niet methaan oxideren. Dit is in overeenstemming met suggesties uit eerder werk. Deze nieuwe MA's in combinatie met ^{13}C-verarmde isotopensignatuur bieden dus nieuwe unieke biomarkers voor het opsporen van ethaan-oxiderende mycobacteriën in bodems van moderne en vroegere milieus.

Samenvattend kan gesteld worden dat de resultaten beschreven in dit proefschrift het aantal lipide biomarkers dat beschikbaar is voor organisch geochemici om methaan- en ethaanoxidatie te detecteren in vroegere en huidige milieus aanzienlijk uitgebreid hebben. Samen met meer gevestigde biomarker-proxy's, kunnen deze nieuw ontwikkelde proxy's, zoals mycocerosinezuren en de AminoBHP-index, in de toekomst verder worden getest en mogelijk worden toegepast om de biogeochemische kringloop van koolwaterstoffengassen in verschillende milieus te bestuderen.

Zusammenfassung

Methan ist ein wirkungsvolles Treibhausgas in der Erdatmosphäre und Änderungen in der atmosphärischen Methankonzentrationen sind stark verknüpft mit Temperaturschwankungen über geologische Zeiträume. Anthropogene Treibhausgasemissionen haben die atmosphärischen Methankonzentrationen in den letzten zwei Jahrhunderten erheblich ansteigen lassen. Methan wird durch diverse anthropogene und natürliche Quellen in die Umwelt eingetragen, z.B. durch fossile Brennstoffe, Feuchtgebiete, schmelzenden Permafrost, marine Sedimente oder natürliche Kohlenwasserstoffaustritte. Die mikrobielle Methanoxidation spielt eine wichtige Rolle als Senke für atmosphärisches Methan in terrestrischen als auch in marinen Ökosystemen. Dort wird Methan entweder anaerob (unter Ausschluss von Sauerstoff) von anaeroben methanotrophen Archaeen oder aerob (mit Sauerstoff) von methanoxidierenden Bakterien chemisch umgewandelt. Neben DNA-basierten Analysemethoden, ist der Einsatz von Lipid-Biomarkern (Moleküle hauptsächlich bestehend aus Kohlenwasserstoffverbindungen) eine vielversprechende und zuverlässigere Methode um methanoxidierende Mikroorganismen in der heutigen Umwelt sowie in der geologischen Vergangenheit zu identifizieren. Böden stellen weltweit die größte mikrobielle Senke für atmosphärisches Methan dar, jedoch beschränken sich bisherige Studien über Lipid-Biomarker in der geologischen Vergangenheit meist auf marine Methanaustritte. Daher wurden bisher nur wenige spezifische Lipid-Biomarker, z.B. spezifische Fettsäuren oder die sogenannten Bacteriohopanepolyole (BHPs), gefunden mit denen aerobe methanoxidierende Mikroorganismen in der Umwelt nachgewiesen werden können. Diese Doktorarbeit zielt daher darauf ab, sowohl neue charakteristische Lipid Biomarker für aerobe und anaerobe methanoxidierende Mikroorganismen zu finden als auch ihre Anwendung speziell in der terrestrischen Umwelt zu untersuchen.

Im ersten Teil der vorliegenden Doktorarbeit werden Lipid-Biomarker spezifisch für die Nitrit/ Nitrat-abhängige Methanoxidation in zwei unterschiedlichen Anreicherungskulturen untersucht. Das intra-aerobe methanotrophe Bakterium ‚*Candidatus* Methylomirabilis oxyfera' (‚*Ca*. M. oxyfera') synthetisiert vier einzigartige demethylierte Hopanoide (22,29,30-trisnorhopan-21-ol, 3-methyl-22,29,30-trisnor hopan-21-one and 3-methyl- 22,29,30-trisnorhopan-21-ol), von denen nur 22,29,30-trisnorhopan-21-one bereits zuvor identifiziert wurde. Mithilfe einer neuentwickelten Analysemethode, welche spezifisch nach den vier neuen Trisnorhopanoiden sucht, wurde ‚*Ca*. M. oxyfera' in einem

Moorgebiet nachgewiesen. Die zweite Anreichungskultur besteht aus vorrangig dem anaeroben methanotrophen (ANME-2d) Archaeon *‚Candidatus* Methanoperedens', diese produzieren hauptsächlich die beiden Komponenten Archaeol und Hydroxyarchaeol als auch GDGTs und einen hohen Anteil von Hydroxy-GDGTs. Zukünftig können die neu identifizierten Trisnorhopanoide und ungewöhnlich hohe Konzentrationen von Hydroxy-GDGTs in Studien zum Nachweis von Nitrit/ Nitrat-abhängigen Methanoxidation in modernen Umweltproben und älteren geologischen Zeiträumen angewandt werden.

Schlammvulkane sind ideale Lebensräume für methanoxidierende Mikroorganismen, da sie kontinuierlich Methan in die Erdatmosphäre ausstoßen. Drei große terrestrische Schlammvulkane wurden auf das Vorhandensein von Lipid-Biomarkern für anaerobe und aerobe Methanoxidation in Sizilien untersucht. Der frisch ausgetretene Schlamm wies einen sehr hohen Anteil an erdölspezifischen Kohlenwasserstoffverbindungen auf wohingegen der Anteil von mikrobiellen Lipid-Biomarkern, z.B. Fettsäuren und Hopansäuren, sehr gering war. Des Weiteren waren diese Lipid-Biomarker isotopisch nicht abgereichert in ^{13}C, was darauf hindeutet, dass diese Lipide nicht von methanoxidierenden Bakterien produziert wurden. Die Analyse von GDGTs in den Schlammvulkanen zeigte jedoch, dass in einem Schlammvulkan (Maccalube di Aragona) GDGT-Verteilungen vorhanden waren, welche von anaeroben methanotrophen Archaeen stammen könnten. Die Ergebnisse zeigen, dass nur wenige bis keine spezifischen Lipid-Biomarker für methanoxidierende Bakterien oder Archaeen in den drei Schlammvulkanen gefunden werden konnten. Daher spielen vermutlich andere mikrobielle Prozesse als Methanoxidation in terrestrischen Schlammvulkanen eine größere Rolle als in früheren Studien für marine Schlammvulkane gezeigt wurde.

Im nächsten Teil der Doktorarbeit werden bekannte und neue Bacteriohopanpolyole (BHPs) und ähnliche Verbindungen untersucht, um ihr Potential als Lipid-Biomarker für aerobe methanoxidierende Bakterien abzuschätzen. Zuerst wurde eine neuentwickelte UHPLC-$HRMS^2$ Methode eingesetzt um nicht derivatisierte BHPs in diversen Bakterienkulturen und einer Bodenprobe von einem natürlichen terrestrischen Methanaustritt in Sizilien (‚Fuoco di Censo' – immer brennendes Feuer) zu identifizieren. Diese Analysen zeigten Massenspektren mit sehr spezifischen Fragmentierungsmustern von weitverbreiteten BHPs und ihren methylierten und ungesättigten verwandten Molekülen, sowie typische AminoBHPs und diverse neue Komposit-BHPs (z.B. EthenolamineBHPs). Die Anwendung dieser Methodik auf zwei Profile über

terrestrische Methanaustritte (Censo und Bissana) zeigte hohe relative Vorkommen von AminoBHPs (z.B. Aminopentol) und neuartigen EthenolamineBHPs direkt an und nahe um die Austrittsstellen. Diese Verbindungen werden wahrscheinlich von Typ I methanoxidierende Bakterien (z.B. *Methylococcaceae*) und möglicherweise auch anderen Spezies von Proteobakterien synthetisiert. Der Censo Methanaustritt zeigt zusätzlich hohe relative Vorkommen eines neuartigen Aminotriol-Isomers, welches auch in Verrucobakterien (*Methylacidimicrobium cyclopophantes* 3B) gefunden wurde und damit als ein neuer Biomarker für Verrucobakterien genutzt werden kann. Ein neuer Index wurde auf Basis ausgewählter BHPs entwickelt (AminoBHP-index), welcher hohe Werte (≥ 0.4) nahe an Methanaustrittsstellen und niedrige Werte (<0.2) in Entfernungen über 3 m von aktiven Austritten zeigt. Dieser neue AminoBHP-index bietet einen neuen Biomarker-Proxy, um moderne und vergangene terrestrische aerobe Methanoxidation zu rekonstruieren.

Im letzten Teil der Arbeit werden Bodenproben von Fuoco di Censo (Sizilien, Italien) untersucht, welche hohe Vorkommen von neuartigen Mykobakterien an und nahe den Austrittsstellen durch 16S rRNA Gensequenzierung und Mycocerosic Acids (MAs), d.h. multi-methylierte verzweigte Fettsäuren, zeigten. Diese neuartigen Boden-Mykobakterien sind phylogenetisch eng verwandt mit dem *Mycobacterium simiae* Komplex jedoch weniger eng verwandt mit den vielfach untersuchten pathogenen Mykobakterien, z.B. *M. tuberculosis*. An und nahe den Methanaustrittsstellen wurden neue MAs entdeckt, zum Teil mit sehr ungewöhnlichen Methylverzweigungen in der Mitte der Moleküle an Position C-12 und C-16. Diese neuen MAs sind abgereichert in ^{13}C, was darauf hindeutet, dass die Mykobakterien die ausgetretenen Gase (Methan und/oder Ethan) aufnehmen. Um die vorherigen Annahmen zu bestätigen, dass die Mykobakterien die ausgetretenen Gase als Kohlenstoffquelle nutzen, wurden Inkubationen mit isotopisch gelabelten ^{13}C Methan und ^{13}C Ethan durchgeführt. Diese Inkubationen zeigen, dass die MAs nur ein ^{13}C Isotopenlabel in den Inkubationen mit Ethan aufzeigen jedoch nicht mit Methan. Diese Ergebnisse implizieren, dass die Mykobakterien in den Censo-Böden ausschließlich Ethan oxidieren und nicht Methan, was vorherige Studien zu anderen Mykobakterienarten bestätigt. Daher stellen die neuen MA-Moleküle in Kombination mit abgereicherten ^{13}C Isotopensignaturen neue einzigartige Lipid-Biomarker dar um Gasoxidation und gaskonsumierende Boden-Mykobakterien zukünftig in modernen Ökosystemen sowie in geologisch vergangenen Lebensräumen nachzuweisen.

Die Ergebnisse dieser Arbeit haben die verfügbaren Lipid-Biomarker-Methoden signifikant zur Detektion von moderner und vergangener Methan- und Ethanoxidation erweitert. Zusammen mit etablierten Biomarker-Proxys können diese neuen Proxys, z.B. Mycocerosic Acids und der AminoBHP-Index, getestet und weiterentwickelt werden, um den Kreislauf von Methan und höheren Kohlenwasserstoffen in der Gegenwart und der geologischen Vergangenheit zu rekonstruieren.

Chapter 1

Introduction

1. Introduction

1.1 The greenhouse gas methane in the carbon cycle

The cycling of carbon is a continuous transformation process between the atmosphere, terrestrial biosphere, hydrosphere and geosphere and directly influences global climate systems (Solomon et al., 2007; Kirschke et al., 2013). Methane (CH_4) is a powerful greenhouse gas and anthropogenic emissions have led to a significant increase in its atmospheric concentration in the last two centuries (Fig. 1). Although methane is the second most abundant greenhouse gas in the atmosphere, its global warming potential is about 25x that of CO_2 over 100 years and 84x over a time period of 20 years (Kirschke et al., 2013; Myhre, 2014). However, the average lifetime of methane in the atmosphere is shorter than that of CO_2 and so its effect on the warming potential varies with time. Notably, variations in greenhouse gas concentrations, especially in methane and CO_2 are strongly correlated with changes in temperature over geological timescales (e.g. Delmotte et al., 2004; Pagani et al., 2006; Frieling et al., 2016). Understanding the sources and sinks of methane is important to mitigate anthropogenic caused climate change.

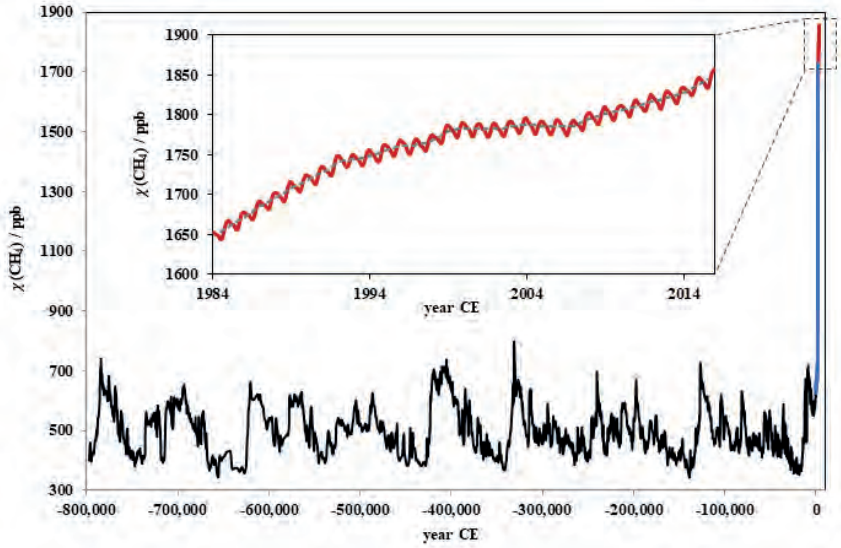


Figure 1 Past and recent fluctuations in atmospheric CH_4 concentrations in the last 800,000 years (modified and adapted from Dean et al., 2018). The enlarged insert demonstrates direct atmospheric CH_4 measurements from 1984 to 2015, and the dashed grey line is a 12 month running mean fit. Data sources: black—data set from the EPICA Dome C ice core record (data from IGBP PAGES/World Data Center for Paleoclimatology, Data Contribution Series # 2008-054) (Loulergue et al., 2008); blue—merged data set from the Law Dome ice core and firn measurements and Cape Grim atmospheric measurements (Etheridge et al., 1998; Trudinger et al., 2002; Ferretti et al., 2005; MacFarling Meure et al., 2006) (data from IGBP PAGES/World Data Center for Paleoclimatology, Data Contribution Series # 2010-070); and red—global average CH4 mole fraction from direct atmospheric measurements (Tsutsumi et al., 2009)(data from World Date Center for Greenhouse Gases, WMO Greenhouse Gas Bulletin, No.12, 2016).

Atmospheric methane concentration has risen substantially to more than 1800 ppb, compared to preindustrial times when the concentrations ranged from 300 to 800 ppb (Fig. 1). Over the industrial period (about 200 years) concentrations increased until the 1990s and stabilized after 2000, which was followed by a new increase in methane concentrations since 2007 until today (Dlugokencky et al., 2009; Nisbet et al., 2014). Past atmospheric methane concentrations are mainly estimated from air which is trapped in polar ice and firn cores (Wolff and Spahni, 2007; Sapart et al., 2013) and the oldest methane record only dates back to the Pleistocene (Loulergue et al., 2008). Natural variations in past atmospheric methane concentration were mainly controlled by the 100,000 year orbital cycle of glacial and interglacial periods, the so-called Milankovitch cycles (Loulergue et al., 2008): Low atmospheric CH_4 concentrations appeared in the glacial periods whereas high atmospheric methane concentrations were seen during warmer interglacial times. The 23,000 year orbital cycle, related to changes in the Earth´s rotational axis, influenced the methane concentrations from 400,000 to 5,000 years ago (Ruddiman, 2003). Unfortunately, quantitative atmospheric methane records prior to the Pleistocene are still unexplored due to the lack of reliable proxies.

To understand the atmospheric methane cycle, it is important to understand the natural sources and sinks which strongly influence global climate feedback mechanisms (Fig. 2). There are various natural and anthropogenic sources emitting methane into the atmosphere. The biggest natural source of methane are wetlands, while other sources include freshwater systems, thawing permafrost, methane hydrates, marine sediments and oceans, diverse fauna and geological sources like hydrocarbon seeps (Reeburgh, 2007; Saunois et al., 2016; Dean et al., 2018). Anthropogenic methane sources include agriculture, waste, biomass burning and fossil fuels (Schwietzke et al., 2016; Dean et al., 2018). Methane is commonly classified into thermogenic or biogenic derived methane, which feature distinct stable carbon isotopic signatures ($\delta^{13}C$) (Bernard et al., 1978; Whiticar, 1999). Thermogenic methane is found mainly in gas and petroleum reservoirs, is formed by thermal cracking of buried organic matter at high temperatures and pressure (Judd, 2004; Peters et al., 2005), and shows $\delta^{13}C$ values of –50 to –20‰. In contrast, the largest source of methane is of biogenic origin and is produced by strictly anaerobic microbes namely methanogenic Euryarchaeota (e.g. from the orders *Methanosarcinales* or *Methanomicrobiales*) (Deppenmeier et al., 1996; Liu and Whitman, 2008; Thauer et al., 2008). Microbial methanogenesis is the terminal step of anaerobic organic matter degradation and is controlled by the abundance of degradable organic matter. Methanogenic archaea convert CO_2 (with H_2), acetate, methanol, methylamines,

or methylsufides into methane via three main biogenic pathways: hydrogenotrophic, methylotrophic or acetoclastic methanogenesis (Deppenmeier et al., 1996; Ferry, 1999; Welte and Deppenmeier, 2014). The biological process of methanogenesis yields strongly ^{13}C-depleted methane values ranging from –110 to –50‰.

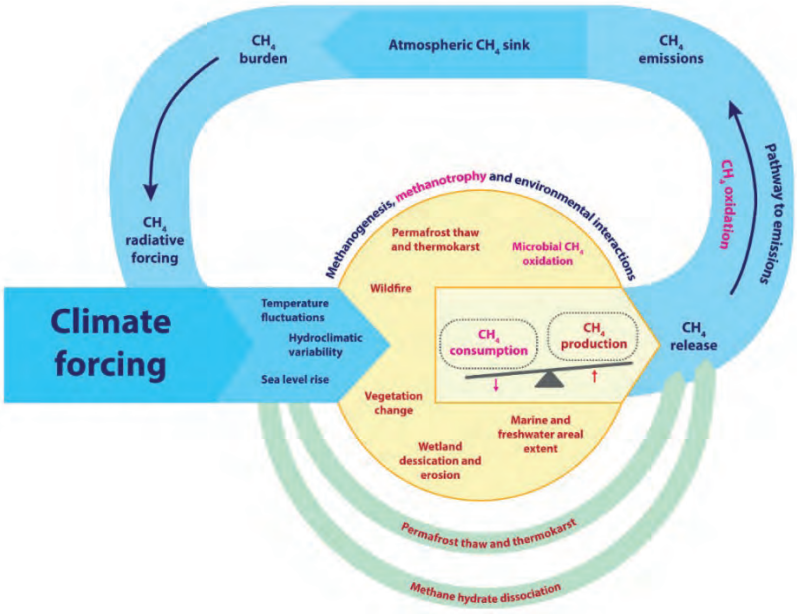


Figure 2. Methane climate feedback schematic (modified after Dean et al., 2018). The schematic shows each feedback mechanism (within the yellow circle) is forced by climate change, affecting CH_4 emissions to the atmosphere. Direct CH_4 feedbacks are demonstrated in green arrows, indirect feedbacks are within the yellow circle, and processes of CH_4 consumption are shown in pink. The positive CH_4 climate feedback cycle is shown in blue.

The largest sink for atmospheric methane is the oxidation by OH radicals in the troposphere where about 85% of methane is oxidized. During this process, methane oxidation leads to O_3 formation, followed by the production of H_2O vapor in the stratosphere leading to radiative forcing and finally the C atom ends up in the form of CO_2 (Myhre, 2014). Besides the oxidation of methane in the stratosphere, both biogenic and thermogenic methane are converted to CO_2 by anaerobic or aerobic microbial oxidation processes in many terrestrial and marine environments (Hanson and Hanson, 1996; Boetius et al., 2000; Valentine and Reeburgh, 2000; Kip et al., 2010). Terrestrial environments, and especially soils, are one of the largest sinks for atmospheric methane, taking up to 10–15% of methane worldwide by aerobic methanotrophic bacteria (Hanson and Hanson, 1996; Bull et al., 2000; Smith et al., 2000; Delgado-Baquerizo et al., 2018). In marine environments, most methane is oxidized microbially by the anaerobic oxidation of methane in anoxic sediments or in the water column via the aerobic

oxidation of methane before reaching the atmosphere (Hinrichs et al., 1999; Valentine and Reeburgh, 2000; Orphan et al., 2001; Reeburgh, 2007). These processes are discussed in more detail below.

1.2 Microbial oxidation of methane

1.2.1 Anaerobic oxidation of methane

In the late 1970s the first indications of anaerobic oxidation of methane (AOM) were discovered from geochemical calculations on porewater profiles in anoxic marine environments, but the microbes involved were only identified two decades later (Hinrichs et al., 1999; Boetius et al., 2000; Pancost et al., 2000). AOM is conducted by anaerobic methanotrophic archaea (ANME type 1, 2 and 3) which belong to the Euryarchaeota and are found primarily in marine environments (Hinrichs et al., 1999; Boetius et al., 2000; Pancost et al., 2000; Orphan et al., 2001). The different clades of ANME-archaea are phylogenetically closely related to methanogenic Euryarchaea and perform the reverse methanogenesis pathway (Hinrichs et al., 1999; Krüger et al., 2003; Hallam et al., 2004). In marine sediments, ANME-archaea form consortia with sulfate-reducing bacteria (SRBs), which create sulfate-methane transitions zones in marine sediments via sulfate-dependent AOM processes (Knittel and Boetius, 2009). Furthermore, ANME-1 and ANME-2 archaea often form consortia with *Desulfosarcina* and *Desulfococcus* while ANME-3 archaea form consortia with *Desulfobulbus*. However, they can also exist without SRB-partners (Boetius et al., 2000; Orphan et al., 2001; Niemann et al., 2006).

Besides sulfate-dependent AOM, nitrate-dependent AOM was more recently discovered in ANME-2d archaea, namely '*Candidatus* Methanoperedens nitroreducens', living in freshwater ecosystems (Haroon et al., 2013; Arshad et al., 2015; Vaksmaa et al., 2017). ANME-2d archaea can coexist with a bacterial partner, '*Candidatus* Methylomirabilis oxyfera' ('*Ca*. M. oxyfera'). Although '*Ca*. M. oxyfera' performs nitrite-dependent methane oxidation under anoxic conditions by reducing nitrite to nitric oxide, it is thought to produce its own intracellular oxygen from nitric oxide for the intra-aerobic oxidation of methane (Raghoebarsing et al., 2006; Ettwig et al., 2008, 2010). Moreover, it has been shown that '*Ca*. M. oxyfera' fixes carbon autotrophically by solely using inorganic carbon (Kool et al., 2012, 2014; Rasigraf et al., 2014). '*Ca*. M. oxyfera' can be found mainly in freshwater sediments or peatlands but was also recently detected in marine environments (Padilla et al., 2016). In addition, some specific types of methane-oxidizing bacteria like *Methylomonas denitrificans* have been shown to be able to oxidize methane under anoxic conditions via nitrate and oxygen reduction (Kits et al., 2015).

1.2.2 Aerobic oxidation of methane

In terrestrial and marine environments, microbial mediated aerobic oxidation of methane plays a key role in oxidizing methane before reaching the atmosphere (Bull et al., 2000; Bodelier et al., 2009; Bowman, 2011). Aerobic methane oxidising bacteria (MOBs) are divided into three phylogenetic groups based on their metabolism and morphology: Type I methanotrophs (*Gammaproteobacteria*), Type II methanotrophs (*Alphaproteobacteria*) and the (thermo)acidophilic *Verrucomicrobia* (Hanson and Hanson, 1996; Pol et al., 2007; Op den Camp et al., 2009).

Type I methanotrophs can be classified into the families *Methylococcaceae* or *Methylothermaceae*, which are only known to contain methanotrophic bacteria (Hirayama et al., 2014; Knief, 2015). Type I methanotrophs live mainly in methane-rich terrestrial and marine environments where they fix carbon via the ribulose monophosphate pathway and solely utilize methane as their carbon and energy source (Hanson and Hanson, 1996; Knief, 2015). Some members of the *Methylococcaceae* family are also able to utilize higher gaseous hydrocarbons (ethane and propane) in environments where gas or petroleum seepage occurs (Kinnaman et al., 2007; Redmond et al., 2010). In addition to Type I methanotrophs, a group of gram-positive organisms in the Corynebacterium–Nocardia–Mycobacterium–Rhodococcus group (Dworkin and Foster, 1958; Ashraf et al., 1994) as well as some gram-negative Pseudomonas species are able to oxidize higher gaseous hydrocarbons in the environment (Hamamura et al., 1997; Takahashi, 1980).

Type II methanotrophs belong to two families, the *Methylocystaceae* and *Beijerinckiaceae*, which also include various non-methanotrophic genera (Knief, 2015). Type II methanotrophs use the serine cycle for carbon fixation and might assimilate only part of their carbon from methane oxidation, while the rest originates from CO_2 (Hanson and Hanson, 1996). Type II methanotrophs are commonly found in the terrestrial realm where they can utilize moderate to low and even atmospheric methane concentrations (Holmes et al., 1999; Bull et al., 2000).

Finally, *Verrucomicrobia* are phylogenetically distinct from the Type I and II methanotrophic *Proteobacteria* and were first discovered in a fumarole where they perform aerobic oxidation of methane under high temperature and low pH conditions (Pol et al., 2007; Op den Camp et al., 2009; van Teeseling et al., 2014). Two genera of methanotrophic *Verrucomicrobia* the "*Methylacidiphilum*" (i.e. *Methylacidimicrobium cyclopophantes*) and the "*Methylacidimicrobium*" (i.e. *Methylacidiphilum fumariolicum*) (Pol et al., 2007; Op den Camp et al., 2009; van Teeseling et al., 2014), have been shown

to use both methane and CO_2 as their carbon source. Similar to Type II methanotrophs, they fix carbon by the ribulose-1,5-bisphosphate carboxylase/oxygenase pathway (Khadem et al., 2011).

The three groups of MOBs all contain the methane monooxygenase enzyme, which uses O_2 for the initial conversion step of methane to methanol (Hanson and Hanson, 1996; Knief, 2015). There are two forms of this enzyme, i.e. the soluble methane monooxygenase (sMMO) and the membrane bound particulate methane monooxygenase (pMMO) (Hanson and Hanson, 1996; McDonald et al., 2008). The gene encoding for the β-subunit of the particulate methane monooxygenase (pmoA) is frequently used as a marker to trace MOBs in modern environments since it is present in most MOB-species (Hanson and Hanson, 1996; McDonald et al., 2008; Ettwig et al., 2010).

1.3 Biomarkers to trace past microbial oxidation of methane

Much is known about the present-day role of aerobic and anaerobic methane oxidizers in marine and terrestrial realms, primarily by using DNA-based techniques. However, much less is known about their role in the past as DNA is usually not preserved well over very long-time scales (Hofreiter et al., 2001; Brocks and Pearson, 2005; Boere et al., 2011). Combined with the lack of proxies for past atmospheric methane concentrations, this leads to a gap in our understanding of the impact of climate on the methane cycle and vice versa.

One approach to solve this is to use molecular fossils such as biomarkers, which are organic compounds revealing specific information about present and past biogeochemical cycles. Lipid biomarkers are synthesized by defined groups of microorganisms and can thus serve as an indicator for the presence of certain classes of microbes (Madigan et al., 1997; Brocks and Pearson, 2005; Peters et al., 2005; Killops and Killops, 2013). These lipid biomarkers are preserved over long timescales and can reflect for example global climate changes, anoxic conditions in the water column or the role of microbial communities in extreme environments (Brassell et al., 1986; Summons and Powell, 1986; Hinrichs et al., 1999; Brocks et al., 2005). Typical lipid biomarkers biosynthesized by eukaryotes are e.g. steroids (Volkman, 2003), while archaea produce archaeol and glycerol dialkyl glycerol tetraether lipids (GDGTs) (Tornabene and Langworthy, 1979; Schouten et al., 2013) and some bacteria synthesize e.g. hopanoids (Ourisson et al., 1979; Zundel and Rohmer, 1985; Ourisson and Albrecht, 1992).

Lipid biomarkers characteristic for anaerobic and aerobic methanotrophs can often be detected through their depleted carbon isotope values ($\delta^{13}C$) (Freeman et al., 1990; Hayes, 1993; Hayes, 2001).

As discussed above, biogenic methane is often characterized by a strong depletion in ^{13}C (Whiticar, 1999). Methane oxidizers often do not only oxidize methane to CO_2 but also utilize this methane and/or CO_2 for synthesis of their biomass, and therefore for their lipids, which aids in the recognition of past processes of methane oxidation. The depletion of the ^{13}C isotopic values of lipids that derive from methanotrophs depend on the source of methane (thermogenic or biogenic) as well as the pathway of carbon fixation (pMMO or sMMO) in the different microbial species (Jahnke et al., 1999). As discussed in section 1.2.2, Type I MOBs assimilate solely CH_4 as their carbon and energy source leading to strongly depleted ^{13}C lipid values, while Type II MOBs, *Verrucomicrobia* and '*Ca*. M. oxyfera' also assimilate CO_2 or inorganic carbon and their lipids can show less depleted ^{13}C values. Stable isotope probing (SIP) experiments with ^{13}C-labeled methane in environmental samples can also give insight into lipid biomarkers derived from methanotrophs and their methane consumption rate (e.g. Boschker et al., 1998; Bull et al., 2000; Evershed et al., 2006; Maxfield et al., 2006; Kellermann et al., 2012). The lipid biomarkers most characteristic for methanotrophs are discussed in more detail below.

Table 1. Characteristic lipid biomarker for microbes involved in anaerobic and aerobic oxidation of methane based on culture and environmental studies. Key: ANME= anaerobic methanotrophic archaea; GDGT= glycerol dialkyl glycerol tetraether; PMI= 2,6,10,15,19-pentamethyleicosane; OH= hydroxy; PH= phosphatidyl hexose; PG= phosphatidyl glycerol; MAGE= non-isoprenoidal mono glycerol ethers; DAGE= non-isoprenoidal dialkyl glycerol ethers; FAs= fatty acids; BHP= bacteriohopanepolyol; aminopentol= 35-aminobacteriohopane-30,31,32,33,34-pentol; aminotetrol= 35-aminobacteriohopane-31,32,33,34-tetrol, 3-Me BHhexol= 3-methyl bacteriohopane-30,31,32,33,34,35-hexol.

Methane oxidation	**Phylogeny**	**Main diagnostic lipids**
Anaerobic[1]	ANME-1	PMI, GDGTs (mainly GDGT-1 to GDGT-3 bound to diglycosidic polar headgroups)
	ANME-2	PMI, GDGTs & OH-GDGTs, archaeol & sn2-OH-archaeol (bound to polar headgroups e.g. PH or dihexose)
	ANME-3	Archaeol & sn2-OH-archaeol bound to e.g. PG or PH polar headgroups
	SRBs–Deltaproteobacteria (i.e. Desulfosarcina)	MAGEs & DAGEs, characteristic FAs (e.g. $iC_{15:0}$ & $aiC_{15:0}$, $C_{16:1\omega5}$, 10Me $C_{16:0}$)

Aerobic[2]	**Type I MOBs**–Gammaproteobacteria (i.e. Methylococcaceae)	BHPs: aminopentol & aminotetrol, 3-Me aminoBHPs, Characteristic FAs ($C_{14:0}$, $C_{16:0}$, $C_{16:1}$= $C_{16:1\omega8c}$ and $C_{16:1\omega5t}$), 3-Me hopanoids (e.g. diplopterol)
	Type II MOBs–Alphaproteobacteria (i.e. Methylocystaceae)	BHPs: aminotetrol Characteristic FAs ($C_{18:0}$, $C_{18:1}$= $C_{18:1\omega8c}$ & $C_{18:1\omega7c}$, $C_{18:2}$= $C_{18:2\omega7c,12c}$)
	Verrucomicrobia (i.e. Methylacidimicrobium)	BHPs: high aminotriol & early eluting aminotetrol, Characteristic FAs ($iC_{14:0}$, $aiC_{15:0}$, $iC_{18:0}$)
	'Ca. M. oxyfera' (intra-aerobic methane oxidation)	3-Me BHhexol, Characteristic FA ($10MeC_{16:1\Delta7}$)

[1](Pancost et al., 2001b; Schouten et al., 2003; Blumenberg et al., 2004; Niemann and Elvert, 2008; Rossel et al., 2011); [2](Bull et al., 2000; Crossman et al., 2005; Dedysh et al., 2007; Talbot and Farrimond, 2007; Coolen et al., 2008; Bodelier et al., 2009; Op den Camp et al., 2009; van Winden et al., 2012; Kool et al., 2012, 2014; Talbot et al., 2014).

1.3.1 Lipid biomarkers for anaerobic oxidation of methane

Sulfate-dependent AOM can be traced by using lipids which derive from ANME-archaea such as ^{13}C depleted 2,6,10,15,19-pentamethyleicosane (PMI), archaeol and hydroxyarchaeol as well as specific isoprenoidal or intact GDGTs (Table 1 and Fig. 3; structures 1–3) (Hinrichs et al., 1999; Pancost et al., 2000; Wakeham et al., 2004). ANME-1 are known to synthesize mainly GDGTs bound to diglycosidic polar headgroups, while ANME-2 species produce also high amounts of archaeol and hydroxy-archaeol next to GDGTs bound to diverse polar headgroups like dihexose (Blumenberg et al., 2004; Niemann and Elvert, 2008; Rossel et al., 2011; Kellermann et al., 2012). ANME-3 synthesize only archaeol and hydroxyarchaeol and no GDGTs (Niemann and Elvert, 2008; Rossel et al., 2011). ANMEs can live in consortia with SRBs, which also produce some specific lipid biomarkers depleted in ^{13}C such as non-isoprenoidal mono and dialkyl glycerol ethers (MAGEs and DAGEs) as well as some characteristic fatty acids (e.g. $iC_{15:0}$ & $aiC_{15:0}$, $C_{16:1\omega5}$, 10Me $C_{16:0}$) (e.g. Pancost et al., 2000, 2001a; Elvert et al., 2003; Blumenberg et al., 2004, 2006).

1.3.2 Lipid biomarkers for aerobic oxidation of methane

Aerobic methanotrophic bacteria synthesize several saturated and unsaturated phospholipid fatty acids (PLFAs) (Bowman et al., 1993; Bull et al., 2000;

Bodelier et al., 2009). Type I methanotrophs synthesize high abundances of saturated and specific unsaturated C_{16} PLFAs (Fig. 3; structures 4 and 4a) of which some species (i.e. *Methylomonas*) also produce C_{14} PLFAs, whereas Type II methanotrophs typically synthesize high amounts of saturated and specific unsaturated C_{18} PLFAs (Fig. 3; structures 5 and 5a) (Hanson and Hanson, 1996; Dedysh et al., 2007; Bodelier et al., 2009). *Verrucomicrobia* of the phylum *Methylacidimicrobium* have been shown to produce $C_{14:0}$, $C_{15:0}$ and $C_{18:0}$ PLFAs (Op den Camp et al., 2009). The intra-aerobic methanotroph '*Ca*. M oxyfera' produces the distinct $10MeC_{16:1\Delta7}$ PLFA (Kool et al., 2012). However, as PLFAs and fatty acids are not specific to MOBs, these lipids are only useful as biomarkers to trace aerobic oxidation of methane when combined with strongly depleted ^{13}C isotopic values. Furthermore, these lipids are generally not well preserved in the geological record, although sometimes fatty acids can be preserved as bound compounds in old sediments such as in the Miocene (Ahmed et al., 2001).

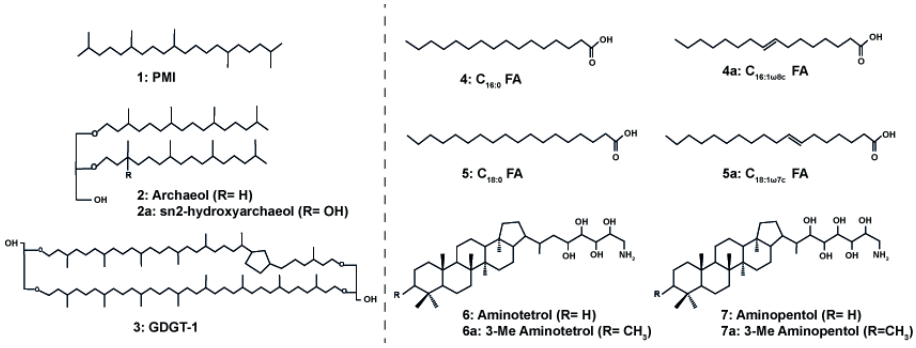


Figure 3. Examples of lipid biomarkers indicative for microbial mediated anaerobic oxidation of methane (left column) and aerobic oxidation of methane (right column).

Some Type I methanotrophs (*Methylococcus capsulatus* and *Methylosphaera hansonii*) have been shown to produce highly characteristic methylated steroids depleted in ^{13}C detected for example in Ace Lake in Antarctica or in the marine Haakon Mosby mud volcano (Schouten et al., 2000; Elvert and Niemann, 2008). However, these sterols have not been frequently reported in ancient sediments and steroids are mostly sourced by eukaryotes.

Hopanoids and their intact precursors bacteriohopanepolyols (BHPs) are pentacyclic triterpenoids that can contain methylations at C-2 or C-3 and unsaturations at C-6 and/or C-11 depending on their bacterial producers and preserve well over time (e.g. Rohmer et al., 1984; Neunlist and Rohmer, 1985; Zundel and Rohmer, 1985; Ourisson and Albrecht, 1992; Cvejic et al., 2000; Brocks et al., 2005). Large structural diversity has been observed in BHPs, especially in the extended polyfunctionalized side chains. Cultivated species of

MOBs have been shown to produce a group of distinct BHPs, namely the C-35 aminoBHPs containing an amine group at position C-35 and 3, 4 or 5 additional hydroxy groups on the side chain (Neunlist and Rohmer, 1985; Cvejic et al., 2000; Talbot and Farrimond, 2007; van Winden et al., 2012). Type I methanotrophs synthesize 35-aminobacteriohopane-30,31,32,33,34-pentol ('aminopentol'; Fig. 3; structure 7) and its unsaturated homologues, which have been detected in different terrestrial and marine environments like peats, geothermal sinters or sediments (Neunlist and Rohmer, 1985; Cvejic et al., 2000; Gibson et al., 2008; van Winden et al., 2012; Talbot et al., 2014). Furthermore, two additional amino-BHPs are produced by Type I and Type II methanotrophs as well as *Verrucomicrobia*: 35-aminobacteriohopane-31,32,33,34-tetrol ('aminotetrol'; Fig. 3; structure 6) and 35-aminobacteriohopane-32,33,34-triol ('aminotriol') (Talbot et al., 2001; Talbot and Farrimond, 2007; van Winden et al., 2012). Aminotetrol is also synthesized in minor amounts by some SRBs of the genus *Desulfovibrio*, whereas aminotriol is produced by a wide range of non-methanotrophic bacteria such as cyanobacteria or SRBs (Neunlist et al., 1985; Blumenberg et al., 2006; Talbot and Farrimond, 2007). Thus, aminotriol is less source-specific than aminotetrol and aminopentol to trace MOBs in the environment.

Methylation at the C-3 position in BHPs and aminoBHPs (Fig. 3; structures 6a and 7a) as well as in hopanoids (e.g. 3-methyl diplopterol) was previously suggested to be highly specific to Type I methanotrophs like *Methylococcus spp*. and *Methylocaldum spp*. (Neunlist and Rohmer, 1985; Cvejic et al., 2000). Furthermore, 3-methyl BHPs have also been identified in the intra-aerobic methanotroph '*Ca*. M. oxyfera', which synthesizes the specific 3-Me BHhexol (3-methyl bacteriohopane-30,31,32,33,34,35-hexol) as well as other more common BHPs (Kool, et al., 2014). However, the hpnR gene responsible for the C-3 methylation in hopanoids was only found in a limited number of methanotrophic bacteria as well as in other bacterial species (Welander and Summons, 2012). Nevertheless, the combination of 3-methylated hopanoids and depleted ^{13}C isotopic signals are often used to trace MOBs (Summons et al., 1994; Jahnke et al., 1999; Birgel and Peckmann, 2008). However, Type II methanotrophs and *Verrucomicrobia* assimilate both CH_4 and CO_2 and hence their PLFAs and hopanoids show less depleted ^{13}C values.

1.4 Scope and outline of this thesis

To this date, there are only a few unambiguous biomarkers for methanotrophs to trace microbes involved in the aerobic oxidation of methane, especially in the terrestrial realm. Paleorecords of lipid biomarkers mainly focused on marine

methane seep systems and much less on terrestrial environments despite soils being the biggest microbial sink for atmospheric methane. Therefore, there is an urgent need for the development of new reliable biomarkers to study the presence of methanotrophs, especially in modern and past terrestrial environments. This thesis aims to expand the current toolbox of characteristic biomarkers used to trace aerobic and anaerobic oxidation of methane and other hydrocarbon gasses.

In **Chapter 2**, the lipid biomarker inventory of the intra-aerobic methanotroph '*Candidatus* Methylomirabilis oxyfera' ('*Ca*. M. oxyfera') performing nitrite-dependent methane oxidation under anoxic conditions was investigated. The gas chromatography-amenable hopanoids were dominated by four demethylated hopanoids, of which only one (22,29,30-trisnorhopan-21-one) had been identified previously. We tentatively identified three novel hopanoids as 22,29,30-trisnorhopan-21-ol, 3-methyl-22,29,30-trisnorhopan-21-one and 3-methyl-22,29,30-trisnorhopan-21-ol as well as potential candidate genes responsible for the demethylation in hopanoids via bioinformatical analysis. A sensitive multiple reaction monitoring method was developed to trace these four trisnorhopanoids in complex environmental sample matrices. This new method successfully detected the new hopanoids in a peatland profile where '*Ca*. M. oxyfera' had been identified before via DNA-based techniques. Thus, the novel trisnorhopanoids may offer a new tool to detect nitrite-dependent methane oxidation in modern and past environments.

Chapter 3 explores for the first time the lipid inventory of intact and core archaeal lipids of ANME-2d archaea ('*Candidatus* Methanoperedens'), which synthesize archaeol and hydroxyarchaeol as well as isoprenoidal GDGTs and hydroxy-GDGTs, and occur with the intact polar headgroups dihexose, phosphatidyl-hexose and the rare monopentose. Moreover, SIP incubations using ^{13}C labeled bicarbonate and methane show that ANME-2d archaea assimilate mainly methane and to a lower extent dissolved inorganic carbon (DIC), which is different to ANME-1 archaea that assimilate solely DIC. The differences in the lipid inventory and the carbon assimilation pathway of ANME-2d archaea distinguishes them from other ANME species in microbial cultures and in environmental settings.

In **Chapter 4,** lipid biomarkers indicative for aerobic and anaerobic oxidation of methane were analyzed in three prominent terrestrial Sicilian mud volcanoes (MVs) (Maccalube di Aragona, Comitini and Santa Barbara). Freshly emitted mud breccias showed a high abundance of petroleum-derived hydrocarbons

compared to lipids from recent microbes in all three MVs. Saturated hydrocarbon biomarkers indicate that the mud breccias from the three MVs originate from a similar mix of mature marine carbonate or marl source strata and immature terrestrial strata with slightly more mature organic matter in the Maccalube di Aragona mud breccias. Only a few microbial lipid biomarkers such as fatty acids and 17β,21β(H) hopanoic acids were found in the MVs. However, they were not substantially depleted in ^{13}C, which would allow to attribute them to MOBs. In contrast, isoprenoidal GDGT distributions and a high GDGT-based Methane index ≥ 0.9 in the Maccalube di Aragona MV suggested the presence of ANME-archaea involved in anaerobic oxidation of methane at this MV. For the first time these results show characteristic GDGT patterns for anaerobic oxidation of methane involved ANME-archaea in a terrestrial MV system. The other two MVs Santa Barbara and Comitini only showed GDGT distribution indicative of Thaumarchaeota and not for microbes involved in the anaerobic oxidation of methane. Our results imply that microbial processes other than CH_4 oxidation may play a major role in terrestrial MVs in contrast to marine MVs, which host a higher abundance and diversity of ^{13}C-depleted lipid biomarkers indicative of aerobic and anaerobic oxidation of methane.

In Chapter 5 a novel method to detect underivatized bacteriohopanepolyol derivatives using UHPLC coupled to electrospray ionization-high resolution multi-stage MS ($HRMS^2$) was developed. This method provides mass spectra, which show typical fragmentation patterns of the BHP core, the functionalized side chain and the head group in case of composite BHPs. Common BHPs and their methylated and (di)unsaturated homologues as well as aminoBHPs and numerous composite BHPs were detected in bacterial cultures applying this novel technique. In addition, over 130 individual BHPs were identified in a soil near a methane seep (Fuoco di Censo everlasting fire in Sicily, Italy). This soil contains a complex distribution of known and novel adenosylhopanes typical for soils, of which the head group compositions of both adenosylhopane type-2 and type-3 could be identified for the first time. Furthermore, we detected a novel series of BHPs that are conjugated to an ethenolamine moiety. These novel ethenolamineBHPs as well as aminoBHPs were identified acylated to a range of fatty acids. This new analytical method allows simultaneous analysis of the full suite of IPLs including BHPs and is an essential step forward in the detection and application of novel BHPs as well as in environmental lipidomics.

In **Chapter 6** we apply the novel underivatized BHP method from **Chapter 5** to two terrestrial methane seep transects (Bissana and Censo seep) in Sicily, Italy.

16S rRNA gene sequencing showed the presence of Type I and Type II methanotrophs at both seeps, while methanotrophic *Verrucomicrobia* were only present at the Censo seep. The Bissana and Censo seeps show high relative abundances of the three aminoBHPs (aminotriol, aminotetrol and aminopentol), as well as the recently identified ethenolamine-BHPs (ethenolamine-BHT, -BHpentol and -BHhexol) and N-acyl-aminotriols at, and close to, the active seep sites. 3-methyl aminoBHPs were detected at the Bissana seep, but not at the Censo seep. AminoBHPs cluster together with the novel ethenolamineBHPs and N-acyl-aminotriols using principal component analysis (PCA) thus suggesting an origin from likely aerobic methanotrophs at both Sicilian seeps. Besides, the Censo seep showed high relative abundances of a novel late-eluting aminotriol isomer, which was also identified in the verrucomicrobial strain, *Methylacidimicrobium cyclopophantes* 3B, suggesting it as a new biomarker for *Verrucomicrobia*. Both transects showed increasing relative abundances of soil-marker BHPs (i.e. adenosylhopane) with increasing distance from the seeps. A new index (AminoBHP-index) was developed based on selected BHPs, which shows high values close to the seeps (≥ 0.4) and drop to < 0.2 at distances >3 m from the active seeps. The novel AminoBHP-index offers a new biomarker proxy to reconstruct present and past aerobic oxidation of methane in the terrestrial realm.

Chapter 7 describes the analysis of fatty acid biomarkers in soils from the natural everlasting fire of Fuoco di Censo (Censo seep) in Sicily, Italy. The soils at and close to the seep show high relative abundances of sequences of novel mycobacteria which are phylogenetically close to the *Mycobacterium simiae* complex and more distant from well-studied pathogenic mycobacteria like *M. tuberculosis*. Interestingly, the soils contain high abundances of mycocerosic acids (MAs), fatty acids typical for mycobacteria, which decrease with increasing distance from the seep. The major MA at this seep was tentatively identified as 2,4,6,8-tetramethyl tetracosanoic acid. Other unusual MAs with mid-chain methyl branches at positions C-12 and C-16 were also present but in lower abundances. The molecular structures of the Censo seep MAs appear to be different from those of well-studied mycobacteria like *M. tuberculosis* and have ^{13}C-depleted values, suggesting a direct or indirect utilization of gases like methane or ethane at the everlasting fire.

To assess the hypothesis that MAs originate from mycobacteria using seep gases, we performed stable isotope probing (SIP) using ^{13}C-labeled methane and ethane (C_2H_6) with the Censo seep soils in **Chapter 8**. SIP incubations with methane

and ethane show a ^{13}C label incorporation into regular and unsaturated C_{16} and C_{18} fatty acids, indicating Type I and Type II methanotrophs are able to not only oxidize methane but also ethane at the Censo seep. In contrast, ^{13}C label incorporation in MAs was only observed in the $^{13}C_2H_6$ incubations and not in those with $^{13}CH_4$. This suggests that mycobacteria present at the Censo seep oxidize ethane, and probably also higher gaseous hydrocarbons, but not methane, which is in agreement with previous studies on mycobacteria. Soil mycobacteria might be an important group of microbes oxidizing gaseous higher hydrocarbons in global terrestrial gas seeps and soils. Therefore, the structurally distinct MAs in combination with ^{13}C depleted signatures offer a new biomarker tool to trace hydrocarbon gas consuming mycobacteria in present and past environments.

In summary, the results of this thesis offer new lipid biomarker and proxies to trace methane as well as higher gaseous hydrocarbon oxidizing microbes, especially in terrestrial environments. The results of this study have therefore substantially expanded the toolbox for organic geochemists and may be tested and applied to trace past methane and higher gaseous hydrocarbon cycles in present and past environments in the future.

Chapter 2

Demethylated hopanoids in '*Ca*. Methylomirabilis oxyfera' as biomarkers for environmental nitrite-dependent methane oxidation

Nadine T. Smit, Darci Rush, Diana X. Sahonero-Canavesi, Monique Verweij, Olivia Rasigraf, Simon Guerrero Cruz, Mike S.M. Jetten, Jaap S. Sinninghe Damsté, Stefan Schouten

Published in Organic Geochemistry

Abstract

Hopanoids are lipids that are widespread in the bacterial domain and established molecular biomarkers in modern and paleo environments. In particular, the occurrence of ^{13}C-depleted 3-methylated hopanoids are characteristic of aerobic bacteria involved in methane oxidation. Previously the intra-aerobic methanotroph '*Candidatus* Methylomirabilis oxyfera' ('*Ca*. M. oxyfera'), which performs nitrite-dependent methane oxidation in anoxic environments, has been shown to synthesize bacteriohopanepolyols (BHPs) and their 3-methylated counterparts. However, since '*Ca*. M. oxyfera' is not utilizing methane as a carbon source, its biomass and lipids do not show the characteristic ^{13}C-depletion. Therefore, the detection of '*Ca*. M. oxyfera' in various environments is challenging, and still underexplored. Here, we re-investigated the hopanoid content of '*Ca*. M. oxyfera' bacteria using enrichment cultures. We found the GC-amenable hopanoids of '*Ca*. M. oxyfera' to be dominated by four demethylated hopanoids of which only one, 22,29,30-trisnorhopan-21-one, had been identified previously. The three novel hopanoids were tentatively identified as 22,29,30-trisnorhopan-21-ol, 3-methyl 22,29,30-trisnorhopan-21-one and 3-methyl 22,29,30-trisnorhopan-21-ol. These unique demethylated hopanoids are most likely biosynthesized directly by '*Ca*. M. oxyfera' bacteria and bioinformatical analysis of the '*Ca*. M. oxyfera' genome revealed potential candidate genes responsible for the demethylation of hopanoids. For the sensitive detection of the four trisnorhopanoid biomarkers in environmental samples, a multiple reaction monitoring (MRM) method was developed and used to successfully detect the trisnorhopanoids in a peatland where the presence of '*Ca*. M. oxyfera' had been confirmed previously by DNA-based analyses. These new biomarkers may be a novel tool to trace nitrite-dependent methane oxidation in various (past) environments.

Key words

demethylated hopanoids; *Candidatus* Methylomirabilis oxyfera; nitrite-dependent methane oxidation; trisnorhopanoids; 3-methyl 22,29,30-trisnorhopan-21-one; methylation

1. Introduction

Hopanoids are pentacyclic triterpenoids that modify the properties of membranes in micro-organisms and can serve as molecular biomarkers in modern and paleo environments (Ourisson et al., 1979; Ourisson and Albrecht, 1992). They are biosynthesized by different bacterial phyla and have been used to trace back the presence of these bacteria as far as the Proterozoic (Brocks et al., 2005). It was

previously believed that hopanoids indicate oxic environmental conditions, however this changed with the discovery of hopanoid biosynthesis in anaerobic bacteria such as fermentative bacteria (Neunlist et al., 1985; Llopiz et al., 1992), anammox bacteria (Sinninghe Damsté et al., 2004), sulfate reducing bacteria (SRB) (*Desulfovibrio*) (Blumenberg et al., 2006; Blumenberg et al., 2012) and bacteria of the *Geobacter* genus (Eickhoff et al., 2013). The molecular structures of hopanoids synthesized by bacteria are structurally very diverse and range from simple C_{30} hopanols, such as diplopterol, to extended C_{35} polyols known as bacteriohopanepolyols (BHPs). During diagenesis, these hopanoids can be abiotically or microbially altered (Rohmer and Ourisson, 1976; Albaiges and Albrecht, 1979) resulting in demethylation, aromatization, sulfurization, skeleton rearrangements and loss or transformation of functional groups. Both BHPs and derived diagenetic products are extensively used as lipid biomarkers for (specific) microbial communities and their metabolic role in modern and ancient ecosystems (e.g. Brocks and Pearson, 2005; Talbot et al., 2007).

One important example is the use of hopanoids to trace microbes involved in methane oxidation. BHPs with penta- and hexa-functionalized side chains containing an amine group at the C-35 position are characteristic of methane-oxidizing bacteria (MOB) from the alpha- and gammaproteobacterial classes (Talbot and Farrimond, 2007; van Winden et al., 2012). Another important structural feature in hopanoids is the methylation at the C-3 position of the A-ring. This methylation was first shown in acetic acid bacteria (Rohmer and Ourisson, 1976) and later in MOBs (Zundel and Rohmer, 1985). Subsequently, C-3 methylated hopanoids were used as a biomarker for the presence of aerobic methanotrophy in ancient sediments (Summons and Jahnke, 1992). However, Welander and Summons (2012) identified the *hpnR* gene responsible for 3-methylation in hopanoids and showed that it is widely distributed among various bacterial taxa, including those not involved in the methane cycle. Nevertheless, the origin of hopanoids from MOB can be inferred from their depletion in ^{13}C ($\delta^{13}C$ below ca. -40 ‰) as usually ^{13}C-depleted CH_4 is used as carbon source for their biomass (Summons et al., 1994; Jahnke et al., 1999). Consequently, C-3 methylated hopanoids can be used as an indicator for methanotrophy if accompanied by additional evidence of ^{13}C-depleted carbon isotope values.

An apparent exception to this ^{13}C depletion is the intra-aerobic methanotroph, '*Candidatus* Methylomirabilis oxyfera' ('*Ca.* M. oxyfera'), which belongs to the NC10 phylum and occurs in clusters with '*Ca.* Methanoperedens nitroreducens' archaea (ANME-2d) (Raghoebarsing et al., 2006; Ettwig et al., 2010; Vaksmaa et al., 2017). '*Ca.* M. oxyfera' lives under

anoxic conditions and reduces nitrite to nitric oxide, and subsequently is believed to produce its own intercellular oxygen from nitric oxide for the intra-aerobic oxidation of methane (Ettwig et al., 2010). Thus, '*Ca*. M. oxyfera' directly connects the methane and nitrogen cycle (Ettwig et al., 2010). In contrast to most other MOB lipids, '*Ca*. M. oxyfera' lipids were shown not to be depleted in ^{13}C from methane (Kool et al. 2012, 2014) as they autotrophically fix carbon dioxide via the Calvin cycle as their carbon source (Rasigraf et al., 2014). These lipids include 3-methyl hopanoids, first recognized through genomic analysis (Welander and Summons, 2012), such as the rare BHP-hexol and a novel BHP identified as 3-methyl-BHP-Hexol (Kool et al., 2014). Further studies of the lipid inventory of '*Ca*. M. oxyfera' showed that they produce rare methylated fatty acids such as $10MeC_{16:0}$ and $10MeC_{16:1\Delta7}$ (Kool et al., 2012). However, the use of fatty acids as biomarkers for nitrite-dependent methane oxidation is limited to more recent geological time periods because these lipids are not resistant to alteration processes (Atlas and Bartha, 1973; Wenger et al., 2002).

Here, we re-investigated the hopanoid content of *'Ca. M. oxyfera'* biomass but now focused on those amenable to gas chromatography (GC) analysis. These biomarkers were then compared to GC-amenable hopanoids detected in a core from a peatland where the occurrence of '*Ca.* M. oxyfera' had been confirmed previously.

2. Material and methods

2.1 '*Ca*. M. oxyfera' enrichment cultures

An enrichment culture of '*Ca*. M. oxyfera' bacteria was obtained from a bioreactor operated under the following conditions described previously by Ettwig et al. (2009). The reactor contained bright fluffy cell material which originated from active '*Ca*. M. oxyfera' bacteria, and black material from the bottom of the bioreactor that was most likely dead '*Ca*. M. oxyfera' biomass. The bioreactor contained 2/3 of '*Ca*. M. oxyfera' bacteria while the rest was composed of a mix of ANME-2d archaea and different uncharacterized bacteria phyla. The steady state concentration of nitrite in this bioreactor was kept at 0-5 mg/L.

Biomass from a second reactor enriched from Italian paddy field soils consisted of less '*Ca.* M. oxyfera' cell material (30%), more '*Ca*. Methanoperedens nitroreducens' (ANME-2d) archaea (30%) and other microbial communities (40%) as described previously (Vaksmaa et al., 2017).

2.2 Peatland core

A peatland core from the Brunsummerheide peatland, Netherlands (BRH) was sampled and divided into nine sections from 51 to 102 cm depth (i.e. 51-60 cm, from 60 to 95 cm, 95-102 cm) (Kool et al., 2012; Kool et al., 2014). Previous quantitative polymerase chain reaction (qPCR) analysis of '*Ca*. M. oxyfera' specific 16S rRNA genes revealed the peak of their abundance at 70 to 90 cm depth. In this zone both methane and nitrate showed depletion in pore water concentrations (Zhu et al., 2012).

2.3 Fluorescence *in-situ* hybridization

Fluorescence *in-situ* hybridization (FISH) of '*Ca*. M. oxyfera' enrichments from the two bioreactors was performed for abundance estimation. Liquid samples of 1.5 mL from an active enrichment culture were centrifuged for 5 min at 7000x*g* (Eppendorf™ 5424, Hamburg) and biomass pellets were washed 2x with 1 mL phosphate-buffered saline (PBS: 130 mM NaCl and 10 mM phosphate buffer pH 7.4). After washing, the samples were fixed with 4% paraformaldehyde at a 3:1 volume ratio to the sample at 4°C overnight. FISH was performed as described previously (Ettwig et al., 2008). Formamide concentration in the hybridization buffer was set to 35%. The following oligonucleotide probes were used: DAMOBACT-0193 (CGC TCG CCC CCT TTG GTC) specific for '*Ca*. M. oxyfera'-like bacteria; DAMOARCH-0641 (GGT CCC AAG CCT ACC AGT) specific for '*Ca*. M. nitroreducens'-like archaea; EUB 338 (S-D-Bact-0338-a-A-18) (Amann et al., 1990), EUB 338 II (S-D-Bact_0338-b-A-18) and EUB 338 III (S-D-Bact-0338-c-A-18) (Daims et al., 1999) for most bacteria; S-D-Arch-0915-a-A-20 for most archaea. Images were collected with a Zeiss Axioplan 2 epifluorescence microscope equipped with a CCD camera and processed with the Axiovision software package (Zeiss, Germany).

2.4 Genomic analysis

Identification of potential genes involved in hopanoid biosynthesis in '*Ca*. M. oxyfera' (NCBI: taxid671143) was carried out on the NCBI (National Center for Biotechnology Information) server with the PSI-BLAST algorithm (Schäffer et al., 2001) using characterized proteins from different organisms as queries (Suppl.Table 1). Operon analysis and visualization was performed by using the Gene Context Tool NG (http://biocomputo2.ibt.unam.mx/gctng/) and protein analysis was conducted with Uniprot (UniProt Consortium, 2018).

2.5 Hopanoid analyses

2.5.1 Extraction of hopanoids

Freeze dried biomass of the '*Ca*. M. oxyfera' enrichments as well as freeze dried BRH peat core material (Kool et al., 2012) were extracted using a modified Bligh and Dyer technique (Schouten et al., 2008; Bale et al., 2013). The samples were ultrasonically extracted with a solvent mixture containing methanol (MeOH), dichloromethane (DCM) and phosphate buffer (2: 1: 0.8, v: v: v). After sonication (10 min) and centrifugation the solvent layer was collected and the residue reextracted twice. The combined solvent layers were separated from the aqueous layer by adding additional DCM and phosphate buffer to achieve a ratio of MeOH, DCM and phosphate buffer (1: 1: 0.9, v: v: v). The separated organic bottom layer was removed and collected while the aqueous layer was washed two more times with DCM. The combined DCM layers were dried under a continuous flow of N_2.

Aliquots of the total lipid extracts (TLEs) (ca. 0.5- 3 mg) were base hydrolyzed (saponified), with 2 mL of 1N KOH in MeOH solution, refluxed for 1 hour at 130 °C, the pH adjusted to 5 with a 2N HCL in MeOH solution, separated with 2 mL bidistilled water and 2 mL DCM and washed with DCM two more times. The combined DCM layers were dried over a Na_2SO_4 column. Afterwards the saponified fraction was methylated with diazomethane in diethyl ether, filtered over a small silica column with ethyl acetate and silylated with pyridine and N, O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) at 60 °C for 20 min. These total lipid extracts (TLEs) were analyzed using GC and GC-mass spectrometry (MS).

2.5.2 Instrumental analysis

GC-MS analyses of the TLEs of different '*Ca*. M. oxyfera' biomass and BRH peat core material were performed on an Agilent Technologies GC/MS Triple Quad 7000C in full scan and multiple reaction monitoring (MRM) mode. The chromatography was carried out using a CP-Sil5 CB column (25 m x 0.32 mm with a film of 0.12 μm, Agilent Technologies) with helium as carrier gas (constant flow 2 mL/ min). For each sample 1 μL was injected on-column at 70 °C, temperature was increased at 20 °C/min to 130 °C, raised further by 4 °C/min to 320 °C, which was held for 20 min for full scan mode and for 10 min for the MRM analysis. GC-MS full scan analysis was conducted over a mass range of *m/z* 50 to 850, the gain on 3 and with a scan rate of 700 ms. Compounds were identified by comparison with those previously published (Zundel and Rohmer, 1985; Sinninghe Damsté et al., 2004; Kool et al., 2012).

A new GC-MS/MS MRM method was developed to detect hopanoid ketones and alcohols in environmental samples using transitions and collision energies shown in Table 1. The MS1 and MS2 resolution was set on wide, the

gain on 10 and the dwell time on 10 ms for all scans. As collision gases in the MRM scan mode He was admitted at a flow of 1.5 mL/ min and N_2 at a flow of 1 mL/ min.

Table 1. Specifications of the MRM method for the four '*Ca*. M. oxyfera' trisnorhopanoids, showing the target transition (m/z) and optimized collision energy (eV).

	Target transition (m/z)	Collision energy (eV)
trisnorhopan-21-one	384.3 → 191.1	10
trisnorhopan-21-ol	458.3 → 237.2	15
3-methyl trisnorhopan-21-one	398.4 → 205.2	10
3-methyl trisnorhopan-21-ol	472.4 → 237.2	15

3. Results and Discussion

3.1 Lipid analysis of '*Ca.* M. oxyfera' reveals novel demethylated hopanoids

The microbial community composition and microstructure of the '*Ca.* M. oxyfera' enrichment culture was assessed by FISH. Based on the average fluorescence signal, we estimated that approximately two thirds of the biomass granules were composed of bacteria related to '*Ca*. M. oxyfera' (Suppl. Fig. 1). The lipid composition of this biomass was dominated by a series of fatty acids in particular the $C_{16:0}$, $10MeC_{16:1\Delta7}$ and $10MeC_{16:0}$ fatty acids, as described previously (Kool et al., 2012). We also detected a series of hopanoids using mass chromatography m/z 191 and m/z 205 revealing diploptene, diplopterol, C_{32} 17β, 21β-hopanoic acid and some of their 3-methylated homologues (Fig. 1). These hopanoids, especially diploptene and diplopterol, are found in a diverse range of aerobic and anaerobic bacteria (Rohmer et al., 1984; Summons et al., 1994; Sinninghe Damsté et al., 2004; Blumenberg et al., 2006). Furthermore, these hopanoids have been found in various environments like oxygen minimum zones (Wakeham et al., 2007), marine hydrocarbon seeps (Niemann et al., 2006), lacustrine sediments (Innes et al., 1997), peatlands (van Winden et al., 2010) or soils (Crossman et al., 2005).

The GC-MS analysis also revealed the presence of four more minor peaks which we tentatively identified as hopanoids (Fig. 1 & 2). Of these four, 22,29,30-trisnorhopan-21-one (compound 1-TNHone in Fig. 1) was tentatively identified based on a published mass spectrum (Simoneit, 1977). This hopanoid was first identified in sediments from the Black Sea (Simoneit, 1977) and has since been found in various marine sediment and water column samples, where it is believed to be formed by aerobic diagenetic alteration processes of BHPs (Conte et al., 1998; Botz et al., 2007). Directly eluting after 22,29,30-

trisnorhopan-21-one is a compound (labelled 2 TNHol in Fig.1) with a diagnostic fragment of m/z 191 suggesting a non-A ring methylated hopanoid (Fig. 2B).

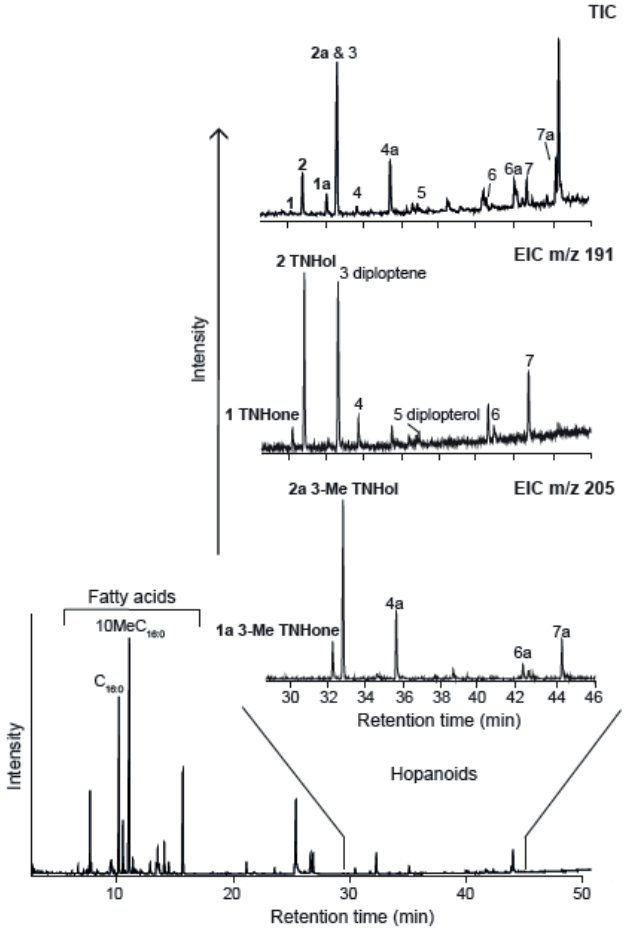


Figure 1. Gas chromatogram of saponified and silylated total lipid extract from '*Ca*. M. oxyfera' biomass. Enlarged inserts show the TIC (total ion chromatogram) of the hopanoids, the EIC (extracted ion chromatogram) of m/z 191 specific for hopanoids and m/z 205 specific for methylated-hopanoids indicating the main hopanoids in '*Ca*. M. oxyfera' biomass. 3 = diploptene, 5= diplopterol, 6= C_{32} 17β, 21β hopanoic acid, 7= C_{32} 17β, 21β-30-hydroxy hopanoic acid and one unknown hopanoid (4) as well as their methylated counterparts in the top EIC of m/z 205. New tentatively identified hopanoids for '*Ca*. M. oxyfera' are 1= 22,29,30-trisnorhopan-21-one (TNHone), 2= 22,29,30-trisnorhopan-21-ol (TNHol), 1a= 3-methyl 22,29,30-trisnorhopan-21-one (3-Me TNHone) and 2a= 3-methyl 22,29,30-trisnorhopan-21-ol (3-Me THNol).

The difference of 74 Da between the molecular ions of 22,29,30-trisnorhopan-21-one and unknown compound 2 TNHol, as well as ions at m/z 73 and 75, suggests the presence of a TMSi-derivatized alcohol. This is confirmed by the second diagnostic fragment of m/z 237, compared to the m/z 163 in the 22,29,30-trisnorhopan-21-one, suggesting that the alcohol is located at the D- or E-ring position. The fragment at m/z 367 indicates that the alcohol is likely at the E-ring position (Fig. 2B). This hopanoid was therefore tentatively identified as

22,29,30-trisnorhopan-21-ol. A similar doublet of peaks (compounds 1a 3-Me TNHone and 2a 3-Me TNHol, in Fig. 1) eluted after the 22,29,30-trisnorhopan-21-one and 22,29,30-trisnorhopan-21-ol, respectively. Both mass spectra (Fig. 2C and 2D) show the diagnostic fragment of m/z 205, which indicates a methylation at the A-ring. Furthermore, both have molecular ions which are 14 Da heavier (m/z 398 and m/z 472) than of 22,29,30-trisnorhopan-21-one and 22,29,30-trisnorhopan-21-ol, respectively (Fig. 2A and 2B, respectively). The first eluting compound 1a 3-Me TNHone shows the diagnostic ion of the D/E-ring fragmentation of 22,29,30-trisnorhopan-21-one (m/z 163; Fig 2C), whereas compound 2a 3-Me TNHol shows that of 22,29,30-trisnorhopan-21-ol (m/z 237; Fig. 2D).

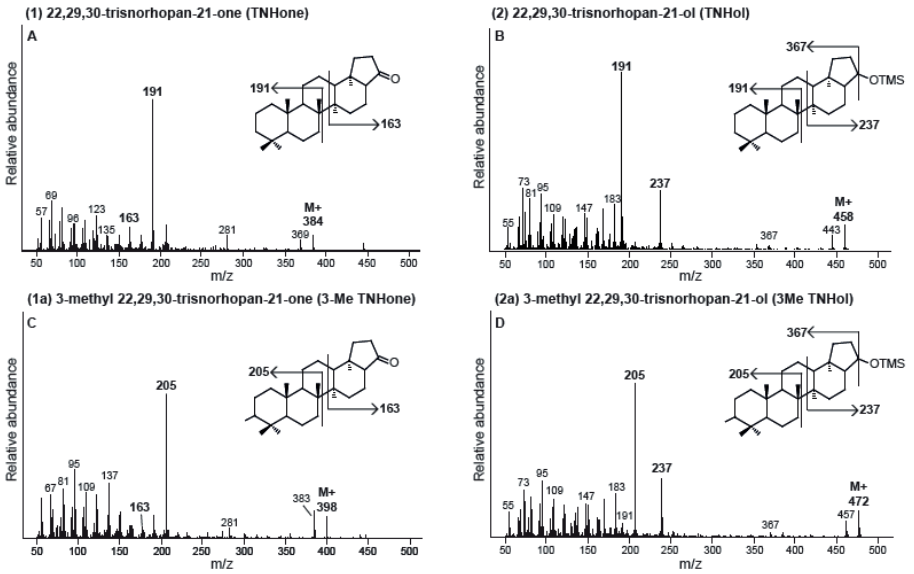


Figure 2. Mass spectra with fragmentation and molecular structures of trisnorhopanoids in '*Ca*. M. oxyfera' biomass. The mass spectra of 22,29,30-trisnorhopan-21-one was described previously by Simoneit, 1977.

Therefore, we tentatively identify these two compounds as the methylated counterparts of 22,29,30-trisnorhopan-21-one and 22,29,30-trisnorhopan-21-ol, respectively. Previous studies have shown the presence of the HpnR gene responsible for the 3-methylation of hopanoids in the '*Ca*. M. oxyfera' genome (Welander and Summons, 2012), as well as the presence of 3-methylated BHPs of BHP-hexol, BHP-pentol and BHP-tetrol in '*Ca*. M. oxyfera' biomass (Kool et al., 2014). Thus, we tentatively identified these two methylated compounds as 3-methyl 22,29,30-trisnorhopan-21-one (Fig. 2C) and 3-methyl 22,29,30-trisnorhopan-21-ol (Fig. 2D).

3.2 Origin of (3-methyl) 22,29,30-trisnorhopan-21-one and 22,29,30-trisnorhopan-21-ol

To the best of our knowledge, the three novel hopanoids (22,29,30-trisnorhopan-21-ol, 3-methyl 22,29,30-trisnorhopan-21-one and 3-methyl 22,29,30-trisnorhopan-21-ol) have not been identified before in any other bacterial cultures nor have they previously been detected in environmental samples. Since '*Ca*. M. oxyfera' biomass derives from highly enriched, but not pure cultures, it cannot be fully excluded that other bacteria in the bioreactor produce these hopanoids. However, this possibility seems to be rather unlikely because of the low amounts of other individual bacterial species (less than 5%) in the total enrichment culture, most of which are not known to produce 3-methylated hopanoids. Furthermore, we investigated another enrichment culture from a methanotrophic bioreactor fed with nitrate and methane which contained a lower proportion of '*Ca*. M. oxyfera' (30 %) and higher abundance of other species, mainly ANME-2D archaea (30%) (Vaksmaa et al., 2017). This showed a much lower abundance of the trisnorhopanoids (Fig. 3), suggesting that '*Ca*. M. oxyfera' is the likely source of these hopanoids.

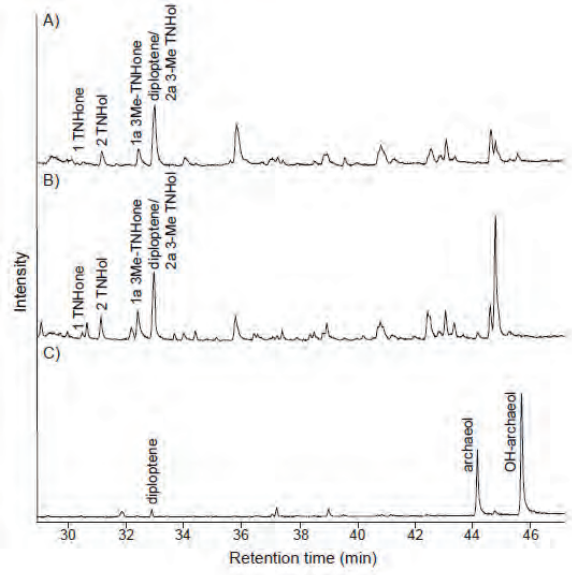


Figure 3. Gas chromatograms of saponified and silylated total lipid extracts from (A) '*Ca*. M. oxyfera' active biomass (bright fluffy cell material), (B) '*Ca*. M. oxyfera' dead biomass (black cell material from bioreactor bottom) and (C) '*Ca*. M. oxyfera' (30%) with enriched ANME-2D archaea (30%) from Italian paddy field soils. The four novel trisnorhopanoids 1= 22,29,30-trisnorhopan-21-one (TNHone), 2= 22,29,30-trisnorhopan-21-ol (TNHol), 1a= 3-methyl 22,29,30-trisnorhopan-21-one (3-Me TNHone) and 2a= 3-methyl 22,29,30-trisnorhopan-21-ol (3-Me THNol) as well as the typical archaeal lipids archaeol and OH-archaeol (sn2-hydroxy-archaeol) are annotated in the chromatograms.

Of the four trisnorhopanoids, only the 22,29,30 trisnorhopan-21-one has been reported previously, i.e. in anaerobic enrichment cultures of planctomycetes by Sinninghe Damsté et al., (2004) and in various environmental samples such as the anoxic brine-filled basins in the Mediterranean (Ten Haven et al., 1987), in coals and shales from Indonesia (Hoffmann et al., 1984) or anaerobic sediments (Pancost and Sinninghe Damsté, 2003). These studies have suggested that

trisnorhopan-21-one is a diagenetic or microbial altered product of BHPs such as bacteriohopanetetrol (BHT) and diplopterol formed under oxic conditions (Santos et al., 1994; Conte et al., 2003). Moreover, trisnorhopan-21-one is believed to be an indicator for high bio-productivity when found in the water column or sediments (Simoneit, 1977; Botz et al., 2007). However, since '*Ca.* M. oxyfera' biomass was incubated under strictly anaerobic conditions, aerobic degradation processes can be most likely excluded. We also analyzed black cell material from the bottom of the '*Ca.* M. oxyfera' enrichment bioreactor, which mostly represents dead cell material from '*Ca.* M. oxyfera'. This material had no elevated abundances of the novel trisnorhopanoids compared to living cell material (Fig. 3), suggesting that the trisnorhopanoids are not formed after cell death. Together, these results suggest that the trisnorhopanoids in '*Ca.* M. oxyfera' bacteria are more likely formed via a direct biosynthesis rather than by degradation processes in the bioreactor itself.

3.3 Potential biosynthesis of demethylated hopanoids in '*Ca.* M. oxyfera'

To further investigate the potential direct biosynthesis of the trisnorhopanoids, we explored biosynthetic genes that could be involved in the demethylation of the hopanoids (i.e., 22,29,30-trisnorhopan-21-one, 22,29,30-trisnorhopan-21-ol and their 3-methyl homologues) in '*Ca*. M. oxyfera'.

This demethylation of the C_{22}, C_{29} and C_{30} carbon atoms can either take place before the cyclization of squalene to a hopane or thereafter. Squalene biosynthesis in bacteria involves a three-step reaction encoded by the HpnD, HpnC and HpnE genes and are usually found in one operon (Pan et al., 2015). We found the gene that potentially codes for HpnD (DAMO_1512) in the genome of '*Ca*. M. oxyfera', but we did not find any homologous sequence for HpnC nor HpnE using BLAST searches (Suppl. Table 1). Previous studies suggested that these steps can alternatively be performed by farnesyl diphosphate farnesyl transferase (FdfT) activities as found in for example acidobacteria, gammaproteobacteria or cyanobacteria (Lee and Poulter, 2008; Ohtake et al., 2014; Sinninghe Damsté et al., 2017). The gene DAMO_2922 is annotated as a farnesyl-diphosphate farnesyltransferase and could be responsible for the squalene formation by this mechanism. Thus, a demethylated intermediate could potentially be formed by an alternative pathway to the three-step reaction by HpnD, HpnC and HpnE genes prior the cyclization reaction by the squalene synthetase DAMO_0045 gene and thus directly form the trisnorhopanoids.

Alternatively, and perhaps more likely, the demethylation of hopanoids could occur after the squalene cyclization in '*Ca*. M. oxyfera'. We speculate that such a process can be caused by oxidative activities similar to the removal of the

C-4 methyl from lanosterol to generate 4-desmethyllanosterol in bacteria catalyzed by SdmA and SdmB proteins (Lee et al., 2018). However, a BLAST search for SdmA and SdmB homologue proteins in the '*Ca*. M. oxyfera' genome did not indicate the presence of these two proteins, suggesting that the SdmA-SdmB proteins are only present in bacteria (e.g. *Methylococcus capsulatus*) that produce the C-4 demethylation in lanosterol. In eukaryotes, the C-4 sterol demethylation is performed by other proteins i.e. ERG 25, ERG26 and ERG27 (Bard et al., 1996; Gachotte et al., 1998; Gachotte et al., 1999). In this process, a C-4 sterol methyl oxidase (ERG 25), a C-3 sterol dehydrogenase (C-4 decarboxylase) (ERG26) and a 3-keto sterol reductase (ERG27) are involved in an iterative mechanism for the demethylation of sterols. However, no homolog sequence for ERG25 and ERG27 were found in the '*Ca*. M. oxyfera' genome, although a BLAST algorithm search using ERG26 (P53199) from the yeast *Saccharomyces cerevisiae* retrieved DAMO_0933 gene, which potentially codes for a NAD-dependent epimerase/dehydratase (Suppl. Table 1). DAMO_0933 is part of the predicted operon mox-DAMO_0930 which is composed of genes that codes for a protein of unknown function (DAMO_0930), a predicted glycosyltransferase (DAMO_0931), a potential methyltransferase (DAMO_0932), the NAD-dependent epimerase/dehydratase (DAMO_0933), a short-chain alcohol dehydrogenase (DAMO_0934) and a potential oxidoreductase (DAMO_0935). This combination of enzymes suggests that the trisnorhopanoids can be formed by a sequential oxidative process of diplopterol catalyzed by the genes in the DAMO_0930 operon (Suppl. Fig. 2). We propose that the oxidoreductase activity coded by the DAMO_0935 gene in combination with the dehydrogenase activity coded by the DAMO_0934 gene and the potential dehydratase coded by DAMO_0933 could be involved in a sequential demethylation process (Suppl. Fig. 2). Future work should involve the verification of the proposed enzymatic processes for the demethylation of hopanoids to better understand the mechanism involved in the synthesis of trisnorhopanoids in '*Ca*. M. oxyfera' and potentially other demethylated hopanoids in bacterial phyla.

3.4 Environmental occurrence of novel '*Ca.* M. oxyfera' trisnorhopanoids

To investigate whether the novel hopanoids detected in '*Ca*. M. oxyfera' are also present in the environment, we developed an MRM method to sensitively detect the novel trisnorhopanoids in environmental samples. For this, product ion scans of the molecular ions for 22,29,30-trisnorhopan-21-one, 22,29,30-trisnorhopan-21-ol and their 3-methylated counterparts were conducted using GC-MS/MS. The selection of the target ions was based on the most selective and abundant

ions in these product scans (Fig. 2). The collision energies were then optimized to provide an optimal abundance for the different transitions within the MRM method (Table 1). The resulting MRM chromatograms show substantially enhanced signal to noise ratio compared to extract ion chromatograms of m/z 191 and m/z 205 from full scan runs and thus provides a substantially enhanced sensitivity (Fig. 4A).

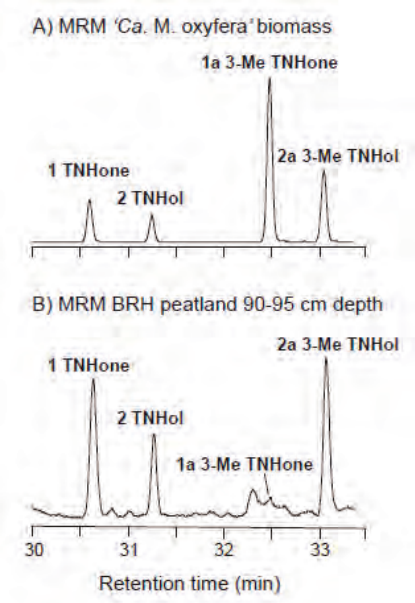


Figure 4. MRM (metastable reaction monitoring) chromatograms of the four distinct trisnorhopanoids (1= 22,29,30- trisnorhopan-21-one; 2= 22,29,30-trisnorhopan-21-ol; 1a= 3-methyl 22,29,30-trisnorhopan-21-one; and 2a= 3-methyl 22,29,30-trisnorhopan-21-ol) in (A) '*Ca*. M. oxyfera' biomass and (B) Brunsummerheide (BRH) peat from 90-95 cm depth.

Application of this MRM method to a peatland core from the Brunsummerheide, which previously revealed the presence of '*Ca.* M. oxyfera' based on DNA and fatty acids (Kool et al., 2012; Kool et al., 2014), showed the presence of the novel hopanoids (Fig. 4B). However, the peat showed a different relative abundance of the four hopanoids than the '*Ca.* M. oxyfera' biomass, i.e. with low abundance of the 3-methyl 22,29,30-trisnorhopan-21-one and a higher relative abundance of 22,29,30-trisnorhopan-21-one and 22,29,30-trisnorhopan-21-ol as well as a dominant 3-methyl 22,29,30-trisnorhopan-21-ol (Fig. 3B). Due to the lack of standards to quantify these hopanoids in the MRM transitions a hopanoid ratio was developed, i.e. the 3-Me trisnorhopanone / (3-Me trisnorhopanone + trisnorhopanone) ratio (Fig. 4D). This is based on the assumption that 22,29,30-trisnorhopan-21-one can also originate from bacteria other than '*Ca*. M. oxyfera' as it has been reported in many environments (Ten Haven et al., 1987; Botz et al., 2007), in contrast to its 3-methylated version. The peak of this hopanoid ratio is at 70-90 cm depth and corresponds to the peak in '*Ca*. M. oxyfera' cell abundance data based on qPCR 16S rRNA gene analysis as well as the specific $10MeC_{16:0}$ and $10MeC_{16:1}$ fatty acids (Fig. 4B; Kool et al., 2012). Thus, this ratio of trisnorhopanoids may be a complimentary biomarker tool for the detection of '*Ca*. M. oxyfera' in the environment. Moreover, these trisnorhopanoids may be better preserved over geological timescales than fatty acids and thus be a tool for the detection of past

intra-aerobic methanotrophy in paleo-environments, in contrast to $\delta^{13}C$ analysis of hopanoids (Kool et al., 2014; Rasigraf et al., 2014).

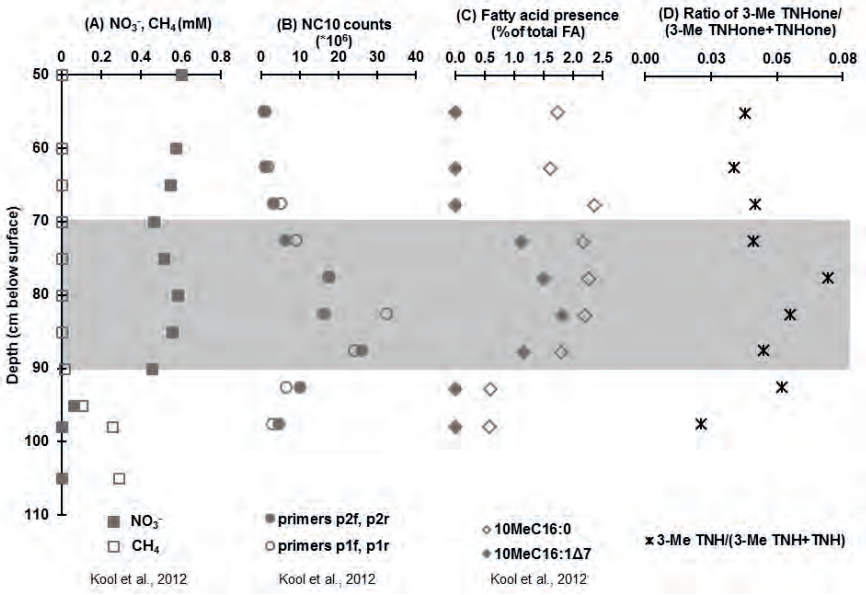


Figure 5. Depth profile of the Brunsummerheide (BRH) peatland core, showing (A) NO_3- and CH_4 concentrations in pore water; (B) NC10 bacteria 16S rRNA gene copy number abundance; (C) specific fatty acids (FAs) of 'Ca. M. oxyfera' (Kool et al., 2012); and (D) the novel hopanoid ratio 3-Me TNH/ (3-Me TNH+TNH) of 'Ca. M. oxyfera'. The maximum in RNA abundance of '*Ca*. M. oxyfera' is indicated with the grey area.

3.5 Implications

Our results may have consequences on how we view the origin and sources of demethylated hopanoids detected in present and past environments. Up to now demethylation processes of hopanoids were mainly attributed to diagenetic alteration processes like the microbial reworking of organic matter for bisnor- and trisnorhopanes or heavy biodegradation of petroleum for the presence of 25-norhopanes (Moldowan et al., 1984; Noble et al., 1985; Moldowan and McCaffrey, 1995). For example, the 18α(H)-22,29,30-trisnorhopane (Ts) and the 17α(H)-22,29,30-trisnorhopane (Tm), are thought to derive from diplopterol by a simple side-chain cleavage followed by an acid-catalyzed methyl shift (Seifert and Moldowan, 1978) and are often used to determine the source and the thermal maturity of petroleum (Peters et al., 2005; Seifert and Moldowan, 1978). It has been suggested that the 25-demethylation of 28,30-bisnorhopane to 25,28,30-trisnorhopane, common biomarker lipids in petroleum and organic rich sediments (e.g. Noble et al., 1985; Peters et al., 2005), occurs during advanced stages of petroleum biodegradation in which the 28,30-bisnorhopane probably derives from a hopene precursor having a 17(18) double bond (Rullkötter et al.,

1982; Volkman et al., 1983). A tentative identification of 28,30-bisnorho-17(18)-ene as a precursor has been made previously in sediments from the Gulf of California (Rullkötter et al., 1982). However, it was also suggested that 28,30-bisnorhopane and 25,28,30-trisnorhopane were biosynthesized directly since they are absent as sulfur bound hydrocarbons (Schoell et al., 1992; Schouten et al., 2001a).The ^{13}C-content of both compounds and their distribution strongly suggest that they are derived from anaerobic bacteria or those living under low oxygen conditions (Schoell et al., 1992; Schouten et al., 2001b). Our finding of demethylated hopanoids in '*Ca*. M. oxyfera' as well as those reported in the anaerobic planctomycetes (Sinninghe Damsté et al., 2004) now strongly suggest that demethylated hopanoids may be indeed biosynthesized as such and not be derived from (aerobic) diagenetic processes.

Conclusion

The intra-aerobic methanotroph '*Ca.* Methylomirabilis oxyfera' synthesizes a series of unique demethylated hopanoids identified as 22,29,30-trisnorhopan-21-one, 22,29,30-trisnorhopan-21-ol, 3-methyl 22,29,30-trisnorhopan-21-one and 3-methyl 22,29,30-trisnorhopan-21-ol. These unique hopanoids suggest a possible demethylation process of hopanoids by '*Ca.* M. oxyfera'. This is further supported by the finding of potential candidate genes responsible for the demethylation of hopanoids in the '*Ca*. M. oxyfera' genome. For the sensitive detection of these four hopanoids, an MRM method was developed and successfully applied to an environmental setting where '*Ca.* M. oxyfera' was previously detected by specific fatty acids and genomic analysis. The novel trisnorhopanoids and the developed MRM method offer a new tool to investigate the presence of '*Ca.* M. oxyfera' and nitrite-dependent methane oxidation in modern and past environments. Moreover, the identification of demethylated hopanoids in living bacterial biomass gives a new perspective on the origin and sources of demethylated hopanoids found in the geological record.

Author contributions

NTS, DR and SS planned research. MJM, OR and SCG provided biomass and FISH analysis data. NTS performed lipid analysis and NTS and SS interpreted the data. MV, SS and NTS developed MRM method. DXSC analyzed genome data for potential biosynthesis pathways of demethylated hopanoids. NTS wrote the paper with input from all authors.

Declaration of Competing Interest

The authors declared that there is no conflict of interest.

Acknowledgements

We thank Irene Rijpstra for laboratory assistance and helpful comments in compound identification as well as Laura Villanueva for discussions about hopanoid biosynthesis pathways. We also thank Annika Vaksmaa for providing methanotrophic ANME-2D/ 'Ca. M. oxyfera' biomass for comparison. We thank Martin Blumenberg and anonymous reviewer, as well as the Associate Editor for constructive comments. This study received funding from the Netherlands Earth System Science Center (NESSC) and Soehngen Institute for Anaerobic Microbiology (SIAM) through Gravitation grants (024.002.001 and 024.002.002) from the Dutch Ministry for Education, Culture and Science. DXSC received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement n◦ 694569 – MICROLIPIDS) and MSMJ by ERC AG ecomom 339880.

Supplemental material (can be found online at https://doi.org/10.1016/j.orggeochem.2019.07.008)

Suppl. Fig. 1 Fluorescent micrograph of the granular biomass sample from '*Ca*. M. oxyfera' enrichment culture.

Suppl. Fig. 2 Predicted operon in '*Ca*. M. oxyfera' potentially involved in hopanoid demethylation.

Suppl. Table 1 Potential genes involved in the synthesis and modification of hopanoids in '*Ca.* M. oxyfera'.

Chapter 3

Anaerobic methanotrophic archaea of the ANME-2d clade feature lipid composition that differs from other ANME archaea

Julia M. Kurth, Nadine T. Smit, Stefanie Berger, Stefan Schouten, Mike S.M. Jetten, Cornelia U. Welte

Published in FEMS Microbiology Ecology

Abstract

The anaerobic oxidation of methane (AOM) is a microbial process present in marine and freshwater environments. AOM is important for reducing the emission of the second most important greenhouse gas methane. In marine environments anaerobic methanotrophic archaea (ANME) are involved in sulfate-reducing AOM. In contrast, *Ca*. Methanoperedens of the ANME-2d cluster carries out nitrate AOM in freshwater ecosystems. Despite the importance of those organisms for AOM in non-marine environments little is known about their lipid composition or carbon sources. To close this gap, we analysed the lipid composition of ANME-2d archaea and found that they mainly synthesize archaeol and hydroxyarchaeol as well as different (hydroxy-) glycerol dialkyl glycerol tetraethers, albeit in much lower amounts. Abundant lipid headgroups were dihexose, monomethyl-phosphatidyl ethanolamine and phosphatidyl hexose. Moreover, a monopentose was detected as a lipid headgroup which is rare among microorganisms. Batch incubations with ^{13}C labelled bicarbonate and methane showed that methane is the main carbon source of ANME-2d archaea varying from ANME-1 archaea which primarily assimilate dissolved inorganic carbon (DIC). ANME-2d archaea also assimilate DIC, but to a lower extent than methane. The lipid characterization and analysis of the carbon source of *Ca*. Methanoperedens facilitates distinction between ANME-2d and other ANMEs.

1. Introduction

Methane is the second most important greenhouse gas on earth with an atmospheric methane budget of about 600 Tg per year (Conrad, 2009; Dean et al., 2018). About 69% of methane emission into the atmosphere is caused by methanogenic archaea (Conrad, 2009). On the other hand, aerobic and anaerobic methanotrophic microorganisms can oxidize methane back to carbon dioxide that is a 25-times less potent greenhouse gas than methane. The anaerobic oxidation of methane (AOM) is a microbial process present in marine and freshwater environments. AOM has first been described to be performed by a consortium of anaerobic methanotrophic archaea (ANME) and sulfate-reducing bacteria in microbial mats in the deep sea or in marine sediments (Hoehler et al., 1994; Hinrichs et al., 1999; Boetius et al., 2000; Hinrichs and Boetius, 2002; Orphan et al., 2002). ANME archaea are related to methanogens and oxidize methane by using the reverse methanogenesis pathway (Hallam et al., 2004; Arshad et al., 2015; McAnulty et al., 2017; Timmers et al., 2017). In addition to sulfate, also oxidized nitrogen compounds (Raghoebarsing et al., 2006; Ettwig et al., 2010; Haroon et al., 2013) as well as iron and manganese (Beal et al., 2009;

Ettwig et al., 2016; Cai et al., 2018) can be used as electron acceptors within the AOM process.

Anaerobic methanotrophic archaea can be assigned to three distinct clusters within the Euryarchaeota, ANME-1, ANME-2 and ANME-3, which are related to the orders Methanosarcinales and Methanomicrobiales (Knittel and Boetius, 2009). The phylogenetic distance between the groups is quite large (16S rRNA gene sequence identity between 75-92%) (Knittel and Boetius, 2009). Most analysed members of the three ANME clades have been described to perform sulfate driven AOM in marine environments (Pancost et al., 2001; Blumenberg et al., 2004; Niemann and Elvert, 2008; Rossel et al., 2008; Wegener et al., 2008; Kellermann et al., 2012). However, members of the ANME-2d cluster have not been found in consortia with sulfate reducers. Instead, ANME-2d archaea are the main players in nitrate-dependent AOM. Microorganisms conducting nitrate AOM have been enriched from anoxic freshwater sediments, digester sludge and rice paddies (Raghoebarsing et al., 2006; Hu et al., 2009; Arshad et al., 2015; Vaksmaa, Guerrero-Cruz, et al., 2017). Denitrifying AOM can either be conducted by a consortium of nitrate-reducing ANMEs, *Ca*. Methanoperedens sp., and nitrite reducing NC10 bacteria, *Ca.* Methylomirabilis sp. (Raghoebarsing et al., 2006; Haroon et al., 2013; Arshad et al., 2015) or by a consortium of those ANME archaea and anammox bacteria (Haroon et al., 2013). In those consortia *Ca*. Methylomirabilis sp. or anammox bacteria are important to reduce the toxic nitrite produced during nitrate AOM by *Ca*. Methanoperedens sp.

To understand the prevalence of anaerobic methane oxidation in past and present environments and identify the key players at different environmental sites, it is necessary to identify biomarkers for those organisms. As core lipids are much more stable than DNA over time, lipid biomarkers are a useful tool to trace microorganisms and therefore specific microbial processes back in time. Moreover, intact polar lipids are crucial to examine present microbial communities and to distinguish between different microorganisms (Ruetters et al., 2002; Sturt et al., 2004). Quite some information is available on core and intact polar lipids as well as on carbon assimilation in marine AOM consortia of ANME archaea and sulfate-reducing bacteria (Pancost et al., 2001; Blumenberg et al., 2004; Niemann and Elvert, 2008; Rossel et al., 2008; Wegener et al., 2008; Kellermann et al., 2012). In contrast, lipids from one of the main players in denitrifying AOM, *Ca*. Methanoperedens sp., have hardly been studied: a preliminary study on the lipids of a culture containing *Ca*. Methanoperedens sp. and *Ca*. Methylomirabilis oxyfera only detected sn2-hydroxyarchaeol as the dominant lipid of the archaeal partner (Raghoebarsing et al., 2006).

Besides the characterization of lipids in ANME archaea it is also pivotal to understand which carbon source those organisms use for biomass production. The main carbon assimilation pathway in methanogenic Euryarchaeota is the reductive acetyl-CoA pathway (Whitman, 1994; Berg et al., 2010). In this pathway a carbonyl group and a methyl group are combined to form acetyl-CoA. In archaea, acetyl-CoA is used for formation of membrane lipids via the isoprenoid compound geranylgeranylphosphate in the mevalonate pathway, although not all of the enzymes involved in this pathway are known with certainty (Koga and Morii, 2007; Matsumi et al., 2011). An ether bond is formed between the glycerol-1-phosphate backbone and the isoprenoid side chains. Subsequently cytidine-diphosphate is attached and finally the unsaturated isoprenoid side chains are reduced to form diphytanylglycerol diether, also known as archaeol (Matsumi et al., 2011).

The isotopic composition of lipids provides information on the carbon source used by the microorganism. The lipids of ANMEs involved in sulfate-driven AOM are usually strongly depleted in ^{13}C, with $\delta^{13}C$ values ranging from –70 to –130‰ (Elvert et al., 1999; Pancost et al., 2000; Niemann and Elvert, 2008). Such low $\delta^{13}C$ values of lipids have been explained by the assimilation of 13C-depleted methane carbon during methane uptake into biomass (Elvert et al., 1999; Hinrichs et al., 1999; Pancost et al., 2000; Orphan et al., 2002). Mixed assimilation of CH_4 and CO_2 has been reported for marine ANME-1, -2a, and -2b strains indicating that at least some ANME strains can use methane-derived carbon for biomass production (Wegener et al., 2008). However, for ANME-1 it has been shown that methane oxidation is decoupled from the assimilatory system and that CO_2-dependent autotrophy is the predominant mode of carbon fixation (Kellermann et al., 2012). In general, ANME archaea seem to be able to assimilate both, methane and dissolved inorganic carbon, and the preferred carbon source for assimilation might vary between the different ANME clusters.

In this study, we performed analysis of the lipids from ANME-2d archaea and compared these with previous studies about different ANME lipids. Moreover, we analysed the incorporation of ^{13}C-labelled methane and bicarbonate in lipids of these archaea to establish the carbon sources used for assimilation.

2. Materials and methods

2.1 ANME-2d bioreactor operation and sampling for lipid analysis

For lipid analysis of *Ca*. Methanoperedens sp. two different bioreactors were sampled. One bioreactor contained archaea belonging to the ANME-2d clade enriched from the Ooijpolder (NL) (Arshad et al., 2015; Berger et al., 2017) and

the other reactor ANME-2d archaea enriched from an Italian paddy field (Vaksmaa, Jetten, et al., 2017; Guerrero-Cruz et al., 2018). The anaerobic enrichment culture dominated by Ca. Methanoperedens sp. strain BLZ2 originating from the Ooijpolder (Berger et al., 2017) was maintained in an anaerobic 10 L sequencing batch reactor (30°C, pH 7.3 ± 0.1, stirred at 180 rpm). The mineral medium consisted of 0.16 g/L $MgSO_4$, 0.24 g/L $CaCl_2$ and 0.5 g/L KH_2PO_4. Trace elements and vitamins were supplied using stock solutions. 1000 × trace element stock solution: 1.35 g/L $FeCl_2 \times 4\ H_2O$, 0.1 g/L $MnCl_2 \times 4\ H_2O$, 0.024 g/L $CoCl_2 \times 6\ H_2O$, 0.1 g/L $CaCl_2 \times 2\ H_2O$, 0.1 g/L $ZnCl_2$, 0.025 g/L $CuCl_2 \times 2\ H_2O$, 0.01 g/L H_3BO_3, 0.024 g/L $Na_2MoO_4 \times 2\ H_2O$, 0.22 g/L $NiCl_2 \times 6$ H2O, 0.017 g/L Na_2SeO_3, 0.004 g/L $Na_2WO_4 \times 2\ H_2O$, 12.8 g/L nitrilotriacetic acid; 1000 × vitamin stock solution: 20 mg/L biotin, 20 mg/L folic acid, 100 mg/L pyridoxine-HCl, 50 mg/L thiamin-HCl × 2 H_2O, 50 mg/L riboflavin, 50 mg/L nicotinic acid, 50 mg/L D-Ca-pantothenate, 2 mg/L vitamin B12, 50 mg/L p-aminobenzoic acid, 50 mg/L lipoic acid. The medium supply was continuously sparged with Ar:CO_2 in a 95:5 ratio. Per day 30 mmol nitrate added to the medium were supplied to the bioreactor and were completely consumed. Methane was added by continuously sparging the reactor content with CH_4:CO_2 in a 95:5 ratio at a rate of 15 mL/min. The reactor was run with a medium turnover of 1.25 L per 12 h. A 5 min settling phase for retention of biomass preceded the removal of supernatant. Under these conditions nitrite was not detectable with a colorimetric test with a lower detection limit of 2 mg/L (MQuant test stripes, *Merck*, Darmstadt, Germany). Growth conditions and operation of the bioreactor containing ANME-2d archaea enriched from an Italian paddy field soil are described by Guerrero-Cruz et al., 2018 (Guerrero-Cruz et al., 2018). Sampled material from both reactors was centrifuged (10000 × g, 20 min, 4°C) and pellets were kept at −80°C until subsequent freeze-drying and following lipid and isotope analysis.

2.2 Analysis of the microbial community

For the Ooijpolder enrichment we performed whole genome metagenome sequencing. DNA extraction, library preparation and metagenome sequencing were performed as described before by Berger and co-workers (Berger et al., 2017). Quality-trimming, sequencing adapter removal and contaminant filtering of Illumina paired-end sequencing reads were performed using BBTools BBDuk 37.76 (BBMap - Bushnell B. - sourceforge.net/projects/bbmap/). All processed paired-end reads were assigned to a taxon using Kaiju 1.6.2 (Menzel et al., 2016) employing the NCBI BLAST non-redundant protein database (NCBI Resource Coordinators, 2016).

2.3 Batch cultivation of ANME-2d bioreactor cell material for ^{13}C-labelling experiment

60 ml bioreactor material of a *Ca*. Methanoperedens sp. BLZ2 culture enriched from the Ooijpolder (Arshad et al., 2015; Berger et al., 2017) were transferred with a syringe to a 120-ml serum bottle that had been made anoxic by flushing the closed bottle with argon gas for 10 min. Afterwards, the culture was purged with 90% argon and 10% CO_2 for 5 min. 2.5 mM bicarbonate and 18 ml methane (*Air Liquide*, Eindhoven, The Netherlands) were added. Except for the negative controls, either ^{13}C-labelled methane (99 atom%; *Isotec Inc*., Matheson Trigas Products Division) or ^{13}C-labelled bicarbonate (*Cambridge Isotope Laboratories Inc*., Tewksbury, USA) was used in the batch incubations. The bottles were incubated horizontally on a shaker at 30°C and 250 rpm for one or three days. All bottles contained sodium nitrate (0.6 mM) at the start of incubation and additional nitrate was added when the concentration in the bottles was close to 0, as estimated by MQuant (*Merck*, Darmstadt, Germany) test strips. The methane concentration in the headspace was measured twice a day by gas chromatography with a gas chromatograph (Hewlett Packard 5890a, Agilent Technologies, Santa Clara CA, US) equipped with a Porapaq Q 100/120 mesh and a thermal conductivity detector (TCD) using N_2 as carrier gas. Each measurement was performed by injection of 50 µl headspace gas with a gas-tight syringe. With this technique a decrease in methane concentration from ~24 % to ~20 % within three days of incubation was observed. After batch incubation, cell material was centrifuged (10000 × g, 20 min, 4°C) and pellets were kept at −80°C until subsequent freeze-drying and following lipid and isotope analysis. It has to be considered that cultures with ^{13}C-labelled bicarbonate also contained ^{12}C derived from CO_2. Having in mind that 10% of CO_2 were added to the culture (pCO_2= 0.1 atm) the CO_2 concentration in the solution was calculated to be about 3.36 mM by use of the equation $[CO_2]aq= pCO_2/kh$ (Henry constant (kh) is 29.76 atm/(mol/L) at 25°C). Therefore, it has to be assumed that about half of the carbon in the cultures where ^{13}C labelled bicarbonate was added derived from ^{12}C-CO_2 dissolved in the medium after gassing with a mixture of 10% CO_2/90% Ar gas.

2.4 Lipid extraction and analysis

2.4.1 Bligh and Dyer extraction

Bligh and Dyer extraction was used to extract Intact Polar Lipids (IPLs) from the ANME-2d enrichment. Lipids of freeze-dried biomass (between 20 and 70 mg) were extracted by a modified Bligh and Dyer method as described by Bale et al.

(Bale et al., 2013) using a mixture of methanol, dichloromethane and phosphate buffer at pH 7.4 (2:1:0.8 v/v/v). After ultrasonic extraction (10 min) and centrifugation the solvent layer was collected. The residue was re-extracted twice. The combined solvent layers were separated by adding additional DCM and phosphate buffer to achieve a ratio of MeOH, DCM and phosphate buffer (1:1:0.9 v/v/v). The separated organic DCM layer on the bottom was removed and collected while the aqueous layer was washed two more times with DCM. The combined DCM layer was evaporated under a continuous stream of nitrogen.

2.4.2 Acid hydrolysis

Head groups of archaeal lipids were removed using acid hydrolysis. Therefore, this method is used to analyse the core lipids of the microorganisms present in the ANME-2d enrichment. About 20 mg freeze-dried biomass was hydrolyzed with 2 ml of a 1.5 N HCl/MeOH solution and samples stirred for 2 h while heated at 130°C with a reflux system. After cooling, the pH was adjusted to pH 4-5 by adding 2 N KOH/MeOH solution. 2 ml DCM and 2 ml distilled H_2O were added. The DCM bottom layer was transferred to a new vial and the MeOH/H_2O layer washed twice with DCM. Combined DCM layers were dried over a Na_2SO_4 column and the solvent removed by evaporation under a stream of nitrogen.

2.4.3 BF_3 methylation and silylation

For gas chromatography (GC) analysis aliquots of the acid hydrolyzed samples were methylated using 0.5 ml of BF_3-methanol and react for 10 min at 60°C in an oven. 0.5 ml H_2O and 0.5 ml DCM were added to the heated mixture to separate the DCM and aqueous layers. Samples were mixed, centrifuged and the DCM-layer taken off and collected. The water layer was washed three more times with DCM. The combined DCM-layer were evaporated under a N_2 stream and water was removed by use of a $MgSO_4$ column. After dissolving the sample in ethyl acetate, the extract was cleaned over a small silicagel column and lipids were eluted with ethyl acetate. The extract was dried under N_2. For GC analysis, extracts (0.3 to 0.5 mg) were dissolved in 10 µl pyridine and 10 µl BSTFA. Samples were heated for 30 min at 60°C and afterwards diluted with ethyl acetate to 1 mg/ml.

2.4.4 GC-MS

This method was used to analyse bacterial fatty acids as well as archaeal diether lipids. Gas chromatography linked to mass spectrometry (GC-MS) was performed with a 7890B gas chromatography system (Agilent) connected to a

7000 GC/MS Triple Quad (Agilent). The gas chromatograph was equipped with a fused silica capillary column (25 m × 0.32 mm) coated with CP Sil-5 CB (0.12 µm film thickness) and a Flame Ionization Detector (FID). Helium was used as the carrier gas. The samples were injected manually at 70°C via an on-column injector. The oven temperature was programmed to a temperature increase from 70 to 130°C with 20°C/min and a further increase to 320°C with 4°C/min to, 320°C was held for 10 min. The mass range of the mass spectrometer was set to scan from m/z 50 to m/z 850.

2.4.5 GC-IRMS

To analyse the incorporation of ^{13}C labelled methane or bicarbonate into lipids present in the ANME-2d enrichment GC-IRMS was conducted. Gas chromatography coupled to isotope-ratio mass spectrometry (GC-IRMS) was performed on a TRACE 1310 Gas Chromatograph (Thermo Fisher Scientific) interfaced with a Scientific GC IsoLink II Conversion Unit connected to an IRMS DELTA V Advantage Isotope-ratio mass spectrometer (Thermo Fisher Scientific). The gas chromatograph was equipped with a fused silica capillary column (25 m × 0.32 mm) coated with CP Sil-5 CB (0.12 µm film thickness). Helium was used as the carrier gas. The acid hydrolyzed samples containing the core lipids were injected at 70°C via an on-column injector. The oven temperature was programmed to a temperature increase from 70 to 130°C with 20°C/min and a further increase to 320°C with 4°C/min, 320°C was held for 10 min. $\delta^{13}C$ values were corrected for methyl group derived from BF_3 methanol in case of carboxylic acid group (bacterial lipids) and methyl groups derived from BSTFA in case of hydroxyl groups (mainly archaeal lipids). Averaged $\delta^{13}C$ values are based on experimental triplicates, but not on analytical duplicates.

2.4.6 UHPLC-APCI-TOF-MS

UHPLC-APCI-TOF-MS analysis of the acid hydrolyzed lipids was conducted in order to obtain information about the tetraether lipids. About 0.4 to 0.8 mg of the acid hydrolyzed lipid extract was dissolved in a mixture of hexane/isopropanol 99:1. Extracts were filtered by use of a 0.45 µm, 4 mm diameter PTFE filter. About 2 mg per ml core lipid containing extracts were used for analysis by ultra-high performance liquid chromatography linked to time-of-flight atmospheric pressure chemical ionization mass spectrometry using a (UHPLC-APCI-TOFMS). Core lipid analysis was performed on an Agilent 1260 Infinity II UHPLC coupled to an Agilent 6230 TOF-MS. Separation was achieved on two UHPLC silica columns (BEH HILIC columns, 2.1 × 150 mm, 1.7 µm; Waters) in series maintained at 25˚C. The injection volume was 10 µl. Lipids were eluted

isocratically for 10 min with 10% B, followed by a linear gradient to 18% B in 15 min, then a linear gradient to 30% B in 25 min, then a linear gradient to 100% B in 30 min, and finally 100% B for 20 min, where A is hexane and B is hexane: isopropanol (9:1). Flow rate was 0.2 ml/min and pressure 400 bar. Total run time was 120 min with a 20 min re-equilibration. Settings of the ion source (APCI) were as followed: gas temperature 200°C, vaporizer 400°C, drying gas 6 l/min, nebulizer 60 psig. The lipids were identified using a positive ion mode (600–1400 m/z).

2.4.7 UHPLC-ESI-MS

This method was used to analyse the intact polar lipids (IPLs) of the ANME-2d enrichment. 0.3 to 0.7 mg of Bligh and Dyer sample was dissolved in an injection solvent composed of hexane/isopropanol/water (72:27:1;v/v/v) and filtered through a 0.45 µm regenerated cellulose filter with 4 mm diameter prior to analysis by ultra-high performance liquid chromatography linked to ion trap mass spectrometry using electrospray ionization (UHPLC-ESI-MS). UHPLC separation was conducted on an Agilent 1200 series UHPLC equipped with a YMC-Pack Diol-120-NP column (250 × 2.1 mm, 5 µm particle size) and a thermostated autoinjector, coupled to a Thermo LTQ XL linear ion trap with Ion Max source with electrospray ionization (ESI) probe (Thermo Scientific, Waltham, MA). Solvent A contained 79% hexane, 20% isopropanol, 0.12% formic acid, 0.04% ammonium and solvent B 88% isopropanol, 10 % H_2O, 0.12% formic acid, 0.04% ammonium. Lipids were eluted with 0% B for 1 min, a linear gradient from 0 to 34% B in 17 min, 34% B for 12 min, followed by a linear gradient to 65% B in 15 min, 65% B for 15 min and finally a linear gradient to 100% B in 15 min. The IPLs were identified using a positive ion mode (m/z 400–2000) and a collision energy of 35 eV.

3. Results and Discussion

3.1 Analysis of microbial community

We performed phylogenetic analysis of the microbial community in the ANME-2d enrichment originating from the Ooijpolder by assigning all processed paired-end reads to a taxon. Twenty-three percent of the reads were assigned to *Ca*. Methanoperedens sp. strain BLZ2, 33% to *Ca*. Methylomirabilis sp. acting as nitrite scavenger, 8% to Alphaproteobacteria, 6% to Gammaproteobacteria, 5% to Betaproteobacteria 1% to Deltaproteobacteria, 3% to Terrabacteria, 3% to Sphingobacteria and 1% to Planctobacteria. The only archaeon in the bioreactor was *Ca*. Methanoperedens sp. strain BLZ2. Analysis of the microbial community in the Italian paddy field ANME-2d enrichment using a bioreactor approach has

been described by Guerrero-Cruz et al., 2018. Metagenome sequencing of the DNA derived from this bioreactor revealed that 83% of 16S rRNA gene reads were assigned to *Ca*. Methanoperedens nitroreducens strain Verserenetto (Guerrero-Cruz et al., 2018). In this study we mainly show the results derived from lipid analysis of the Ca. Methanoperedens sp. BLZ2 enrichment originating from the Ooijpolder (Arshad et al., 2015; Berger et al., 2017). However, the results deriving from a *Ca*. Methanoperedens sp. Versenetto enrichment originating from Italian paddy field soil (Guerrero-Cruz et al., 2018) look very similar, indicating that our results are not dependent on the strain or the environment from which the strain was enriched.

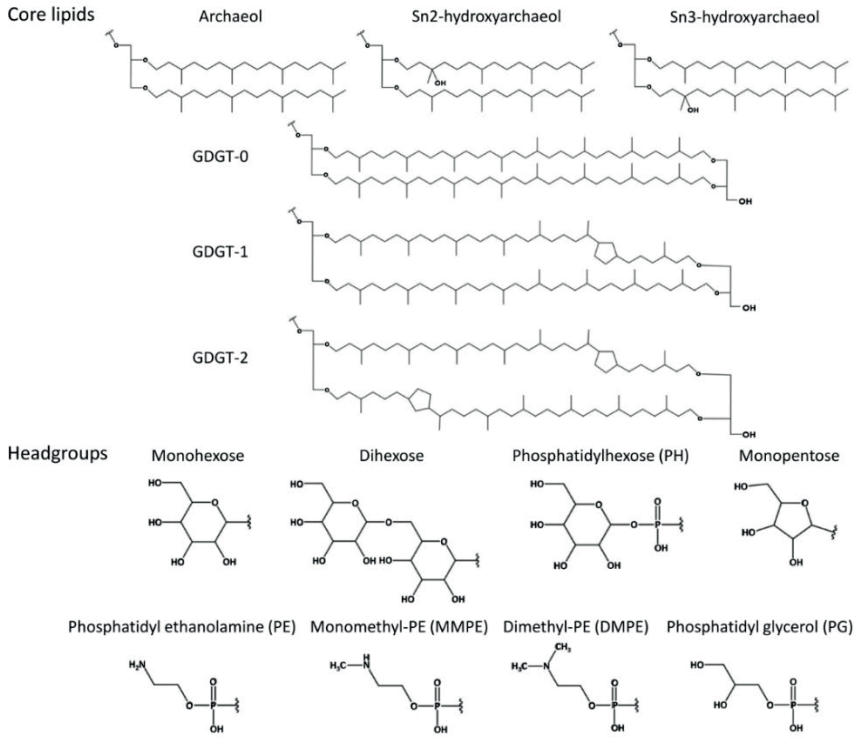


Figure 1. Structures of archaeal core lipids and headgroups. The main core lipids and headgroups that we found to be present in ANME-2d archaea are shown. The exact structure of monounsaturated archaeol as well as the pentose headgroup is not known. Intact polar lipids consist of a core lipid and one or two headgroups (only GDGTs can contain two headgroups). GDGT = glycerol dialkyl glycerol tetraethers.

3.2 Core lipids of *Ca*. Methanoperedens sp.

To analyse the lipids of ANME-2d archaea, biomass from a bioreactor containing *Ca*. Methanoperedens sp. BLZ2 enrichment was sampled and core lipid analysis with GC-MS and UHPLC-APCI-TOF-MS was performed. Shorter archaeal lipids like archaeol (Fig. 1) can be detected via GC-MS, whereas longer archaeal lipids like GDGTs can only be detected via UHPLC-APCI-TOF-MS.

GC analysis of the core lipids released by acid hydrolysis showed that of the microbial community harbored bacterial fatty acids and isoprenoidal archaeal lipids (Fig. 2). We detected the typical membrane lipids of *Ca*. Methylomirabilis sp., namely 10-methylhexadecanoic acid (10MeC16:0) and its monounsaturated variant (10MeC16:1Δ7) (Kool et al., 2012). The archaeal isoprenoids were predominantly composed of archaeol with lower amounts of sn2-hydroxyarchaeol and sn3-hydroxyarchaeol as well as two monounsaturated archaeols (Fig. 1) which could be identified with help of literature references (Nichols and Franzmann, 1992). Monounsaturated archaeol has already been described to be present in environmental samples containing lipids of archaea associated with anaerobic methane oxidation in marine environments (Pancost et al., 2001; Blumenberg et al., 2005). However, the monounsaturated archaeol might be produced from hydroxyarchaeol during acidic treatment of the lipids and therefore might not be part of native membrane lipid structures (Ekiel and Sprott, 1992). Alteration of the hydroxyarchaeol structure caused by different reaction conditions during lipid treatment has also been shown by Hinrichs and co-workers (Hinrichs et al., 2000). On the other hand, monounsaturated archaeols have also been described for *Halorubrum lacusprofundi* (Franzmann et al., 1988; Gibson et al., 2005), *Methanopyrus kandleri* (Nishihara et al., 2002), *Methanococcoides burtonii* (Nichols and Franzmann, 1992; Nichols et al., 1993), even if using mild alkaline hydrolysis instead of acidic treatment for lipid extraction (Nishihara et al., 2002).

One possibility to distinguish between the different ANME groups is the sn2-hydroxyarchaeol to archaeol proportion (Blumenberg et al., 2004). For ANME-1 this ratio is described to be 0-0.8, for marine ANME-2 1.1 to 5.5 and for ANME-3 within the range of ANME-2 (Blumenberg et al., 2004; Niemann et al., 2006; Nauhaus et al., 2007; Niemann and Elvert, 2008). In our study with the non-marine ANME-2d archaea we observed a sn2-hydroxyarchaeol to archaeol ratio of around 0.2. As mentioned before, the monounsaturated archaeol species might be an artefact of hydroxyarchaeol. If the monounsaturated archaeols are added to that of the sn2-hydroxyarchaeol abundance, the ratio would still be only around 0.3. That means that the hydroxyarchaeol to archaeol ratio of ANME-2d is more similar to that of ANME-1 archaea than to that of other ANME-2 or ANME-3 archaea. Members of the related methanogen order *Methanomicrobiales* only contain archaeol and GDGT-0 in their membranes but not hydroxyarchaeol (Koga et al., 1998). In the order *Methanosarcinales* the lipid composition varies between the different members. Most strains produce archaeol and hydroxyarchaeol, but the ratio differs and also the type of hydroxyarchaeol isomer varies. *Methanosarcinaceae* mainly produce the sn2-

isomer, whereas *Methanosaetaceae* mainly produce the rare sn3-isomer (Koga et al., 1998).

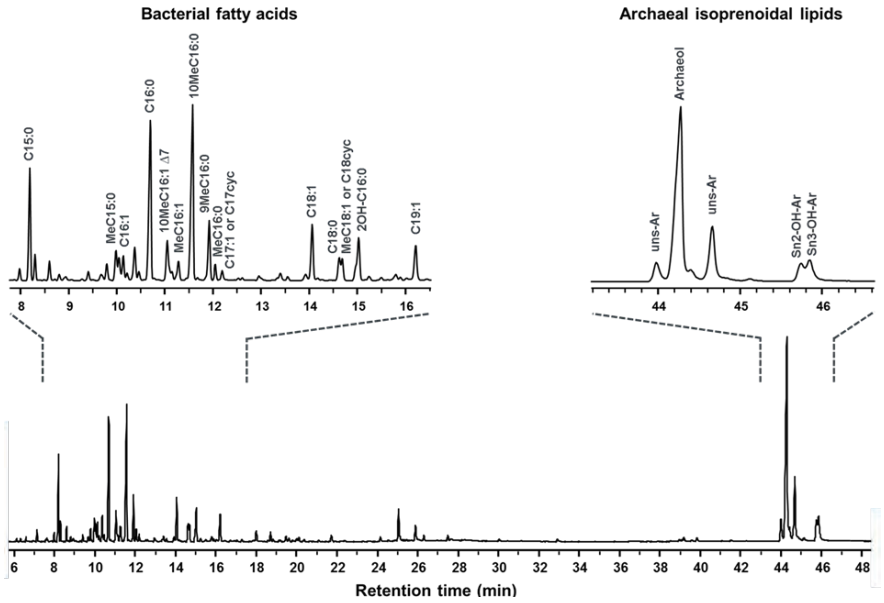


Figure 2. Gas chromatogram of core lipids released by acid hydrolysis from *Ca*. Methanoperedens sp. (ANME-2d) enrichment. The enriched biomass of ANME-2d originates from the Ooijpolder (NL) (Arshad et al., 2015). Enlarged inserts show the TIC (total ion chromatogram) of the bacterial and archaeal lipids. The most abundant compounds are annotated with their compound name and following abbreviations: Uns-Ar = monounsaturated archaeol, OH-Ar = hydroxyarchaeol.

Subsequently, UHPLC-APCI-TOF-MS analysis of the lipid extract was conducted in order to obtain information about the tetraether lipids (Fig. 1). This revealed that the relative abundance of archaeol was two times higher than that of glycerol dialkyl glycerol tetraethers (GDGTs) (Table 1). Moreover, several types of GDGTs were present in the enrichment. GDGTs contained either no (GDGT-0), one (GDGT-1) or two (GDGT-2) cyclopentane rings and about 64% of the GDGTs were hydroxylated (OH-GDGTs). The most abundant GDGTs were GDGT-0 with 6% and di-OH-GDGT-2 with 5% of total lipids. In conclusion, ANME-2d archaea synthesize various core-GDGTs, however archaeol and its homologues are the main isoprenoidal core-lipids in this enrichment.

Environmental samples from Mediterranean cold seeps with marine AOM associated archaea mainly contained GDGTs with 0 to 2 cyclopentane rings (Pancost et al., 2001). In a study on distinct compartments of AOM-driven carbonate reefs growing in the northwestern Black Sea, GDGTs could only be found in samples when ANME-1 archaea were present, but not when only

ANME-2 archaea were found, which led to the conclusion that ANME-2 archaea are not capable of synthesizing internally cyclized GDGT (Blumenberg et al., 2004). Later on in a study on methanotrophic consortia at cold methane seeps, samples associated with ANME-2c were shown to contain relatively high amounts of GDGTs (Elvert et al., 2005). In general, GDGTs are dominant in ANME-1 communities, while in marine ANME-2 and ANME-3 communities archaeol derivatives are most abundant (Niemann and Elvert, 2008; Rossel et al., 2008). GDGTs are not only present in marine archaea but are also produced by soil microbiota. For example, several members of the phylum Thaumarchaeota from soil environments have been shown to produce GDGTs with crenarchaeol as the major core GDGT similar to the aquatic thaumarchaeota (Sinninghe Damsté et al., 2012). Next to crenarchaol GDGTs with 0 to 4 cyclopentane moieties and GDGTs containing an additional hydroxyl group were detected in by Sinninghe Damsté and co-workers. In comparison to ANME-2d archaea, members of the related methanogen order *Methanomicrobiales* produce relatively high amounts of GDGT-0 (Koga et al., 1998; Schouten et al., 2012), whereas other members of the *Methanosarcinales* produce no or only minor amounts of GDGTs, mainly GDGT-0 (De Rosa and Gambacorta, 1988; Nichols and Franzmann, 1992; Schouten et al., 2012). Hydroxylated GDGTs seem to be relatively rare. In marine sediment samples the hydroxy-GDGT to total core GDGT ratio has been shown to vary between 1 and 8 % and the dihydroxy-GDGT to total core GDGT ratio is below 2% (Liu et al., 2012). Hydroxylated GDGTs have so far only been identified in the methanogenic Euryarchaeon *Methanothermococcus thermolithotrophicus* (Liu et al., 2012) and in several Thaumarchaeota (Schouten et al., 2012; Sinninghe Damsté et al., 2012). Until now only hydroxylated GDGTs with 0 to 2 cyclopentane rings have been found (Liu et al., 2012; Schouten et al., 2012; Sinninghe Damsté et al., 2012). In conclusion, the substantial abundance of GDGTs, especially hydroxylated GDGTs, is a distinct feature of the non-marine ANME archaeon *Ca*. Methanoperedens sp., as GDGTs have so far mainly been described for marine archaea like those of the ANME-1 clade as well as Thaumarchaeota present in marine and non-marine environments.

Comparing the results obtained in this study and lipid characterizations of marine ANMEs, it is apparent that the ratio of archaeol and GDGTs are distinctive in the different ANME groups: ANME-1 and partially ANME-2c contain substantial amounts of GDGTs and especially in ANME-1, GDGTs are the predominant membrane lipids (Niemann and Elvert, 2008). In contrast to ANME-1, but similar to other ANME-2 and ANME-3, we found that the dominating lipids in the membrane of clade ANME-2d archaea were archaeol

variants and not GDGTs. However, about 30% of the membrane lipids in ANME-2d archaea were GDGTs. Most strikingly, the majority of those GDGTs were hydroxylated, which is quite rare and has not been observed for other ANMEs so far.

3.3 Intact polar lipids of *Ca.* Methanoperedens sp.

Although intact polar lipids (IPLs) degrade more quickly than core lipids, IPLs are of higher taxonomic specificity and therefore useful to study especially present environments (Ruetters et al., 2002; Sturt et al., 2004). To identify IPLs of Ca. Methanoperedens archaea, UHPLC-ESI-MS was performed.

Table 1. Abundance of archaeol and GDGTs of *Ca*. Methanoperedens sp. Lipid extraction was performed in quadruplicates, error is given as standard deviation. For calculation of the relative abundance of archaeol also peaks derived from archaeol artefacts created during the experimental procedure were used.

Lipid	Relative abundance (%)	Relative abundance (%)
Archaeol	68 ± 5	68 ± 5
GDGT-0	6 ± 1	
GDGT-1	3 ± 1	
GDGT-2	2 ± 1	
OH-GDGT-1	3 ± 1	
OH-GDGT-2	1 ±1	32 ± 5
di-OH-GDGT-1	3 ± 2	
di-OH-GDGT-2	5 ± 2	
other GDGT-2 derivatives	9 ± 4	

The three most abundant archaeal IPLs detected were archaeol with a dihexose headgroup and hydroxyarchaeol with either a monomethyl phosphatidyl ethanolamine (MMPE) or a phosphatidyl hexose (PH) headgroup. Further headgroups attached to archaeol were monohexose, MMPE, dimethyl phosphatidyl ethanolamine (DMPE), phosphatidyl ethanolamine (PE) and PH. Next to MMPE and PH, hydroxyarchaeol based IPLs also contained dihexose, monopentose, DMPE, PE, pentose-MMPE, hexose-MMPE and pentose-PE (Fig. 1). Headgroups of GDGTs were found to be diphosphatidyl glycerol and dihexose phosphatidyl glycerol. The identification of a pentose as a headgroup of hydroyxarchaeol (mass loss of m/z 132) was unexpected. To our knowledge, this is the first description of a pentose as headgroup for microbial IPLs.

ANME-1 archaea mainly produce diglycosidic GDGTs, whereas lipids of marine ANME-2 and ANME-3 are dominated by phosphate-based polar derivatives of archaeol and hydroxyarchaeol (ANME-2: phospatidyl glycerol, phosphatidyl ethanolamine, phosphatidyl inositol, phosphatidyl serine, dihexose; ANME-3: phospatidyl glycerol, phosphatidyl inositol, phosphatidyl serine) (Rossel et al., 2008). Furthermore, marine ANME-2 archaea produce only minor amounts of GDGT-based IPLs and ANME-3 archaea produce no GDGT-based IPLs at all (Rossel et al., 2008). Intact GDGTs are assumed to be synthesised by head-to-head condensation of two intact archaeol molecules and substitution of the headgroups (De Rosa et al., 1980; Nishihara et al., 1989; Kellermann et al., 2016). In a study of Wegener and co-workers IPLs of marine ANME-2 and ANME-1 enrichments were analysed (Wegener et al., 2016). In the marine ANME-2 archaea mainly archaeol with a diglycosyl, monoglycosyl or phoshatidyl glycerol headgroup and hydroxyarchaeol with a monoglycosyl or phoshatidyl glycerol headgroup were detected. The ANME-1 enrichment contained mainly GDGTs with a diglycosyl headgroup (Wegener et al., 2016). IPLs of ANME-2d archaea can be distinguished from those of ANME-1 archaea by the prevalence of phosphate containing headgroups as well as archaeol and hydroxyarchaeol based IPLs. Furthermore, ANME-2d can be distinguished from other ANME-2 and ANME-3 archaea by the high abundance of dihexose as headgroup, the rare MMPE and DMPE headgroups and putatively also the pentose headgroup, which so far has not been described in the literature. In contrast to ANME-3 archaea, ANME-2d and marine ANME-2 archaea produce GDGT-based IPLs, albeit only in minor amounts.

In marine environments, a variety of archaeal lipids including those identified in ANME archaea can be found, e.g. those of the abundant Thaumarchaeota (GDGTs with hexose or phosphohexose headgroups, Sinninghe Damsté et al., 2012) and uncharacterized archaea (mainly GDGTs with glycosidic headgroups and in subsurface sediments also archaeol with glycosidic headgroups, Sturt et al., 2004; Lipp et al., 2008). In freshwater environments, IPLs of methanotrophic archaea have hardly been studied. Two studies on peat samples identified GDGTs with a glucose or glucuronosyl headgroup (Liu et al., 2010) and with a hexose-glycuronic acid, phosphohexose, or hexose-phosphoglycerol head group (Peterse et al., 2011). GDGTs with a hexose-phosphoglycerol head group were also identified in our study for ANME-2d archaea. Therefore, ANME-2d together with other archaea might be part of the peat microbial community based on the IPL profile. Using DNA biomarkers, most notably the 16S rRNA gene, *Ca*. Methanoperedens sp. has been detected in

various peat ecosystems (Cadillo-Quiroz et al., 2008; Zhang et al., 2008; Wang et al., 2019).

Other members of the order *Methanosarcinale*s mainly produce archaeol and hydroxyarchaeol with the headgroups glucose, phosphatidyl glycerol (only *Methanosarcinaceae*), phosphatidyl inositol, phosphatidyl ethanolamine, galactose (only *Methanosaetaceae*) (Koga et al., 1998). On the other hand, members of the related order *Methanomicrobiales* contain GDGT-0 and archaeol with the lipid headgroups glucose, galactose, phosphatidyl aminopentanetetrols, phosphatidyl glycerol (Koga et al., 1998). Therefore, IPLs from *Ca*. Methanoperedens sp. differ from methanogen IPLs by the high abundance of dihexose, MMPE and phosphatidyl hexose as lipid headgroup and the absence of the quite common headgroup phosphatidyl serine.

3.5 Incorporation of carbon derived from methane and bicarbonate in lipids

We were not only interested in characterizing the lipids of *Ca*. Methanoperedens sp., but also in answering the question if the organism incorporates carbon derived from methane or from dissolved inorganic carbon (DIC) in its lipids. In a labelling experiment from 2006 with an ANME-2d enrichment culture, incorporation of carbon derived from methane could hardly be detected for archaeal lipids (Raghoebarsing et al., 2006). To establish the carbon sources for *Ca.* Methanoperedens sp. we incubated the enrichment culture with ^{13}C labelled bicarbonate and methane and analysed lipid extracts for $\delta^{13}C$ depletion by GC-IRMS (Fig. 3).

Analysis of the isotopic composition of archaeol and its derivatives showed that ANME-2d archaea incorporated carbon derived from both methane and bicarbonate into their lipids. However, the main carbon source for biomass production seemed to be methane and not DIC as the former shows more label in the archaeal lipids. However, it has to be considered that the cultures to which ^{13}C labelled bicarbonate was added did not exclusively contain ^{13}C-DIC. About half of the DIC in the cultures derived from ^{12}C-CO_2 dissolved in the medium after gassing with a mixture of 10% CO_2/90% Argon gas (calculations in the methods part). Considering this, the $\delta^{13}C$ values of the archaeol isomers without ^{12}C-DIC in the incubations would vary most probably between -40 and -60‰. Nevertheless, the respective lipids were still quite depleted in $\delta^{13}C$ in comparison to the incubations with labelled methane (-2 to 12‰; 3 days incubation). Therefore, we concluded that mainly methane and not DIC is incorporated in the lipids of *Ca*. Methanoperedens sp.. Supporting this result, cultures containing marine ANME-1 and ANME-2 were shown to incorporate carbon derived from labelled methane into archaeol, monounsaturated archaeol and biphytanes

(Blumenberg et al., 2005). In another study it was found that ANME-1 archaea assimilated primarily inorganic carbon (Kellermann et al., 2012). Incubations with sediments containing ANME-1, 2a & 2b archaea showed that both, labelled methane and inorganic carbon, were incorporated into the archaeal lipids (Wegener et al., 2008).

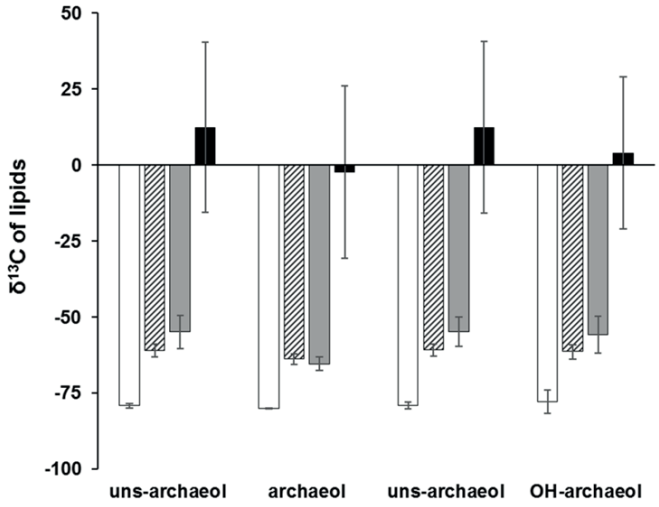


Figure 3. $\delta^{13}C$ values of *Ca.* Methanoperedens sp. lipids after batch cultivation with labelled bicarbonate or methane. ANME-2d reactor material originating from the Ooijpolder was incubated in anaerobic batch cultures with either ^{13}C labelled bicarbonate for three days (striped columns) or ^{13}C labelled methane for one (light grey columns) or three days (black columns). Controls contained only non-labelled carbon sources (white columns). Incubations were performed in triplicates, error bars = standard deviation. $\delta^{13}C$ values were obtained by analysing acid hydrolysed samples that only contained the core lipids by GC-IRMS. Peak identification was conducted with the help of GC-MS analysis of the same samples, showing that lipid extracts contained archaeol, hydroxyarchaeol and two monounsaturated archaeols (Fig. 1). Uns-archaeol = monounsaturated archaeol.

Incubations with freshwater sediments including ANME-2d archaea followed by RNA stable isotope probing demonstrated that those microbes mainly incorporated methane into their lipids but may have the capability of mixed assimilation of CH_4 and dissolved inorganic carbon (Weber et al., 2017). Our data confirmed that ANME-2d archaea are capable of mixed assimilation of CH_4 and DIC, but that methane is the preferred carbon source. In contrast to anaerobic methanotrophs, aerobic methanotrophs require oxygen for methane oxidation and the first step of methane oxidation to methanol is catalysed with the enzyme methane monooxygenase (Dalton, 1980). Also the nitrite-dependent intra-aerobic methanotroph *Ca*. Methylomirabilis sp. which is assumed to use nitric oxide to generate internal oxygen to oxidize methane encodes enzymes for the conventional aerobic methane oxidation pathway, including the monooxygenase (Ettwig et al., 2010). Aerobic methanotrophs have different pathways of carbon fixation: proteobacterial methanotrophs assimilate C1 compounds deriving from

methane oxidation via the ribulose monophosphate (RuMP) and/or the serine pathway (Dalton, 1980) while verrucomicrobial methanotrophs and NC10 bacteria like Ca. Methylomirabilis sp. use the Calvin-Benson-Bassham cycle, mainly assimilating dissolved inorganic carbon (Khadem et al., 2011; Rasigraf et al., 2014). As mentioned before, the reductive acetyl-CoA pathway is the main carbon assimilation pathway in methanogenic Euryarchaeota (Whitman, 1994; Berg et al., 2010) and most likely also in ANME-1 and ANME-2d archaea (Hallam et al., 2004; Haroon et al., 2013). In this pathway a carbonyl group and a methyl group are combined to form acetyl-CoA.

Table 2. Lipids of different ANME groups. For ANME-2d lipid analysis we used *Ca*. Methanoperedens sp. enriched bioreactor material. For the other ANME groups information was based on publication about the specific lipid characteristic (Blumenberg *et al*., 2004; Niemann and Elvert, 2008) or ^{13}C labelling experiments (Blumenberg *et al*., 2005; Wegener *et al*., 2008; Kellermann *et al*., 2012). GDGT: glycerol dialkyl glycerol tetraether, PE: phosphatidyl ethanolamine, MMPE: monomethyl phosphatidyl ethanolamine, DMPE: dimethyl phosphatidyl ethanolamine, PG: phosphatidyl glycerol, MH: monohexose, DH: dihexose, PH: phosphatidyl hexose, PC: phosphatidyl choline.

	ANME-1	**ANME-2a/b**	**ANME-2c**	**ANME-2d**	**ANME-3**
Environment	marine	marine	marine	**freshwater**	marine
Core lipids	GDGT	(OH-) archaeol	(OH-) archaeol, GDGTs	**(OH-) archaeol, (OH)-GDGTs**	(OH-) archaeol
Sn-2-OH-archaeol / archaeol ratio	0 - 0.8	1.1 - 5.5	1.1 - 5.5	**0.1 - 0.3**	1.1 - 5.5
IPLs	GDGT + dihexose	(OH-) Archaeol + PG, PE, PH, PS, dihexose	(OH-) Archaeol + PG, PE, PH, PS, Dihexose	**(OH-) archaeol + dihexose, hexose, pentose, PH, PE, MMPE, DMPE**	(OH-) Archaeol + PG, PH, PS
Main carbon source	DIC			**CH_4**	

We were able to show in this study, that ANME-2d archaea are capable of mixed assimilation of CH_4 and DIC, but preferably incorporate methane in their biomass. For this reason, it can be concluded that the C1 compounds required for the reductive acetyl-CoA pathway derive from oxidation of methane as well as from DIC, whereby methane is the primary carbon source. Although ANME-1 archaea are assumed to use the same carbon fixation pathway (Hallam et al., 2004), they have been shown to primarily assimilate inorganic carbon (Kellermann et al., 2012). Therefore, the type of carbon fixation pathway does not directly allow conclusions on the preferred carbon source used for carbon assimilation of a microorganism.

Conclusion

In this study, we analysed the lipids from the main player in nitrate AOM, Ca. Methanoperedens sp. We found several lipid characteristics that enable distinction between ANME-2d and other ANME groups (Table 2).

ANME-2d archaea therefore can be distinguished from ANME-1 by the higher ratio of archaeol and hydroxyarchaeol instead of GDGTs as well as phosphate containing headgroups. Furthermore, ANME-2d can be distinguished from other ANME-2 and ANME-3 archaea by the high abundance of dihexose as headgroup, the rare MMPE and DMPE headgroups and putatively also the pentose headgroup, which so far has not been described in the literature. The appearance of a monopentose as headgroup of ANME-2d lipids is an interesting observation and might be further analysed in the future.

In contrast to other ANME groups ANME-2d archaea have been shown to produce relatively rare hydroxylated GDGTs. ANME groups do not only differ in their membrane lipids itself, but also in the way they incorporate carbon into their biomass. For ANME-1 it has been shown that primarily carbon derived from DIC is incorporated into the lipids (Kellermann et al., 2012). In case of ANME-2d archaea, we were able to demonstrate that both, carbon derived from DIC and from methane, are incorporated into their lipids, with methane as the preferred carbon source.

Funding

CUW and MSMJ were supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek through the Soehngen Institute of Anaerobic Microbiology Gravitation Grant 024.002.002 and the Netherlands Earth System Science Center Gravitation Grant 024.002.001. MSMJ was supported by the European Research Council Advanced Grant Ecology of Anaerobic Methane Oxidising Microbes 339880. JK was supported by the Netherlands Earth System

Science Center Gravitation Grant 024.002.001 and the Deutsche Forschungs Gesellschaft Grant KU 3768/1-1. SB and CW were supported by the Nederlandse Organisatie voorWetenschappelijk Onderzoek through Grant ALWOP.293.

Acknowledgements

We thank Michel Koenen for IPL analysis, Ronald van Bommel for technical assistance with the GC-IRMS, Denise Dorhout and Monique Verweij for technical assistance with GC-MS and UHPLC systems. Moreover, we thank Annika Vaksmaa for supplying bioreactor material, Theo van Alen for technical assistance with metagenomics sequencing and Jeroen Frank for helping with metagenomics analysis.

Conflict of interest

The authors have no conflict of interest to declare

Chapter 4

Lipid biomarker insights into hydrocarbon sources and microbial communities in the terrestrial mud volcanoes of Sicily, Italy

Nadine T. Smit, Fausto Grassa, Anchelique Mets, Jaap S. Sinninghe Damsté, Stefan Schouten

Under review at Organic Geochemistry

Maccalube di Aragona mud volcano in Sicily, Italy

Comitini mud volcano in Sicily, Italy

Abstract

Mud volcanoes (MVs) are one of the most prominent geological sources of the greenhouse gas methane and release about 33Tg yr^{-1} of methane into the Earth's atmosphere, directly effecting global warming. However, the sources and thermal maturity of organic matter as well as microbial gas utilization processes have been rarely investigated in terrestrial MV systems. Here, three prominent terrestrial Sicilian MVs (Maccalube di Aragona, Comitini and Santa Barbara) were investigated to analyze the composition and origin of organic matter, microbial aerobic methane oxidation (AMO) and anaerobic methane oxidation (AOM) processes using lipid biomarker techniques. Lipid analysis was conducted on freshly emitted mud breccias which showed a high abundance of petroleum-derived hydrocarbons compared to lipids from extant microbes in all three MV sites. The *n*-alkane distribution as well as saturated hydrocarbons (e.g. CPI and Pr/Ph) and hopane indices (Ts/Tm and 22S/(22S+22R) isomerization) indicate that all mud breccias derive from a similar mix of mature marine carbonate or marl source strata and immature terrestrial strata. The Maccalube di Aragona mud breccias seem to originate from slightly more mature, possibly marine, source organic matter with an admixture of less immature material than the mud breccias from the other two MVs. Microbial biomarkers including saturated, and unsaturated short chain fatty acids (C_{14}, C_{16} and C_{18}), 10-methyl C_{16} fatty acid and 22R C_{31}, C_{32} and C_{33} 17β,21β(H)-hopanoic acids did not reveal depleted $\delta^{13}C$ isotopic signals attributable to methane utilizers and suggest a mixture of different bacterial sources. These results suggest no major abundance of at least AMO processes in the three Sicilian MVs probably due to limited oxygen availability in the mud breccias. However, mud breccias from the Maccalube di Aragona MV indicate the presence of ANME-archaea involved in AOM based on the relatively high abundance of GDGT-1 to -4 and a high GDGT-based Methane index $\geq$ 0.9, similar to marine AOM influenced environments. To the best of our knowledge, this is the first time that isoprenoidal GDGTs characteristic for AOM could be identified in a terrestrial MV system. In contrast, the Comitini and Santa Barbara MVs demonstrate signals of marine or terrestrial Thaumarchaea with high relative abundances of crenarchaeol and GDGT-0 as well as the presence of the crenarchaeol isomer, all probably originating from immature marine or terrestrial MV subsurface strata. Overall, the terrestrial Sicilian mud volcanoes reveal organic matter which derives from several source strata with different degrees of thermal maturity. Moreover, recent microbial lipids suggest the presence of AOM in the Maccalube di Aragona MV but not in the other two MVs implying that microbial

processes other than methane oxidation may play a more important role in these terrestrial MVs.

Keywords
terrestrial mud volcanoes; anaerobic methane oxidation (AOM); aerobic methane oxidation (AMO); source organic matter; biomarker indices; isoprenoidal-GDGTs; Methane index

1. Introduction

The greenhouse gas methane (CH_4) is an important contributor to natural and anthropogenic global climate changes in present and past environments (Reeburgh, 1996, 2007; Valentine and Reeburgh, 2000). A better understanding of the sources, sinks and chemical reaction pathways in methane cycling is important to constrain the impact of methane on global warming. Several natural marine and terrestrial sources of methane are known, e.g. biogenic methane produced in marine sediments, peatlands and thawing permafrost soils, as well as thermogenic derived methane from seeps and mud volcanos (Dimitrov, 2002; Etiope et al., 2009; O'Connor et al., 2010; Dean et al., 2018).

Mud volcanoes (MVs) are one of the main contributors of atmospheric methane releasing about 33 Tg methane annually from about 1800 prominent individual MV systems (Dimitrov, 2003; Kopf, 2003; Milkov, 2005). MVs are remarkable geological structures formed by semi-liquid and gas-enriched mud breccias, which derive from deeper sedimentary layers often connected to petroleum systems (Kopf et al., 2001; Dimitrov, 2003). Their formation is mainly driven by abnormally high pore-fluid pressures that can be caused by a combination of in-situ gas generation, high sedimentation and structural or tectonic compression. Besides, the morphologically structures of MVs are very diverse occurring from plano-conical shapes rising some hundred meters high to irregular, negative funnel-shapes, which range up to areas of 100 km^2 to only several hundred m^2 of size in submarine or terrestrial environments (Kopf et al., 2001; Dimitrov, 2002).

The MVs are gas- and often petroleum-rich environments that feature diverse biogeochemical processes fueling different microbial communities of bacteria and archaea (Stadnitskaia et al., 2005; Alain et al., 2006; Niemann et al., 2006a; Niemann et al., 2006b). Two of the most investigated processes are the aerobic (AMO) and anaerobic oxidation of methane (AOM), and the latter was previously found to be significant in marine MVs (Boetius et al., 2000; Pancost et al., 2000; Stadnitskaia et al., 2005; Niemann et al., 2006a; Niemann et al.,

2006b). In general, AMO is performed by type I and II methanotrophs (Gammaproteobacteria and Alphaproteobacteria) (Dedysh et al., 2007; Bodelier et al., 2009) or *Verrucomicrobia* (Op den Camp et al., 2009), while AOM is thought to be mediated by several clades of methanotrophic archaea (ANME type 1, 2 and 3), often in syntrophic consortia with bacteria such as sulfate-reducing bacteria (SRBs) (Hinrichs et al., 1999; Boetius et al., 2000; Pancost et al., 2000; Orphan et al., 2001). AMO can be traced using specific lipid biomarkers like ^{13}C-depleted unsaturated C_{16} and C_{18} fatty acids, hopanes and bacteriohopanepolyols (e.g. Elvert and Niemann, 2008; Bodelier et al., 2009; van Winden et al., 2012). Sulfate-dependent AOM can generally be traced by ^{13}C-depleted archaeal lipids like hydroxyarchaeol and glycerol dibiphytanyl tetraethers (GDGTs) as well as SRB characteristic fatty acids and dialkyl glycerol diethers (DAGE) with depleted ^{13}C contents (Pancost et al., 2000; Blumenberg et al., 2004; Niemann and Elvert, 2008; Rossel et al., 2011).

In contrast to marine MVs, only a few studies have examined the lipid biomarker inventory and microbial communities in terrestrial MVs (Alain et al., 2006; Chang et al., 2012; Cheng et al., 2012; Heller et al., 2012). These studies found similar biogeochemical processes as in marine MVs like AMO, AOM or methanogenesis but with lower abundances of characteristic lipid biomarkers than in marine MV systems (Alain et al., 2006; Chang et al., 2012; Heller et al., 2012). Biomarker and genetic data showed that methane oxidation mainly occurred at the surface and in biofilms at Northern Italian and Taiwanese MVs, while microbial genes associated with methanogenesis are generally more abundant in the deeper layers of the Taiwanese MVs (Chang et al., 2012; Cheng et al., 2012; Wrede et al., 2012). Furthermore, the terrestrial MVs in Romania and Northern Italy featured mainly lipid biomarkers typical for SRBs (e.g. various DAGE or 10-methyl C_{16} fatty acid), methanogenic Euryarchaea (archaeol) and minor abundances of hydroxyarchaeol characteristic for ANME-archaea performing AOM. However, these terrestrial MV systems showed less depleted ^{13}C lipid biomarkers then observed in typical sulfate-dependent AOM lipids next to low AOM rates in incubation experiments, probably due to the restricted abundance of sulfate in terrestrial MVs compared to marine MVs (Alain et al., 2006; Heller et al., 2012). Therefore, it seems likely that petroleum biodegradation and associated methanogenesis play a more important biogeochemical role than AOM in terrestrial MVs in contrast to marine MV systems.

High amounts mature organic matter are often emitted by MVs originating from deep subsurface reservoirs of up to 20 km of depths (Ivanov et

al., 1996; Dimitrov, 2002). Hence, the petroleum like hydrocarbons can serve as a carbon source for microbes involved in the biodegradation of oil as shown by numerous studies of petroleum seeps, for example, offshore in the Gulf of Mexico or Santa Barbara Basin (Sassen et al., 2004; Farwell et al., 2009; Schubotz et al., 2011). Besides, the analysis of certain hydrocarbons (e.g. *n*-alkanes, acyclic isoprenoids or hopanes) can give insight into the source organic matter and maturity degree (Seifert and Moldowan, 1978; Peters et al., 2005) of the erupted material and provide key information about the geological history of the area. However, to the best of our knowledge only a few studies on marine MVs (Stadnitskaia et al., 2007; López-Rodríguez et al., 2014) and none on terrestrial MVs explored the source organic matter and its thermal maturation history as they mostly focus on the gas composition or the microbial lipid biomarker inventory and microbial communities.

Here we study terrestrial MVs located in Sicily, Italy. Italian MVs release about 10^5 T yr^{-1} of methane into the atmosphere of which the Sicilian onshore MVs are responsible for 400 T yr^{-1} (Etiope et al., 2002, 2007). We investigated the biomarker composition of three major terrestrial MVs on Sicily named Maccalube di Aragona, Comitini and Santa Barbara. The biomarker results were used to gain insights into the sources and thermal maturity degree of organic matter in the Sicilian MVs as well as to identify microbial communities, in particular those involved in methane oxidation processes (AMO and AOM).

2. Material and methods

2.1 Geological setting and study area

More than half of the total number of global MVs are located along the Alpine Himalayan active tectonic belt where approximately 110 deep-sea MVs and 52 onshore MVs are situated on the Mediterranean Ridge with a high number of MVs located in Italy (Dimitrov, 2003; Etiope et al., 2009). Sicily is located along the boundary between the convergent African and European plates and is largely controlled by the compressional regime of the Maghreb-Apennine fold and thrust belt (Lickorish et al., 1999; Catalano et al., 2008). Three main geological complexes mark the collisional complex of Sicily and its offshore continuation (Fig.1): the Hyblean Foreland; the Caltanisetta Basin, an intensively faulted syntectonic accretion foredeep basin from Late Miocene to the Quaternary; and a complex chain composed of the Calabrian Arc and the Maghrebian thrust belt (Catalano et al., 2008; Madonia et al., 2011). Onshore Sicilian MVs occur over the accretionary wedge, which was developed in front

of the Sicilian-Maghrebian fold and thrust belt mainly in the Caltanisetta Basin. These MVs originate from clastic sediments deposited in thrust-top basins, which were progressively shortened and displaced during the Late Miocene to Pleistocene (Monaco and Tortorici, 1996; Lickorish et al., 1999). The MVs in Sicily are smaller than those typically occurring in other hydrocarbon-prone areas like Romania or Azerbaijan (Etiope et al., 2002, 2007) and feature infrequent hazardous paroxysmal eruptions (Bonini, 2009; Madonia et al., 2011). Some of these MVs form water pools of several meters in diameter, where gas actively bubbles out while others form cone shaped craters with bubbling fluid mud. In the Caltanisetta Basin, the emitted gas is comprised of a mix of thermogenic and biogenic methane. In contrast, the Paterno mud volcano near Mount Etna is emitting a major amount of carbon dioxide originating from igneous volcanic activity (Etiope et al., 2002, 2007; Grassa et al., 2004). Daily emission rates from single methane emitting vents in the different MV sites range from 0.01 to 6.8 kg per day (Etiope et al., 2002, 2007).

The first site investigated here is Maccalube di Aragona, (MDA; Fig. 1) the biggest MV of Italy (Etiope et al., 2007), located in Southern-Central Sicily next to the town of Aragona where the MV spreads over a large area of about 1.4 km² and releases a mix of thermogenic and biogenic methane into the atmosphere (Etiope et al., 2002). The emitted gas consists of 90–99% methane ($\delta^{13}C$ = –48‰) and minor amounts of other gasses such as CO_2, O_2, N_2 and H_2 (Grassa et al., 2004). The MDA site features different active gas and mud emitting vents including an active field with mud craters emitting fluid mud ranging from a few centimeters to half a meter height, 3 active bubbling mud ponds and a huge area covered with altered clay from past eruptions. Furthermore, the site is characterized by a clayey body, with marls (Miocene) and stone blocks of different composition and age (Cretaceous to Miocene) and clastic sediments of the Terravecchia Formation (Tortonian), with no soil coverage (Etiope et al., 2002).

North to the MDA site, another active MV was investigated, named Comitini (COM; Fig. 1). This site is located on a forested hillside, which features several small active vents combined with watery pools with visible mud fluid and gas bubbling and no conical shaped crater structures. The emitted gas consists of 91% methane ($\delta^{13}C$ = –45‰) and minor amounts of other gases like N_2, CO_2 and higher gaseous hydrocarbons (e.g. ethane and propane) (Heller, 2011; Cangemi and Madonia, 2014).

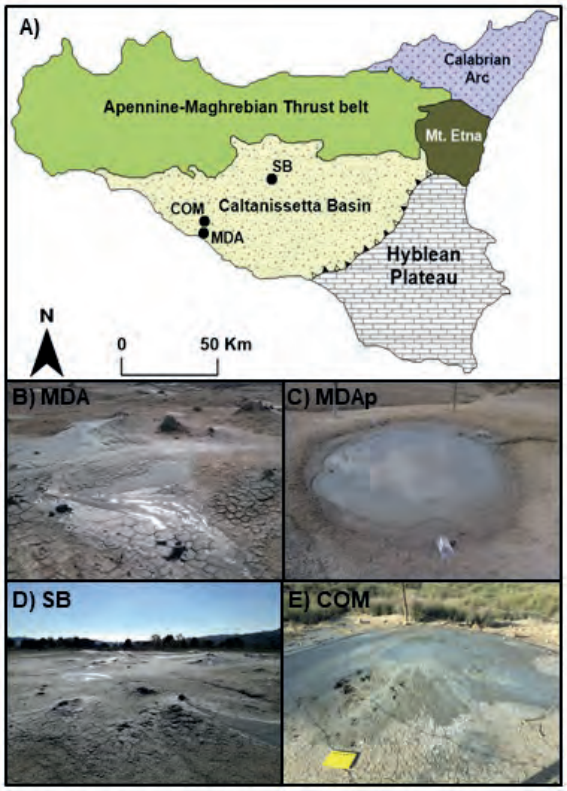


Figure 1. A) Map of the Sicily region showing the locations of the three investigated mud volcano sites Maccalube di Aragona (MDA), Santa Barbara (SB) and Comitini (COM). B–E) Images of the three MV sites, Italy. Key: B) = Maccalube di Aragona MV field (MDA), C) = Maccalube di Aragona mud pond (MDAp), D) = Santa Barbara MV field (SB) and E) = Comitini MV crater (COM).

Finally, in Central Sicily, a large MV field named Santa Barbara (SB) was investigated, located in the village of Santa Barbara in the north-easternmost part of the city of Caltanisetta (Fig. 1). The active Santa Barbara MV field covers an area of about 0.5 km^2 and hosts a multitude of small conical mud-emitting structures up to tens of centimeters high. The main emitted gas is thermogenic methane ($\delta^{13}C$ = –50‰) with minor amounts of other gases like CO_2, N_2 and higher gaseous hydrocarbons. Furthermore, the site is characterized by a clayey body, with elements from different geological formations which span from the Late Miocene to the Quaternary (Madonia et al., 2011; Ogniben, 1954). The bottom is represented by Tortonian clays, which are covered by evaporites (diatomites, limestones and gypsum interbedded with clay levels) from the Messinian salinity crisis in the Mediterranean, and marls and marly limestones from the Trubi Formation (Pliocene) (Ogniben, 1954; Madonia et al., 2011).

Deposits from these formations outcrop the SB field, with no soil coverage. First historical notifications of paroxysmal eruptions of this MV are reported back to the early 1820s by a Sicilian abbot (Bonini, 2009; Madonia et al., 2011).

2.2 Sample collection

Freshly emitted mud samples and mud samples from past eruptions (altered, dry mud) were recovered during field campaigns in October 2017 and October 2018 from the three different MV sites. At all three sites the fresh fluid mud breccias were directly taken from active mud and gas emitting mud craters (conical mud-emitting structures up to about 30 cm high), while altered dry muds from past eruptions were retrieved from several meters distance to the active area where plant vegetation started (Table 1). At MDA, two samples were also recovered from the edges of one of the active gas and mud bubbling ponds (Maccalube di Aragona mud pond = MDAp) up to 3 m in diameter which are located about 100 m Northwest from the main active mud volcano field. All mud breccias were collected from a horizon 5 to 10 cm below the surface with *in-situ* temperatures of about 20–25 °C at time of collection. The mud breccias were directly transferred into clean geochemical sampling bags and stored frozen at –20 °C until freeze drying and lipid extraction.

2.3 Extraction, saponification and derivatization

Ca. 10 g of freeze dried mud breccias from the three MV sites were extracted using a modified Bligh and Dyer method (Schouten et al., 2008; Bale et al., 2013). Mud samples were ultrasonically extracted with a solvent mixture of methanol (MeOH), dichloromethane (DCM) and phosphate buffer (2:1:0.8, v/v/v) for 10 min and centrifuged. Afterwards the solvent layer was collected, combined and the residues re-extracted twice. The combined solvent layers were separated from the aqueous layer by the addition of DCM and phosphate buffer to a ratio of MeOH, DCM and phosphate buffer (1:1:0.9, v/v/v). The bottom layer (DCM) was removed by pipetting and collected, while the remaining aqueous layer was washed twice with DCM. The combined DCM layers were dried under a continuous flow of N_2.

Aliquots of the total lipid extract (TLE) were saponified using 2 ml of a 1 N KOH MeOH solution and refluxed for 1 h at 130 °C. After cooling the pH was adjusted to 5 with a 2 N HCL MeOH solution. The layers were separated by 2 ml bidistilled water and 2 ml DCM, and the organic bottom layer was collected. The aqueous layer was washed two more times with DCM and the combined organic layers were dried first over a Na_2SO_4 column and then by N_2.

Table 1 Sample description as well as source and maturity related biomarker indices of the three Sicilian mud volcano sites Maccalube di Aragona, Santa Barbara and Comitini based on n-alkanes, acyclic isoprenoids and hopanes. Key: n.p.= not present.

Site	Sample ID	Sample description	Pr/Ph	CPI	OEP1	C_{29}nH/ C_{30}H	Ts/ (Ts+Tm)	22S C_{31}H	22S C_{32}H
Maccalube di Aragona	1MDA 1A (17)	Fresh wet mud, dark grey	1.04	0.97	1.03	0.73	0.35	0.57	0.59
Maccalube di Aragona	1MDA 3A (17)	Altered dry mud with vegetation, light grey	0.84	1.04	1.05	0.86	0.34	0.59	0.57
Maccalube di Aragona	MDA 1A (18)	Fresh wet mud, dark grey	0.71	0.92	1.00	0.62	0.48	0.56	0.60
Maccalube di Aragona mud pond	3MDAp 1A (17)	Fresh wet mud from bubbling mud pond, dark grey	0.58	1.34	1.10	0.67	0.35	0.56	0.59
Maccalube di Aragona mud pond	MDAp 4A (18)	Fresh wet mud from bubbling mud pond, dark grey	0.94	1.43	1.10	0.74	0.39	0.53	0.57
Santa Barbara at Caltanisetta	SB 1A (17)	Fresh wet mud, dark grey	1.00	1.51	1.10	0.72	0.31	0.50	0.55

Santa Barbara at Caltanisetta	SB 1C (17)	Fresh wet mud, dark grey	0.90	1.51	1.21	0.81	0.29	0.49	0.55
Santa Barbara at Caltanisetta	SB 4A (17)	Altered dry mud with vegetation, light grey	0.69	2.96	1.01	0.58	0.30	0.55	0.57
Santa Barbara at Caltanisetta	SB 1A (18)	Fresh wet mud, dark grey	1.58	1.45	1.11	0.69	0.30	0.59	0.57
Santa Barbara at Caltanisetta	SB 3A (18)	Fresh wet mud, dark grey	1.62	1.57	1.08	0.71	0.29	0.57	0.54
Comitini	COM 1A	Fresh wet mud, dark grey	n.p.	1.53	1.14	0.72	0.48	0.52	0.66
Comitini	COM 1B	Fresh wet mud, dark grey	0.67	1.54	1.20	0.65	0.37	0.50	0.50

Pr/Ph= pristane/phytane; CPI= $((C_{25}+C_{27}+C_{29}+C_{31}+C_{33})/(C_{24}+C_{26}+C_{28}+C_{30}+C_{32})+ (C_{25}+C_{27}+C_{29}+C_{31}+C_{33})/ (C_{24}+C_{26}+C_{28}+C_{30}+C_{32}))/2$; OEP1= $(C_{21}+6*C_{23}+C_{25})/(4*C_{22}+4*C_{24})$; $C_{29}NH/C_{30}H$= 17α(H),21β(H)-C_{29}-norhopane/17α(H),21β(H) 22S C_{30} hopane; Ts/(Ts+Tm)= 18α(H)-22,29,30-trisnorneohopane/(18α(H)-22,29,30-trisnorneohopane+17α(H)-22,29,30-trisnorhopane); 22S/(22S+22R)= 17α(H),21β(H) and 17β(H),21α(H) 22S and 22R of C_{31} and C_{32} homohopanes.

The saponified TLEs of the MV samples were subsequently esterified by heating with 0.5 ml of a boron trifluoride-methanol solution (BF_3 solution) for 10 min at 60 °C, before adding 0.5 ml bidistilled water and 0.5 ml DCM and mixing to separate the aqueous and organic layer. The organic layer was collected, and the water layer was extracted twice with DCM. The combined DCM layers were dried over an $MgSO_4$ column and afterwards eluted over a small silica gel column with ethyl acetate to remove polar high molecular weight molecules.

Extracts were subsequently separated using a small column packed with activated aluminum oxide into three fractions. The first fraction (saturated hydrocarbon fraction) was eluted with 4 column volumes of hexane, followed by a second fraction (fatty acid methyl ester fraction) eluted with 4 column volumes of DCM, and a third fraction (polar fraction) eluted with 3 column volumes of DCM/MeOH (1:1, v/v). The fractions were dried under a continuous flow of N_2 and aliquots of the polar fraction were silylated with pyridine and N, O bis(trimethylsilyl) trifluoroacetamide (BSTFA) at 60 °C for 20 min. Afterwards the fractions were analyzed using gas chromatography-mass spectrometry (GC-MS), GC- ion ratio mass spectrometry (GC-irMS) and ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS).

2.4 Instrumental analysis

2.4.1 Gas chromatography-mass spectrometry (GC-MS) and gas chromatography-isotope ratio mass spectrometry (GC-irMS)

GC-MS was performed using an Agilent Technologies GC-MSD in full scan mode with a scan time of 700 ms and a mass range of *m/z* 50 to 850. GC-IRMS was carried out with a Thermo Scientific Trace 1310 with a GC-Isolink II, a ConFlo IV and a Delta Advantage IRMS. GC was performed on a CP-Sil 5 column (25 m x 0.32 mm with a film thickness of 0.12 μm) with He as carrier gas (constant flow 2 ml min^{-1}). The BF_3 methylated samples (dissolved in ethyl acetate) were on-column injected at 70 ºC and subsequently, the oven was programmed to 130 ºC at 20 ºC min^{-1}, and then at 4 ºC min^{-1} to 320 ºC, which was held for 10 min. Stable carbon isotope ratios are reported in delta-notation against VPDB ^{13}C standard. Values were determined by triplicate analysis, corrected for the BF_3-methanol methylation and results averaged to a mean value.

2.4.2 Ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS)

UHPLC-MS was carried out on an Agilent 1260 UHPLC coupled to a 6130 quadrupole MSC in selected ion monitoring mode following a modified method of (Hopmans et al., 2016). The UHPLC was performed on two BEH HILIC silica columns (2.1 x 150 mm; 1.7 μm, Waters) in series, equipped with a 2.1 x 5 mm pre-column of the same material as the main columns and operated at 30 °C. Isocratic elution was used for GDGT separation, starting with 90% of eluent A (hexane) and 10% of eluent B (hexane: isopropanol, 9:1) for 10 min, followed by 70% eluent A and 30% eluent B for 40 min, then a linear gradient to 100% eluent B in 30 min and 100% eluent B for 20 min at a flow rate of 0.2 ml/min. Ionization was achieved by atmospheric pressure chemical ionization (APCI) with following settings: probe heater temperature, 350 °C; sheath gas (N_2) pressure, 50 AU (arbitrary units); auxiliary gas (N_2) pressure, 5 AU; spray current, 5 μA; capillary temperature, 275 °C; S-lens, 100 V. GDGTs and archaeol were identified by detecting the $[M+H]^+$ ions in selected ion monitoring (SIM) mode for *m/z* 651.6, 653.6, 743.8, 1286.3, 1288.3, 1290.3, 1292.3, 1294.3, 1296.3, 1298.3, 1300.3, 1302.3, 1304.3, 13016.3, 1330.3 and 1344.3.

3. Results

3.1 Distribution of hydrocarbons

Analysis of the saturated hydrocarbon fractions of the mud breccias from the three different MVs showed that they are dominated by saturated *n*-alkanes (C_{16} to C_{35}), pristane and phytane as well as a series of hopanes (Fig. 2 and Table 1). Furthermore, the presence of an unresolved complex mixture (UCM) was observed in all samples. Below, we describe the distributions of the different saturated hydrocarbon classes for representative samples of the three MVs in more detail.

3.1.1 N-alkanes and acyclic isoprenoids

The fresh mud breccia MDA 1A taken from a mud cone from the Maccalube di Aragona MV, has a unimodal distribution of *n*-alkanes maximizing around C_{22}/C_{24} and a weak odd over even carbon number predominance (Fig. 2A) as well as a high relative abundance of pristane and phytane. In contrast, the MDAp 4A sample (Fig. 2B) from the same MV site but from a mud pond reveals a bimodal distribution with a strong odd over even predominance of the longer chain *n*-alkanes (C_{25} to C_{31}). The shorter chain *n*-alkanes, maximize around C_{19} with high relative abundances of pristane and phytane. A multimodal distribution

of saturated *n*-alkanes is visible in the chromatograms of freshly emitted mud breccias from the Santa Barbara MV (SB 1C; Fig. 2C) and Comitini MV (COM 1B; Fig. 2D) with dominating C_{25} to C_{31} *n*-alkanes featuring an odd over even carbon number predominance. Furthermore, SB 1C and COM 1B also contain high relative abundances of pristane and phytane.

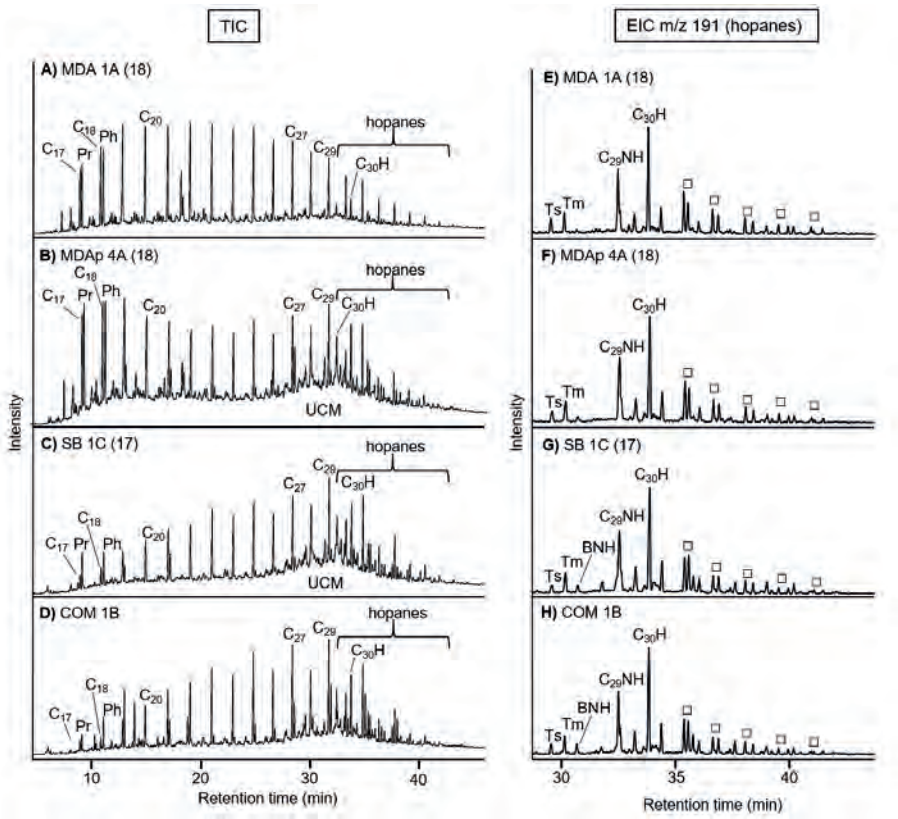


Figure 2. Total ion chromatograms (TIC) (A–D) and mass chromatograms of m/z 191 (E–H) of the saturated hydrocarbon fractions of fresh mud breccias from the three investigated MVs Maccalube di Aragona (MDA), Santa Barbara (SB) and Comitini (COM) in Sicily. The chromatograms shown in the left panels (A–D) reveal the distribution of the n-alkanes (e.g. C_{20}), isoprenoid hydrocarbons (Pr = pristane, Ph = phytane) and hopanes. The right panel mass chromatograms (E–H) show the distribution of the hopanes. Key: Ts = 18α(H)-22,29,30-trisnorhopane, Tm = 17α(H)-22,29,30-trisnorhopane, BNH = 29,30-bisnorhopane, C_{29}NH = 17α(H),21β(H)-29-norhopane, C_{30}H = 17α,21β(H) C_{30} hopane, open squares (□) = C_{31} to C_{35} 22S and 22R 17α,21β(H) and 17β,21α(H) homohopane series.

To compare *n*-alkane and acyclic isoprenoid distributions of all MV samples we calculated the biomarker ratio of the isoprenoids pristane/phytane (Pr/Ph) as well as the carbon preference index (CPI) and the odd-to-even predominance (OEP) of saturated *n*-alkanes (Bray and Evans, 1961; Scalan and Smith, 1970) (Table 1). The Maccalube di Aragona MV samples demonstrate Pr/Ph values of 0.58 to 1.04, while the Santa Barbara MV shows a greater

variation in this ratio ranging from 0.69 to 1.62. Interestingly, the SB MV samples taken in 2018 feature much higher values (≥ 1.5) of the Pr/Ph ratio than the samples taken in 2017 (≤ 1) (Table 1). Only one of the Comitini MV samples has sufficient amounts of Pr and Ph in their saturated fraction to calculate a Pr/Ph value of 0.67. The CPI index values are around 1 for the Maccalube di Aragona fresh mud breccias, whereas the mud pond samples from this area show higher values about 1.3 to 1.4 in both sampling years. The OEP ratio is ranging from 1.0 to 1.1 in the Maccalube di Aragona MV and does not indicate much difference between the different sample types and years. For the other two MVs, Santa Barbara and Comitini, the CPI is relatively constant around 1.5 to 1.6, with the exception of a high CPI of 2.9 in the altered mud breccia SB 4A (17) taken further away from the main active MV area and from an area covered with vegetation. The OEP ranges from 1.0 to 1.2 in all sample types and years in the Santa Barbara and Comitini MVs.

3.1.2 Distribution of hopanes

The composition of hopanes was determined using mass chromatograms of *m/z* 191 (Fig. 2 E–H). In the MVs, the most abundant hopane is the 17α,21β(H) C_{30} hopane (C_{30}H) followed by a series of C_{31} to C_{35} 22S and 22R 17α,21β(H)- and 17β,21α(H)-homohopanes of decreasing abundance with increasing carbon number. The hopane distribution pattern is strikingly similar between the three different MV sites with minor changes only in the relative abundance of the 22S versus 22R hopanes. For example, in the Maccalube di Aragona MV and mud pond samples (Fig. 2E and F) only the 22R 17β,21α(H)-C_{31} homohopane is detected, while in the other two MVs Santa Barbara and Comitini the both the 22S and 22R 17β,21α(H)-C_{31} hopanes are present. Another relatively abundant hopane is the 17α,21β(H)-29-norhopane (C_{29}NH), while also the trisnorhopanes Ts (18α(H)-22,29,30-trisnorneohopane) and Tm (17α(H)-22,29,30-trisnorhopane) were detected of which Tm is more abundant than Ts in all MV sites. Finally, the Santa Barbara and Comitini MVs feature 17α,18α(H) 29,30-bisnorhopane (BNH), whereas the Maccalube di Aragona MV samples do not reveal this compound.

To compare the source and thermal maturity of the expelled organic matter from the three MVs, we calculated characteristic biomarker indices of hopanes (Peters et al., 2005) (Table 1). The C_{29}NH/C_{30}H hopane ratio is an indicator for the source of organic matter (Peters et al., 2005) and shows relatively low values ranging from 0.58 to 0.81 in all three MVs. The Ts/(Ts+Tm) index gives insight into the source and thermal maturity of organic

matter (Seifert and Moldowan, 1978; McKirdy et al., 1984; Moldowan et al., 1986) and ranges from 0.35 to 0.48 for the Maccalube di Aragona MV, around 0.3 for the Santa Barbara MV and approximately 0.4 to 0.5 for the Comitini MV. The 22S/(22S+22R) isomerization ratio of C_{31} and C_{32} homohopanes is a proxy for the thermal maturity of organic material (Seifert and Moldowan, 1980; Schoell et al., 1983) and varies from 0.49 to 0.66 for both C_{31} and C_{32} homohopanes in the three MV sites.

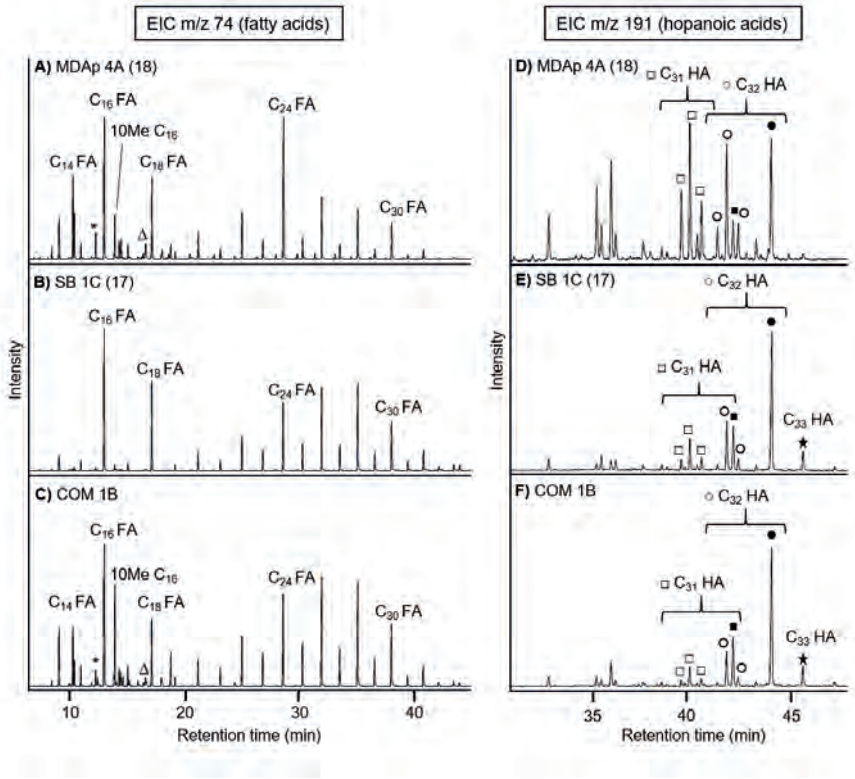


Figure 3. Mass chromatograms of m/z 74 showing the distribution of fatty acids (A–C) and m/z 191 hopanoic acids (D– F) of the fatty acid methyl ester fractions of the three MV sites. The left panel displays fatty acids (FAs) e.g. C_{16} FA or C_{24} FA, and some unsaturated counterparts *=unsaturated C_{16} FAs, Δ= unsaturated C_{18} FAs and 10Me C_{16} = 10-methyl hexadecanoic acid. The right panel shows the hopanoic acid (HA) distribution ranging from □ = C_{31} 22S and 22R 17α,21β(H), C_{31} 22S and 22R 17β,21α(H)-hopanoic acids, ○= C_{32} 22S and 22R 17α,21β(H), C_{32} 22S and 22R 17β,21α(H)-hopanoic acids, ▪= C_{31} 22R 17β,21β(H)-hopanoic acid (homohopanoic acid), ●= C_{32} 22R 17β,21β(H)-hopanoic acid (bishomohopanoic acid) and filled star = C_{33} 22R 17β,21β(H)-hopanoic acid (trishomohopanoic acid).

3.2 Distribution and $\delta^{13}C$ values of fatty and hopanoic acid biomarkers

In the fatty acid methyl ester fractions, the main lipids identified were fatty (FAs) and hopanoic acids (analyzed as their methyl ester derivatives), whose relative distributions are shown by mass chromatograms of *m/z* 74 (typical for FAs; Figs.

3A–C) and *m/z* 191 (typical for hopanoic acids; Figs. 3D–F), respectively. The Maccalube di Aragona fresh mud breccias from the active MV field hardly contained FAs or hopanoic acids and thus only a representative mass chromatogram of the fatty acid methyl ester fraction of the Maccalube mud pond (MDAp; Fig. 3D) is shown.

The Comitini mud breccia and Maccalube mud pond sample contain a range of FAs with major relative abundances of saturated C_{14}, C_{16} and C_{18} FAs as well as some of their unsaturated counterparts, whereas in the Santa Barbara mud breccia only the saturated C_{16} and C_{18} FA were detected. The Comitini and Maccalube sites also show a major relative abundance of 10-methyl C_{16} FA which is absent at the Santa Barbara MV site. Furthermore, longer chain FAs are also present in the MV sites, the Maccalube di Aragona mud pond is dominated by C_{24} FA with a decreasing abundance of the longer chain FAs C_{26}, C_{28} and C_{30} with an even over odd predominance. In contrast, the other two MV sites Santa Barbara and Comitini feature higher abundances of C_{26} and C_{28} FAs with lower relative abundance of C_{24} and C_{30} FAs with an even over odd FA predominance (Figs. 3A–C).

Table 2 Stable **c**arbon isotope values $\delta^{13}C$ of present fatty acids (FAs) and hopanoic acids (HAs) in the fatty acid methyl ester fraction extracted from the three MV sites Maccalube di Aragona, Santa Barbara and Comitini with standard deviation calculated from triplicate $\delta^{13}C$ measurements. n.p.= not present.

Sample ID	C_{14} FA	C_{16} FA	10MeC_{16} FA	C_{18} FA	17β,21β(H) C_{32} HA
MDAp 4A (18)	–5.0 ± 0.6	–14.5 ± 0.1	–11.7 ± 0.1	–19.1 ± 0.3	n.p.
3MDAp 1A (17)	–9.2 ± 1.2	–19.3 ± 0.3	–21.4 ± 0.5	–22.9 ± 0.7	n.p.
SB 1A (17)	n.p.	–29.6 ± 0.4	–31.1 ± 0.5	–27.9 ± 0.8	–29.5 ± 0.4
SB 1C (17)	n.p.	–27.5 ± 0.5	n.p.	–28.5 ± 0.6	–29.3 ± 0.1
COM 1B (18)	–11.5 ± 0.9	–26.7 ± 0.2	–21.8 ± 0.6	–25.1 ± 0.7	–28.6 ± 0.3

The *m/z* 191 mass chromatograms of the fatty acid methyl ester fractions show a series of hopanoic acids (Fig. 3D–F). The mud breccias of the Santa Barbara and Comitini MVs (Fig. 3E and F) feature a series of low present signals of 22S and 22R 17α,21β(H)- and 17β,21α(H)-homohopanoic acids as well as a higher signal of 22R 17β,21β(H)-homohopanoic acid. Furthermore, 22R 17α,21β(H)- and 17β,21α(H)-bishomohopanoic acids, a major 22R 17β,21β(H)-

bishomohopanoic acid as well as a low signal of 22R 17β,21β(H)-trishomohopanoic acid were detected. In contrast, the Maccalube di Aragona mud pond site (Fig. 3D) shows a relatively high abundance of C_{31} and C_{32} 22R 17α,21β(H)-hopanoic acids in tandem with a high 22R 17β,21β(H)-bishomohopanoic acid peak with a lower presence for the other stereochemical configurations of the C_{31} and C_{32} hopanoic acids. Overall, the Santa Barbara and Comitini MVs feature the same hopanoic acid distribution with a major 22R 17β,21β(H)-bishomohopanoic acid, whereas the Maccalube mud pond reveals a higher relative abundance of all 22S and 22R configurations and the whole suite of C_{31} and C_{32} hopanoic acid configurations.

The stable carbon isotopic compositions ($\delta^{13}C$) could only be determined for the major FAs (C_{14}, C_{16}, C_{18} and 10-methyl C_{16} FA) due to the low abundance of other FAs and the presence of an UCM. Based on the lipid biomarker distribution, we picked representative samples for ^{13}C measurements, which are reported in Table 2. These major FAs show $\delta^{13}C$ values ranging from –14 to –30‰ in all three MV sites although the C_{14} FA in the Maccalube di Aragona mud pond is substantially more enriched ($\delta^{13}C$ value = –5‰). Unfortunately, due to the low abundance of most of the hopanoic acids, only stable carbon isotopic compositions of the 22R 17β,21β(H)-bishomohopanoic acid could be determined, revealing $\delta^{13}C$ values around –28 to –30‰ in the Santa Barbara and Comitini MV (Table 2).

3.3 Lipid biomarker signatures of polar biomarkers

The GC-amenable polar fraction of the three MV sites contained some beta-hydroxy fatty acids (C_{16}, C_{18}, C_{24} and C_{26}) and C_{27}–C_{29} sterols (e.g. cholesterol) as well as a large UCM. We were not able to unambiguously detect biomarkers such as hopanols, archaeol and hydroxyarchaeol, possibly because of the large UCM in the polar fraction.

Analysis of the polar fraction using UHPLC-MS for archaeal derived ether lipids revealed their presence in all three MV sites (Table 3 and Fig. 4). The Maccalube di Aragona mud breccia shows high fractional abundances of GDGT-0 ranging from 21.6 to 32.9% and of GDGT-1 to -4 which varies from 11.4% (GDGT-4) up to 25.9% (GDGT-3) in combination with low amounts of crenarchaeol (2.5 to 3%) and the absence of the crenarchaeol isomer (Table 3). Different GDGT distributions are observed in the Maccalube mud pond with high fractional abundances of GDGT-0 (89%), low abundances of GDGT-1 to -4 (1.9 to 3%) as well as very low crenarchaeol (0.8%) and the non-detection of the crenarchaeol isomer. Furthermore, we detected high fractional abundances

of crenarchaeol ranging from 25 to 46% in the Santa Barbara mud breccias in combination with high GDGT-0 (26 to 39%) and low crenarchaeol isomer (4 to 5%), while the other GDGTs (GDGT-1 to -4) were present in lower abundances (≤ 14%). The Comitini mud breccias feature high fractional abundances of GDGT-0 (33 and 37%) and crenarchaeol (21 and 28%) with lower abundances of GDGT-1 to -4 (1 to 19%) and crenarchaeol isomer (ca. 2%).

Additionally, archaeol was also detected in the three MV sites and the ratio of archaeol/GDGT-0 was calculated ranging from 0.3 to 9.5 in the Maccalube di Aragona mud breccias and mud pond while the Santa Barbara mud breccias had values ranging from 0.5 to 2.4 and the Comitini mud breccias 0.5 and 0.6 (Table 3). The fact that archaeol was detected by UHPLC-MS and not GC-MS might have to do with the higher sensitivity of the UHPLC-MS where selective ion monitoring was used, in contrast to the full scan monitoring of the GC-MS.

The isoprenoidal GDGT based Methane index (MI index = (GDGT-1+ GDGT-2+ GDGT-3) / (GDGT-1+ GDGT-2+ GDGT-3+Cren+ Cren´)) has been used to differentiate the GDGT distributions of marine ANME-archaea involved in AOM from marine Thaumarchaeota (Zhang et al., 2011). In general, high MI values (≥ 0.5) indicate methane or hydrate influenced marine environments, while low values reflect normal marine signals. Here, we applied this index for the three terrestrial MV sites to determine AOM-related GDGT signals. The Maccalube di Aragona site indicates high MI values from 0.90 to 0.96, while the Santa Barbara MV (0.26 to 0.48) and Comitini MV (0.54 and 0.63) show lower values for the MI (Table 3).

Fig. 4 shows a ternary diagram with the fractional abundances of GDGT-0, GDGT-1 to -4 and crenarchaeol among the three Sicilian MV. The Maccalube di Aragona mud breccias cluster together at the bottom of the ternary plot and are clearly separated from the other Sicilian MVs due to their minor abundance of crenarchaeol. Interestingly, the Maccalube mud pond samples show differences in the fractional abundance of GDGT-0 which is very high in these two samples with low GDGT-1 to -4 and very low crenarchaeol. The other two MV sites Comitini and Santa Barbara cluster together in the middle of the ternary plot with a high abundance of crenarchaeol, especially in the Santa Barbara MV samples (Fig. 4). Overall, the Santa Barbara MV and Comitini MV differ from the Maccalube di Aragona MV by the higher abundances of GDGT-0 and crenarchaeol versus high GDGT-1 to -4 and minor crenarchaeol.

Table 3 Fractional abundance (%) of isoprenoidal-GDGTs (-1 to -4), crenarchaeol (Cren) and crenarchaeol isomer (Cren´) as well as GDGT- and archaeol-based ratios in the polar fraction of the three investigated MV sites Maccalube di Aragona, Santa Barbara and Comitini. Key: MI index (Methane index) = (GDGT-1+ GDGT-2+ GDGT-3) / (GDGT-1+ GDGT-2+ GDGT-3+Cren+ Cren´).

Sample ID	GDGT-0	GDGT-1	GDGT-2	GDGT-3	GDGT-4	Cren	Cren´	Archaeol/ GDGT-0	MI Index
1MDA 1A (17)	21.6	14.2	20.1	25.9	15.3	3.0	0.0	9.52	0.95
1MDA 3A (17)	32.9	16.0	19.2	17.7	11.4	2.8	0.0	0.95	0.95
MDA 1A (18)	24.2	16.4	21.3	23.1	12.5	2.5	0.0	0.32	0.96
3MDAp 1A (17)	89.2	2.4	3.0	2.5	2.0	0.8	0.0	4.15	0.91
MDAp 4A (18)	89.9	2.2	2.8	2.5	1.9	0.8	0.0	2.73	0.90
SB 1A (17)	38.7	10.3	13.4	4.4	3.0	25.0	5.2	2.37	0.48
SB 1C (17)	31.1	7.0	9.0	3.1	1.1	45.0	3.8	0.47	0.28
SB 4A (17)	25.8	9.2	9.5	3.9	2.8	45.1	3.7	0.56	0.32
SB 1A (18)	30.4	6.8	8.7	2.6	1.0	45.9	4.6	0.77	0.26
SB 3A (18)	32.4	7.6	9.7	2.6	0.5	43.2	4.0	1.90	0.30
COM 1A	33.2	9.7	17.6	7.8	1.4	28.3	2.0	0.62	0.54
COM 1B	37.0	10.3	18.6	9.6	2.2	20.6	1.6	0.54	0.63

4. Discussion

4.1 Source and thermal maturity of organic matter

Materials expelled from MVs are a mixture of mud breccias, waters and gas which may not be all co-genetic in their source strata (Stadnitskaia et al., 2007, 2008; López-Rodríguez et al., 2014). Therefore, lipid biomarker distributions and indices can help to characterize the organic geochemical signatures of potential source strata (Peters et al., 2005). Below, we reconstruct and discuss the potential source and thermal maturity of organic matter in the investigated mud breccias from the three Sicilian MVs.

4.1.1 Source of organic matter in the expelled mud breccias

The Maccalube di Aragona MV and mud pond are characterized by a high abundance of short chain *n*-alkanes (C_{17} to C_{21}) with a slight even over odd predominance and Pr/Ph values <1 (Table 1 and Fig. 2), suggesting that the source organic matter derives from mature anoxic marine deposited strata connected to carbonates or marls (Bray and Evans, 1961; Didyk et al., 1978; Volkman et al., 1992). This is further supported by the presence of C_{29}NH and a major C_{30}H together with the extended C_{31}–C_{35} homohopane series. These biomarker signatures are usually found in marine strata which were deposited in a shelf environment often forming carbonates, marls or other evaporates (Connan et al., 1986; Seifert and Moldowan, 1978; Peters and Moldowan, 1991). Similar hopane patterns are present in the other two MVs Santa Barbara (SB) and Comitini (COM) implying similar marine deposited source strata for the organic matter in the mud breccias as the Maccalube MV. However, the C_{29}NH/C_{30}H index is <0.9 in all three MV sites which slightly disagrees with the origin from marine carbonate deposits as values >1 would be expected (Zumberge, 1984; Connan et al., 1986). One reason could be that the organic matter in the mud breccias acts similar to mature petroleum derived compounds from carbonate source rocks which have been found to generate low C_{29}NH/C_{30}H values like marine evaporitic and carbonate source rocks located in the Brazilian marginal basins (Mello et al., 1988a, b), marls in the Northern Apennines in Italy (Ten Haven et al., 1985; Sinninghe Damsté et al., 1986) or in carbonate source rocks from the Dead Sea in Israel (Rullkötter et al., 1985). Another more likely reason is that the organic matter represents a mixture from different geological strata in the MVs subsurface at which the biomarker signal from carbonate reefs gets diluted with hydrocarbons from other strata, e.g. shales producing lower C_{29}NH/C_{30}H values. This is further supported by the fresh Santa Barbara mud breccias which indicate noticeable differences in their Pr/Ph ratio in the two

different sampling years 2017 and 2018 (Table 1) proposing a change in the predominant source strata of the expelled organic matter and MV activity over one year time.

Additionally, the Maccalube mud pond site, the Santa Barbara and Comitini MVs feature a relatively high abundance of long chain *n*-alkanes with an odd over even carbon number predominance (C_{25} to C_{29}), CPI and OEP values generally above 1.3 (Table 1 and Fig. 2). Here, the odd longer chain *n*-alkanes might derive from input of land-plant material produced from epicuticular waxes in higher plants from the surrounding vegetation or from buried terrestrial strata in the MV subsurface (Eglinton and Hamilton, 1967; Volkman et al., 1992). Moreover, the high abundance of even carbon numbered long-chain fatty acids (Fig. 3) also implies a major input of immature allochthonous terrestrial material (Eglinton and Hamilton, 1967) into these mud breccias, as these compounds would not be expected to be in high abundances in mature petroleum like organic matter. Only at the Maccalube di Aragona MV, we did not detect long-chain even carbon numbered fatty acids, which coincides with a low CPI of long chain *n*-alkanes (Table 1). Hence, it seems to be most likely that the mud breccias from the Maccalube di Aragona MV derive from marine more mature source strata whereas the organic material from the Maccalube mud pond originates from a mix of mature marine carbonate strata and more immature terrestrial strata with a high input of plant material. Similar assumptions can be drawn for the Santa Barbara and Comitini mud breccias which possibly derive from a mixture of thermally mature and immature marine (e.g. carbonates or marls) and terrestrial organic matter from different source strata. Indeed, previous studies mainly about marine MVs also suggested that the lipid biomarkers (e.g. *n*-alkanes and hopanes) in the mud breccias originate from different mature geological sections in the subsurface (Stadnitskaia et al., 2007; López-Rodríguez et al., 2014).

4.1.2 Thermal maturity

The degree of thermal maturation of organic matter can be determined by a suite of lipid biomarkers. The 22S/(22S+22R) homohopane isomerization index of C_{31} and C_{32} homohopanes ranges between 0.50 to 0.66 in the three different MVs (Table 1) suggesting that at least part of the organic matter in the mud breccias experienced a moderate to high degree of thermal maturation (Seifert and Moldowan, 1980; van Duin et al., 1997; Peters et al., 2005). Besides, the calculated values of the 22S/(22S+22R) homohopane isomerization index are close to the thermodynamic equilibrium values of ca. 0.6 (Seifert and Moldowan, 1980; van Duin et al., 1997), thus indicating that most of the organic matter in

the Maccalube di Aragona mud breccias experienced slightly higher maturations than in the other two MVs due to higher values in this index. Another hopane maturity indicator is the Ts/(Ts/Tm) index which shows low values of about 0.30 to 0.48 in the three MV sites indicating relatively immature source organic matter (Peters et al., 2005). However, certain factors such as lithology are known to affect the rate of isomerization of hopanes. The presence of high amounts of clay in the MVs could also cause clay-catalyzed reactions which produce mature signals by 17α,21β(H)-hopanes or influence Ts and Tm distribution (McKirdy et al., 1983; Kolaczkowska et al., 1990). Additionally, the occurrence of both the biological 22R 17β,21β(H) configuration in C_{31} and C_{32} hopanoic acids in tandem with the diagenetically altered 22R and 22S 17α,21β(H)-hopanoic acids indicate the presence of at least two sources of organic matter with different maturity degrees in the three MVs (van Duin et al., 1997). It seems likely that the organic matter from the Maccalube di Aragona MV derives from slightly more mature source strata than the other two MVs since they feature more abundant 22R and 22S 17α,21β(H)-hopanoic acids as well as slightly higher 22S/(22S+22R) homohopane isomerization values. Overall, the coincidence of 17β,21β(H) and 17α,21β(H) configurations next to thermally mature 22S/(22S+22R) homohopane values and immature signals of the Ts/(Ts/Tm) index clearly suggest the presence of more than one source of organic matter in the mud breccias as observed in other studies about MVs e.g. in the Mediterranean or Black Sea (López-Rodríguez et al., 2014; Stadnitskaia et al., 2007, 2008). The aforementioned *n*-alkanes, pristane and phytane distributions as well as the CPI and OEP indices also confirm this hypothesis, having both immature organic matter and moderately mature or petroleum derived compounds in the investigated mud breccias.

4.1.3 Potential source strata of mud breccias

The three MVs are located in the Caltanisetta Basin in Central Sicily in which potential source strata have rarely been investigated yet. Potential source strata for the mud breccia organic matter in the Maccalube di Aragona and Comitini MVs might be of Miocene age (Burdigalian and Aquitarian) and related to carbonate reefs or platform carbonates of the Trapanese-Saccense units and/or basinal facies of Triassic age of the Sicani units (Granath and Casero, 2004). The pristane/phytane and the $C_{29}NH/C_{30}H$ biomarker indices (Table 1) as well as the distribution of *n*-alkanes relate partly the organic matter from the Maccalube di Aragona and Comitini MV sites to oils from the Gela 93 well at the Ragusa platform (Pieri and Mattavelli, 1986; Mattavelli and Novelli, 1990) but in

contrast to these mature oils we could not detect gammacerane in the muds of the MVs. Moreover, the Santa Barbara MV is located more northern and closer to the thrust belt where potential source strata could be related to a Lower Jurassic carbonate facies (Mattavelli and Novelli, 1990). In general, the organic matter from the three investigated MV systems seems to be linked to rather unidentified source strata which share some of the characteristics of known geological sections and petroleum characteristics in Sicily.

4.2 Microbial lipid signatures

4.2.1 Aerobic methane oxidation

The fatty acid composition of the mud breccias revealed abundant C_{14}, C_{16} and C_{18} FAs in all three MV sites, which likely derive from microbes involved in the utilization of the released gases or petroleum hydrocarbons. For example, Type I or Type II methanotrophs (Gammaproteobacteria or Alphaproteobacteria) are known to synthesize C_{14}, C_{16} and C_{18} FAs as well as their unsaturated counterparts (Bull et al., 2000; Dedysh et al., 2007; Bodelier et al., 2009). Furthermore, the $C_{18:0}$ FA in tandem with the presence of C_{14} and C_{15} FAs have been identified previously in *Verrucomicrobia* (Op den Camp et al., 2009) and in terrestrial MVs in Taiwan (Chang et al., 2012; Cheng et al., 2012). However, the $\delta^{13}C$ values of the major FAs (C_{14}, C_{16} and C_{18} FAs) (Table 2) rather exclude an origin from aerobic methanotrophic bacteria since one would expect $\delta^{13}C$ values more depleted than the one of the released methane, which generally ranges from –45 to –50‰ in the three MV sites (Grassa et al., 2004; Heller, 2011).

Various hopanoic acids were also detected; they can originate from a range of methanotrophic Alpha-and Gammaproteobacteria but are also known to be synthesized by other bacterial phyla including Planctomycetes or Acidobacteria (Thiel et al., 2003; Sinninghe Damsté et al., 2004, 2017; Birgel and Peckmann, 2008). The mud breccias and mud pool samples from the three MVs contain different stereoisomeric configurations of hopanoic acids the 22S and 22R 17α,21β(H) and 17β,21β(H) suggesting at least two different sources of hopanoic acids since they cannot coexist with each other (Seifert and Moldowan, 1980; van Duin et al., 1997; Bennett and Abbott, 1999). Here, the 17α,21β(H)-hopanoic acids probably derive from thermally mature source strate in the deeper layers of the MVs and not from recent bacterial biomass as mentioned earlier in section 4.1.2 (van Duin et al., 1997; Bennett and Abbott, 1999; Farrimond et al., 2002). In contrast, 22R 17β,21β(H)-hopanoic acids could derive from recent bacterial biomass in the upper MV layers or directly in the mud breccias and represent a signal from e.g. aerobic methanotrophic bacteria. Although the acid

group of these hopanoids suggest a certain diagenetic alteration, which could be caused by clay-catalyzed reactions in the mud or through other processes, 22R 17β,21β(H)-hopanoic acids are regularly found in recent sediments and soils where they derive from the early steps of bacteriohopanepolyol diagenesis (Rohmer et al., 1984; Bennett and Abbott, 1999; Crossman et al., 2005; Inglis et al., 2018). Unfortunately, due to minor abundances of the C_{31} and C_{33} 22R 17β,21β(H)-hopanoic acids and the UCM, it was only possible to measure $\delta^{13}C$ on one hopanoic acid, namely the 17β,21β(H)-bishomohopanoic acid. It displays regular $\delta^{13}C$ values of about –28 to –30‰ in the mud breccias (Table 2). These relatively high $\delta^{13}C$ values compared to the released methane confirm that it is probably not derived from microbes related to AMO and might originate from other bacterial sources such as oil degrading bacteria or SRBs. However, it should be noted that some aerobic methanotrophs, for examples '*Ca*. Methylomirabilis oxyfera' do not use methane as their carbon source as they autotrophically fix carbon dioxide, and thus its biomass and lipids do not feature the characteristic $\delta^{13}C$ depletion (Kool et al., 2012; Rasigraf et al., 2014). Additionally, the hopanoic and fatty acids seem to derive from more than one source of organic matter which could also dilute the typical $\delta^{13}C$ signatures of aerobic methanotrophs.

4.2.2 Anaerobic oxidation of methane

The high abundance of CH_4 in tandem with the probably marine derived organic matter in the mud breccias suggest that the Sicilian MVs could also host sulfate dependent AOM as found previously in other terrestrial and marine MVs via molecular and lipid biomarker based techniques (e.g. Stadnitskaia et al., 2005; Alain et al., 2006; López-Rodríguez et al., 2014; Lee et al., 2018). The relatively high abundant 10-methyl C_{16} FA in the Maccalube di Aragona and Comitini MVs (Fig. 3A and C) suggests a possible presence of sulfate reducing bacteria (SRB) of the genera *Desulfobacter* and *Desulfobacula* (Dowling et al., 1986; Kuever et al., 2001; Bühring et al., 2005) or the intra-aerobic methanotroph "*Ca*. Methylomirabilis oxyfera" (Kool et al., 2012). In SRBs, 10-methyl C_{16} FA is found to make up approximately 5 to 25% of the total lipids, while in "*Ca*. Methylomirabilis oxyfera" it accounts for 35 to 45% of total FAs. Since we did not find the typical demethylated hopanoids and 10-methyl $C_{16:1\Delta7}$ FA for "*Ca*. Methylomirabilis oxyfera" (Kool et al., 2012; Smit et al., 2019), this source can likely be excluded, although their signals might be suppressed by the high presence of petroleum like hydrocarbons and UCM in the mud breccias. This suggests that 10-methyl C_{16} FA could originate from SRBs. However, we did not

observe the usually depleted $\delta^{13}C$ isotopic values (Table 2), which hints to a heterotrophic origin of this FA (Elvert et al., 2003; Blumenberg et al., 2004). Furthermore, the $\delta^{13}C$ signal could be diluted by signals from other bacteria since the 10-methyl C_{16} FA is also occasionally found in other bacterial species like Actinobacteria (Campbell and Naworal, 1969; Park et al., 2008), *Geobacter* sp. (Lovley et al., 1993) or Planctomycetes (Sinninghe Damsté et al., 2005) but generally in low abundances. Previous studies at the terrestrial Salse di Nirano MV in Northern Italy found partly ^{13}C-depleted non-isoprenoidal dialkyl glycerol diethers (DAGE) and archaeol indicating active AOM (Heller et al., 2012), while Alain et al. (2002) found hydroxyarchaeol, archaeol, some iso and anteiso C_{17} FAs plus 10-methyl C_{16} FA in MVs in the Carpathian Mountains, Romania, suggestive of AOM. In contrast, only traces of archaeol were detected besides the 10-methyl C_{16} FA in the GC-amenable compounds of the three Sicilian MVs, indicating that SRBs might be more involved in other microbial processes as methanogensis than AOM (Parkes, 1999; Zengler et al., 1999).

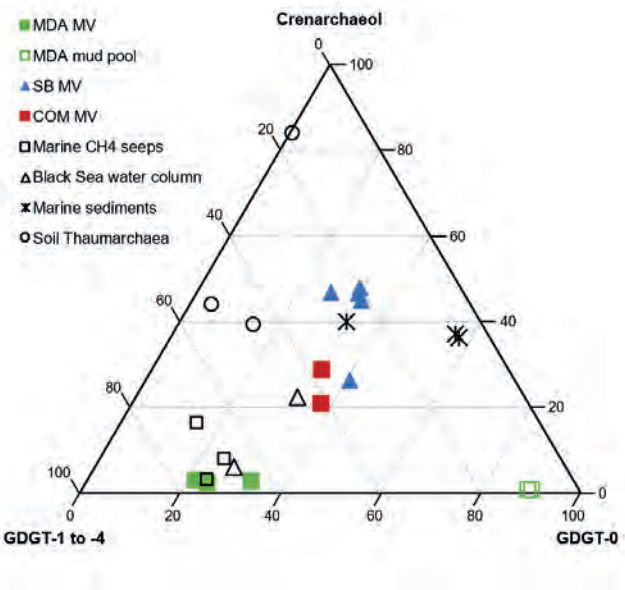


Figure 4. Ternary diagram showing the distribution of GDGT-0, crenarchaeol and sum of GDGT-1 to 4 in the three Sicilian MV sites Maccalube di Aragona (MDA), Santa Barbara (SB) and Comitini (COM). Black data points indicate literature data from previous studies from marine CH_4 seeps, the Black Sea water column, marine sediments and soil Thaumarchaea (Schouten et al., 2000; Wakeham et al., 2003; Wakeham et al., 2004; Stadnitskaia et al., 2005; Sinninghe Damsté et al., 2012).

Additionally, UHPLC-MS analysis revealed the presence of archaeol and GDGTs-0 to -4 plus crenarchaeol in the three MV sites with differences in their

distribution depending on the MV location (Table 3 and Fig. 4). Archaeol is synthesized by most archaeal species including methanotrophic ANME-archaea or methanogenic Euryarchaea as well as Thaumarchaea (Koga et al., 1993; Hinrichs et al., 1999; Pancost et al., 2001b; Schouten et al., 2013). Therefore, the GDGT distribution can give a more distinctive insight on the archaeal species present in the MVs. High abundances of GDGT-1 to -4 in combination with very low abundances of crenarchaeol in the Maccalube di Aragona mud breccias indicate a presence of ANME-archaea at this site, as they match distributions found in these microbes (Pancost et al., 2001a; Wakeham et al., 2003; Blumenberg et al., 2004). Furthermore, considerably high values in the GDGT-based methane index $\geq$ 0.9 (Zhang et al., 2011) and the clustering of the Maccalube MV data in the ternary plot (Fig. 4), with other AOM influenced environments e.g. marine MVs and Black Sea water column strongly suggests the activity of ANME-archaea (Schouten et al., 2000; Wakeham et al., 2003, 2004; Stadnitskaia et al., 2005). It is unclear whether the ANME-archaea are present at the surface of the MV or are derived from the subsurface. Interestingly, besides high amounts of GDGT-1 and GDGT-2 there are also high amounts of GDGT-3 and GDGT-4 suggesting an origin of thermophilic ANME-archaea and thus thermophilic AOM processes (Schouten et al., 2003). This would indicate that the AOM signal is either derived from the subsurface or represents a mixed signal of thermophilic and non-thermophilic AOM from different MV depths in the Maccalube di Aragona MV.

In comparison, the Maccalube di Aragona mud pool samples show less abundant GDGT-1 to -4 and a high presence of GDGT-0 with almost absent crenarchaeol suggesting probably less present ANME-archaea but methanogenic Euryarchaeota since methanogens usually produce high abundances of GDGT-0 and archaeol with less GDGT-1 to -4 (Blumenberg et al., 2004; Schouten et al., 2013). In contrast, the high relative abundances of crenarchaeol in combination with the presence of the crenarchaeol isomer and low GDGT-1 to -4 abundances in the Santa Barbara and Comitini MVs (Table 3 and Fig. 4) imply that these lipids likely originate from marine pelagic Thaumarchaea or from terrestrial Thaumarchaea from buried immature source strata (Sinninghe Damsté et al., 2012; Schouten et al., 2013; Elling et al., 2017). Moreover, the GDGT abundance of these two MVs cluster together with normal marine sediments (Schouten et al., 2000) and are close to soil thaumarchaeal signals (Sinninghe Damsté et al., 2012) in the ternary plot (Fig. 4) which is in accordance with the low values of the calculated methane index. In general, it seems that the Maccalube di Aragona MV reveals evidence for ANME-archaea involved in AOM processes due high

abundances of GDGT-1 to -4 and high methane index values next to presence of SRB typical 10-methyl C_{16} FA. Interestingly, the Comitini MV indicates a moderate methane index around 0.5, GDGT distributions clustering in the middle of the ternary plot and the presence of 10-methyl C_{16} FA as well as moderate levels of crenarchaeol, suggesting a mix of AOM biomarker signals and immature marine or terrestrial thaumarchaeal biomarkers from the MV subsurface strata. The Santa Barbara MV only demonstrates signals of marine or terrestrial Thaumarchaea, probably from immature source strata thus suggesting no active AOM at this MV. This could potentially be related to the constant fluid seepage and episodic eruptions preventing the settlement of a stable microbial community which is also suggested by variations in the saturated biomarker composition in the two different sampling campaign years 2017 and 2018 in the Santa Barbara mud breccias (Table 1).

4.3 Comparison between terrestrial and marine MV systems

Up to now, only a few studies have investigated the composition of lipid biomarkers as well as the microbial community especially in terrestrial MV systems. All investigated mud breccias from the three Sicilian MVs seem to be a mixture of immature and mature organic matter originating from at least two different source strata. Besides, we did find evidence for AOM in the Maccalube di Aragona MV through characteristic GDGT distributions clustering with data from other AOM rich environments. The Comitini MV might also host AOM but to a lesser extent than the Maccalube MV, whereas the Santa Barbara MV did not indicate AOM processes. Furthermore, we could not detect unambiguous evidence for microbes involved in AMO processes in all three Sicilian MVs. Other terrestrial MVs in the Northern Apennines and Romania did show also other archaeal and SRB specific biomarkers e.g. hydroxyarchaeol, DAGE and MAGE suggestive of AOM, besides high amounts of petroleum like hydrocarbons (Alain et al., 2006; Heller et al., 2012). However, the $\delta^{13}C$ enriched values of SRB specific lipids at these sites suggested that mainly heterotrophic SRBs are active in MVs of the Northern Apennines and to a lesser extent involved in AOM. Although, genetic studies in the Northern Apennines (Salse di Nirano) MVs did show the presence of aerobic methanotrophic bacteria from the phylum Methylococcales and found ANME-2 archaea (e.g. *Methanosarcinales*) (Wrede et al., 2012). Terrestrial Taiwanese MVs also indicated the presence of ANME-2 archaea (*Methanosarcinales*) next to methanogenic Euryarchaeota (*Methanomicrobiales*) and various bacterial phyla like Gammaproteobacteria and Firmicutes (Chang et al., 2012; Cheng et al.,

2012). However, it was also found that microbial processes are compartmentalized into discrete stratified niches in the MVs themselves, where methanogenesis is the most abundant metabolism and AOM only occurs in the bubbling fluids. Pore-water profiles in these other terrestrial MVs showed that methane concentrations were inversely correlated with Fe^{2+}/Mn^{2+} mineral concentrations, thus suggesting AOM being coupled to metal reduction instead of sulfate reduction in the Taiwanese MVs (Chang et al., 2012; Cheng et al., 2012). Similar conclusions were also drawn for the Romanian MVs where the microbial community of sulfate dependent AOM was present although less well developed than in marine MV and cold seep systems (Alain et al., 2006). The reason why only the Maccalube di Aragona MV indicates clear markers for AOM and all three MVs lack signals for AMO to other terrestrial MVs is not completely clear. We speculate that especially the Santa Barbara MV could lack sufficient concentrations of oxidizing agents necessary for AOM like sulfate, nitrate or iron in the mud breccias generated in the deep reservoirs, limiting the presence of microbes involved in AOM. Moreover, AMO is possibly limited to the upper few centimeters of the mud breccias where oxygen is available and therefore might only be present in very stratified niches as suggested for the Taiwanese MVs (Chang et al., 2012; Cheng et al., 2012)

In comparison to terrestrial MVs, marine MVs feature a higher abundance and larger diversity of archaeal and bacterial lipid biomarker signatures with distinct $\delta^{13}C$ values characteristic for anaerobic and aerobic methane oxidation including biomarkers like hydroxyarchaeol, 2,6,10,15,19-pentamethyleicosane (PMI), DAGE, diplopterol or fatty acids (Stadnitskaia et al., 2005; Niemann et al., 2006a, b; Lee et al., 2018). Possibly, communities involved in AMO or AOM in terrestrial MVs have a different lipid composition and occur in much lower abundances then in marine MVs. It seems likely that it is connected to environmental factors such as oxygen availability for AMO or the more limited abundance of certain oxidizing agents connected to AOM like sulfate, iron and manganese than in marine MV systems. Alternatively, petroleum like hydrocarbons might be more dominant in terrestrial MVs compared to marine MVs serving as a more constant and longer lasting carbon source for the living microbes than the released gases. One reason could be that the gases are more rapidly released into the atmosphere at terrestrial MVs, while the petroleum compounds stay in the muds after deposition. However, our study suggests at least for one investigated MV (Maccalube di Aragona MV) the presence of AOM characteristic lipid biomarkers (10-methyl C_{16} FA, high abundance of GDGT-1 to -4 and a high MI of ≥ 0.9), similar to those found in

marine MVs. Furthermore, this study is, to best of our knowledge, the first one demonstrating specific GDGT patterns for ANME-archaea in a terrestrial MV system.

Future work should investigate gas and mineral concentrations as well as microbial communities via DNA-based techniques in depth profiles in the Sicilian MVs but also in other terrestrial MVs to gain new insights into the microbial metabolisms in these extraordinary gas rich environments.

Conclusion

The mud breccias of the three investigated terrestrial Sicilian MVs (Maccalube di Aragona, Santa Barbara and Comitini) reveal signatures of saturated petroleum like hydrocarbons (*n*-alkanes, acyclic isoprenoids and hopanes) that dominate over those of lipid biomarkers from living bacteria and archaea. The saturated hydrocarbon distribution as well as source- and maturity-related biomarker indices suggest that the organic matter in the mud breccias derived from a mix of at least two different source strata with different maturity degrees in all three Sicilian MVs. Some source strata seem to be connected to mature anoxic marine deposited marls or carbonates while other strata reflect immature terrestrial plant signals. The Maccalube di Aragona source organic matter probably experienced slightly higher thermal maturities than the other two MV sites as suggested by distinctive *n*-alkanes distributions, a higher 22S/(22S+22R) ratio of C_{31} and C_{32} homohopanes and trends of 22S and 22R 17α,21β(H)-hopanoic acids compared to the other two MVs. Furthermore, the three MVs contain regular saturated and unsaturated C_{14}, C_{16} and C_{18} FAs, a major 10-methyl C_{16} FA in the Comitini and Maccalube MVs and C_{31} and C_{32} 22R 17β,21β(H)-hopanoic acids. However, none of the fatty and hopanoic acids was substantially depleted in $\delta^{13}C$ suggesting either no active methane oxidation processes or signals from a mixture of different heterotrophic and autotrophic bacterial producers. In contrast, the isoprenoidal-GDGT distributions with a high MI $\geq$ 0.9 and high abundances of GDGT-1 to -4 clustering with those from AOM rich environments suggesting AOM at the Maccalube di Aragona MV. The Comitini MV and especially the Santa Barbara MV revealed relatively high abundances of crenarchaeol and the presence of the crenarchaeol isomer next to low amounts of GDGT-1 to -4 suggesting a major input of immature marine or terrestrial thaumarchaeal biomass into these MVs. These results imply the first detection of AOM involved ANME-archaea through characteristic GDGT patterns in terrestrial MVs. Furthermore, our results clearly suggest that terrestrial MVs differ from marine MVs which host a higher abundance and

diversity of isotopically depleted lipid biomarkers linked to aerobic/anaerobic methane oxidation, possibly due to limitations of oxygen and specific oxidizing agents (e.g. sulfate or nitrate) necessary for AMO and AOM.

Acknowledgements

We thank Caitlyn R. Witkowski for helping with sample collection in 2017 as well as Fiona Middleton for helping with sample collection in 2018 and initial lipid analysis on some MV samples. We also thank Jort Ossebaar and Marcel van der Meer for assistance in isotope analysis. This study received funding from the Netherlands Earth System Science Center (NESSC) through a grant to JSSD and SS (024.002.001) from the Dutch Ministry for Education, Culture and Science.

Chapter 5

Analysis of non-derivatized bacteriohopanepolyols using UHPLC-HRMS reveals great structural diversity in environmental lipid assemblages

Ellen C. Hopmans, Nadine T. Smit, Rachel Schwartz-Narbonne, Jaap S. Sinninghe Damsté, Darci Rush

Under review at Organic Geochemistry

Abstract

Bacteriohopanepolyols (BHPs) are lipids with great chemotaxonomic potential for microbial populations and biogeochemical processes in the environment. The most commonly used methods for BHP analysis are chemical degradation followed by gas chromatography-mass spectrometry (MS) or derivatization followed by high performance liquid chromatography (HPLC)- atmospheric pressure chemical ionization/MS. Here we report on significant advances in the analysis of non-derivatized BHPs using U(ltra)HPLC-electrospray ionization-high resolution MS^2. Fragmentation mass spectra provided information on the BHP core, functionalized side chain, as well as the conjugated moiety of composite BHPs. We successfully identified the common bacteriohopanepolyols and their (di)methylated and (di)unsaturated homologues, aminoBHPs, and composite BHPs (e.g., cyclitol ethers and methylcarbamateBHPs) in biomass of several known BHP-producing micro-organisms. To show how the method can be exploited to reveal the diversity of BHPs in the environment, we investigated a soil from an active methane seep, in which we detected ca. 130 individual BHPs, including a complex distribution of adenosylhopanes. We identified the nucleoside base moiety of both adenosyl type-2 and type-3. Adenosyl hopane type-3 contains a methylated adenine as its nucleobase, while type-2 appears to contain a deaminated and methylated adenine, or N1-methylinosine. In addition, we detected novel adenosylhopanes. Furthermore, we identified a novel series of composite BHPs comprising of bacteriohopanepolyols conjugated to an ethenolamine moiety. The novel ethenolamineBHPs as well as aminoBHPs were also detected acylated to fatty acids. The analytical approach described allows for simultaneous analysis of the full suite of IPLs, now including BHPs, and represents a further step towards environmental lipidomics.

Keywords

non-derivatized bacteriohopanepolyols, UHPLC-HRMS, novel composite BHPs, adenosylhopanes

1. Introduction

Bacteriohopanepolyols (BHPs) are membrane lipids with great chemotaxonomic potential with regards to microbial populations as well as biogeochemical processes in the environment. BHPs are composed of a pentacyclic triterpenoid ring system with an extended side chain containing 4, 5, or 6 functional groups (Fig. S1) comprised of either all hydroxyl moieties (bacteriohopanetetrol (BHT),

bacteriohopane pentol (BHpentol), and bacteriohopane hexol (BHhexol); Rohmer et al., 1984; Talbot et al., 2003a; Talbot and Farrimond, 2007) or 3, 4, or 5 hydroxyl moieties and a single terminal amino-group (i.e., aminotriol, -tetrol, and -pentol; e.g., Neunlist and Rohmer, 1985; Rohmer, 1993). The basic, unmodified ring system with a linear side chain is usually the dominant BHP in both culture and environmental samples (see also review of BHP sources in (Talbot and Farrimond, 2007; Talbot et al., 2008). BHPs can be further modified by methylation at the C-2 (2MeBHPs; e.g., Talbot et al., 2008 and references therein), C-3 (3MeBHPs; e.g., Cvejic et al., 2000a), both C-2 and C-3 (2,3diMeBHPs; Sinninghe Damsté et al., 2017), C-31 (Simonin et al., 1994), or C-12 (Costantino et al., 2000) carbon positions. BHPs can also include unsaturation in the ring system located at Δ^6, Δ^{11} or both (see section 3.3.3 and Fig. 3; e.g., Talbot et al., 2007b and references therein). A BHP with an unsaturation in the side chain was identified in a methanotrophic *Methylovulum* bacterium (van Winden et al., 2012). A large variety of so-called composite BHPs, where the terminal functional group is bound to complex, often polar moiety, have also been identified (Talbot et al., 2007a and references therein), such as the adenosylhopanes, acylated BHPs, and cyclitol ether BHPs (BHP-CE) (see Results and Discussion and accompanying figures for structures). BHPs with novel side chains are continually being discovered (e.g., Kool et al., 2014; Rush et al., 2016). The diverse nature of BHP lipid structures has led to particular BHPs being used as biomarkers for unique bacterial source organisms (van Winden et al., 2012; Kool et al., 2014; Rush et al., 2014), environmental conditions (Ricci et al., 2014; Welander and Summons, 2012), and organic matter origin (Zhu et al., 2011).

The first reports of BHPs in the natural environments were based on analysis with gas chromatography coupled to mass spectrometry (GC-MS; Rohmer et al., 1980). Traditionally, the GC-MS identification and quantification of BHPs in organic extracts are based on a degradation with periodic acid which converts the intact BHPs into C_{30} – C_{32} primary alcohols, followed by acetylation. Analyses of these derivatized hopanols provide information about the number of functional groups present in the original intact molecule (Rohmer et al., 1984). The analysis of hopanoids by way of GC-MS has been a big step forward in understanding the distribution of BHPs in modern systems and the geological archive. However, by removing all but one functional group from the side chain using the Rohmer reaction, much of the source-, environment-, or process-specific information is lost. In addition to this being a laborious method,

it is also completely unable to detect composite BHPs, leading to potential underestimation of BHP abundance and complexity. To fully elucidate the array of BHP structures, alternative methodologies have been developed. In 2013, Sessions et al. published a high temperature (HT)GC-MS method that achieves elution and separation of more complex acetylated intact BHPs on two different GC columns (BHT, BHpentol and aminotriol on DB-5HT; 2MeBHPs on DB-XLB stationary phase). Though HTGC-MS shows promise, the vast majority of work on intact BHPs, however, has been performed using HPLC-MS.

Schulenberg-Schell et al. (1989) developed a reversed phase HPLC method for analysis of BHPs after acetylation. This method was modified by Talbot et al. (2001), where its applicability was demonstrated in a study of the BHP profiles from a group of methanotrophic bacteria. Advances followed with the application of atmospheric pressure chemical ionization (APCI)/ion trap multi-stage MS, which allowed for more precise control of the fragmentation of the precursor ions. Since 2003, most environmental and culture studies of BHPs have used a version of this reversed phase chromatographic method (e.g., Blumenberg et al., 2007; Saenz et al., 2011; Talbot et al., 2003a, b). The subsequent investigation of a wider range of hopanoid-producing bacterial cultures (Talbot et al., 2003b, c; Talbot et al., 2007a, b; Talbot et al., 2008) led to the improved understanding of the fragmentation pathways in a greater diversity of BHP structures. This allowed for the identification of known BHPs, and related unknown BHPs (e.g., Van Winden et al., 2012, Rush et al., 2016). Recently, several studies were published successfully applying ultra-high pressure liquid chromatography (UHPLC) for improved separation of acetylated BHPs (Kusch et al., 2018, Hemingway et al., 2018), especially isomers of BHT.

The analysis of derivatized BHPs using HPLC-MS has its own disadvantages. The acetylation efficiencies and response factors of individual BHPs vary, making quantitation difficult. Some BHPs, e.g., BHT-CE, acetylate incompletely resulting in the production of several acetylomers complicating data interpretation. With the introduction of improved UHPLC-MS instruments and advances in the quality and diversity of the stationary phases, it became possible to introduce improved methods for analyzing non-derivatized BHPs. Based on an application note (Isaac et al., 2011), non-derivatized BHPs were successfully identified in bacterial isolates and purified culture material using a UHPLC-tandem MS system (Malott et al. 2014; Wu et al, 2015) but these studies did not show the comprehensiveness and sensitivity necessary to cover a wide range of BHPs. Malott et al. (2014) did not report BHPs known to be synthesized

by their investigated organism (i.e., BHT and unsaturated BHT-CE in *Burkholderia* spp. (Cvejic et al., 2000b)) and Wu et al. (2015) reported a reduction in ionization efficiencies of non-acetylated BHPs compared to their acetylated counterparts. Recently, Talbot et al. (2016) reported on the development of an UHPLC method coupled to APCI/triple Quadrupole MS in multi reaction monitoring (MRM) mode for non-derivatized BHPs. However, this work remained limited to a restricted number of already identified BHPs.

Here we report on significant advances in the analysis of non-derivatized BHPs using UHPLC coupled to electrospray ionization (ESI)-high resolution dual-stage MS ($HRMS^2$). Using an approach, which has been successfully applied to the analysis of intact polar lipids (IPLS), we analyzed a number of bacterial species, a.o. *Candidatus* 'Scalindua profunda', *Ca.* 'Methylomirabilis oxyfera', *Methylococcus capsulatus*, and *Methylomarinum vadi*, as well as a soil from an active terrestrial methane seep. We discuss elution and fragmentation behavior of a wide range of known BHPs, including N-containing BHPs and composite BHPs, and the tentative identification of an extensive set of novel BHPs.

2. Materials and Methods

2.1 Sample description

Komagataeibacter xylinus strain R-2277 (formerly *Gluconacetobacter xylinus* and *Acetobacter aceti* ssp. *xylinum*) was obtained as frozen cells in culture medium from an industrial culture (Hoffmann-La Roche, Basel). This culture has been used in previous BHP studies (Peiseler and Rohmer, 1992; Schwartz-Narbonne et al., 2020).

Ca. 'Methylomirabilis oxyfera'. An enrichment culture of *Ca*. 'M. oxyfera' bacteria was obtained from a bioreactor operated under conditions described previously by Ettwig et al. (2009). The bioreactor population consisted of ca. 67% *Ca*. 'M. oxyfera', while the remainder was composed of a mix of ANME-2d archaea and different minor bacteria phyla (see Smit et al., 2019, for details). This culture has been used in previous BHP studies by Kool et al., (2014).

Methylococcus capsulatus. *M. capsulatus* (strain Bath) was obtained from the University of Warick culture collection (as described in Talbot et al., 2001). *M. capsulatus* has been studied in previous BHP studies by Neunlist and Rohmer (1985) and Talbot et al. (2001).

Ca. 'Scalindua profunda'. An enrichment culture of *Ca.* 'S. profunda' was grown in a sequencing batch reactor ase described by van de Vossenberg et al. (2008) and consisted of 80-90% *Ca.* 'S. profunda'. BHPs have been previously studied in this enrichment culture by Rush et al. (2014) and Schwartz-Narbonne et al. (2020).

Methylomarinum vadi (strain IT-4) was isolated from a microbial mat of a shallow (~23 m water depth) marine hydrothermal system in a coral reef off Taketomi Island, Okinawa, Japan (Hirayama et al., 2007; 2013). Cultivation of this strain was performed at JAMSTEC, Japan, using MJmet mediumat pH 6.6 at 37 °C. This culture has been used in previous BHP studies (Rush et al., 2016).

Censo 0 m soil. The Fuoco di Censo seep (37°37'30.1''N, 13°23'15.0''E), in the mountains of Southwestern Sicily, Italy, is a typical example of a natural 'Everlasting Fire' (Etiope et al., 2002; Grassa et al., 2004; Smit et al., 2021). The Censo seep gas consists of mainly thermally generated methane (76-86%). A soil sample was recovered from a horizon 5-10 cm below soil surface directly at the main gas seep (Censo 0 m). The soils were stored in a clean geochemical sampling bag and kept frozen at -20 °C until freeze drying and extraction. Further details can be found in Smit et al. (2021).

2.2 Lipid extraction

Freeze-dried bacterial biomass and the soil from the Censo seep were extracted using a modified Bligh and Dyer method (Bligh and Dyer, 1959; Bale et al., 2013). The samples were ultrasonically extracted (10 min) with a solvent mixture containing methanol (MeOH), dichloromethane (DCM) and phosphate buffer (2:1:0.8, v:v:v). Solvent was collected after centrifugation and the residues re-extracted twice. A biphasic separation was achieved by adding additional DCM and phosphate buffer to the combined extracts in a ratio of MeOH, DCM and phosphate buffer (1:1:0.9, v:v:). After the DCM layer was collected, the aqueous layer was washed twice with DCM. Combined DCM layers were dried under a continuous flow of N_2. Prior to injection, extracts were redissolved in MeOH:DCM (9:1) and filtered through a 0.45 µm regenerated cellulose syringe filter (4 mm diameter; Grace Alltech, Deerfield, IL).

2.3 UHPLC/HRMS

The method described here (adapted from Wörmer et al. (2013)) is the final method used. Steps in method development are discussed below in Results and

Discussion. All analysis were performed using an Agilent 1290 Infinity I UHPLC, equipped with thermostatted autosampler and column oven, coupled to a quadrupole-orbitrap HRMS equipped with an Ion Max source and heated ESI probe (HESI) (ThermoFisher Scientific, Waltham, MA). Separations were achieved using an Acquity C_{18} BEH column (2.1x 150 mm, 1.7 µm particle; Waters), fitted with a pre-column, and a solvent system consisting of (A) methanol:H_2O (85:15) and (B) methanol:isopropanol (1:1), both containing 0.12% (v/v) formic acid and 0.04% (v/v) aqueous ammonia. Compounds were eluted with 5% B for 3 min, followed by a linear gradient to 40% B at 12 min and then to 100% B at 50 min, with a total run time of 80 min. The flow rate was 0.2 mL min^{-1}. Positive ion HESI settings were: capillary temperature, 300 °C; sheath gas (N_2) pressure, 40 arbitrary units (AU); auxiliary gas (N_2) pressure, 10 AU; spray voltage, 4.5 kV; probe heater temperature, 50 °C; S-lens 70 V. Detection was achieved using positive ion monitoring of *m/z* 350-2000 (resolving power 70,000 ppm at *m/z* 200), followed by data dependent MS^2 (isolation window 1 *m/z*; resolution 17,500 ppm at *m/z* 200) of the 10 most abundant ions, and dynamic exclusion (6 s), with 3 ppm mass tolerance. In addition, an inclusion list of 165 calculated exact masses of BHPs from literature was used. Optimal fragmentation of BHPs was achieved using a stepped normalized collision energy of 22.5 and 40. A mass calibration was performed every 48 h using the Thermo Scientific Pierce LTQ Velos ESI Positive Ion Calibration Solution. All mass chromatograms were produced within 3 ppm mass accuracy, unless otherwise stated.

3. Results and Discussion

3.1 UHPLC method development

The suitability of C_{18} columns for the retention and separation of several non-derivatized BHPs was previously demonstrated by Talbot et al. (2016). Two types of C_{18} column were tested in that study: an Acquity BEH UHPLC column and an Ace base-deactivated Excel UHPLC column, of which the latter showed the best chromatographic behavior, particularly for N-containing BHPs. Previously, the Acquity BEH C_{18} column was selected by Wörmer et al. (2013) in a new UHPLC-ESI/MS method for the analysis of IPLs. As BHPs, like IPLs, consist of a large apolar moiety and a polar functionalized side chain with or without an additional, often polar, moiety or head group, we set out to test whether the chromatographic system, combined with positive ion ESI, as described by Wörmer et al. (2013) for IPLs, is suitable for the analysis of non-

derivatized BHPs. For this purpose, we investigated extracts of biomass of various known BHP-producing micro-organism. The exact distribution of the detected BHPs and their fragmentation spectra are discussed in detail in section 3.3. Here we only discuss the general features of the chromatographic method. BHT was easily identified in several extracts based on its exact mass and diagnostic spectra, however aminotriol and -tetrol produced very broad peaks and were thus difficult to detect, similarly as was described by Talbot et al. (2016). The buffering system according to Wörmer et al. (2013) consists of 0.04% formic acid and 0.1% NH_3, producing a mobile phase with a pH 6. The pH is often a driving factor in the chromatographic behavior of N-containing compounds and therefore a more acidic buffering system with 0.12% formic acid and 0.04% NH_3, which has also been frequently used in IPL HPLC-MS analysis (cf. Sturt et al., 2004) was tested. This change in the pH of the mobile phase to 4 resulted in excellent peak shape for the N-containing BHPs using the Acquity BEH C_{18} column, allowing detection of e.g., aminotriol and -tetrol, as well as adenosylhopane BHPs. Use of the base deactivated Ace Excel C_{18} column in combination with the mobile phase with pH 4 did not result in improved chromatography and thus, for all experiments described below, a combination of the Acquity BEH C_{18} column combined with the pH 4 buffering system was used.

3.2 MS method development

MS detection of derivatized BHPs, in general, has been done by using positive ion APCI combined with ion trap MS^2 (Talbot et al. 2003a,b, 2007a,b). For non-derivatized BHPs, Talbot et al. (2016) employed APCI combined with MRM using a triple quadrupole MS. Here we explored the use of positive ion ESI combined with a quadrupole-Orbitrap MS for untargeted detection and identification of BHPs with data dependent MS^2 analysis using HRMS, similar to the approach that has successfully been applied to the analysis of IPLs (Moore et al., 2016; Besseling et al., 2018; Bale et al., 2019). For non-targeted IPL analysis we apply a stepped collision energy of 15, 22.5 and 30 to produce informative and diagnostic spectra (e.g., Bale et al., 2018; Sollai et al., 2018). For relatively simple BHPs, such as BHT, BHpentol and BHhexol, these settings resulted in good quality MS^2 spectra, but especially for N-containing BHPs such as aminotriol and -tetrol it gave insufficient fragmentation to obtain an informative MS^2 spectrum. Talbot et al. (2016) also reported the necessity for a higher fragmentation energy for the N-containing BHPs. However, exploring

this option led to the MS^2 spectra of non N-containing BHPs having too much fragmentation, and thus loss of information on the structure of the side chains. A workable compromise was achieved by using a stepped collision energy of 22.5 and 40, resulting in MS^2 spectra with diagnostic losses for the different possible functional groups in the tail and diagnostic fragmentation for the hopanoid core. All fragmentation spectra discussed below were produced using these settings.

3.3. Commonly detected BHPs in various microorganisms

Talbot et al. (2016) reported MS^2 spectra under various fragmentation conditions for non-derivatized BHT. However, no spectra were reported for methylated BHTs, unsaturated BHTs or BHpentol and BHhexol. To establish the chromatographic behavior and fragmentation of various BHPS we analyzed biomass of several known BHP producers.

3.3.1. Bacteriohopanetetrols

Biomass of *K. xylinus*, which has been reported to produce multiple (3Me)BHT isomers with up to two unsaturations at Δ^6 and/or Δ^{11} (Peisler and Rohmer,1992, Talbot et al., 2007b), was analyzed to study the various isomers and homologues of BHT. Schwartz-Narbonne et al. (2020) already showed the separation of different stereoisomers of non-derivatized BHTs (i.e., BHT-34*S*, BHT-34*R* and BHT-x) using the method described here. Analysis of biomass of *K. xylinus* provided further evidence that various stereoisomers of BHT are (partially) separated under the UHPLC conditions in this study (Fig. 1). Table S1 lists all BHPs discussed below and provides details on preferred ionization, in-source fragmentation and diagnostic fragments.

The summed mass chromatogram of the calculated exact masses of the protonated ($[M+H]^+$), ammoniated ($[M+NH_4]^+$) and sodiated molecules ($[M+Na]^+$) of BHT (*m/z* 547.472, *m/z* 564.499, *m/z* 569.454, respectively) revealed the presence of multiple isomers of BHT (Fig. 1A). The two peaks at 20.8 and 21.1 min (peaks b and c) match the retention time for BHT with the 22*R*,34*S* and 22*R*,34*R* stereochemistry, respectively, as established by Schwarz-Narbonne et al. (2020). The most abundant isomer of BHT (i.e., labeled a), eluting early at 19.3 min, may be BHT-22*S*,34*S*, which was reported in relatively high abundance (13% of all BHTs) in *K. xylinus* by Peiseler and Rohmer (1992). In addition to these abundant isomers, there are several isomers present at trace levels (not visible at the scale of Fig. 1) eluting after both the early eluting BHT isomer and the BHT-22*R*,34*S*/*R*.

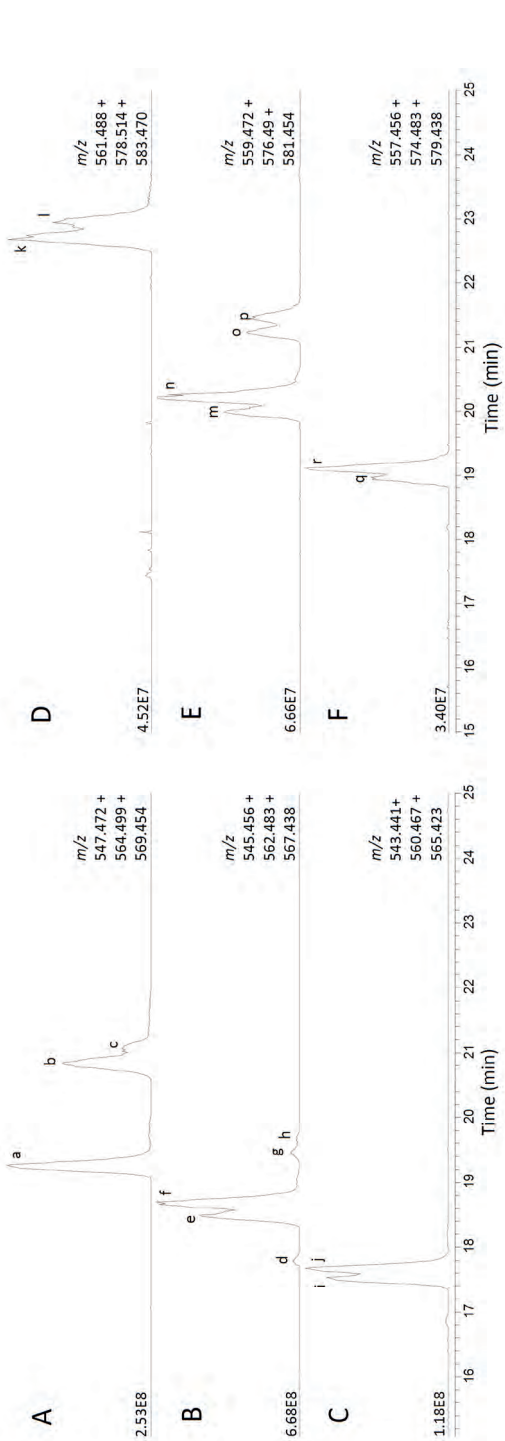


Figure 1. Partial mass chromatograms of A: Bacteriohopanetetrol (BHT), B: unsaturated BHT, C: diunsaturated BHT, D: 3MeBHT, E: unsaturated 3MeBHT, and F: diunsaturated 3MeBHT in *K. xylinus*. All chromatograms represent the summed signals of $[M+H]^+$, $[M+NH_4]^+$ and $[M+Na]^+$. Each trace is labeled with the exact masses used for searching, and the intensity in arbitrary units (AU) of the signal.

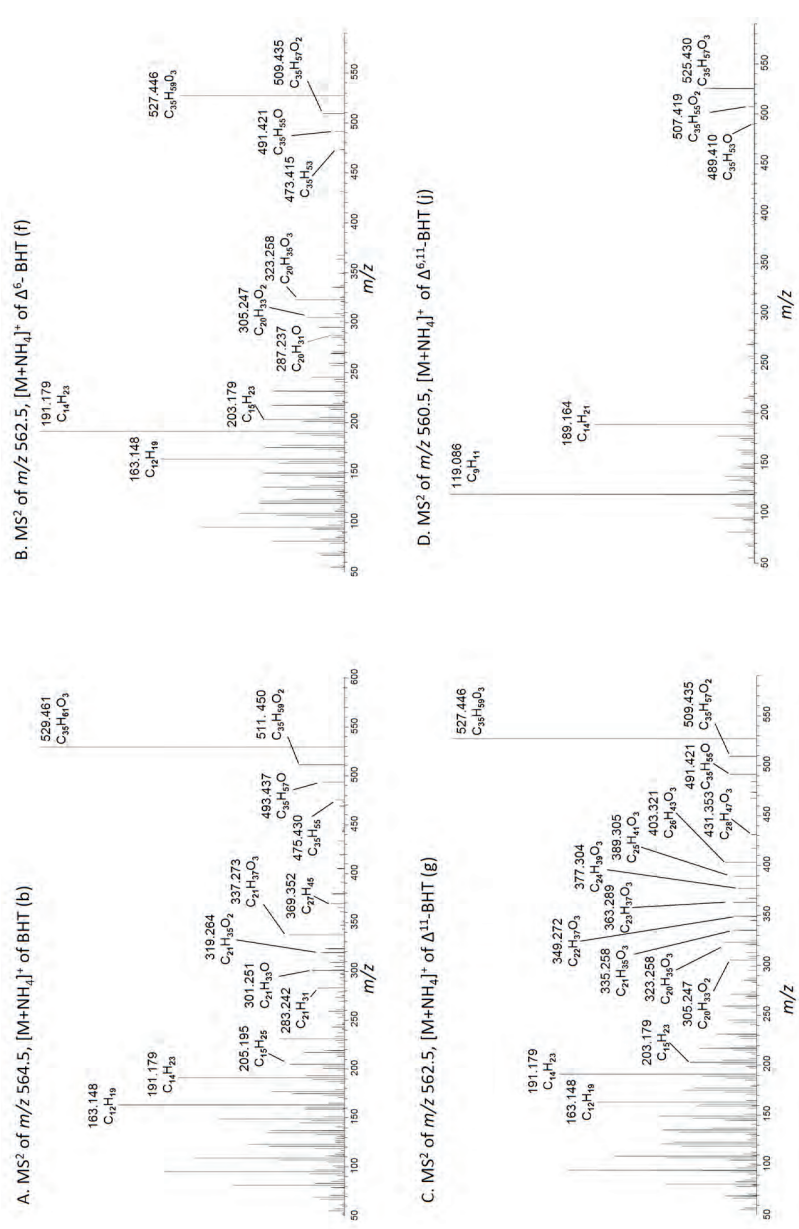


Figure 2. MS^2 spectra of A: Bacteriohopanetetrol (BHT) (peak b, Fig. 1), B: Δ^6-BHT (peak f, Fig. 1), C: Δ^{11}-BHT (peak g, Fig. 1), and D: $\Delta^{6,11}$-BHT (peak j, Fig. 1). Assigned elemental composition of diagnostic mass peaks are listed in Table S1.

All the BHT isomers described above had similar MS^2 spectra, therefore we only discuss the MS^2 spectrum of BHT-22*R*,34*S* (peak b, Fig. 2A), obtained from the ammoniated molecule (*m/z* 564.5). The MS^2 spectrum obtained here is similar to that reported by Talbot et al. (2016) for non-derivatized BHT upon MS^2 fragmentation of *m/z* 529 (formed after in source loss of H_2O from protonated BHT ($[M+H-H_2O]^+$)). Here, *m/z* 529.461 ($C_{35}H_{61}O_3^+$, Δ ppm -0.65) is the base peak. Fragment ions at *m/z* 511.450 ($C_{35}H_{59}O_2^+$, Δ ppm -1.56), *m/z* 493.437 ($C_{35}H_{57}O^+$, Δ ppm -6.73), and *m/z* 475.430 ($C_{35}H_{55}^+$, Δ ppm -0.80) represent consecutive losses of hydroxyl moieties as H_2O (see Table S1 for assigned elemental composition (AEC) of fragments). In the lower mass range, *m/z* 163.148 ($C_{12}H_{19}^+$, Δ ppm -0.17) is the dominant fragment, but the universal diagnostic fragment ion for hopanoids (*m/z* 191.179, $C_{14}H_{23}^+$, Δ ppm -0.56) (Peters and Moldowan, 1993) is clearly present and more dominant than observed in the spectrum of Talbot et al. (2016). Unlike electron ionization (EI) fragmentation, collision-induced dissociation (CID) fragmentation typically does not produce radicals but one protonated fragment and a neutral. Any cleavage results in proton rearrangements and the formation of a double bond equivalent in one of the two fragments (for discussed fragmentation pathways and fragments, see Fig. 3). In case of BHT, the *m/z* 191 fragment represents the protonated A and B ring of the hopanoid structure with 2 double bonds (unlike the structure proposed by Talbot et al. (2003b)), generated by the double cleavage at C-9/C-11 and C-8/C-14 following fragmentation pathway A. Additional diagnostic fragments are found at *m/z* 369.352 (complete hopanoid ring system after loss of side chain; $C_{27}H_{45}^+$, Δ ppm -0.21) and at *m/z* 283.242 ($C_{21}H_{31}^+$, Δ ppm -1.58), *m/z* 301.251 ($C_{21}H_{33}O^+$, Δ ppm -3.63), 319.264 ($C_{21}H_{35}O_2^+$, Δ ppm 3.52), and 337.273 ($C_{21}H_{37}O_3$+, Δ ppm -1.19). These latter C_{21} fragments are complementary to the fragment at *m/z* 191 and represent the D and E rings and side chain with 0 to 3 alcohol moieties remaining (fragmentation A', Fig. 3). Talbot et al. (2003a,b) observed the equivalent fragments for acetylated BHT.

3.3.2 Methylated BHTs

The distribution of the 3MeBHTs in *K. xylinus* in general followed the distribution of the BHTs and they elute roughly 2 min after their BHT-counterparts but were ca. an order of magnitude less abundant. Two isomers of 3MeBHT were detected (peaks k and l, Fig. 1D) and based on their relative retention time were identified as 3MeBHT-22*R*,34*S* and 3MeBHT-22*R*,34*R*,

respectively. Peiseler and Rohmer (1992) reported only the 3MeBHT-22*R*,34*S* in *K. xylinus*, but did observe the *22R*,34*R* stereoisomer in the acetic acid bacterium *A. pasteurianus*. The MS^2 spectra (Fig. S2A) showed the expected offset of +14 Da in the fragments containing the A ring with the methylation (e.g., *m/z* 177.164 ($C_{13}H_{21}^+$, Δ ppm -0.55), *m/z* 205 ($C_{15}H_{25}^+$, Δ ppm -0.33), *m/z* 219 ($C_{16}H_{27}^+$, Δ ppm -0.40) and *m/z* 383 ($C_{28}H_{47}$, Δ ppm 1.52), while fragments related to the D and E rings with side chain did not show the offset (e.g., *m/z* 283.242 ($C_{21}H_{31}^+$, Δ ppm 1.21), *m/z* 301.253 ($C_{21}H_{35}O_2^+$, Δ ppm 1.86), *m/z* 319.263 ($C_{21}H_{35}O_2^+$, Δ ppm 0.48), *m/z* 337.274 ($C_{21}H_{37}O_3^+$, Δ ppm -0.54).

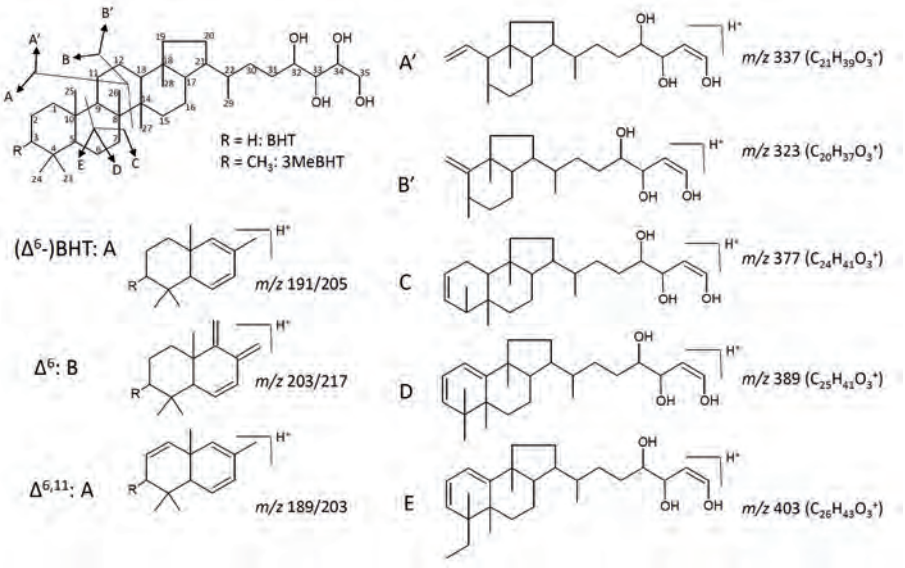


Figure 3. Fragmentation pathways and structures for proposed diagnostic fragments after collision induced fragmentation for BHT, Δ^6- BHT, Δ^{11}- BHT , $\Delta^{6,11}$-BHT, and their 3Me-counterparts.

3.3.3. Unsaturated BHTs

We also detected several isomers of unsaturated BHTs in biomass of *K. xylinus* (Fig. 1B). The summed mass chromatogram of the calculated exact masses of the protonated, ammoniated, and sodiated molecules of unsaturated BHT (*m/z* 545.456 + 562.483 + 567.438) showed 2 main peaks at 18.5 and 18.7 min (e and f), as well as a minor early eluting isomer at 17.8 min (d) and a pair of late eluting isomers at 19.5 and 19.6 min (g and h). Peiseler and Rohmer (1992) reported two isomers of Δ^6-BHT (22*R*,34*S* and 22*R*,34*R*) in relatively high abundance in *K. xylinus* (20 and 36% respectively of total BHTs). Based on this distribution it is likely that peaks e and f represent Δ^6-BHT-22*R*,34*S* and Δ^6-BHT-22*R*,34*R*,

respectively. The MS^2 spectra of the ammoniated molecules (*m/z* 562.5) associated with peaks e and f (e.g., Fig. 2B) are almost identical. The HPLC-MS^2 analysis of acetylated unsaturated BHTs was extensively discussed by Talbot et al. (2007b), but not for non-derivatized BHTs (Talbot et al. 2016). The unsaturation in Δ^6-BHT-22*R*,34*R* (peak f, Fig. 2B) is clearly observed in several fragments formed after consecutive losses of H_2O from the side chain, that are offset by 2 Da from the spectrum of BHT, e.g., *m/z* 527.466 ($C_{35}H_{59}O_3^+$, Δ ppm -0.36), 509. 435 ($C_{35}H_{57}O_2^+$, Δ ppm -0.37), 491.421 ($C_{35}H_{55}O^+$, Δ ppm -6.76) and 473.415 ($C_{35}H_{53}^+$, Δ ppm 1.88). Interestingly, and similar as described by Talbot et al. (2007), no such offset is observed in the diagnostic fragment ion at *m/z* 191.179 ($C_{14}H_{23}^+$, Δ ppm -0.88). While in case of BHT the fragment at m/z 191 is produced by cleavage at C-9/C-11 and C-8/C-14 and a proton rearrangement resulting in 2 double bonds in the A and B Ring fragment (as discussed above), it appears that the presence of a pre-existing double bond at Δ^6 results in an alternative proton rearrangement: only one additional double bond is formed in the A and B ring fragment, thus generating the *m/z* 191 also seen in BHT, and the other double bond is formed in the neutral loss fragment (fragmentation A, Fig. 3). Talbot et al. (2007b) also observed evidence for an alternative fragmentation pathway in acetylated Δ^6-BHT, with cleavage of the C-11/C-12 and C-8/C-14 bonds (pathway B, Fig. 3). With proton rearrangement to form two additional double bonds on the A and B ring fragment this fragmentation would produce the here observed *m/z* 203.179 ($C_{15}H_{23}^+$, Δ ppm -0.233; Fig. 3). A complementary series of C_{20} fragments with 1 to 3 hydroxyl moieties are observed at *m/z* 287.237 ($C_{20}H_{31}O^+$, Δ ppm 0.41), 305.247 ($C_{20}H_{33}O_2^+$, Δ ppm -0.42) and 323.258 ($C_{20}H_{35}O_3^+$, Δ ppm -0.44) and represent the D and E rings with the (partially dehydroxylated) side chain. The MS^2 spectrum of the early eluting isomer (d) is almost identical to those of peaks e and f. Based on this similarity in spectra and the off-set in retention time between the 22S,34R and 22*R*,34*R* isomers observed for BHT, we tentatively identify the early eluting isomer (d) as the 22*S*,34*S* isomer of Δ^6-BHT.

The MS^2 spectra of peaks g and h (Fig. 1B) are largely similar to those of peaks d-f, however there are some notable differences in the middle mass range (Fig. 2C), which shows a series of C_{20} to C_{28} fragments with up to 3 alcohol moieties. It is possible that the presence of a double bond in the C ring leads to a multitude of alternative fragmentation pathways involving the A and B rings. For example, the larger fragments (>C_{24}) appear to originate from cleavage of the C-9/C-10 bond in combination with cleavage at C-5/C-6, C-6/C-7, or C-7/C-

8 (Fig. 3, pathways C, D, E). We, therefore, tentatively identify peaks g and h as Δ^{11}- BHT-22*R*,34*S* and Δ^{11}-BHT-22*R*,34*R*. Peiseler and Rohmer (1992) did not report Δ^{11}-BHT in *K. xylinus*, although two isomers of Δ^{11}-3MeBHT were reported in that culture.

The summed mass chromatogram revealing the distribution of di-unsaturated BHTs (Fig. 1C) shows two main peaks at 17.5 and 17.7 min (peaks i and j), most likely the 34*S* and 34*R* stereoisomers of $\Delta^{6,11}$-BHT as described by Peiseler and Rohmer (1992). MS^2 spectra of the ammoniated molecule of di-unsaturated BHTs (*m/z* 560.5, Fig. 2D) reveal double unsaturations in the fragments that result from the loss of hydroxyl moieties from the side chain (e.g., *m/z* 525.430 ($C_{35}H_{57}O_3^+$, Δ ppm 0.03), *m/z* 507.419 ($C_{35}H_{55}O_2^+$, Δ ppm -0.98), and *m/z* 489.410 ($C_{35}H_{53}O^+$, Δ ppm 1.61)). However, the MS^2 is dominated by fragments at *m/z* 119.086 ($C_9H_{11}^+$, Δ ppm 0.03) and *m/z* 189.164 ($C_{14}H_{21}^+$, Δ ppm -0.36), similar to what was observed by Talbot et al. (2007b) for MS^2 spectra of acetylated di-unsaturated BHT. Rohmer and Ourisson (1986) attributed a fragment at m/z 119, formed after EI from both Δ^6- and $\Delta^{6,11}$-BHPs, to the B-ring and assigned a trimethylbenzene structure. However, a trimethylbenzene would yield, if protonated, a fragment at *m/z* 121 ($C_9H_{13}^+$) and thus is seems the fragment at *m/z* 119 formed upon CID fragmentation, represents a different, and as yet unknown, fragment. The fragment at *m/z* 189 likely represents the A and B rings. Although formation of an A and B ring fragment with 3 double bonds does not appear to be favored in case of a Δ^6 -unsaturation, an additional unsaturation at Δ^{11} apparently forces this proton rearrangement, thus producing a fragment at *m/z* 189 (pathway A, Fig. 3). It is therefore important to recognize that with CID, *m/z* 189 is not diagnostic for a Δ^6-unsaturation (as is the case with EI fragmentation), but instead indicates a $\Delta^{6,11}$ -unsaturation. The fragments in the middle mass range in the MS^2 spectrum of the di-unsaturated BHT are less abundant (<1%), with minor fragments representing cleavages at C9-C11 (pathway A, Fig. 3) and C-11/C-12 (pathway B, Fig. 3) both present.

3.3.4 Unsaturated methylated BHTs

The summed mass chromatogram for the unsaturated 3MeBHT showed two pairs of peaks at 20.0 and 20.2 min (peaks m and n) and at 21.2 and 21.4 min (peaks o and p) (Fig. 1E). The MS^2 spectra of peaks m and n (Fig. S2b) showed very similar fragmentation to the Δ^6-unsaturated BHTs, with the expected +14 Da offsets in the A and B ring fragments. The fragments in the middle mass

range were largely absent. The spectra of peaks o and p (Fig. S2c) showed the same series of C_{20} to C_{26} fragments as observed for the two unsaturated BHT peaks (g and h) designated as Δ^{11}. Based on the relative retention times and similarities of their spectra to the unsaturated BHTs we identify these peaks as (m) Δ^{6}-3MeBHT-22*R*,34*S*, (n) Δ^{6}-3MeBHT-22*R*,34*R*, (o) Δ^{11}-3MeBHT-22*R*,34*S*, and (p) Δ^{11}-3MeBHT-22*R*,34*R*. All but Δ^{11}-3MeBHT-22*R*,34*R* were also reported by Peiseler and Rohmer (1992) in *K. xylinus*.

We also detected two isomers of di-unsaturated 3Me-BHT at 19.00 and 19.12 min (Fig. 1F, peaks q and r). Peiseler and Rohmer (1992) also detected two isomers of $\Delta^{6,11}$ -3MeBHT in *K. xylinus*, differing in stereochemistry at the C-34. Based on this and the relative retention times we identify peaks q and r as $\Delta^{6,11}$-3MeBHT-22*R*,34*S* and $\Delta^{6,11}$-3MeBHT-22*R*,34*R*, respectively. The observed fragmentation spectrum (Fig. S2d) closely matched those of $\Delta^{6,11}$-BHT. It is noteworthy that the fragment at *m/z* 203 (analogous to *m/z* 189 in the MS^2 spectrum of $\Delta^{6,11}$-BHT) is diagnostic for a di-unsaturated MeBHT and not for Δ^{6}-MeBHT as is the case with EI ionization. However, the *m/z* 203 fragment alone is not sufficient to identify a $\Delta^{6,11}$-MeBHT as a fragment with identical *m/z* is also present in the fragmentation spectrum of Δ^{6}-BHT, although it is generated via a different fragmentation pathway.

3.3.5. Bacteriohopanepentols and -hexols

In order to establish elution and fragmentation patterns for bacteriohopanepentol and -hexol (BHpentol and BHhexol, respectively), as well as their 3Me-homologues, we analyzed an extract of *Ca.* 'M. oxyfera' (Kool et al., 2014). Fig. 4A shows summed mass chromatograms revealing the presence of BHT (peak a), BHpentol (peak b) and BHhexol (peak c) and their methylated analogues (peaks d, e, and f). The BHPs elute in reversed order of number of hydroxylations on the side chain. Whereas ammoniation is the preferred ionization for BHT, the balance shifts towards protonation for BHhexol (Table S1). Fig. 4D shows the MS^2 spectra of the protonated molecule ($[M+H]^+$) of BHhexol. The observed fragmentation pattern is comparable to the one described above for BHT, with major fragments representing losses of 1 to 5 hydroxyl moieties at *m/z* 561.451 ($C_{35}H_{61}O_5^+$, Δ ppm -0.86), 543.441 ($C_{35}H_{59}O_4^+$, Δ ppm -0.14), 525.431 ($C_{35}H_{57}O_3^+$, Δ ppm 1.65), 507.421 ($C_{35}H_{55}O_2^+$, Δ ppm 0.16), and 489.406 ($C_{35}H_{53}O^+$, Δ ppm -6.65), respectively. The base peak is the diagnostic ion for BHPs at *m/z* 191.179 ($C_{14}H_{23}^+$, Δ ppm -1.50). The fragment at *m/z* 369.353 ($C_{29}H_{49}^+$, Δ ppm -2.51) representing the ring system after loss of the side chain is only present in very low abundance (<5%), but instead a fragment at *m/z*

397.383 ($C_{29}H_{49}^+$, Δ ppm 0.31) is formed with the cleavage apparently occurring at the C-22/C-30 bond. Similar to what was observed for BHT, we observe several fragments at *m/z* 297.222 ($C_{21}H_{29}O$, Δ ppm 3.56), *m/z* 315.233 ($C_{21}H_{31}O_2$, Δ ppm 2.96), *m/z* 333.243 ($C_{21}H_{33}O_3$, Δ ppm 1.59) and *m/z* 351.253 ($C_{21}H_{35}O_4$, Δ ppm -0.33), which based on their assigned elemental composition, appear to be complimentary to the ion at *m/z* 191 (A and B ring) and represent the remainder of the ring system and the side chain with 0 to 4 hydroxyl moieties.

The MS^2 spectrum of the ammoniated molecule for BHpentol (Fig. 4B) shows the same characteristics as discussed for BHT and BHhexol. 3MeBHpentol and 3MeBHhexol elute ca. 2 min after and fragment similarly to their non-methylated counterparts, generating a.o. the diagnostic fragment at *m/z* 205 and with all ions containing the A-ring shifted by +14 Da (see Fig. S2 for mass spectra).

3.3.6. Aminobacteriohopanepolyols

Mass chromatograms of the protonated molecules of aminotriol, -tetrol, and -pentol (Fig. 5A; peaks a, b, c, and d, respectively), as well as their 3Me-counterparts (Fig. 5B; peaks e, f, and g, respectively) show their distribution in *M. capsulatus*, a well-studied producer of these aminoBHPs (e.g., Neunlist and Rohmer, 1985; Talbot et al., 2001). As expected, the aminoBHPs elute in reversed order of the number of functional groups. The 3Me-aminoBHPs elute ca. 1 to 2.5 min after their corresponding non-methylated counterparts (see Table S1 for exact retention times). Fig. 5A shows a partially resolved, late-eluting isomer of aminotetrol (peak c). Although Talbot et al. (2001) observed a late eluting isomer to aminotetrol in *Methylocystis parvus*, it has to the best of our knowledge not been observed in *M. capsulatus.*

Talbot et al. (2016) discussed the fragmentation characteristics of non-derivatized aminotriol only and, based on those observations, predicted suitable MRM target ions for (3Me)aminotetrol and (3Me)aminopentol. Here we show MS^2 spectra of aminopentol (Fig. 5C) and 3Me-aminopentol (Fig. 5D) (additional spectra for the aminotriols and -tetrols are shown in Fig. S4; the late eluting isomer of aminotetrol shows similar fragmentation as the main isomer). As was observed by Talbot et al. (2003a) for acetylated aminotriol and Talbot et al. (2016) for non-derivatized aminotriol, fragmentation of the amino-BHPs only occurs with increased collision voltage. Under the conditions used here, i.e., with a stepped collision energy, fragmentation was still limited, but yielded sufficient diagnostic features for positive identification of the full series of aminoBHPs.

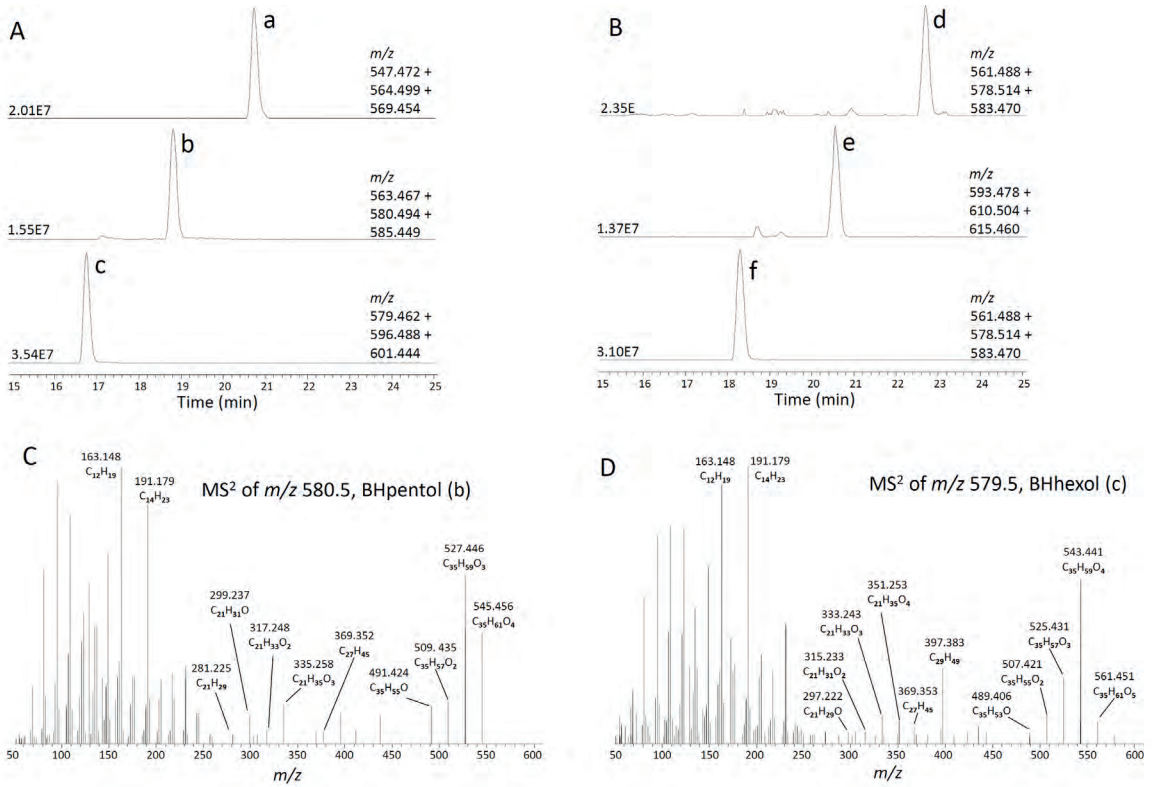


Figure 4. Bacteriohopanepolyols in *Ca*. 'M. Oxyfera'. A: Partial mass chromatograms of BHT (a), Bhpentol (b), and BHhexol (c). B: Partial mass chromatograms of 3MeBHT (d), 3MeBHpentol (e), and 3MeBHhexol (f). All chromatograms represent the summed signals of $[M+H]^+$, $[M+NH_4]^+$ and $[M+Na]^+$. Each trace is labeled with the exact masses used for searching and the intensity in arbitrary units (AU) of the signal. C: MS^2 spectrum of $[M+NH_4]^+$ of BHpentol, and D: MS^2 spectrum of $[M+H]^+$ of BHhexol.

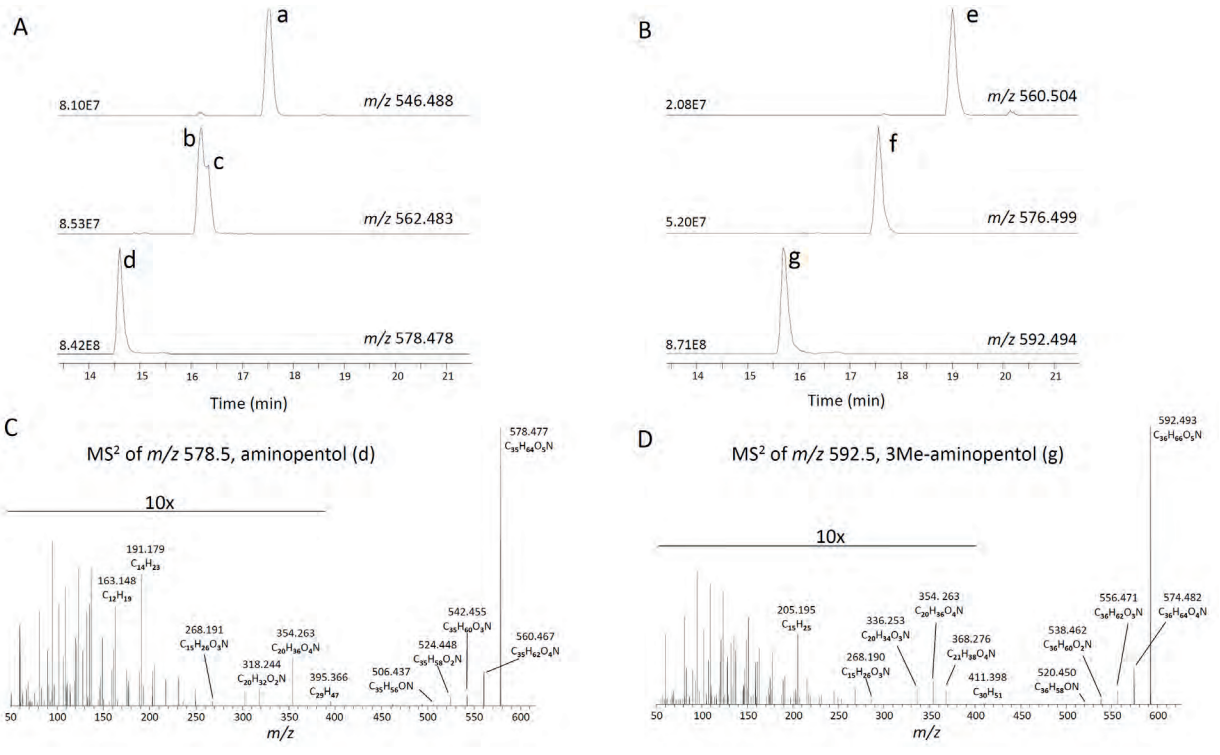


Figure 5. Aminobacteriohopanepolyols in M. capsulatus. A: Partial mass chromatograms of aminotriol (a), aminotetrol (b, c), and aminopentol (d). B. Partial mass chromatograms of 3Me-aminotriol (e), 3Me-aminotetrol (f), and 3Me-aminopentol (g). Each trace is labeled with the exact masses used for searching and the intensity in arbitrary units (AU) of the signal. All compounds are detected as their $[M+H]^+$. C: MS^2 spectrum of aminopentol, and B: MS^2 spectrum of 3Me-aminopentol. Spectra are partially magnified by the factor indicated.

As is the case with the previous discussed BHPs, the alcohol moieties in the side chain are readily lost as H_2O, yielding a series of fragments at *m/z* 506.437 ($C_{35}H_{62}O_4^+$, Δ ppm 2.19), *m/z* 524.448 ($C_{35}H_{58}O_2N^+$, Δ ppm 0.48), *m/z* 542.455 ($C_{35}H_{60}O_3N^+$, Δ ppm -2.99), and *m/z* 560.467 ($C_{35}H_{62}O_4N^+$, Δ ppm 0.65) for aminopentol and *m/z* 520.450 ($C_{36}H_{58}ON^+$, Δ ppm -2.58), *m/z* 538.462 ($C_{36}H_{60}O_2N^+$, Δ ppm -0.12), *m/z* 556.471 ($C_{36}H_{64}O_2N^+$, Δ ppm -1.98), and *m/z* 574.482 ($C_{36}H_{66}O_4N^+$, Δ ppm -1.09) for 3Me-aminopentol. In the lower mass range the diagnostic fragments at *m/z* 191.179 ($C_{14}H_{23}^+$, Δ ppm -3.75) for BHPs and *m/z* 205.195 ($C_{15}H_{25}^+$, Δ ppm -1.40) for MeBHPs are relatively abundant. In addition, there are several N-containing fragments in the mass range between *m/z* 300 and 400 (Figs. 5C and D, Table S1) that represent the D and E rings with side chain, equivalent to the fragmentation observed for BHPs described above.

3.4 Composite BHPs

3.4.1 Cyclitol ether bacteriohopanetetrol

BHT-CE is a commonly detected, so-called composite BHP (Talbot et al., 2007a). Composite BHPs consist of a linear functionalized side chain bound to a more complex, often polar moiety or head group. In case of BHT-CE, the BHT is ether bound to an amino sugar on the C-35 position (Fig. 6B). Here we analyzed cell material of *Ca.* 'Scalindua profunda', in which BHT-CE was previously detected by Rush et al. (2014), to establish elution and fragmentation of this BHP. Fig. 6A shows the mass chromatogram for the $[M+H]^+$ of BHT-CE (*m/z* 708.541). Whereas Rush et al. (2014) detected 3 isomers of BHT-CE in *Ca.* 'S. profunda', eluting closely together, we detected 2 isomers (Fig. 6A, peak a and b). Fig. 6B shows the MS^2 spectrum of the most abundant isomer, peak b. Talbot et al. (2016) also discussed the fragmentation of non-derivatized BHT-CE, which is very similar to the fragmentation pattern observed here. Fragmentation is limited, but several fragments resulting from loss of up to three hydroxyl moieties are observed at *m/z* 690.529 ($C_{41}H_{72}O_7N^+$, Δ ppm -1.32), *m/z* 672. 529 ($C_{41}H_{70}O_6N^+$, Δ ppm -0.65) and *m/z* 654.508 ($C_{41}H_{68}O_5N^+$, Δ ppm -2.60). In the lower mass range several fragments representing the intact headgroup (*m/z* 180.087 ($C_6H_{14}O_5N^+$, Δ ppm -0.72)) and the headgroup after loss of several alcohol moieties (*m/z* 162.076 ($C_6H_{12}O_4N^+$, Δ ppm -1.45), *m/z* 144.066 ($C_6H_{10}O_3N^+$, Δ ppm -0.62) and *m/z* 126.065 ($C_6H_8O_2N^+$, Δ ppm 0.28)) are observed. Fragments at *m/z* 222.097 ($C_8H_{16}O_6N^+$, Δ ppm -1.37) and m/z 204.86 ($C_8H_{14}O_5N^+$, Δ ppm -0.98) appear to consist of the headgroup with two additional carbons from the side chain after fragmentation at the C-33/C-34 bond.

3.4.2 Methylcarbamate-aminoBHPs

A more recently described series of composite BHPs are the methylcarbamate-aminoBHPs (MC-aminoBHPs; Rush et al., 2016). Fig. 7A shows the series detected in *M. vadi*. The MS^2 spectrum of MC-aminotriol (Fig. 7B; spectra of MC-aminotetrol and –pentol are shown in Fig. S5) showed the predicted losses of up to 3 hydroxyl moieties as H_2O generating ions at *m/z* 586.483 ($C_{37}H_{64}O_4N^+$, Δ ppm -0.23), 568.474 ($C_{37}H_{62}O_3N^+$, Δ ppm 2.38), and 550.461 ($C_{37}H_{60}O_2N^+$, Δ ppm -2.99). A loss of 32 Da (CH_3OH, representing the methoxy moiety) from the ion at *m/z* 586 to produce m/z 554.456 ($C_{36}H_{60}O_3N^+$, Δ ppm -1.64) was also observed. Losses of 75 Dalton ($C_2H_5O_2N$), representing the methylcarbamate moiety, were observed from the mass peaks at *m/z* 568 or 550 leading to fragment ions at *m/z* 493.438 ($C_{35}H_{57}O^+$, Δ ppm -4.81) and *m/z* 475.430 ($C_{35}H_{55}$, Δ ppm 0.17). These losses were also observed by Rush et al. (2016) after fragmentation of acetylated MC-aminoBHPs. N-containing product ions were observed at *m/z* 76.040 ($C_2H_6O_2N^+$, Δ ppm 9.40), 88.040 ($C_3H_6O_2N^+$, Δ ppm 5.85) and 118.050 ($C_4H_8O_3N^+$, Δ ppm 1.61) and likely represent the methylcarbamate moiety after fragmentation between the nitrogen and C-35, at C-34/C-35, and at C-33/C-34, respectively.

3.5 BHPs in a soil from a terrestrial methane seep

To further examine the performance of the UHPLC-HRMS method for environmental samples with more complex matrices than biomass, we analyzed a soil from a terrestrial methane seep in Sicily (Censo 0 m). A base peak chromatogram is shown in Fig. S6. We were able to detect several of the above discussed BHPs. All detected BHPs are listed in Table S2 and additional MS^2 spectra of BHPs that are not further discussed in the text are shown in supplemental figures. Here we will focus on the description of not previously discussed and new composite BHPs.

3.5.1. AdenosylBHPs

Adenosylhopanes occur ubiquitously in soils (Cooke et al., 2008a,b). Their occurrence and relative abundance have been used to trace soil organic matter during riverine transport and deposition into the marine environment (Zhu et al., 2011). Three types of adenosylhopanes have been described: type-1, -2, and -3 (see Fig. 9A for general structure; Talbot et al., 2007a; Rethemeyer et al., 2010) of which only the polar head group of adenosylhopane type-1 has been fully identified (Neunlist and Rohmer, 1985) as adenosine.

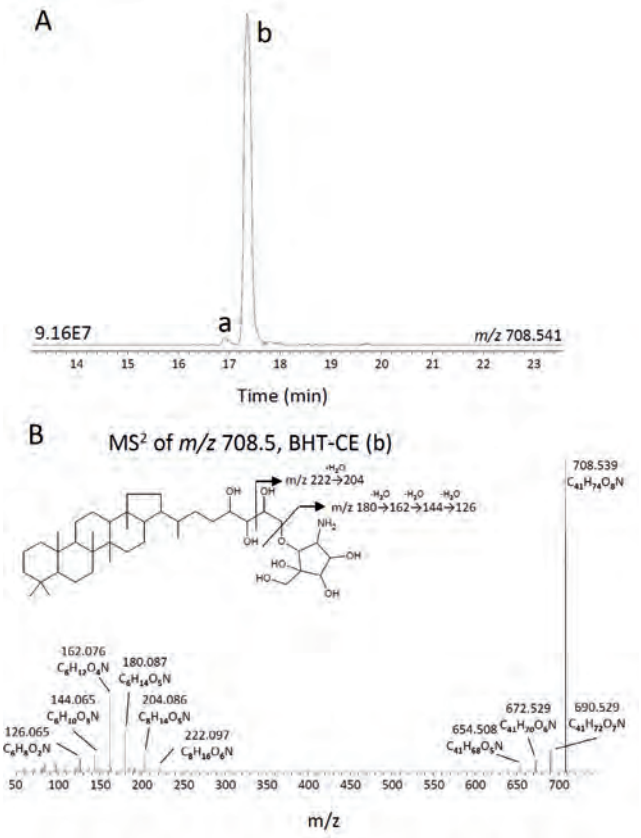


Figure 6. Cyclitol ether BHT (BHT-CE) in *Ca*. 'Scalindua profunda'. A: Partial mass chromatogram of the $[M+H]^+$ of BHT-CE. The trace is labeled with the exact mass used for searching and the intensity in arbitrary units (AU) of the signal. B: MS^2 spectrum of BHT-CE. Also shown is the structure of BHT-CE with diagnostic fragmentations indicated.

Adenosylhopane type-2 and -3 show similar fragmentation behavior to type-1, but have an unknown, presumably, nucleoside-type polar head group. Upon fragmentation of adenosylhopanes, a diagnostic fragment, representing the nucleoside, is formed with *m/z* 136 for type-1 (adenosine), *m/z* 151 for type-2, and *m/z* 150 for type-3 (Talbot et al., 2016). Methylated homologues have been observed for each of the three adenosylhopane types (Talbot and Farrimond, 2007; Rethemeyer et al., 2010). Adenosylhopane type-1 and its (di)methylated homologues are the only adenosylhopanes with a known elemental composition and, therefore, searchable based on their exact mass. For adenosylhopanes type-2 and -3 we initially mined the data using a nominal mass approach. This revealed several peaks, of which we then further investigated both the MS^1 and corresponding MS^2 spectra and which are discussed below.

Fig. 8 shows the full distribution of adenosylhopanes with increasing degree of methylation, as detected in the Censo 0 m soil. The mass chromatogram of the calculated exact mass of the protonated molecule for adenosylhopane type-1 (EC = $C_{40}H_{64}O_3N_5^+$; *m/z* 662.500) revealed a single peak (a) at 22.0 min (Fig. 8A). The MS^2 spectrum (Fig. 9B) contains a single dominant ion at *m/z* 136.062 with an assigned elemental composition of $C_5H_6N_5^+$ (Δ ppm -0.16), confirming the identity of the adenosyl head group. The mass chromatogram of *m/z* 676.516, i.e., the exact mass of the methylated homologue of adenosylhopane type-1, showed a series of 4 peaks (b through e) (Fig. 8B). Peaks b, c and e all produced the same fragment at *m/z* 136 as found for adenosylhopane type-1, thus indicating the position of the methylation to be on the BHP core. Peaks b, c, and e are therefore likely Me-adenosylhopane type-1, similar to the cluster of three isomers detected by Talbot et al. (2016) in a sediment of River Tyne. Both acetylated and non-derivatized 2MeBHPs elute very closely after their non-methylated counterparts using reversed phase chromatography, while 3MeBHPs elute later (Talbot et al. 2003a,b; 2007a,b; 2016) and as observed here, as discussed above). Based on the elution order and retention time differences compared to adenosylhopane type-1, we identify peak b as 2Me-adenosylhopane type-1, peak e as 3Me-adenosylhopane type-1, and peak c as an unknown Me-adenosylhopane type-1 isomer. This distribution of Me-adenosylhopane type-1 isomers mirror the distribution of methylated BHTs detected in this soil (see Table S.2). Peak d (Fig. 8B) showed an MS^2 spectrum with a single fragment at *m/z* 150.077 with an assigned elemental composition of $C_6H_8N_5$ (Δ ppm -0.01; Fig. 9C). A fragment ion at *m/z* 150 is diagnostic for adenosylhopane type-3, and in fact Talbot et al. (2016) used the predicted MRM transition of *m/z* 676 to 150 to successfully detect this compound in the River Tyne sediments. We, therefore, identify peak d as adenosylhopane type-3. Based on the AEC, we propose that the polar head group of adenosylhopane type-3 contains a methylated adenine (see Fig. 7C for proposed structure; placement of the methylation is arbitrary).

Fig. 8C shows the mass chromatogram of *m/z* 690.532 (EC = $C_{42}H_{68}O_3N_5^+$) and shows one dominant peak at 25.5 min (h) and two minor earlier eluting peaks (f and g). The MS^2 spectrum of peak f (Fig. 9D) contains two fragments related to the head group, i.e., the base peak at *m/z* 150.077 ($C_6H_8N_5^+$; Δ ppm -0.01) and a fragment at *m/z* 164.093 ($C_7H_{10}N_5^+$, Δ ppm -0.07). This suggests that this peak represents a co-elution of Me-adenosylhopane type-3 and (a minor) adenosylhopane type-3 with a second methylation on the

adenine. Talbot et al. (2016) used an MRM transition from *m/z* 690 to 150 and detected a single peak in a sediment from the River Tyne. It is likely that this BHP is similar to the here observed peak f. The MS^2 spectrum of peak g (Fig. 9E) showed a base peak at *m/z* 136.061 (EC = $C_5H_6N_5^+$) and a minor fragment at *m/z* 150.077 ($C_6H_8N_5^+$). We, therefore, propose peak g to be a co-elution of diMe-adenosyl type-1 and Me-adenosylhopane type-3. Peak h again appears to be an adenosylhopane type-3, with an additional methylation on the adenine, based on the product ion at *m/z* 164.093 ($C_7H_{10}N_5^+$; Δ ppm -0.99; Fig.9F), similar to the minor co-elution in peak f.

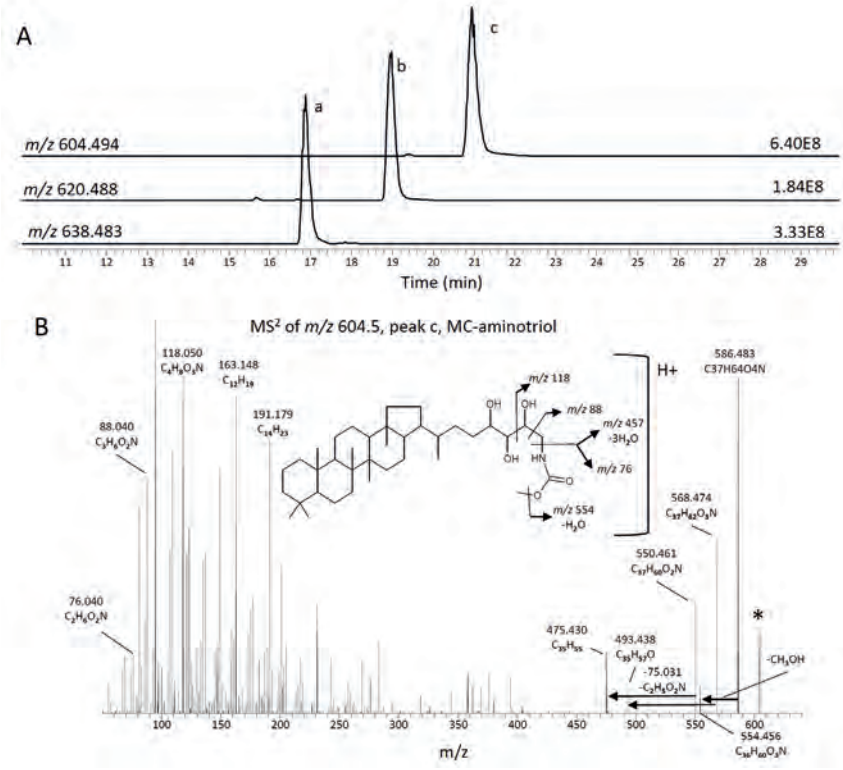


Figure 7. Methylcarbamate-aminoBHPs (MC-aminoBHPs) in *M. vadi*. A: partial mass chromatograms of the $[M+H]^+$ of MC-aminopentol (a), -aminotetrol (b) and -aminotriol (c). Each trace is labeled with the exact mass used for searching, and the intensity in arbitrary units (AU) of the signal. B. MS^2 of MC-aminotriol (mass peak labeled with * result from co-trapping of a co-eluting compound and is not related to this BHP). Also shown is the structure of MC-aminotriol with diagnostic fragmentations indicated.

The mass chromatogram of adenosylhopanes with an EC of $C_{43}H_{70}O_3N_5^+$ (*m/z* 704.547; Fig. 8D) revealed a series of peaks, i, j and k. The MS^2 spectra of all three peaks (shown for peak i in Fig. 9G) show a single fragment at *m/z* 164.093, and therefore we tentatively identify these BHPs as methyl-

adenosylhopanes type-3, with an additional methylation on the adenine head group. Peak j also shows minor fragments related to fragmentation in the hopanoid ring system, including *m/z* 191 (Fig. S7A). These BHPs appear to be the core-methylated homologues of peak h, and based on the offset in retention times to peak h, are tentatively identified as a having a methylation at C-2, methylation at unknown position, and a methylation at C-3; a similar distribution as observed for the MeBHTs and the Me-adenosylhopane type-1 peaks.

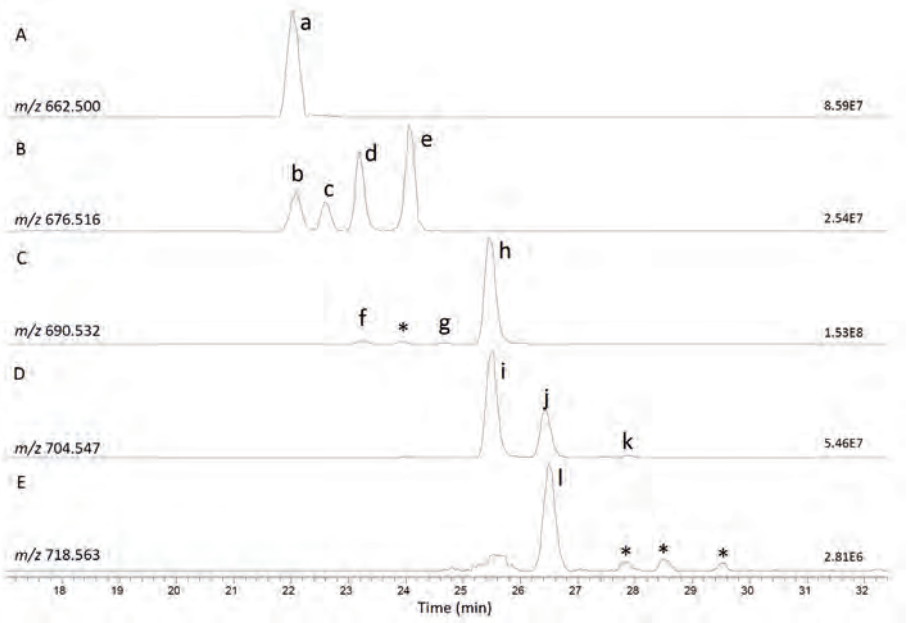


Figure 8. Partial mass chromatograms of adenosylhopanes, with increasing number of methylations, in a soil from a terrestrial methane seep (Censo 0 m). Each trace is labeled with the exact mass used for searching, and the intensity in arbitrary units (AU) of the signal. MS^2 spectra are shown in Fig. 9 if discussed in text; additional MS^2 spectra are shown in Fig. S7 Peaks labeled with '*' are potential isomers but have no MS^2 spectrum associated.

A search for adenosylhopanes with an EC of $C_{44}H_{72}O_3N_5^+$ (*m/z* 718.563) showed a series of peaks (Fig. 8E) with relative low abundance (two orders of magnitude less than the adenosylhopanes with *m/z* 690.532). Only peak m had an associated MS^2 spectrum, which showed one fragment ion related to the head group at *m/z* 164.093 and minor fragments related to fragmentation in the ring structure (Fig. 7H). We have tentatively identified this BHP as dimethyl-adenosylhopane type-3 with an additional methylation on the adenine head group. Evidence for BHPs methylated at C-2 and C-3 was previously seen in *Ca.*

'Koribacter versatilis', isolated from a pasture soil (Sinninghe Damsté et al., 2017). Adenosylhopanes with *m/z* 732.579 (EC $C_{45}H_{74}O_3N_5^+$) were not detected.

As the elemental composition of adenosyl type-2 is unknown, we searched the MS^2 data for the diagnostic fragment ion with a nominal *m/z* of 151 (Talbot et al., 2016). Several signals were found, mostly associated with amino acid lipids such as ornithines, but one MS^2 spectrum (Fig. 10B) clearly showed an adenosylhopane signature, with a single dominant fragment ion at *m/z* 151.061 ($C_6H_7ON_4^+$, Δ ppm -0.38 ppm). Interestingly, this elemental composition matches the EC for N1-methylinosine, which is formed from adenosine via inosine in transfer RNAs, and is found in the RNA of eukaryotes and halo- and thermophilic archaea (Grosjean and Constantinesco, 1996). The EC of the protonated molecule of adenosylhopane type-2 was determined to be $C_{41}H_{65}N_4O_4^+$ (*m/z* 677.500). Fig. 10A shows the mass chromatogram of *m/z* 677.500 from Censo 0 m with the relatively low abundance peak at 24.24 min from which the MS^2 spectrum was derived. A homologue of the tentatively identified N1-methylinosylhopane, methylated on the BHP core (m/z 691.516; $C_{42}H_{67}N_4O_4^+$), was detected at 24.30 min (peak b, Figs. 10A and S7B). Based on the retention time, we tentatively identify this BHP as 2Me-N1-methylinosylhopane. As N1-methylinosine is formed from adenosine via an initial hydrolytic deamination to inosine, we also searched for the proposed intermediate between adenosylhopane and N1-methylinosylhopane, i.e., inosylhopane ($C_{40}H_{63}O_4N_4^+$; *m/z* 663.484). At 20.63 min a peak was identified with an associated MS^2 spectrum showing the predicted head group fragment at *m/z* 137.0458 ($C_5H_5ON_4^+$, Δ ppm -0.20; peak c, Fig. 10A). Minor fragments at *m/z* 529.461 ($C_{35}H_{61}O_3^+$) and *m/z* 511.450 ($C_{35}H_{59}O_2^+$) further confirmed the anhydro-BHT core structure (Fig. 10C).

3.5.2. A novel composite BHP with an N-containing moiety

During a broad search for known BHPs in the Censo 0 m soil, two peaks matching the exact mass and EC of protonated MC-aminotriol and -tetrol (*m/z* 604.494, $C_{37}H_{66}O_5N^+$ and *m/z* 620.488, $C_{37}H_{66}O_6N^+$, respectively) were encountered (Fig. 11A, peaks a and b). However, these peaks elute later than the equivalent BHPs described for *M. vadi* (see above). Although the MS^2 spectra of the peaks detected in the Censo 0 m soil share many characteristics with those of the MC-aminoBHPs, there are several distinct differences.

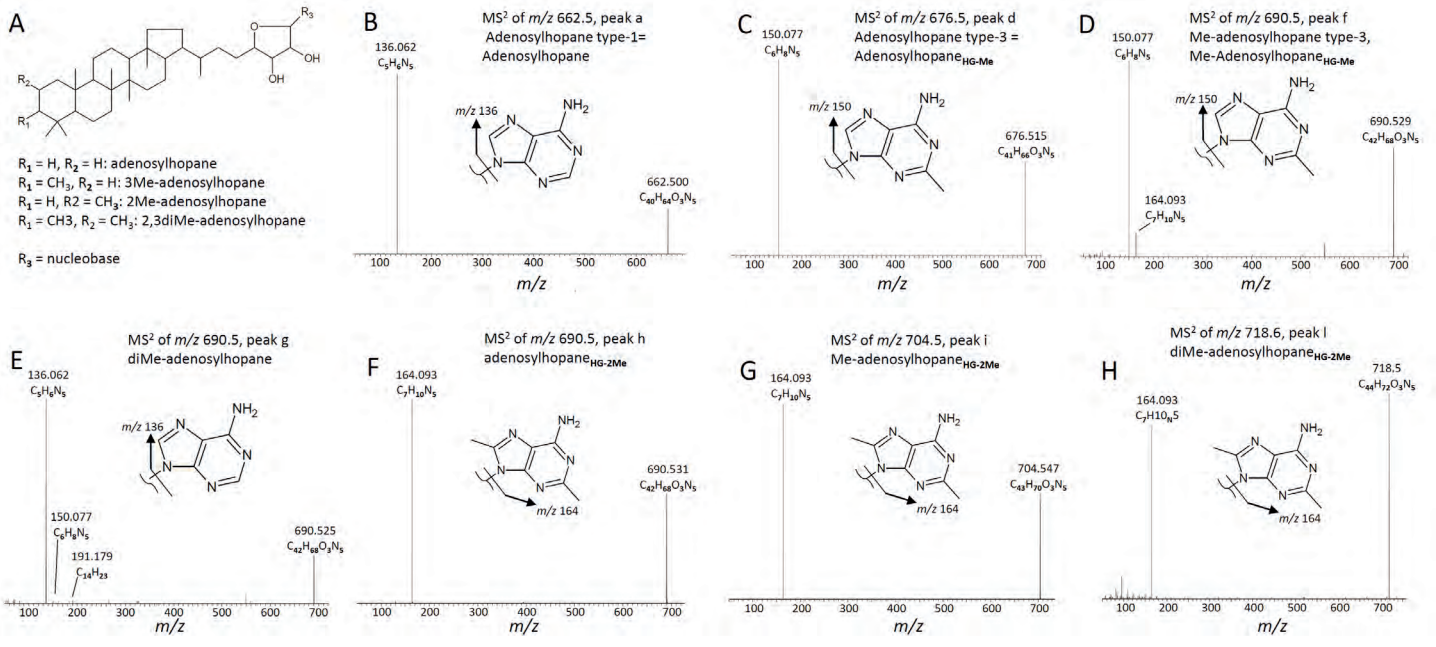


Figure 9. Adenosylhopanes detected in Censo 0 m soil. A: General structure of adenosylhopanes, B through H: MS^2 spectra of adenosylhopanes. Proposed structure of the nucleobase is shown with each spectrum. Spectra of peak f (panel D) and peak g (panel E) appear to be mixed spectra of 2 co-eluting adenosylhopanes. The structure shown with the spectra is the proposed nucleobase of the major component. Placement of methylations on the nucleobase is arbitrary as the position is unknown. As, in fact, it appears that all detected adenosylhopanes are adenosylhopane type-1 with one or two methylations either on the BHP core and/or on the adenine head group, we propose the following new nomenclature for this extended family of adenosylhopanes: methylations on the BHP core are indicated as "(di)Me-adenosylhopane", while methylations on the adenine head group are indicated by a subscript. Adenosylhopane type-1 would thus be simple be named adenosylhopane. Adenosylhopane type-3 would be named adenosylhopane$_{HG\text{-}Me}$. The Me-adenosylhopanes with an additional methylation on the adenine (peaks i,j, and k) would be named Me-adenosylhopane$_{HG\text{-}2Me}$, and peak l would be named diMe-adenosylhopane$_{HG\text{-}2Me}$.

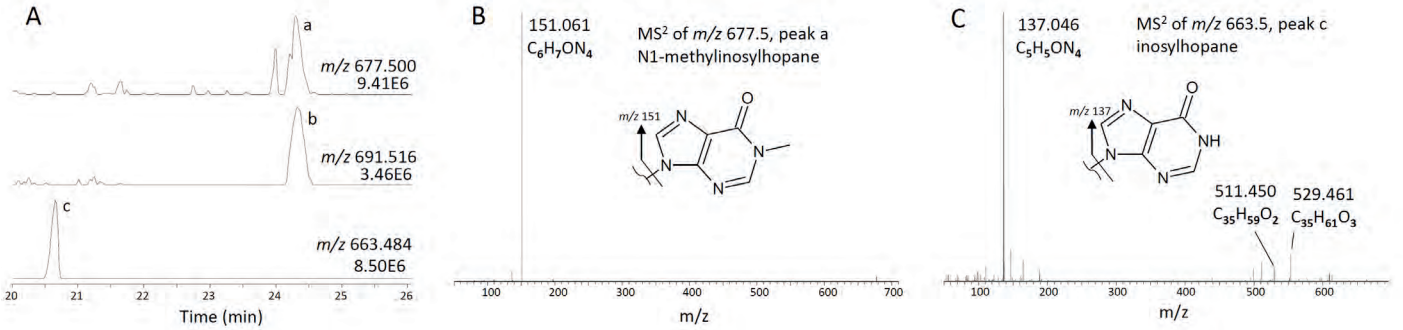


Figure 10. Adenosylhopane type-2 in Censo 0 m soil. A: Partial mass chromatograms, (within 2 ppm mass accuracy) of the tentatively identified N1-methylinosinylhopane (peak a), 2-methyl-N1-methylinosinylhopane (peak b), and inosylhopane (peak c), from a soil from an active terrestrial methane seep in Sicily (Censo 0 m). Each trace is labeled with the exact mass used for searching, and the intensity in arbitrary units (AU) of the signal. B: MS^2 spectrum of N1-methylinosylhopane. C: MS^2 spectrum of inosylhopane. Proposed structures of the inosine based headgroups are show

The MS^2 spectrum of peak b (Fig. 11B) is characterized by a series of initial losses of H_2O (-18) from the protonated molecule producing ions at *m/z* 586.483 ($C_{37}H_{64}O_4N^+$, Δ ppm -0.33) , *m/z* 568.474 ($C_{37}H_{62}O_3N^+$, Δ ppm 2.23), *m/z* 550.462 ($C_{37}H_{60}O_2N^+$, Δ ppm -0.23), and *m/z* 532.453 ($C_{37}H_{58}ON^+$, Δ ppm 2.64). However, instead of the characteristic loss of 75 Da, representing the loss of the methylcarbamate, losses of 41 Da (C_2H_3N) and 59 Da (C_2H_5ON) were observed here, generating fragment ions at *m/z* 491.424 ($C_{35}H_{55}O^+$, Δ ppm -2.43) and m/z 473.414 ($C_{35}H_{55}^+$, Δ ppm -1.41), respectively. Dominant N-containing product ions are observed at *m/z* 60.045 ($C_2H_6ON^+$, Δ ppm 4.52) and *m/z* 102.055 ($C_4H_8O_2N^+$, Δ ppm 1.90). The middle *m/z* region shows similar minor fragments, representing the D and E ring and the side chain after loss of the functionalities, as observed for BHpentol. Based on the assigned EC and the fragmentation pattern, we propose this BHP is a composite BHP based on BHpentol bound to an ethenolamine moiety (C_2H_4N) via an ether bond. To the best of our knowledge this composite BHP has not been observed before.

After having tentatively identified this ethenolamine-BHpentol, we identified the BHT and BHhexol homologues (peak a and c, respectively, in Fig. 11A), based on calculated exact mass and MS^2 fragmentations (Table S3, Figs. S8A and B). Although methylated ethenolamine-BHPs were not detected in Censo 0 m, a BHP matching the calculated exact mass and EC of methylated ethenolamine-BHT ($C_{38}H_{68}O_4N^+$) was detected at 21.44 min (for details see Table S3). The fragmentation pattern was almost identical to what was observed for ethenolamine-BHT (Fig. 11B). However, both diagnostic N-containing product ions were offset by +14 Da resulting in fragments at *m/z* 74.061 ($C_3H_8ON^+$, Δ ppm 9.446), and *m/z* 116.071 ($C_5H_{10}O_2N^+$, Δ ppm 3.575) (Fig. S8C). No diagnostic losses could be observed in this case. This BHP was, therefore, tentatively identified as a propenolamine-BHT. A butenolamine homologue was not detected.

3.5.3 Acylated ethenolamines

Using the exact mass of the most common diagnostic ion for BHPs, i.e., *m/z* 191.179 ($C_{14}H_{23}^+$), the MS^2 data was investigated for other potential unknown BHPs. This revealed the presence of a series of late eluting (35-40 min) compounds with *m/z* values >800 (Fig. S9). One of the most abundant of these unknowns was a compound with *m/z* 840.743 and an assigned EC of $C_{54}H_{98}O_5N^+$ (Δ ppm -1.54). A mass chromatogram for this EC showed a cluster of peaks consisting of at least 8 isomers (peaks a to h, Fig. 12A) with near identical MS^2 spectra.

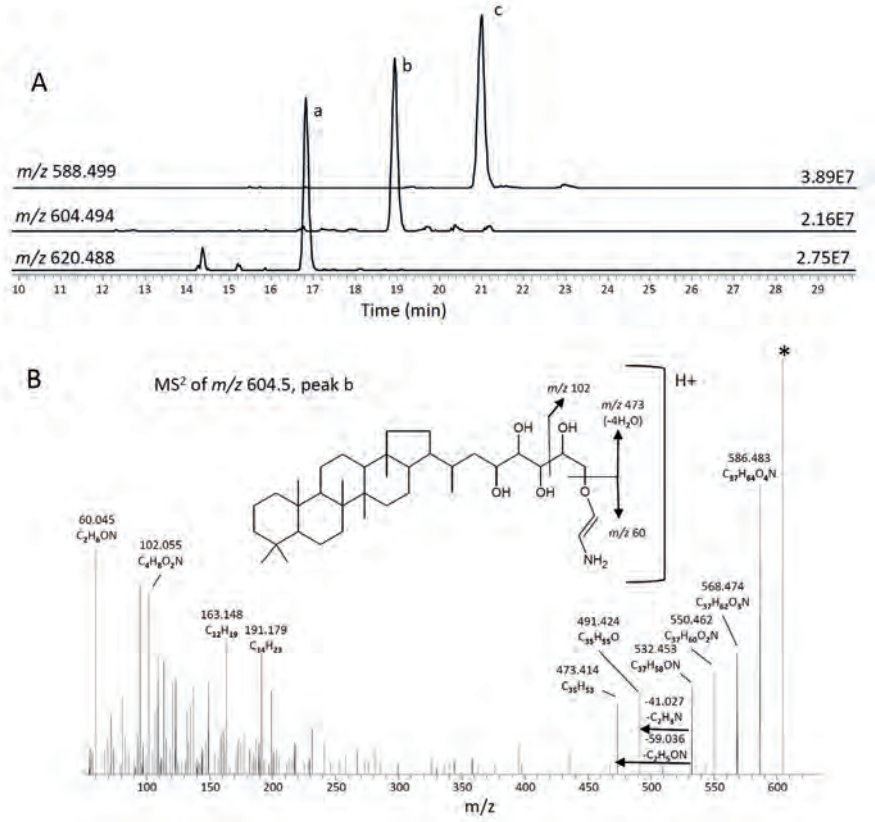


Figure 11. Novel N-containing composite BHPs in Censo 0 m soil. A. Partial mass chromatograms of a series of novel composite BHPs (a, b, c). Each trace is labeled with the exact mass used for searching, and the intensity in arbitrary units (AU) of the signal. B. MS^2 spectrum associated with peak b. (mass peak labeled with * results from co-isolation of a co-eluting compound and is not related to this BHP). Also shown is the proposed structure of ethenolamine-BHpentol with diagnostic fragmentations indicated.

Fig. 12B shows the MS^2 spectrum of the most abundant peak e. Two losses of H_2O are observed from the parent ion resulting in fragment ions at *m/z* 822.733 ($C_{54}H_{96}O_4N^+$, Δ ppm -0.26) and *m/z* 804.720 ($C_{54}H_{94}O_3N^+$, Δ ppm -0.46). In the middle region of the mass spectrum a cluster of fragment ions are observed, which are formed after an initial loss of 252.246 ($C_{17}H_{32}O$, Δ ppm 0.81) from the parent ion and further losses of -18 Da (H_2O) and, interestingly, -41 Da (C_2H_3N) and -59 Da (C_2H_5ON), similar to the fragmentation pattern observed for the earlier described ethenolamine BHPs. Indeed, the diagnostic product ions of this novel class of BHPs at *m/z* 60 and 102 were also present. Based on the resemblance of the MS^2 spectrum to that of ethenolamine-BHT and the loss of a $C_{17}H_{32}O$ moiety, we tentatively identified this compound as a $C_{17:0}$-N-acyl-ethenolamine-BHT (Fig. 12B). The fatty acid moiety can, based on the data here, only be identified to the level of carbon number and double bond equivalent.

Several causes are possible for the multitude of isomers. The head group can be bound to the polyol tail of the BHP at a different positions (C-32, C-33, C-34, or C-35), and/or by isomery within the BHT core structure, or by structural differences (linear vs. branched) in the fatty acid tail.

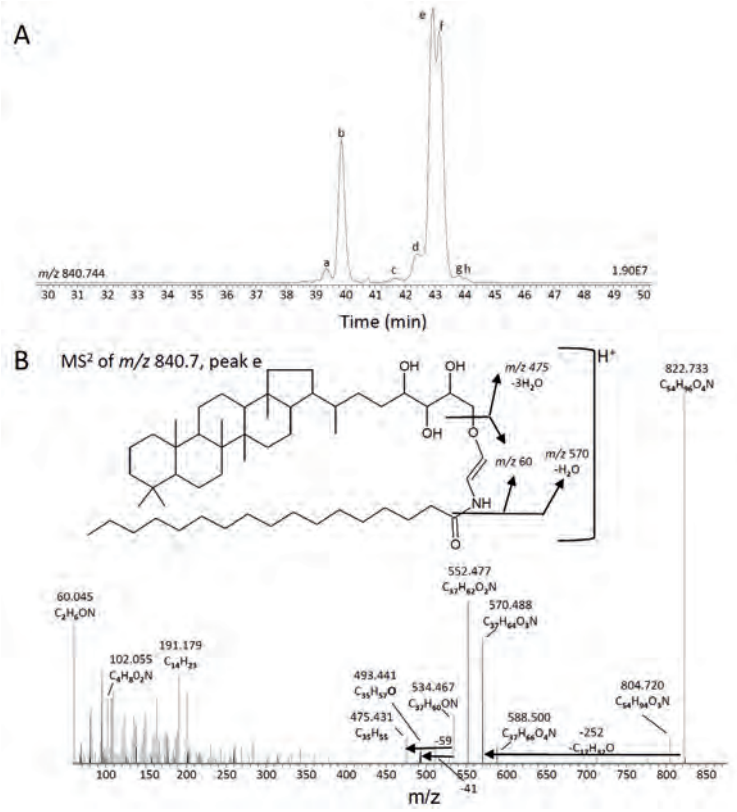


Figure 12. An example of an N-acyl-ethenolamine-BHT in Censo 0 m, a soil from a terrestrial methane seep in Sicily. A. Partial mass chromatogram of $C_{17:0}$-N-acyl-ethenolamine-BHT. The trace is labeled with the exact mass used for searching, and the intensity in arbitrary units (AU) of the signal. B. MS^2 spectrum of peak e from panel A with the proposed structure for $C_{17:0}$-N-acyl-ethenolamine-BHT with diagnostic fragmentations indicated. Position of the conjugation is arbitrary as the exact position is unknown. The fatty acid moiety is shown with a linear carbon chain, as the exact structure cannot be determined here.

In addition to the $C_{17:0}$-N-acyl-ethenolamine-BHTs we detected complex distributions of $C_{15:0}$ to $C_{18:0}$-N-acyl-ethenolamine-BHTs (Fig. S9). $C_{14:0}$- and $C_{19:0}$-N-acyl-ethenolamine-BHTs appeared only present at trace levels and could not be confirmed by obtaining MS^2 spectra. A full listing by retention time, for those isomers confirmed by MS^2, is given in Table S2. Ethenolamine-BHTs bound to $C_{15:0}$ and $C_{17:0}$ fatty acids were most abundant in the Censo 0 m soil, but the $C_{16:0}$ bound ethenolamine BHTs showed the most complex distribution with 12 isomers confirmed by MS^2 spectra. A search for unsaturated homologues

resulted in the detection of ethanolamine-BHTs bound to $C_{17:1}$ and $C_{18:1}$ fatty acids (Table S2, Fig. S10), which were only present at trace levels. Acylated ethenolamine-BHTs comprising an unsaturated hopanoid core were not detected. Acylated ethenolamine-BHPs based one BHpentol and BHhexol were also detected (Table S2, Fig. S10) and comprised $C_{15:0}$ to $C_{18:0}$, and $C_{17:1}$ and $C_{18:1}$ N-bound acyl moieties, i.e., comparable with the distribution of the ethenolamine-BHTs. Interestingly, many more isomers were identified for ethenolamine-BHhexol than for ethenolamine-BHpentol, while BHhexol itself was not detected.

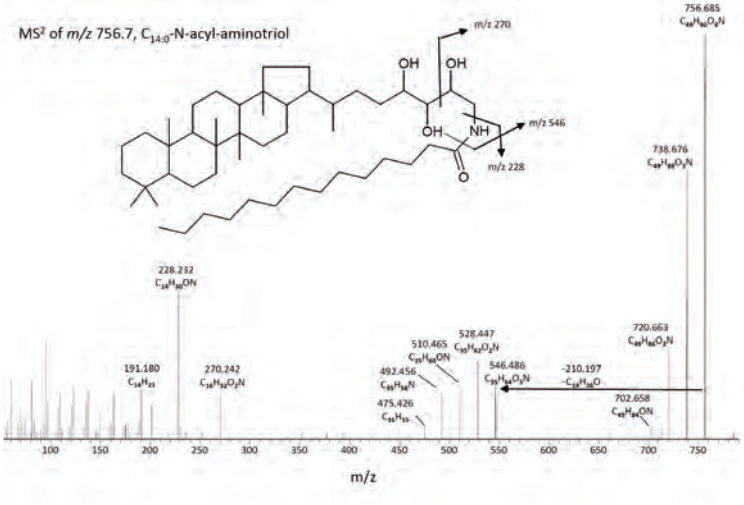


Figure 13. MS^2 spectrum of $C_{14:0}$-N-acyl-aminotriol from Censo 0 m, a soil from a terrestrial methane seep in Sicily. Proposed structure and diagnostic fragmentations are shown. The fatty acid moiety is shown with a linear carbon chain, as the exact structure cannot be determined here. Further details can be found in table S2.

3.5.4 Acylated aminotriols

A further search for acylated BHPs revealed a series of N-acyl-aminotriols in the Censo 0 m soil (Fig. S11). HPLC-MS detection of derivatized N-acyl-aminotriols was previously reported by Talbot et al. (2007a) in *N. europaea* and *R. vannielii*. However, the MS^2 spectrum (shown for $C_{14:0}$-N-acyl-aminotriol, Fig. 13) of the non-derivatized molecule proves to be much more diagnostic than that produced from the derivatized molecule. Here, three losses of H_2O are observed from the parent ion resulting in fragment ions at *m/z* 738.676 ($C_{49}H_{86}O_2N^+$, Δ ppm 0.04), *m/z* 720.663 ($C_{49}H_{86}O_2N^+$, Δ ppm -3.078) and *m/z* 702.658 ($C_{49}H_{84}ON^+$, Δ ppm -3.17). Loss of 210.199 Da ($C_{14}H_{26}O$, Δ ppm 4.07) leads to a product ion at *m/z* 546.486 ($C_{35}H_{64}O_3N^+$, Δ ppm 4.07), matching the $[M+H]^+$ of aminotriol. Further fragmentation indeed matches the fragmentation for amino BHPs as described earlier, although fragment ions in the lower mass

region appear more abundant here. The acyl moiety is easily defined by a prominent fragment ion at *m/z* 228.232 ($C_{14}H_{30}ON^+$, Δ ppm -4.12). An additional N-containing fragment ion is observed at *m/z* 270.243 ($C_{16}H_{32}O_2N^+$, Δ ppm -3.01) and is likely formed after cleavage of the C-33/C-34 bond.

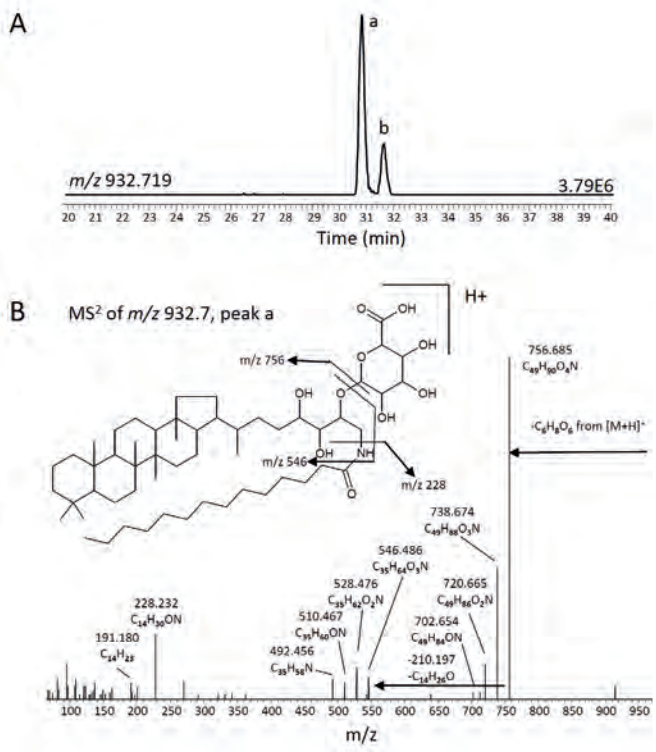


Figure 14. Novel multi-substituted aminotriol in Censo 0 m, a soil from a terrestrial methane seep in Sicily. A: Partial mass chromatogram of $C_{14:0}$-N-acyl-aminotriol with additional glucuronic acid substitution. Trace is labeled with the exact mass used for searching, and the intensity in arbitrary units (AU) of the signal. B. MS^2 spectrum associated with peak a with proposed structure of $C_{14:0}$-N-acyl-glucuronyl-aminotriol with diagnostic fragmentations indicated. Position of the glucuronyl moiety is arbitrary. The fatty acid moiety is shown with a linear carbon chain, as the exact structure cannot be determined here.

Whereas Talbot et al. (2007a) reported aminotriol acylated to $C_{16:0}$ and $C_{16:1}$ FAs in *N. europaea* and to $C_{18:0}$, $C_{18:1}$, and $C_{19:0}$ in *R. vannieli*, we detect aminotriols acylated to much shorter chain fatty acids ranging from $C_{8:0}$ to $C_{18:0}$ FA, with the $C_{11:0}$-N-acyl-aminotriol being the most abundant homologue (peak d, Fig. S11). In addition, aminotriols acylated to $C_{11:1}$, $C_{12:1}$, $C_{14:1}$ and $C_{16:1}$ fatty acids were detected (Table S2). N-acyl-aminotetrols were not detected, and only trace levels of a $C_{16:0}$-N-acyl-aminopentol were detected in the Censo 0 m soil (Fig. S12). A complicating factor in identifying the full spectrum of acyl-BHPs is the fact that there is a considerable overlap in elemental composition between the acyl-ethenolamine-BHPs with the N-acyl-aminoBHPs. For illustration, the

$[M+H]^+$ of an aminotetrol bound to a $C_{17:1}$ FA has an elemental composition of $C_{52}H_{94}O_5N^+$, which is identical to that of an ethenolamineBHT bound to a $C_{15:0}$ FA. It is, therefore, important to evaluate the fragmentation of each detected compound.

3.5.5 Multi-conjugated composite aminotriols

While charting the full inventory of N-acyl-aminotriols, we also encountered several compounds that were clearly related to the N-acyl-aminotriols as evident from their MS^2 spectra (e.g., Fig. 14B). The most abundant of these was a compound revealed in a mass chromatogram of *m/z* 932.719 (Fig. 14A, peak a) and an assigned EC of the protonated molecule $C_{55}H_{98}O_{10}N^+$ (Δ ppm -1.07). An initial loss of 176 ($C_6H_8O_6$) yields a base peak at *m/z* 756.685 with an assigned EC matching that of $C_{14:0}$-N-acyl-aminotriol ($C_{49}H_{90}O_4N^+$, Δ ppm -1.42). The initial loss of 176 Da matches the predicted loss and EC for a glucuronic acid moiety. Further fragmentation was identical to that observed for $C_{14:0}$-N-acyl-aminotriol, revealing the aminotriol core at *m/z* 546 and the $C_{14:0}$ fatty acid moiety at *m/z* 228. The MS^2 spectrum of peak b is identical to that of peak a, and peak b thus probably reflects an isomer. Although glucuronic acid is not a very common head group in intact polar lipids, it has been observed in bacteria and fungi (e.g., Burugupalli et al., 2020; Fontaine et al., 2009; Wang et al., 2020). We therefore tentatively identify these compounds as C_{14}:0-N-acyl-glucuronosyl-aminotriols (Fig. 14B). In addition to the $C_{14:0}$-N-acyl-glucuronosyl-aminotriol, a $C_{15:0}$-N-acyl-glucuronyl-aminotriol was also detected (Table S2). To the best of our knowledge this is the first report of BHPs with conjugations on more than one position on the BHP core.

Conclusions

We have shown the applicability of UHPLC-ESI/$HRMS^2$ for the analysis of non-derivatized BHPs in both microbial cultures as well as environmental samples. The chromatographic system used here allows separation of a broad range of BHPs, ranging from the relatively simple BHPs to nitrogen-containing BHPs and complex composite BHPs. Furthermore, isomers are readily separated. Identification is achieved based on diagnostic spectra, that contain information on the core BHP core structure, the functionalized tail, as well as bound moieties. For the first time we established the elemental composition of the nucleobase of adenosylhopanes type-2 and type-3, showing that in fact, all adenosylhopanes identified so far are modifications of adenosyl either by one or two methylations on the adenine head group (type-3) or by deamination followed by methylation

(type-2). Furthermore, we have demonstrated the usefulness of HRMS in the identification of novel composite BHPs. We have tentatively identified several new composite BHPs in a soil (e.g., the (N-acyl-)ethenolamine-BHPs), showing a previously unobserved diversity and complexity in existing BHP structures. The analytical approach described here allows for simultaneous analysis of the full suite of IPLs, now including BHPs, and represents a further step towards environmental lipidomics. With this method a more complete view of the full assembly of BHPs will be possible. Connecting specific intact BHPs to specific sources and/or geochemical cycles will further aid in the interpretation of their diagenetic products, the geohopanoids, in the geological record. Future work will aim to establish a quantitative protocol for this method using isolated BHPs as well as synthetic internal standards.

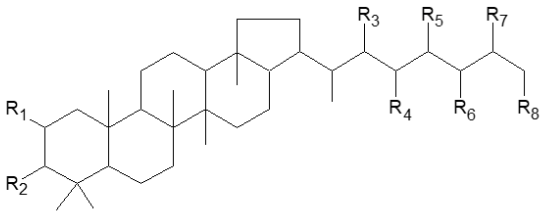
Acknowledgements

This work was funded by Netherlands Earth System Science Center (NESSC) through a Gravitation grant to JSSD (grant no. 024.002.001) from the Dutch Ministry for Education, Culture and Science, NWO middelgroot grant no. 834.13.004 to ECH, and a Natural Environment Research Council (NERC; United Kingdom) grant to DR (project ANAMMARKS (NE/N011112/1)). Cultures were kindly provided by H. Hirayama (JAMSTEC), G.H.L. Nuijten, O. Rasigraf, and M.S.M. Jetten (Radboud University), M. Rohmer and P. Schaeffer (Strasbourg University). We thank F. Grassa (INGV) for assistance to NS during field work.

Supplementary material (available upon request)

Table S1. BHPs detected in biomass. Diagnostic fragments are only shown for the main isomer.

Table S2. BHPs detected in Censo 0 m, a soil from a terrestrial methane seep in Sicily.



BHP	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8
BHT	H	H	H	H	OH	OH	OH	OH
2MeBHT	CH_3	H	H	H	OH	OH	OH	OH
3MeBHT	H	CH_3	H	H	OH	OH	OH	OH
2,3-diMe-	CH_3	CH_3	H	H	OH	OH	OH	OH
BHpentol	H	H	H	OH	OH	OH	OH	OH
2Me-	CH_3	H	H	OH	OH	OH	OH	OH
3Me-	H	CH_3	H	OH	OH	OH	OH	OH
2,3-diMe-	CH_3	CH_3	H	OH	OH	OH	OH	OH
BHhexol	H	H	OH	OH	OH	OH	OH	OH
2Me-	CH_3	H	OH	OH	OH	OH	OH	OH
3Me-	H	CH_3	OH	OH	OH	OH	OH	OH
2,3-diMe-	CH_3	CH_3	OH	OH	OH	OH	OH	OH
aminotriol	H	H	H	H	OH	OH	OH	NH_2
2Me-	CH_3	H	H	H	OH	OH	OH	NH_2
3Me-	H	CH_3	H	H	OH	OH	OH	NH_2
2,3-diMe-	CH_3	CH_3	H	H	OH	OH	OH	NH_2
aminotetrol	H	H	H	OH	OH	OH	OH	NH_2
2Me-	CH_3	H	H	OH	OH	OH	OH	NH_2
3Me-	H	CH_3	H	OH	OH	OH	OH	NH_2
2,3-diMe-	CH_3	CH_3	H	OH	OH	OH	OH	NH_2
aminopentol	H	H	OH	OH	OH	OH	OH	NH_2
2Me-	CH_3	H	OH	OH	OH	OH	OH	NH_2
3Me-	H	CH_3	OH	OH	OH	OH	OH	NH_2
2,3-diMe-	CH_3	CH_3	OH	OH	OH	OH	OH	NH_2

Figure S1. Core structure of BHPs discussed in text. Complex structures are shown in figures where applicable.

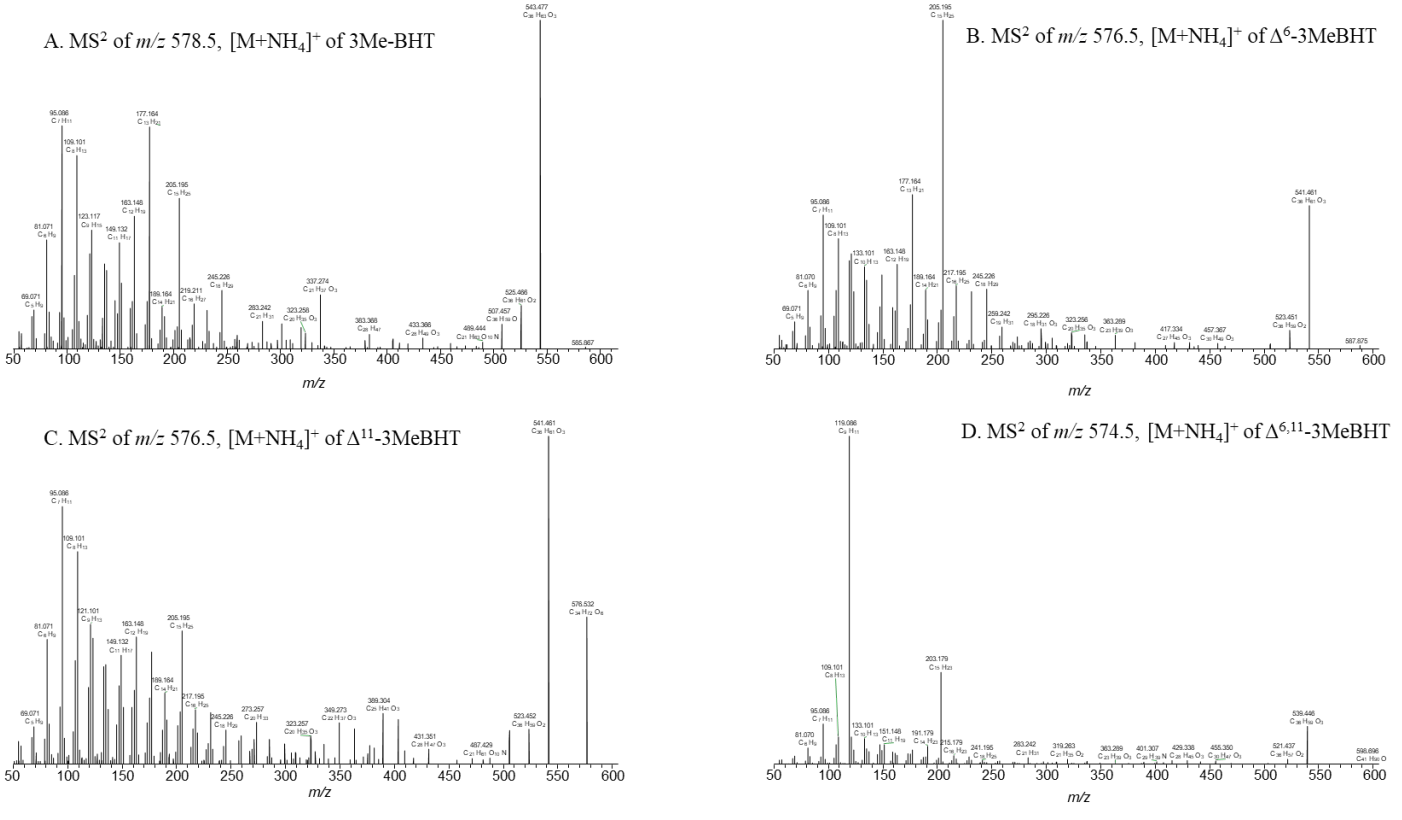


Fig. S2. MS2 spectra of 3MeBHTs in *K. xylinum*. A: MeBHT (peak k, Fig. 1D), B: Δ6-MeBHT (peak m, Fig. 1E), C: Δ11-MeBHT (peak o, Fig. 1E), and D: Δ6,11-MeBHT (peak r, Fig. 1F). Assigned elemental composition of diagnostic mass peaks are listed in Table S1.

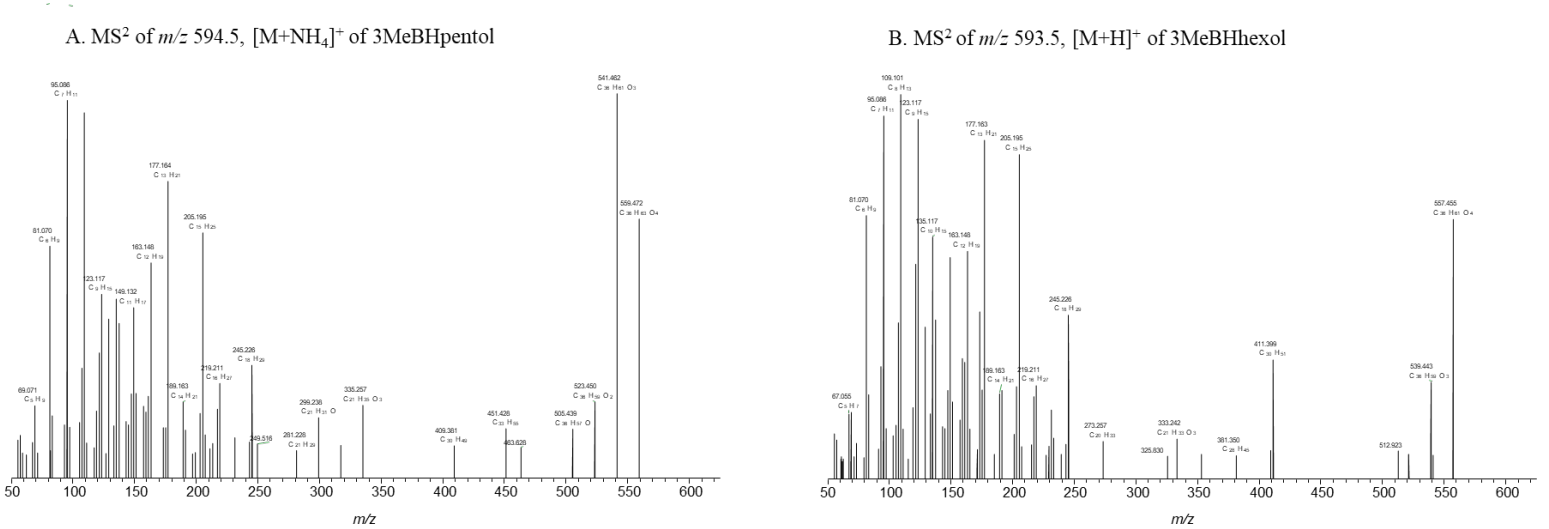


Fig. S3. MS2 spectra of 3MeBHpentol and 3MeBHhexol in *Ca*. 'M. oxyfera'. A: 3MeBHpentol (peak e, Fig. 4B), B: 3MeBHhexol (peak f, Fig. 4B). Assigned elemental composition of diagnostic mass peaks are listed in Table S1.

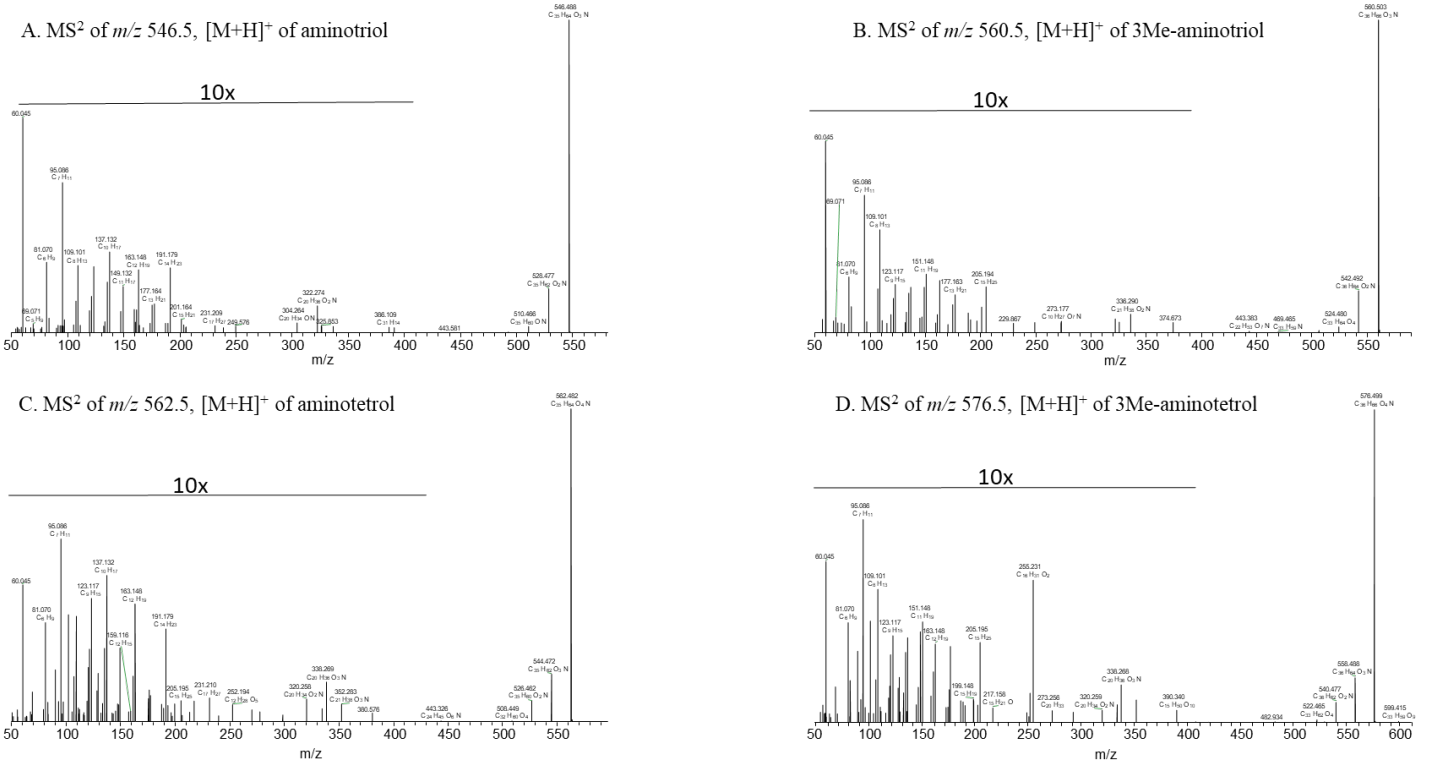


Fig. S4. Additional MS2 spectra of amino-BHPs in M. capsulatus. A: aminotriol (peak a, Fig. 5A), B: 3Me-aminotriol (peak e, Fig. 5B), C: aminotetrol (peak b, Fig. 5A), D: 3Me-aminotetrol (peak g, Fig. 5B). Assigned elemental composition of diagnostic mass peaks are listed in Table S1.

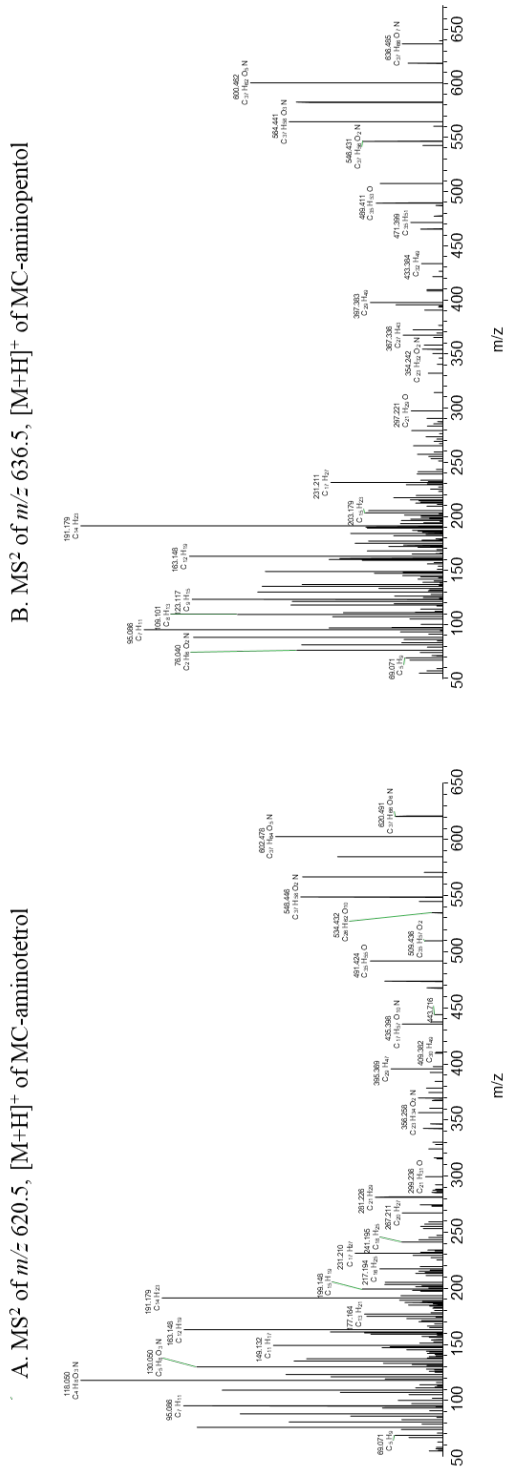


Fig. S5. MS2 spectra of MC-aminotetrol and MC-aminopentol in *M. vadi*. A: MC-aminotetrol (peak a, Fig. 7a), B: MC-aminopentol (peak b, Fig. 7a). Assigned elemental composition of diagnostic mass peaks are listed in Table S1.

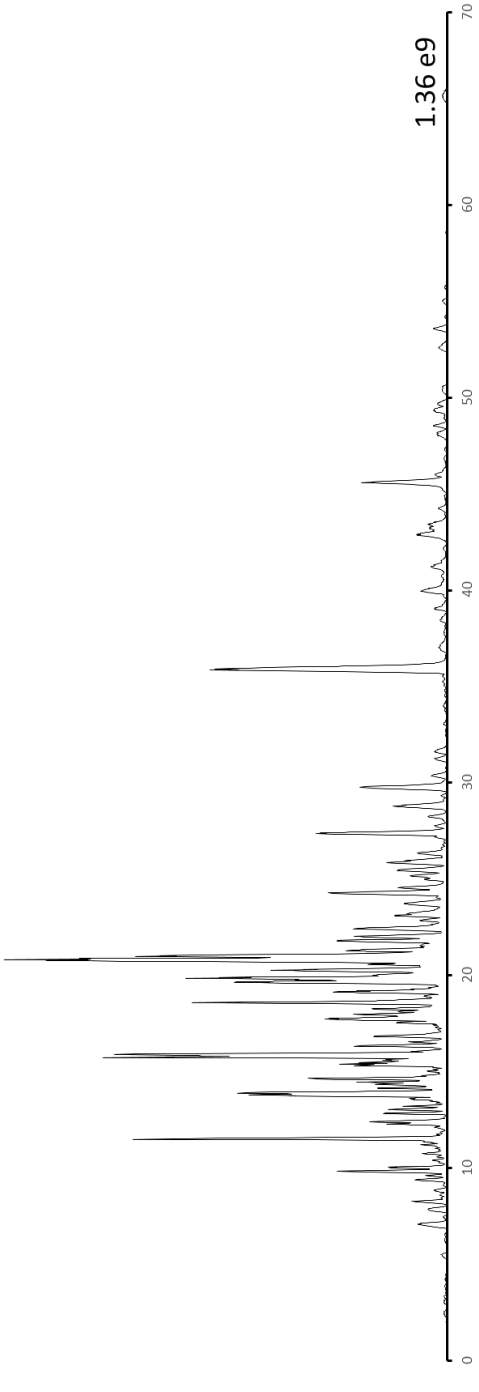


Fig. S6. Base peak chromatogram (m/z 350-2000) of a BDE of a soil from a Sicilian methane seep (Censo 0 m). Intensity (AU) of the signal is indicated.

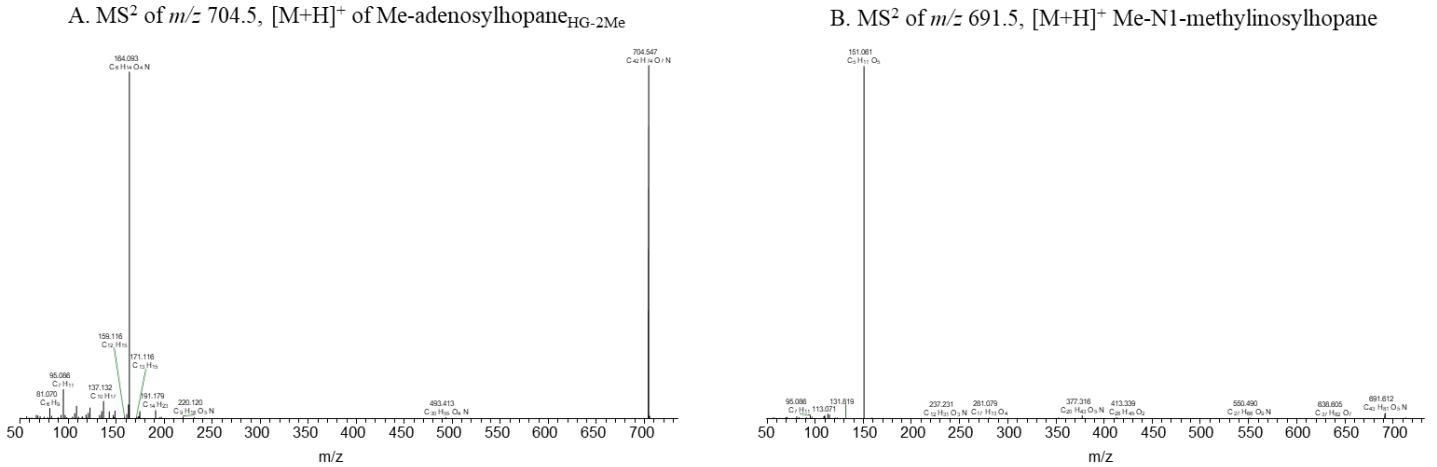


Fig. S7. Additional MS2 spectra of adenosyl hopanes in Censo 0 m soil. A: Me-adenosylhopaneHG-2Me (peak j, Fig. 8D), B: Me-N1-methylinosylhopane (peak b, Fig. 10A).

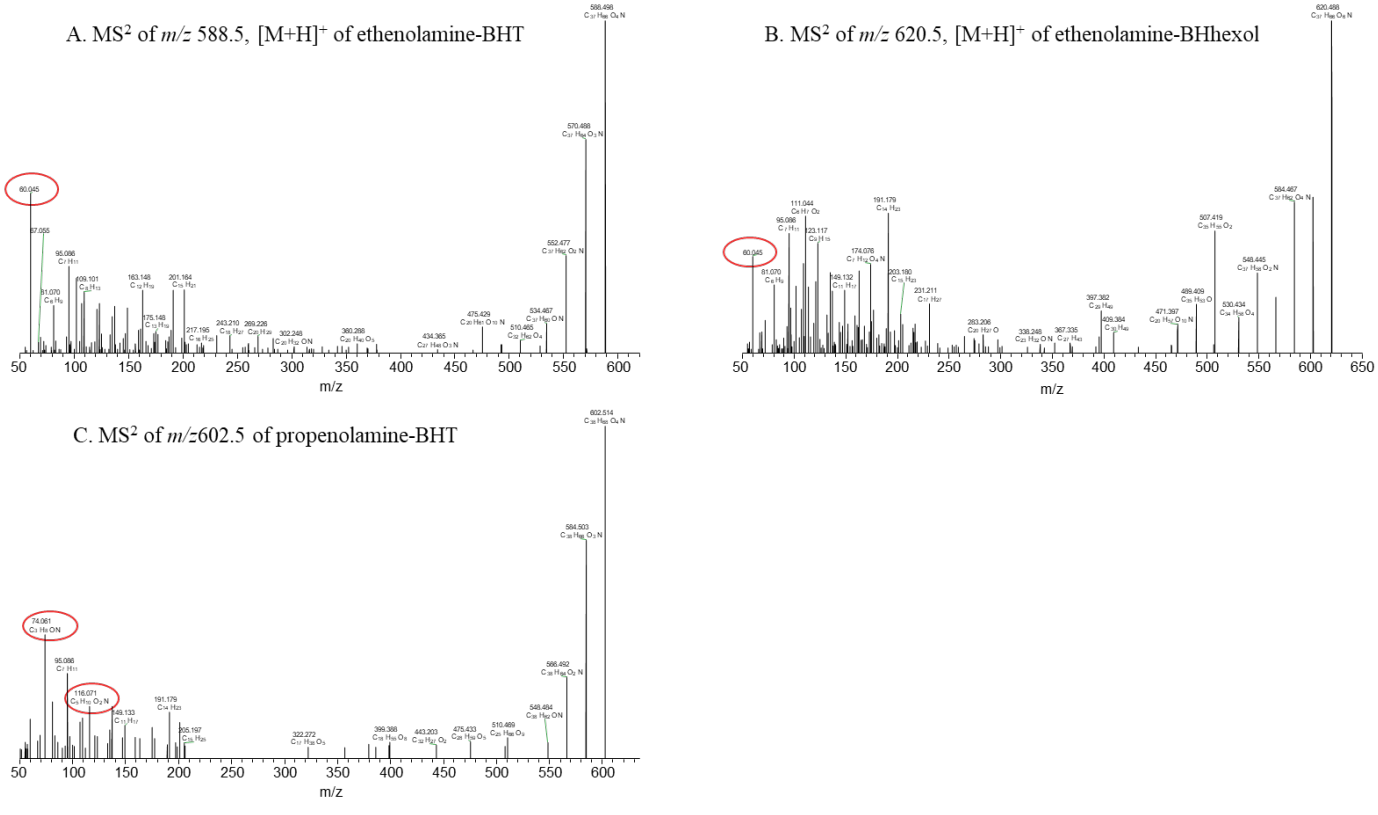


Fig. S8. Additional MS2 spectra of novel N-containing BHPs in Censo 0 m soil. A: ethenolamine-BHT (peak c, Fig. 11A), B: ethenolamine-BHhexol (peak a, Fig. 11A), C: propenolamine-BHT.

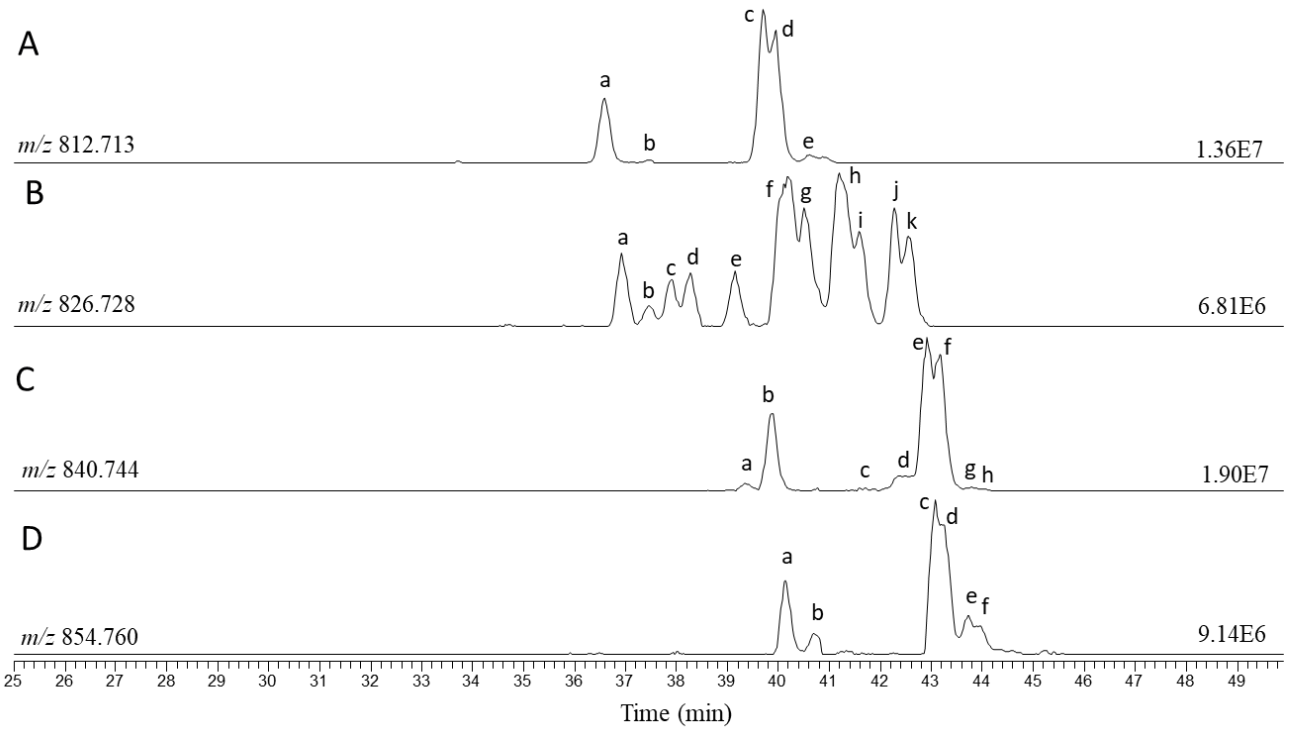


Fig. S9. Partial mass chromatograms of a series of N-acyl-ethenolamineBHTs in a soil from an active terrestrial methane seep in Sicily (Censo 0 m). A. C15:0-N-acyl-ethenolamine BHT. B. C16:0N-acyl-ethenolamine BHT. C. C17:0-N-acyl-ethenolamine BHT. D. C18:0-N-acyl-ethenolamine BHT. Each trace is labeled with the calculated exact mass, and the intensity (AU) of the signal. The fatty acids moiety is indicated by 'carbon number: double bond equivalent".

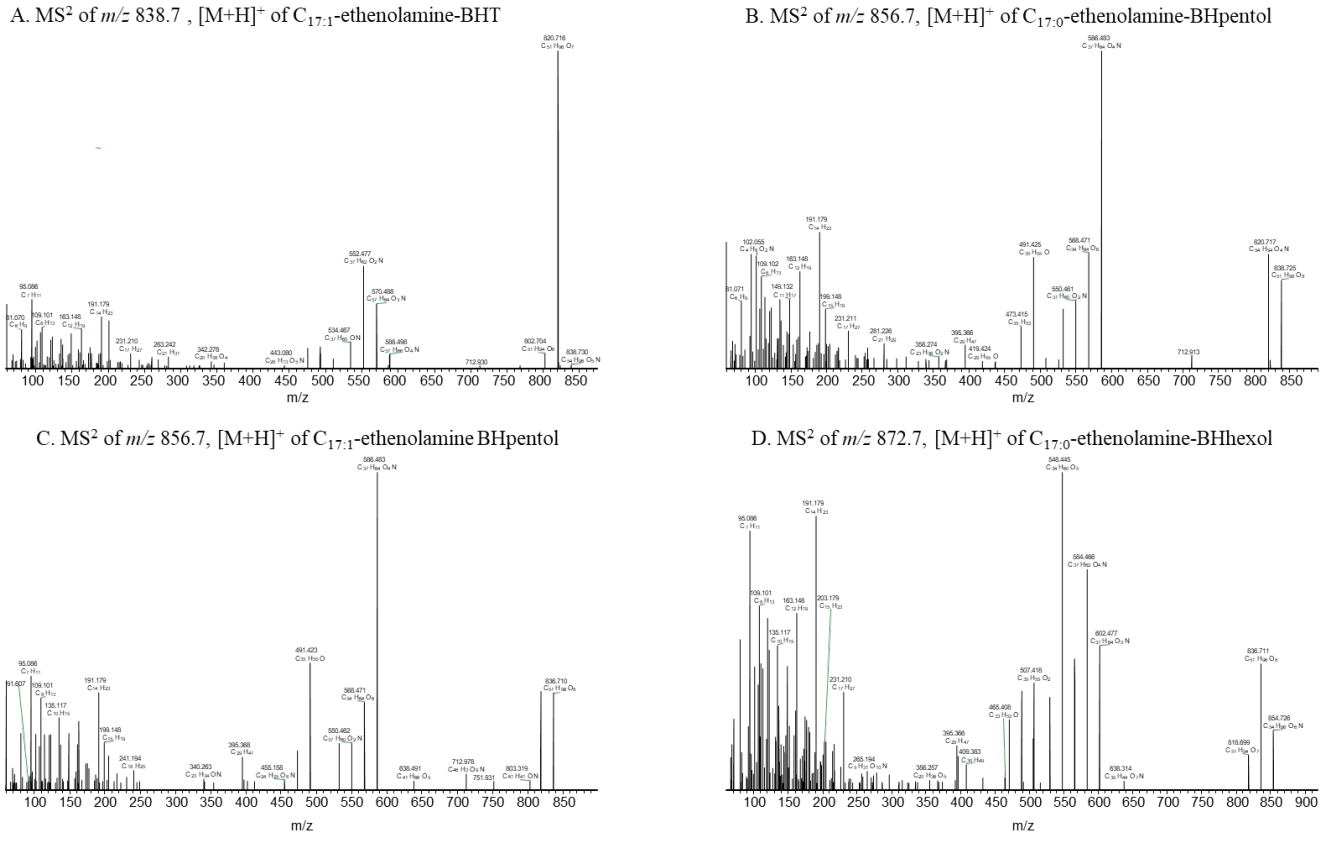


Fig. S10. Additional MS2 spectra of N-acyl-ethenolamineBHPs in Censo 0 m soil. A: C:17:1-N-acyl-ethenolamine-BHT, B: C17:0-N-acyl-ethenolamine-BHpentolexol, C: C17:1-N-acyl-ethenolamine-BHpentol, D: C17:0-N-acyl-ethenolamine-BHhexol.

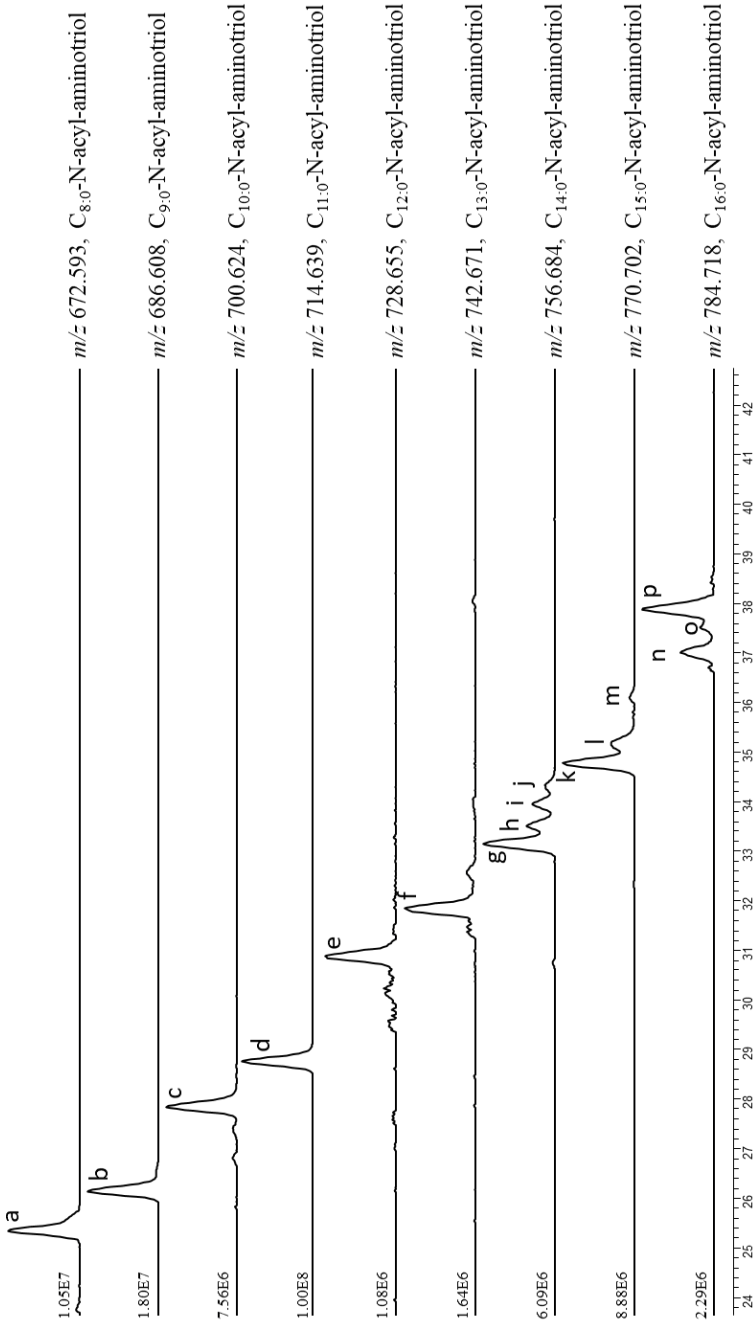


Fig. S11. Distribution of N-acyl-aminotetrols, with fatty acids with different chain lengths, in Censo 0 m, a soil from a terrestrial methane seep in Sicily. Traces are partial mass chromatograms. Indicated are the exact mass ($[M+H]^+$) used for searching, and the intensity (AU) of the signal.

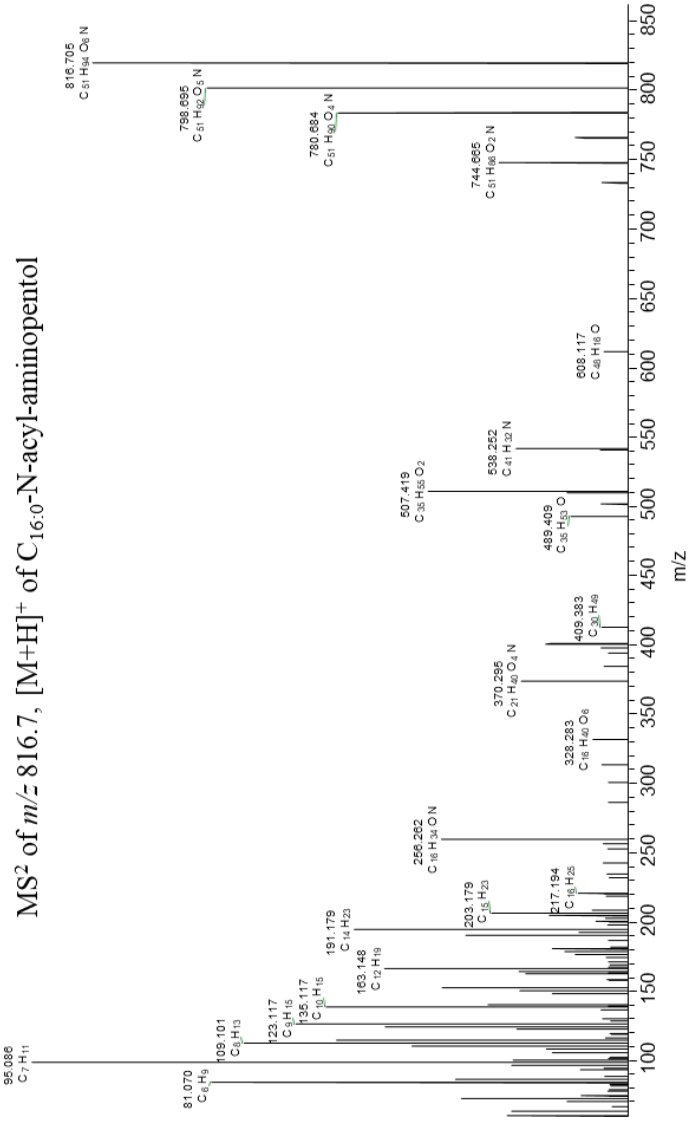


Fig. S12. MS^2 spectrum of $C_{16:0}$-N-acyl-aminopentol in Censo 0 m soil.

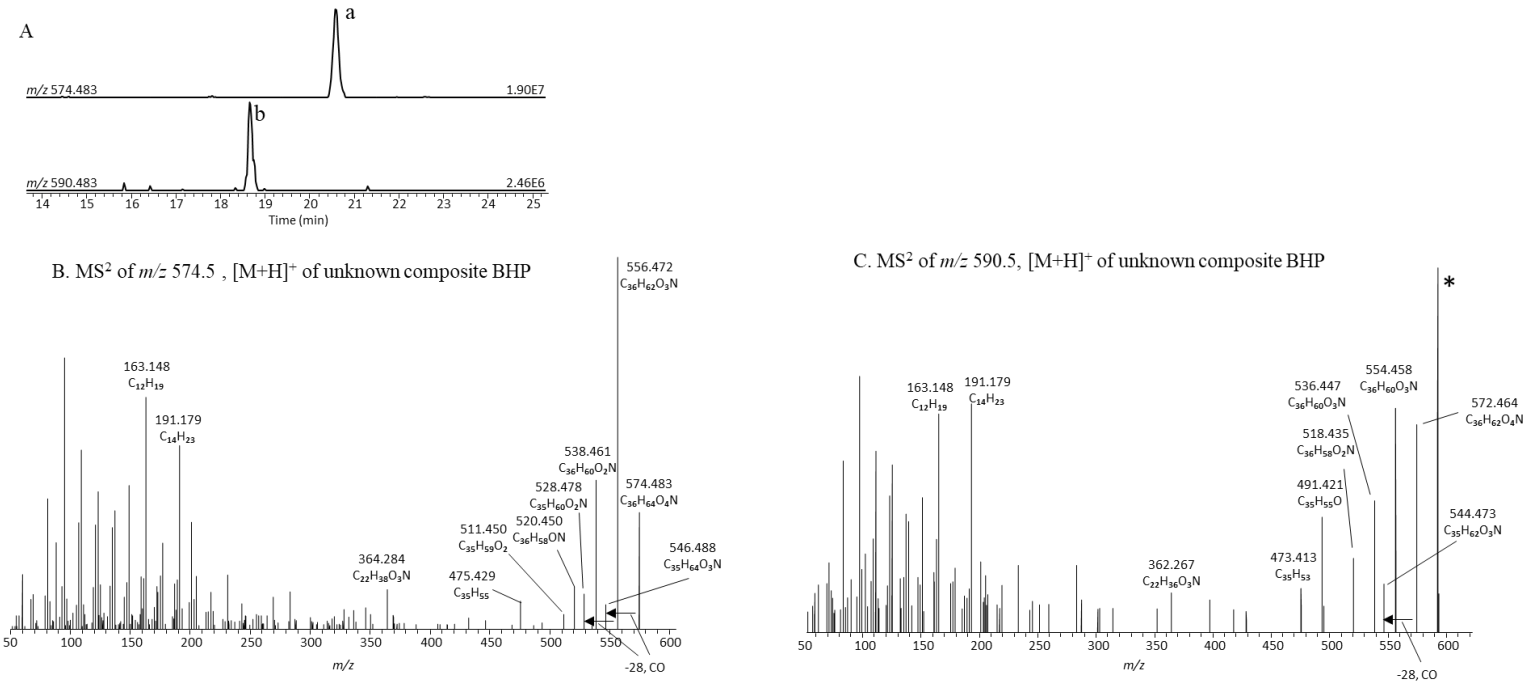


Fig. S13. Unknown composite BHPs in Censo 0 m soil. A: Partial mass chromatograms of two unknown composite BHPs (within 1 ppm mass accuracy), labeled with the calculated exact mass, and the intensity (AU) of the signal. B: MS^2 spectrum of an unknown composite BHP with an AEC of $C_{36}H_{64}O_4N^+$ and *m/z* 574.483 (peak a, panel A), C: MS^2 spectrum of an unknown composite BHP with an AEC of $C_{36}H_{64}O_5N^+$ and *m/z* 590.483 (peak b, panel A; mass peak labeled with * results from co-isolation of a co-eluting compound and is not related to this BHP). In addition to several consecutive losses of water representing 4 (panel A) or 5 (panel B) hydroxyl moieties on the side chain, the fragmentation spectra of both unknown composite BHPs is characterized by a loss 28 Da (CO) as indicated in panel B and C. The ions at *m/z* 364 (panel B) and 362 (panel C), respectively, appear to represent the complimentary fragment to the ion at *m/z* 191 (A and B ring).

Chapter 6

The abundance of nitrogen-containing bacteriohopanepolyols reflect aerobic methane oxidation at two terrestrial hydrocarbon seeps in Sicily

Nadine T. Smit, Ellen C. Hopmans, Laura Villanueva, Dina Boukhchtaber Castillo, Fausto Grassa, Carmen Hogendoorn, Huub J. M. Op den Camp, Jaap S. Sinninghe Damsté, Stefan Schouten, Darci Rush

In preparation for Geochimica et Cosmochimica Acta

Bissana seep near Cattolica Eraclea in Sicily, Italy

Fuoco di Censo everlasting fire in Sicily, Italy

Abstract

Bacterial lipid biomarkers such as bacteriohopanepolyol (BHP) derivatives that contain a nitrogen atom, e.g. in a terminal amine group (aminoBHPs), can be used to trace methanotrophic bacterial communities, and thus aerobic oxidation of methane, in modern and ancient environments. However, the distributions of N-containing BHPs as potential indicators for aerobic oxidation of methane have not yet been fully explored. Here, we investigate the distribution of N-containing BHPs for two terrestrial methane seep transects (Censo seep and Bissana seep) in Sicily, Italy. The Censo and Bissana seeps show high relative abundances of three aminoBHPs (aminotriol, aminotetrol and aminopentol) as well as recently identified ethenolamineBHPs (ethenolamine-BHT, -pentol and -hexol) and N-acyl-aminotriols at, and close to, the active seep sites. Principal component analysis (PCA) demonstrate that the aminoBHPs cluster together with the novel ethenolamineBHPs and N-acyl-aminotriols thus suggesting an origin likely from aerobic methanotrophs. 16S rRNA gene sequencing showed that Type I methanotrophs comprise up to 8% of the bacterial community at the Censo seep while they comprised only up to 0.2% at the Bissana seeps. Possibly, some groups within the Gammaproteobacteria, which comprised a major part of the bacterial 16S rRNA gene reads at the Bissana seeps, were also a source of aminoBHPs and their 3-methylated homologues. The Censo seep contained high relative abundances of a novel late eluting aminotriol, which was also identified in a verrucomicrobial strain, *Methylacidimicrobium cyclopophantes* 3B, suggesting it as a potential biomarker for these methanotrophic *Verrucomicrobia*. Both seep sites show increasing relative abundances of soil-marker BHPs (i.e. adenosylhopane and N1-methylinosylhopane) with increasing distance from the seeps and an opposite trend in the abundance of the N-containing BHPs. An index for aerobic oxidation of methane based on selected BHPs was developed which showed high values close to the seeps (≥ 0.4) and drop to < 0.2 at distances >3 m from the active seeps. Further development and application of this novel AminoBHP-index may offer a new biomarker tool to reconstruct present and past aerobic oxidation of methane in terrestrial environments.

1. Introduction

The greenhouse gas methane (CH_4) plays an important role in the Earth's carbon cycle (Reeburgh, 1996, 2007; Myhre et al., 2014; Dean et al., 2018). Atmospheric methane concentrations have substantially increased since preindustrial times, but terrestrial methane sources, sinks and chemical reaction

pathways remain poorly constrained (Reeburgh, 1996; Myhre et al., 2014; Schwietzke et al., 2016). Various natural sources are known to emit methane into the atmosphere, e.g. biogenic methane is emitted from wetlands, thawing permafrost soils, and marine sediments, while thermogenic methane is emitted from marine and terrestrial hydrocarbon seeps (Reeburgh, 1996; Etiope et al., 2009; O'Connor et al., 2010; Dean et al., 2018).

In many terrestrial and marine environments, biogenic and thermogenic methane is subject to microbial oxidation processes, which convert methane to CO_2 thus reducing methane emission to the atmosphere and affecting biogeochemical carbon cycling (Hanson and Hanson, 1996). The anaerobic oxidation of methane (AOM) is conducted by anaerobic methanotrophic archaea (ANMEs) found in marine (Hinrichs et al., 1999; Boetius et al., 2000; Pancost et al., 2000a) and freshwater environments (Takeuchi et al., 2011; Haroon et al., 2013). Aerobic oxidation of methane plays a key role in removing methane not only in marine environments (e.g. oxic water columns) but also in the terrestrial realm (e.g. soils or peatlands) (Bull et al., 2000; Bodelier et al., 2009; Bowman, 2011). The main aerobic methane-oxidizing bacteria (MOBs) are from three different phylogenetic groups, i.e. *Gammaproteobacteria* (Type I methanotrophs), *Alphaproteobacteria* (Type II methanotrophs) and the (thermo)acidophilic *Verrucomicrobia* (Hanson and Hanson, 1996; Pol et al., 2007; Op den Camp et al., 2009; Schmitz et al., 2021).

In addition to DNA-based techniques, lipid biomarkers offer a tool to investigate bacterial communities involved in methane oxidation in the environment. Lipids are more stable over long timescales than DNA, which favours their use to reconstruct past biogeochemical cycles, including the methane cycle (Hofreiter et al., 2001; Brocks and Pearson, 2005; Boere et al., 2011). Besides fatty acids and hopanoids (Bull et al., 2000; Crossman et al., 2005; Dedysh et al., 2007; Bodelier et al., 2009), bacteriohopanepolyols and their derivatives (here abbreviated as BHPs) have been used to trace MOBs and thus the aerobic oxidation of methane (van Winden et al., 2012; Talbot et al., 2014; Wagner et al., 2014). BHPs are a diverse group of compounds with a large structural diversity especially in the length, position and complexity of the functional groups located on the side chain (e.g. Neunlist et al., 1985; Talbot and Farrimond, 2007; Talbot et al., 2007a, b; Cooke et al., 2008a; Hopmans et al., 2021).

Previous studies suggested that methylation at the C-3 position in hopanoids, BHPs and especially aminoBHPs (structures **a-c**; see Fig. 1) is specific to MOBs (Summons et al., 1994; Farrimond et al., 2004). However,

Welander and Summons (2012) identified the *hpn*R gene responsible for the methylation at C-3 in hopanoids and revealed it is distributed amongst various bacterial taxa other than methanotrophs. Moreover, the 3-methylation was only found in a limited number of Type I methanotrophic bacteria (Welander and Summons, 2012). In general, 3-methyl hopanoids and BHPs can be used as biomarkers for aerobic oxidation of methane when accompanied by depleted ^{13}C isotopic signatures mainly in marine environments (Collister et al., 1992; Summons et al., 1994; Jahnke et al., 1999). However, recent investigations demonstrated that not all MOB-derived lipids have substantially depleted ^{13}C values, especially when derived from Type II methanotrophs and NC10 bacteria in terrestrial environments such as modern peat bogs, past lignite deposits or wetlands (Pancost et al., 2000b; van Winden et al., 2010; Ettwig et al., 2010; Rasigraf et al., 2014; Inglis et al., 2018).

Aerobic methanotrophs in culture produce a range of distinct N-containing BHPs, such as those with an amine group at position C-35 and 3, 4 or 5 hydroxy groups on the side chain of the hopanoid skeleton (structures **a**-**c**; Fig. 1) (Cvejic et al., 2000; van Winden et al., 2012; Talbot et al., 2014; Rush et al., 2016). The 35-aminobacteriohopane-30,31,32,33,34-pentol (‘aminopentol’; structure **c**) is produced almost solely by Type I methanotrophs (Neunlist and Rohmer, 1985; Cvejic et al., 2000). Furthermore, aminopentol and its unsaturated homologues were detected in different environments like soils, peats, geothermal sinters, the water column, or marine and lake sediments (Gibson et al., 2008; Sáenz et al., 2011; van Winden et al., 2012; Talbot et al., 2014; Rush et al., 2016; Kusch et al., 2019). Previous studies have used aminopentol as biomarker to trace aerobic oxidation of methane in the marine realm (Talbot and Farrimond, 2007; Sáenz et al., 2011; Talbot et al., 2014). However, a more recent study (Rush et al., 2016) showed that aminopentol might be a more reliable marker for MOBs in methane-influenced terrestrial environments than marine environments as aminopentol was only found to be present in one marine Type I methanotroph species and in low abundance in sediments and authigenic carbonates from marine methane-rich environments.

Finally, two additional amino-BHPs are synthesized by Type I and Type II methanotrophs as well as by methanotrophic *Verrucomicrobia*: 35-aminobacteriohopane-31,32,33,34-tetrol (‘aminotetrol’; **b**) and 35-aminobacteriohopane-32,33,34-triol (‘aminotriol’; **a**) (Talbot et al., 2001, 2014; Talbot and Farrimond, 2007; van Winden et al., 2012). However, aminotetrol is also biosynthesized in low relative abundances by some sulfate-reducing bacteria of the Genus *Desulfovibrio* (Neunlist et al., 1985; Blumenberg et al., 2006;

Talbot and Farrimond, 2007), while aminotriol is produced by a wide variety of non-methanotrophic bacteria like sulfate-reducing bacteria or cyanobacteria (Talbot et al., 2008; Blumenberg et al., 2012). Therefore, these BHPs, and especially aminotriol, appear less suitable for tracing MOBs and aerobic oxidation of methane.

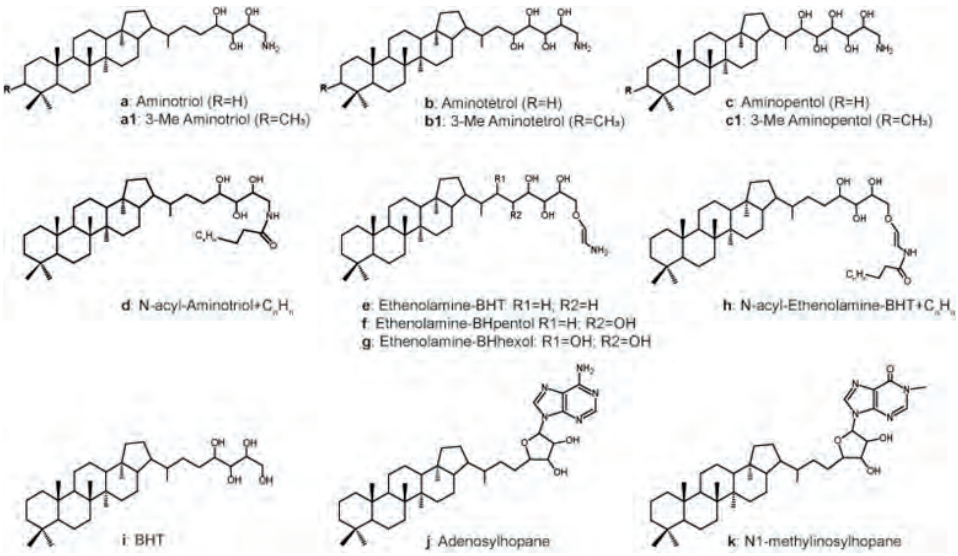


Figure 1. Chemical structures of the most relevant bacteriohopanepolyol derivatives in the Bissana and Censo seep soils (for indicative mass spectra see Hopmans et al., 2021).

Despite soils being one of the biggest microbial sinks for atmospheric methane (Fierer et al., 2012; Dean et al., 2018), only few studies in terrestrial environments have used BHPs to trace the presence of MOBs. To date, only peatlands geothermal sinters and lake sediments were investigated for MOB specific BHPs (Coolen et al., 2008; Gibson et al., 2008; van Winden et al., 2012; Talbot et al., 2016) which showed the presence of all three aminoBHPs and some of their 3-methylated homologues. In these environments, aminopentol and its 3-methyl homologue were suggested to derive from Type I methanotrophs, while aminotetrol originates from Type I and Type II methanotrophs and its 3-methyl homologue from Type I methanotrophs. Aminotriol and its 3-methyl counterpart are known to derive from Type II methanotrophs but also from other more ubiquitous bacterial species such as Acidobacteria (Gibson et al., 2008; van Winden et al., 2012; Talbot et al., 2016; Sinninghe Damsté et al., 2017).

In this study, we used an ultra-high pressure liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS) method (Hopmans et al., 2021) to analyse the diversity of BHPs along transects of two continuous terrestrial CH_4

seeps, Bissana and Fuoco di Censo (everlasting fire) in Sicily, Italy (Etiope et al., 2002; Smit et al., 2021a). Previous studies on a few soils from the Censo seep showed the presence of methanotrophic bacteria and ^{13}C depleted fatty acids (Smit et al., 2021a) as well a large diversity in BHPs (Hopmans et al., 2021), making them excellent candidates to investigate the presence of BHPs selective for aerobic oxidation of methane. The changes in distribution of BHPs were evaluated in tandem with those in the prokaryotic community as determined by 16S rRNA gene amplicon sequencing. Our results led to the establishment of an index that can potentially be used to trace back aerobic oxidation of methane in modern and past terrestrial environments.

2. Material and methods

2.1 Study sites

The Bissana seep (37°29'00.4"N, 13°23'17.5"E) near Cattolica Eraclea and the Fuoco di Censo seep (37°37'30.1"N, 13°23'15.0"E), in the following referred to as Bissana seep and Censo seep, respectively, are both located in the mountains of Southwestern Sicily, Italy (Etiope et al., 2002, 2007; Grassa et al., 2004). The investigated area is part of the Alpine orogenic belt in the Mediterranean and located along the boundary of the African and European plates (Basilone, 2012). The Bissana and Censo seeps are located in the Bivona and Cattolica Eraclea areas, which are characterized by a complex geological setting including formations with sandy clays, marls or carbonates from the Tortonian-Messinian often covered by thrusting limestones or other Pliocene/Pleistocene deposits (Trincianti et al., 2015).

The Bissana seep is characterized by a wide pool (about 10 m in diameter) in which gases, mud and salty waters gurgle with varying intensities from different smaller gas and mud seeps (Piepoli, 1931; Etiope et al., 2002). The pool contains dry and wet areas with active seeps at its edges. The gas composition at the Bissana seep mainly contains methane (96.2%) with minor proportions of other gases like N_2, CO_2, He or H_2 (Etiope et al., 2002). Soil degassing occurs with an average methane flux of 1 x 10^5 mg $m^{-2}d^{-1}$ and a total methane emission of 2.7 tons yr^{-1} (Etiope et al., 2002). The released methane has a $\delta^{13}C$ value of –47‰ and C1/(C2+C3) gas ratios greater than ca.130 (Grassa et al, unpublished results).

The Censo seep is a typical example of a natural ‘everlasting fire’ at which the released gas consists mainly of methane (76-86%) and N_2 (10-17%) as well as some other minor gases like CO_2, O_2, ethane, propane, He and H_2 (Etiope et al., 2002; Grassa et al., 2004). A diffuse soil degassing was present

within an area of 80 m^2 with an average methane flux of 7 x 10^6 mg m^{-2} d^{-1} and a total methane emission of 6.2 tons yr^{-1} (Etiope et al., 2002, 2007). The methane is suggested to be generated by the thermal alteration of organic matter from overmature marine source rocks (kerogen type II) and is characterized by a $\delta^{13}C$ value of –35‰ and C1/(C2+C3) ratios greater than 100 (Grassa et al., 2004).

2.2 Samples

Soil and mud samples of the Bissana seep and Censo seep (Smit et al., 2021a) were taken during a field campaign in October 2017. The samples were obtained in triplicate within 0.5 m^2 from a horizon 5 to 10 cm below the soil surface and at different distances from the active seeps (Table 1). At the Censo seep, the area with the major gas flux is determined as the main seep site (Censo 0 m), whereas at the Bissana seep three active seep sites (Bis seep 1, 2 and 3) were discovered, which all actively bubbled a mixture of gas and liquid mud (Table 1). The in-situ temperature of the soil and mud from both seep sites at the time of collection were ranging between 18–20 ºC. The soil and mud from both seep sites were directly transferred into a clean geochemical sampling bag and stored frozen at –20 ºC until freeze drying and extraction.

A mesophilic acidophilic verrucomicrobial strain (*Methylacidimicrobium cyclopophantes* 3B*)* was isolated from a volcanic soil taken from the Solfatara crater, which is at the centre of the Campi Flegrei caldera, near Naples (Italy) (van Teeseling et al., 2014). The strain was grown under optimal conditions as described by van Teeseling et al. (2014) at 44 °C and a pH of 2.5.

2.3 Lipid extraction

Freeze-dried Bissana (ca. 5–6 g) and Censo (ca. 12 g) seep soil and mud as well as freeze-dried bacterial biomass of *M. cyclopophantes* 3B were extracted with a modified Bligh and Dyer extraction (Bale et al., 2013). The samples were ultrasonically extracted for 10 min with a solvent mixture containing methanol (MeOH), dichloromethane (DCM) and phosphate buffer (2: 1: 0.8, v: v: v). After centrifugation, the solvent was collected, combined and the residues re-extracted twice. A biphasic separation was achieved by adding additional DCM and phosphate buffer to a ratio of MeOH, DCM and phosphate buffer (1: 1: 0.9, v: v: v). The aqueous layer was washed two more times with DCM and the combined organic layers dried over a Na_2SO_4 column followed by drying under N_2. Aliquots of the Bligh and Dyer extracts (BDEs) were obtained and 0.4 µg/ml (Censo) or 0.04 µg/ml (Bissana) DGTS-D9 (1,2-dipalmitoyl-sn-glycero-3-O-4'-[N,N,N-trimethyl(D9)]-homoserine, Avanti Polar Lipids (Alabaster, AL)) was

added as internal standard before they were filtered through a 0.45 µm regenerated cellulose filter with 4 mm diameter and evaporated until dryness. Finally, the filtered BDEs were dissolved in MeOH:DCM (9:1) and analysed using UHPLC-HRMS.

2.4 UHPLC-HRMS of BHPs

Intact BHP analysis on the extracts of the soil samples and the *M. cyclopophantes* 3B biomass was performed by UHPLC-HRMS according to Hopmans et al. (2021) using an Agilent 1290 Infinity I UHPLC, equipped with thermostatted autosampler and column oven, coupled to a quadrupole-orbitrap HRMS equipped with an Ion Max source (Thermo Fisher Scientific, Waltham, MA) with heated electrospray ionization (HESI) probe. A C_{18} BEH column (2.1x 150 mm, 1.7 µm particle; Waters) fitted with a pre-column, and a solvent system consisting of eluent A methanol:H_2O (85:15) and eluent B methanol:isopropanol (1:1) both containing 0.12% (v/v) formic acid and 0.04% (v/v) aqueous ammonia were used for separation. Compounds were eluted with 5%B for 3 min, followed by a linear gradient to 60% B at 12 min and then to 100% B at 50 min, with a total run time of 80 min. The flow rate was constant at 0.2 mL min^{-1}. Positive ion HESI settings were: capillary temperature, 300 °C; sheath gas (N2) pressure, 40 arbitrary units (AU); auxiliary gas (N2) pressure, 10 AU; spray voltage, 4.5 kV; probe heater temperature, 50 °C; S-lens 70 V. Detection was achieved using positive ion monitoring of *m/z* 350-2000 (resolving power 70,000 ppm at *m/z* 200), followed by data dependent MS^2 (isolation window 1 *m/z*; resolving power 17,500 ppm at *m/z* 200) using a top 10 approach, dynamic exclusion (6s) and an inclusion list of calculated exact masses of ca. 165 BHPs (from literature), with 3 ppm mass tolerance. Optimal fragmentation of all BHPs was achieved using a stepped normalized collision energy of 22.5 and 40. A mass calibration was performed every 48 h using the Thermo Scientific Pierce LTQ Velos ESI Positive Ion Calibration Solution.

BHPs were identified based on the exact mass of the protonated or ammoniated molecular ion (see Tables S1 and S2), relative retention time and MS^2 fragmentation spectra conform Hopmans et al. (2021). Individual BHPs were (semi-)quantified based peak integrations of selected mass chromatogram of the most abundant molecular ion (see abundance in Tables S1 and S2 for the exact mass of ions used in integration) using a 3 ppm mass accuracy window. The peak area of the DGTS-d9 internal standard was used to control for MS variability and ion suppression.

Table 1. Sampling sites and description of the soil and mud samples from the two terrestrial gas seeps Bissana and Censo in Sicily, Italy.

Site	Distance from seep (m)	Sample ID	Sample description
Bissana			
	At seep 1	Bis seep 1	Mud soil mix; directly from seep 1 with active gas and mud bubbling
	1.0	Bis1 +1 m	Dry soil with little vegetation;
	3.0	Bis1 +3 m	Dry soil with vegetation
	8.0	Bis1 +8 m	Dry soil with vegetation
	14.0	Bis1 +14 m	Dry soil with vegetation (control site seep 1)
	Near seep 2	Bis near seep 2	Mud from dried mud pool; next to bubbling mud and gas pool (seep 2)
	At seep 3	Bis seep 3	Mud from bubbling gas and mud seep 3
	3.0	Bis3 +3 m	Dry soil with vegetation
Fuoco di Censo			
	At seep	Censo 0 m	Dry soil; directly from main gas seep
	2	Censo +2 m	Dry soil with vegetation; uphill from main seep site
	1.5	Censo +1.5 m	Dry soil with little vegetation; uphill from main seep site
	0.8	Censo –0.8 m	Dry soil; downhill from main seep site
	1.2	Censo –1.2 m	Dry soil; downhill from main seep site
	1.8	Censo –1.8 m	Dry soil; downhill from main seep site
	2.2	Censo –2.2 m	Dry soil; downhill from main seep site
	13.2	Censo –13.2 m	Dry soil with vegetation; downhill 13.2 m from main seep site (control site)

As no authentic standards were available for the BHPs, abundances are reported based on peak areas of the BHP (i.e. in arbitrary units (AU)/ g soil). The

abundances of specific BHPs were calculated relative to total integrated BHPs and BHPs in the triplicate samples for each distance were averaged assuming identical MS response factors and reported as percentages. BHPs included in the BHP inventory are listed in Table S1 for Censo seep soils, Table S2 for Bissana seep soils and Table S3 for biomass of *M. cyclopophantes* 3B.

Principal Component Analysis (PCA) was performed on the combined data set of the Bissana and Censo seep soil samples where only BHPs were used which occurred at both seeps, fourteen in total. PCA was performed using Sigmaplot 14.0.

2.5 DNA extraction and 16S rRNA gene amplification

DNA was extracted from the Censo and Bissana soils and muds with the PowerMax soil DNA isolation kit (Qiagen) and DNA extracts were stored at -80 °C until further analysis. The general 16S rRNA archaeal and bacterial primer pair 515F and 806RB targeting the V4 region was used for the 16S rRNA gene amplicon sequencing (Caporaso et al., 2012; Besseling et al., 2018). Polymerase chain reaction (PCR) products were gel purified using the QIAquick Gel-Purification kit (Qiagen), pooled and diluted. Sequencing was performed using an Illumina MiSeq sequencing platform at the Utrecht Sequencing Facility (Utrecht, the Netherlands). The 16S rRNA gene amplicon sequences were analysed by an in-house pipeline (Abdala Asbun et al., 2020) that includes quality assessment by FastQC (Andrews, 2010), assembly of the paired end reads with Pear (Zhang et al., 2013), and assignment of taxonomy (including picking representative set of sequences with 'longest' method) with blast by using the Silva 132 release as a reference database (https://www.arb-silva.de/). The relative abundances of bacterial phyla and methanotrophic bacterial species are reported as a percentage of total bacterial 16S rRNA gene reads and average for the triplicate samples of each distance (Table S4).

3. Results

The Censo seep contains one active seep site, at which dry gas is emitted through the soils, although soil degassing was present within an area of 80 m^2 around the site and thus likely affecting soils within a few meters of the seep site. Soil samples (in triplicate) were investigated uphill and downhill from the seep site with increasing distance from the main seep (Table 1). In contrast to the Censo seep area, at the Bissana seep three separate seep sites were discovered that showed active mud and gas bubbling around and in a mud pond. For seep 1, mud and soils were sampled with increasing distance from the active seep (Table 1).

For seep 3, the seep itself and a soil from 3 m away were sampled. For seep 2 only a soil from close to the seep was obtained.

3.1 BHP composition

The BHP composition of the samples taken at the two Sicilian gas seep sites was analysed using UHPLC-HRMS and relative abundances, normalized on the sum of all quantified BHPs, were calculated based on the integrated peak areas of mass chromatograms of either the protonated or ammoniated molecular ion (Tables S1 and S2). Soils are known to be heterogenous and therefore the BHP distributions of the triplicate samples, taken within 0.5 m^2, were averaged.

3.1.1 Censo seep.

The Censo seep soils contain a diversity of N-containing BHPs, whose structural identification was described in detail by Hopmans et al. (2021) for a Censo soil sampled directly at the seep (Censo 0 m). Here we describe the BHPs relevant to MOBs and the most commonly occurring BHPs (Table S1). Directly at the seep (Censo 0 m), there are high relative abundances of the aminoBHPs aminotriol (13.4%), a late-eluting aminotriol isomer (5.7%), aminotetrol (1.1%) and aminopentol (6.9%) (Fig. 2a). High relative abundances of aminotriol and its isomer are found in all soils but are particularly high close to the seep in –0.8 m and –1.2 m, together comprising over 50% of total BHPs. Aminotetrol is only present in high relative abundance directly at the seep and occurs in minor abundances at –0.8 m distance from the seep, while aminopentol is detected in all soils, including the soil at –13.2 m from the seep, although with the highest relative abundances at and close to the main seep. With respect to 3-methylated aminoBHPs, only minor amounts of the 3-methyl aminotriol were detected in soils of the Censo transect.

We also detected the novel composite N-containing BHPs, ethenolamine-BHT, ethenolamine-BHpentol and ethenolamine-BHhexol (structures **e-g**; Fig. 1), which were recently identified in a Censo seep soil (Hopmans et al., 2021). In general, almost all soils contain the three ethenolamineBHPs of which especially high abundances occur between 0 and –2.2 m distance from the seep with exception of the soil at –0.8 m distance, while the remaining three soils more remote from the seep had low (<5%) relative abundances (Fig. 2b). Ethenolamine-BHT shows the highest relative abundance ranging from 1.2 to 11%, followed by ethenolamine-BHpentol with 0.2 to 13% and ethenolamine-BHhexol with 0.5 to 2.5% of total BHPs.

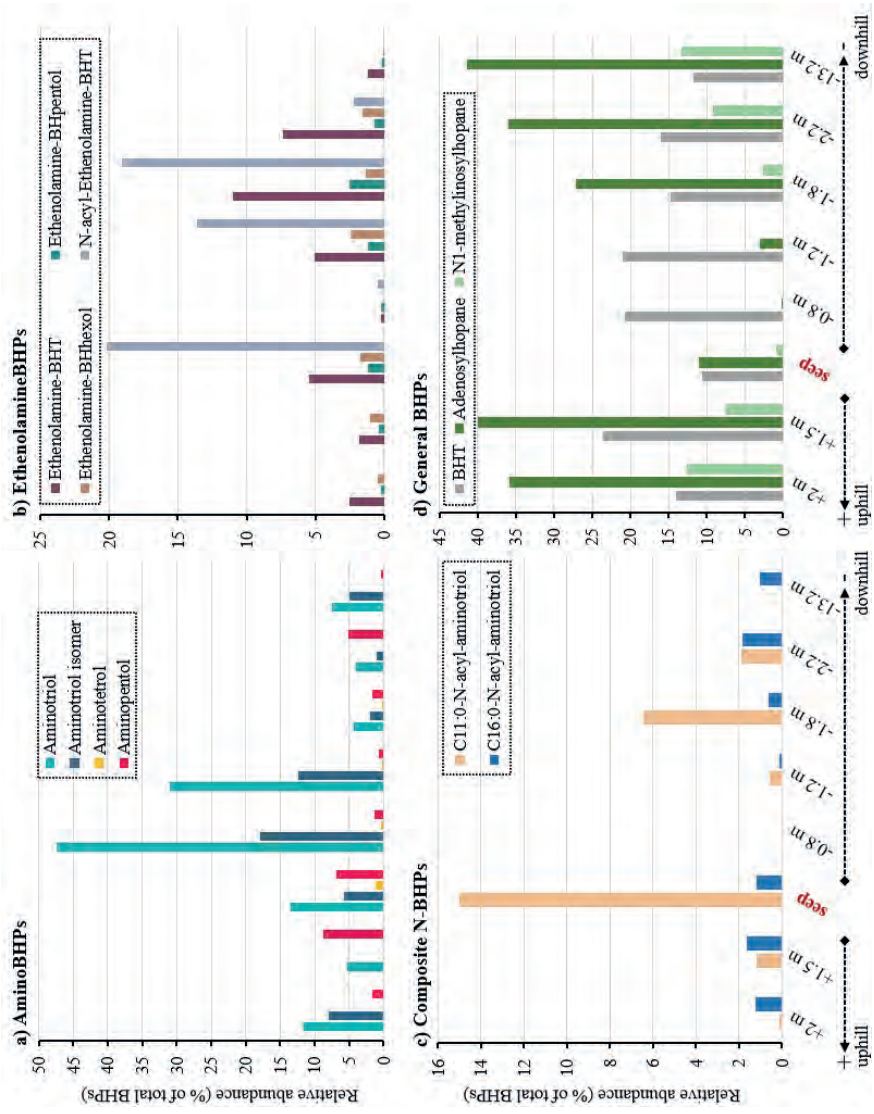


Figure 2. Relative abundances of BHPs in the Censo seep soils at increasing distance uphill (+) and downhill (–) from the main seep site. The bar plots show a) aminoBHPs, b) ethenolamineBHPs, c) composite N-containing BHPs (composite N-BHPS) and d) general BHPs.

Additionally, directly at the seep (Censo 0 m), a high relative abundance (20%) of N-acyl-ethenolamine-BHT bound to fatty acids $C_{15:0}$ to $C_{18:0}$ (structure **h**) were observed, which remained high at –1.2 m (14%) and –1.8 m (19%) distance from the main seep and then decreased with further distance (Fig. 2b). Furthermore, N-acyl-aminotriols (structure **d**) were found to be abundant at the Censo seep of which the most abundant one, the $C_{11:0}$-N-acyl-aminotriol, shows high abundances at the seep (15%) and at –1.8 m (6.4%) distance but decreases in relative abundance further away from the seep (at +1.5 m, –1.8 m and –2 m, 0.1 to 1.9%) (Fig. 2c). In contrast, the less abundant $C_{16:0}$-N-acyl-aminotriol showed no particular trend over the seep transect.

The relative abundance of BHT (structure **i**) ranges from 11 to 24% in the seep transect but shows not trend with increasing distance from the seep (Fig. 2d). Furthermore, directly at the seep, and at –0.8 m and –1.2 m distance from the seep, low relative abundances of adenosylhopane (structure **j**; previously adenosylhopane type 1) and N1-methylinosylhopane (structure **k**; previously named adenosylhopane type 2; see Hopmans et al., 2021 for structural description) are observed (from 0.2 to 10.9%), whereas the soils at increased distance from the seep contain high relative abundances, with adenosylhopane and N1-methylinosylhopane making up more than 50% of total BHPs (Fig. 2d).

3.1.2 Bissana seeps.

Analysis of the BHP inventory showed that directly at Bissana seep 1 (Bis seep 1), there are high relative abundances of aminotriol (26% of total BHPs), aminotetrol (6.8%) and aminopentol (15%) (Fig. 3a; Table S2). A similar distribution is also observed near seep 2 (Bis near seep 2) and at seep 3 (Bis seep 3), where high relative abundances of total BHPs of aminotriol (15.3 and 19.9%), aminotetrol (2.4 and 1.9%) and aminopentol (2.9 and 2.9%) are detected, but with slightly lower relative abundances than at seep 1. At seep 1, the relative abundances of these BHPs decrease with increasing distance from the seep (Fig. 3a). This is also observed at seep 3, where aminoBHPs decrease at 3 m from seep 3 (Bis3 +3 m) to only 0.1% (aminopentol), 0.9% (aminotetrol) and 3.2% (aminotriol) of total BHPs. Furthermore, 3-methyl homologues of the three aminoBHPs occur relatively abundant only directly at or near the three seeps and disappear with increasing distance from the seeps (Fig. 3b). 3-Methyl aminotriol is the most abundant methylated BHP with relative abundances ranging from 1.7 to 5.2%, while 3-methyl aminotetrol shows values of 0.1 to 1.6% and 3-methyl aminopentol 0.4 to 1.4%.

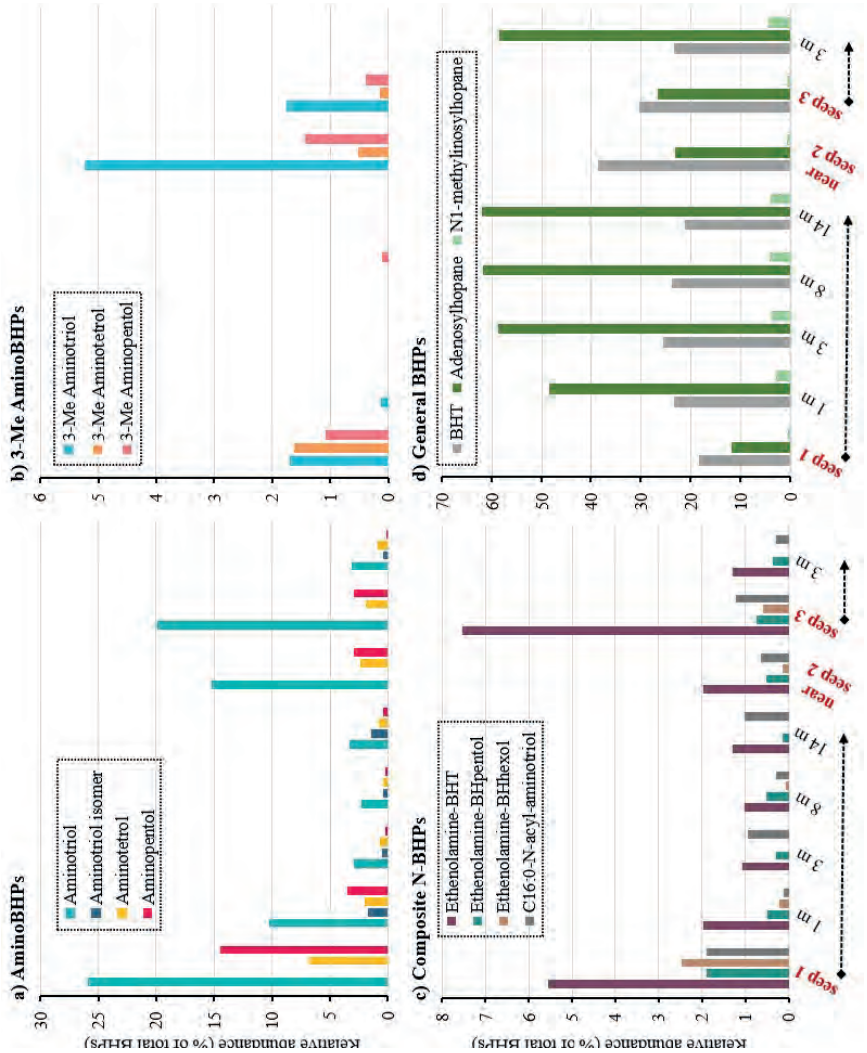


Figure 3. Relative abundances of BHPs in the Bissana seeps soils at increasing distance from seep 1 and seep 3 as well as near seep 2. The bar plots show a) aminoBHPs, b) 3-methyl aminoBHPs (3-Me aminoBHPs), c) composite N-containing BHPs (composite N-BHPS) and d) general BHPs.

In the seeps and some soils close to the seeps (i.e. Bis1 +1 m), there are high relative abundances of the three novel ethenolamineBHPs (Fig. 3c). Ethenolamine-BHT shows the highest relative abundance ranging from 2 to 7.5%, followed by ethenolamine-BHpentol with 0.5 to 1.9% and ethenolamine-BHhexol with 0.1 to 2.5% of total BHPs. Additionally, another previously identified N-containing BHP, namely $C_{16:0}$-N-acyl-aminotriol (Talbot et al., 2007a), was identified at the Bissana seeps in relative abundances up to 1.9 % but this BHP was also detected in some soils away from the three seeps (Fig. 3c).

At the soil transect, the relative abundance of BHT varies from 18 to 39% but does not follow a particular trend with respect to distance from the seep (Fig. 3d). The general BHPs adenosylhopane and N1-methylinosylhopane were detected in all Bissana soils and show a clear increase in their relative abundances with increasing distance from the seeps (Figs. 3d). At Bissana seep 1, adenosylhopane shows a relative abundance of 12%, which increases with increasing distance from seep 1 until it reaches its maximum at 14 m distance from Bissana seep 1 (62%). Near Bissana seep 2 and at seep 3 there is a relatively lower abundance of adenosylhopane (23–27%) and increases with 3 m distance to seep 3 to 59%. A similar trend is found for N1-methylinosylhopane, albeit with lower relative abundances compared to adenosylhopane (0.6–4.5%; Fig. 3d).

3.1.3 Principal component analysis (PCA) of the BHP distributions.

To investigate the variance in the BHP distributions in the seep transects, PCA analysis was performed on the fractional abundances of the fourteen BHPs which are found at both sites (Fig. 4). For this analysis, we used the BHP distributions of the triplicate samples as separate input. Three principal components (PCs) account for 26.3% (PC 1), 19.9% (PC 2) and 18.6% (PC 3) of the total variance in the dataset. All N-containing BHPs (aminoBHPs, ethenolamineBHPs and N-acyl-aminotriol) score positively on PC1, whereas adenosylhopane, N1-methylinosylhopane and 2-Me BHT have a negative loading on PC 1 (Fig. 4a). BHT and 3-Me BHT show a high loading at PC 2, while aminotriol and the aminotriol isomer show a negative loading on PC 2.

The sample loading plot (Fig. 4b) shows that the triplicates samples from within 0.5 m^2 score sometimes quite differently, indicating the heterogeneity of the BHP distributions within short distances. The seeps score more positive on PC1, while the soils more remote from the seeps tend to have more negative scores on PC1. The Bissana seep soils score generally higher on PC 2 than the Censo seep soils.

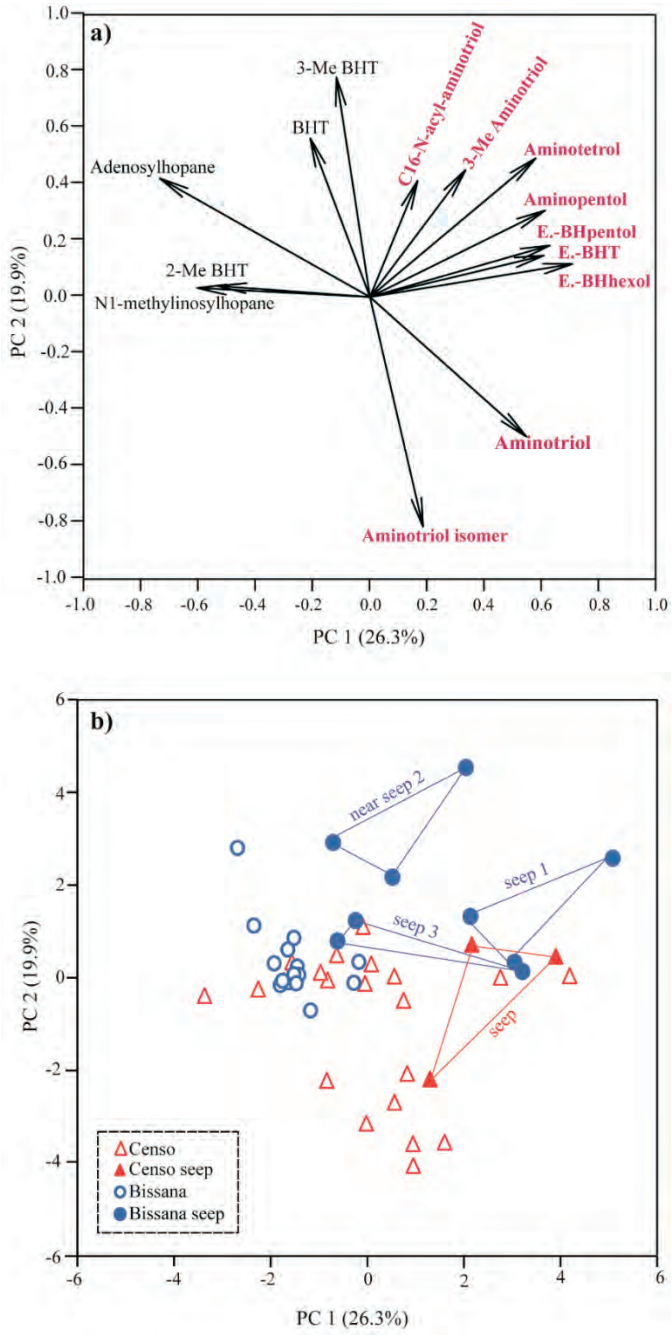


Figure 4. Principal component analysis (PCA) of the fractional abundances of different BHPs and triplicate samples in the Censo seep and Bissana seeps soils (n=48). The first principal component (PC) explains 26.3% and the second 19.9% of the variance. Panel a) shows PC 1 versus PC 2 of the 14 different BHPs abundant at both seep sites. Panel b) shows the sample loading plot with the triplicate samples (taken within 0.5 m^2) at both seeps.

3.2 Bacterial diversity

16S rRNA gene amplicon sequencing performed on the DNA extracted from the two Sicilian gas seep sample sets was used to investigate the bacterial diversity in the seep transects. On average 87,200 reads per sample were obtained of which on average 94% were attributed to bacteria. The archaeal 16S rRNA gene reads were not considered here as Archaea are not sources of BHPs. The remaining bacterial 16S rRNA sequences were assigned into operational taxonomic units (OTUs) and further phylogenetically classified. The number of bacterial reads of the different bacterial phyla were then normalized as a percentage to total number of bacterial 16S rRNA gene reads (Table S4). Finally, the bacterial 16S rRNA gene distributions of the triplicate samples, taken within 0.5 m^2, were averaged.

3.2.1 Censo seep.

A variety of diverse bacterial phyla was detected in the Censo seep soils with high relative abundances of Actinobacteria at (14 %) and close to (99 %) the seep site (Fig. 5a). At and close to the Censo seep, representatives of the phylum Actinobacteria are mainly novel ethane-oxidizing Mycobacteria (Smit et al., 2021a, b). High relative abundances of Gammaproteobacteria ranging from 9 to 29 % are found at and close to the Censo seep site and are lower at the –13.2 m control site (5 %). Furthermore, the high relative abundances of Patescibacteria (4 %) and Cyanobacteria (4 %) at the seep is noteworthy; they are not detected in most other samples. In the –1.8 m soil, the high relative abundance of Chloroflexi (49 %) stands out since they are much less predominant in the other soils. The high relative abundances of Verrucomicrobia (38 %) in the soil at –1.2 m distance from the seep is also remarkable. At the seep site and also close to the seep (–1.2 m and –1.8 m), there are low relative abundances (<3 %) of Acidobacteria, but their relative abundance increases with increasing distance uphill and downhill from the seep (≥ 5%). Additionally, other bacterial phyla such as Deltaproteobacteria, Alphaproteobacteria and Planctomycetes were also present in the soils but did not show any particular trend over the seep transect.

High percentages of bacterial reads of known MOBs, i.e. Type I and Type II methanotrophs and Verrucomicrobia, are found at the Censo seep site (Fig. 5b). An exceptionally high relative abundance of about 38% of reads related to the verrucomicrobial methanotroph family *Methylacidiphilaceae* is revealed at –1.2 m distance from the Censo seep, while its relative abundance is only minor in the seep soils (≤ 0.2%). Furthermore, 16S rRNA gene reads attributed to the Type I methanotroph families *Methylomonaceae* and

Methylococcaceae, ranging from 0.5 to 6 %, were present at the seep as well as in the soils at +1.5 m, -1.2 m, –1.8 m and –2.2 m distance from the seep site. Minor relative abundances of 16S rRNA gene reads assigned to the Type II methanotrophs (≤ 0.4%), from the genera *Methylovirgula*, *Methylosinus* and *Methylobacterium*, were detected at and around the Censo seep site but remained undetected at 13.2 m distance.

3.2.2 Bissana seeps.

The 16S rRNA gene amplicon sequencing showed high relative abundances of Gammaproteobacteria directly at seep 1, seep 3 and near seep 2 ranging from 49 to 56 %, while their relative abundances decreased with increasing distance to the seeps to <6% of total bacterial reads (Fig. 6a). An opposite distribution is shown by the 16S rRNA gene reads of Actinobacteria and Acidobacteria, which show low relative abundances at and near the three seeps (≤ 2%) and increase to up to 43 % (Actinobacteria) and 12 % (Acidobacteria) with increasing distance from the seep sites. Other relatively abundant bacterial phyla (≥ 5%) are Bacteriodetes, Chloroflexi, Gemmatimonadetes, Planctomycetes and Alphaproteobacteria in the Bissana seep soils of which only the Bacteriodetes showed some higher relative abundances at and close to seep 1 and seep 3.

With respect to bacterial OTUs falling in groups of known MOBs, the Bissana seep soils showed the presence, albeit in very low relative abundance (< 0.5%), of some common Type I (Gammaproteobacteria) and Type II methanotrophs (Alphaproteobacteria) (Fig. 6b). At seep 1 and seep 3 slightly increased levels of *Methylohalobiaceae* are present (0.11 and 0.05%, respectively). At 1 m distance from seep 1, there is slightly elevated abundance of representatives from the families *Methylomonaceae* (0.1%) and *Methylococcaceae* (0.1%) accompanied with low abundances of *Methylohalobiaceae* and *Methylobacteriaceae*. Some low read percentages of *Methylococcaceae* and *Methylobacteriaceae* are also visible until 14 m (≤ 0.05%) distance to seep 1.

4. Discussion

In this study we examined both the relative abundances of diverse BHPs as well as the relative abundances of bacterial reads at the two Sicilian gas seeps. A quantitative comparison of these two data sets is hampered by the fact that both datasets represent relative abundances, rather than absolute amounts, and suffer from different biases such as e.g. different ionization efficiencies for BHPs and a varying 16S rRNA gene copy number in bacteria. Furthermore, it is well known

that a considerable part of the bacteria species do not have the capability to synthesise BHPs (Rohmer et al., 1984; Kannenberg and Poralla, 1999; Pearson and Rusch, 2009; Rezanka et al., 2010; Sinninghe Damsté et al., 2017). This also applies to uncultivated bacteria which BHP composition has not been characterised at all. These considerations should be kept in mind when inferring sources of BHPs using 16S rRNA gene data and making quantitative comparisons.

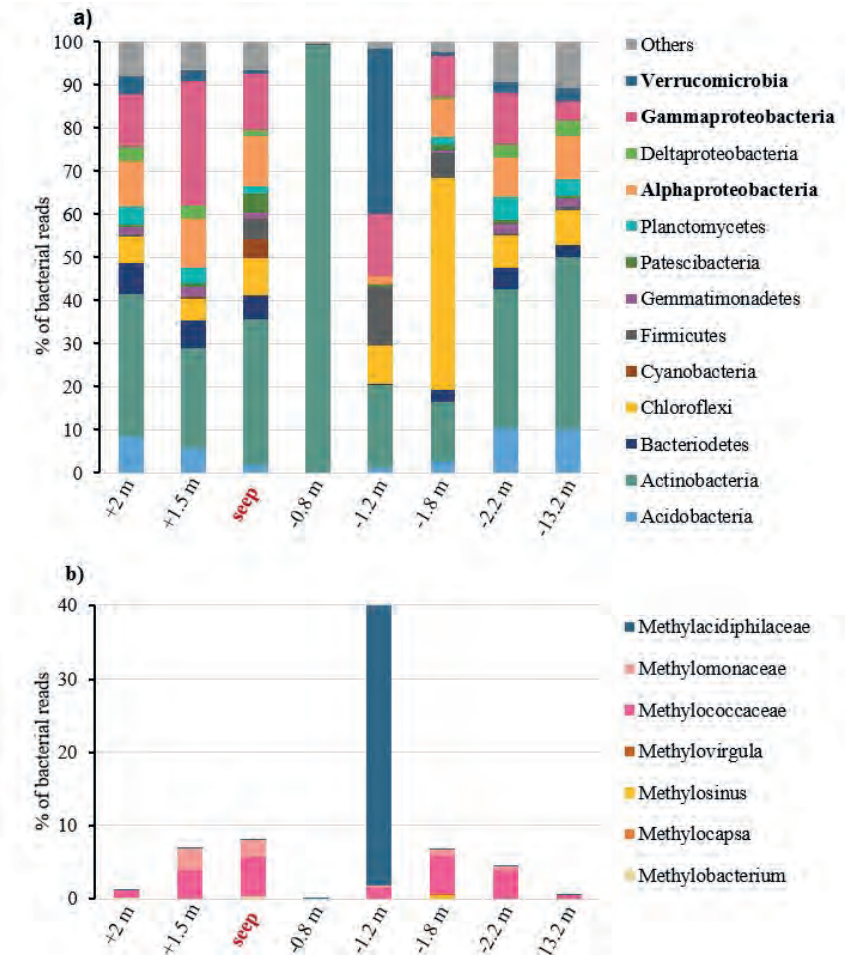


Figure 5. Relative abundance (as percentage of total bacterial reads) of the a) main bacterial phyla and b) known aerobic methanotrophic bacteria (in percentage of assigned bacterial reads) based on 16S rRNA gene amplicon sequencing in the Censo seep soils uphill (+) and downhill (–) from the main seep site.

4.1 AminoBHPs as biomarkers for aerobic methanotrophic bacteria

The Censo and Bissana seeps show both high relative abundances of typical aminoBHPs, namely aminopentol, aminotetrol, and the less specific aminotriol

(Figs. 2 and 3), which are thought to be characteristic for MOBs and thus aerobic oxidation of methane (e.g. Neunlist and Rohmer, 1985; Cvejic et al., 2000; Talbot et al., 2001; van Winden et al., 2012). Indeed, PCA analysis show that seep samples are generally characterized by elevated abundances of these aminoBHPs (Fig. 4a) compared to more general BHPs like BHT. These high abundances of aminoBHPs at CH_4-rich sites agree with previous studies which showed also high abundances of these BHPs where aerobic oxidation of methane was likely prevalent in peatlands, estuaries, lakes, or the water column (Coolen et al., 2008; Zhu et al., 2010; Berndmeyer et al., 2013; van Winden et al., 2012; Talbot et al., 2014, 2016). Isotopic analysis of degradation products of BHPs in these environments have shown they are relatively depleted in ^{13}C confirming their origin from MOBs (e.g. Jahnke et al., 1999; Coolen et al., 2008; Talbot et al., 2014). Furthermore, 16S rRNA gene analysis in lake sediments from the Antarctic Ace Lake further supported an origin of the three ^{13}C depleted aminoBHPs from Type I MOBs (Coolen et al., 2008).

At the Censo seep, aminotetrol is only present at the seep site while aminopentol is abundant at the seep and also around the seep sites (Fig. 2a). These results suggest that CH_4 flux might stimulate the MOB community not only at the seep but also in the soils further away from the Censo seep since only Type I methanotrophs are known to synthesize aminopentol (Neunlist and Rohmer, 1985; Cvejic et al., 2000; Talbot et al., 2001; van Winden et al., 2012). Indeed, a previous study found diffuse soil degassing being abundant within an area of 80 m^2 at the Censo seep site (Etiope et al., 2002). Furthermore, the presence of aminopentol at and around the Censo seep site agrees with the presence of 16S rRNA gene reads related to Type I methanotrophs from the *Methylomonaceae* and *Methylococcacea*e families in most Censo soils. At much longer distance (13.2 m away), 16S rRNA gene reads of these methanotrophs are present in low relative abundance, consistent with the low abundances of aminopentol compared to other BHPs (Fig. 2a and 5b). At the Bissana seep, aminopentol and aminotetrol are also abundant directly at the three seep sites and at 1 m from seep 1, while they are almost absent further away from the seeps (Fig. 3a). This also coincides with the presence of Type I MOB 16S rRNA gene sequences at these sites. Surprisingly, however, the relative abundances of reads of Type I MOBs at Bissana are much lower compared to Censo (~0.2% versus ~8%), suggesting that known MOBs do not comprise a large part of the bacterial community (Fig. 6b). This contrasting result suggests that possibly a large part of the aminoBHPs at the Bissana seeps are not derived from methanotrophic bacteria and may be sourced by other bacterial groups. Interestingly, high

abundances of Gammaproteobacteria are observed directly at the three seeps, i.e. they are making up over 50% of the bacterial reads (Fig. 6a), and Proteobacteria are known to be a significant source of BHPs in the environment (Pearson and Rusch, 2009; Kharbush et al., 2013). Thus, besides the known Type I MOBs, other groups within the Gammaproteobacteria might also be a source of the high amounts of aminopentol, and also other aminoBHPs, at the Bissana seeps.

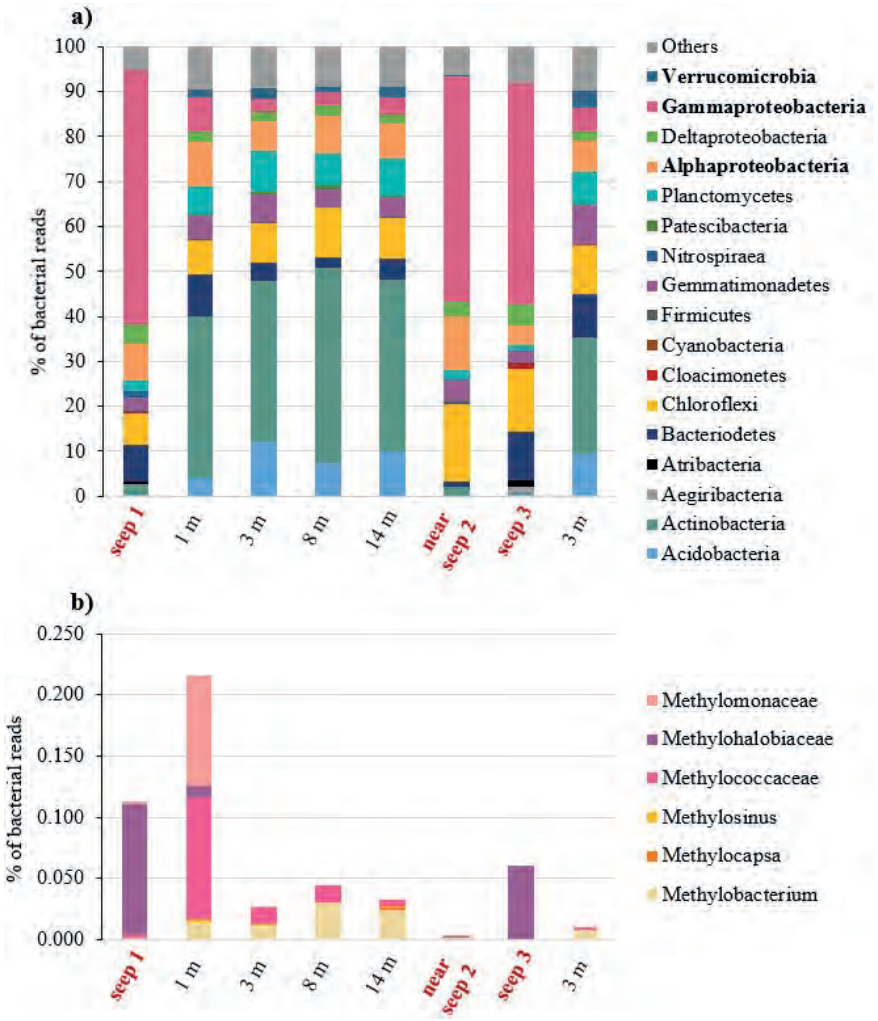


Figure 6. Relative abundance (as percentage of total bacterial reads) of the a) main bacterial phyla and b) known aerobic methanotrophic bacteria (in percentage of assigned bacterial reads) based on 16S rRNA gene amplicon sequencing in the different Bissana seeps soils at increasing distance from seep 1 and seep 3 as well as near seep 2.

Aminotetrol seems to behave similar to the aminopentol at both seep sites. However, at Censo it is only abundant directly at the seep (Figs. 2, 3, 5 and

6). Aminotetrol can also be synthesized by Type II MOBs (Talbot et al., 2001; van Winden et al., 2012) and sequences of Type II MOBs were detected though they make a small portion (<0.3%) of the bacterial reads (Figs. 5b and 6b). The high abundance of the less specific aminotriol at both seeps and its decrease with increasing distance from the seeps suggest that this BHP might derive from MOBs. However, PCA analysis shows that it plots somewhat differently from aminotetrol and aminopentol (Fig. 4a). Aminotriol is known to be produced by some Type I and Type II methanotrophs but it is also synthesized by various other bacterial species (Neunlist et al., 1985; Neunlist and Rohmer, 1985; Talbot and Farrimond, 2007), potentially explaining its somewhat different behavior compared to aminotetrol and aminopentol. Nevertheless, aminotriol still seems largely derived from MOBs at both seeps as it plots opposite to more general BHPs like BHT (Fig. 4a).

Finally, 3-methyl aminoBHPs are solely abundant at the three Bissana seeps (Fig. 3b) and, except for 3-methyl aminotriol, not detected or in low abundance at the Censo seep site. The presence of 3-methyl aminopentol is noteworthy as it has only been observed rarely in the environment, i.e. in a lake sediment (Talbot and Farrimond, 2007) and a geothermal silica sinter (Gibson et al., 2008), where aerobic oxidation of methane was also shown to occur. Similar to their non-methylated homologue, the 3-methyl aminopentol likely originates from Type I methanotrophs (Neunlist and Rohmer, 1985; Cvejic et al., 2000). It should be noted that the *hpn*R gene responsible for the 3-methylation was only found in a limited number of Type I methanotrophs but also in other bacteria such as Actinobacteria, Nitrospirae, and Acidobacteria (Welander and Summons, 2012). However, the presence of 3-methyl aminoBHPs predominantly at the Bissana seeps and the low relative abundances of the other bacterial groups known to carry the *hpn*R gene suggest that these compounds likely derive from MOBs or other bacterial involved in gas oxidation at the seep. Interestingly, the 3-methyl aminotriol clusters with aminotetrol and aminopentol, in contrast to its non-methylated homologue aminotriol (Fig. 4a), suggesting it has a more predominant origin from MOBs compared to the latter.

4.2 Exploring aminotriol isomer as new marker for Verrucomicrobia

Several Censo soils contain high relative abundances of aminotriol together with a new late-eluting aminotriol isomer, representing 40-60% of total BHPs in two Censo seep soils close to the seep (Fig. 2a). Interestingly, they plot differently in the PCA analysis from the other amino BHPs, suggesting potentially different bacterial sources. In an earlier study, van Winden et al. (2012) reported that the

verrucomicrobial methanotrophic species *Methylacidiphilum fumariolicum* strain SolV contained high abundances of aminotriol, and close examination of their LC-MS data reveals what appears to be a double peak for aminotriol, possibly suggesting an incomplete separation of aminotriol isomers. To further explore this, the BHP inventory of the verrucomicrobial methanotrophic species *M. cyclopophantes* 3B (van Teeseling et al., 2014) was analyzed, which shows high abundances of two aminotriol isomers, namely aminotriol (61%) and a late-eluting aminotriol isomer (34%), together making up about 95% of all BHPs in the culture (Fig. 7a and Table S3). Comparisons show that the two aminotriol isomers elute at the same retention time as those in the Censo seep and show similar proportions (Fig. 7b). This suggests a different origin of aminotriol, and especially aminotriol isomer, compared to the other aminoBHPs, and seems to be connected to the presence of methanotrophic Verrucomicrobia. This is further supported by the high relative abundances of Verrucomicrobia phyla in the 16S rRNA gene sequences in –1.2 m which are affiliated to *Ca.* 'Methylacidiphilum' (Fig. 5). These *Ca.* 'Methylacidiphilum' verrucomicrobia are the closest related methanotrophic bacterial genus (Fig. S1) and show about 63% sequence similarity with the investigated *M. cyclopophantes* 3B culture (van Teeseling et al., 2014; Schmitz et al., 2021). Overall, this suggests that high amounts of aminotriol accompanied by a later eluting aminotriol isomer might represent a new BHP-biomarker to detect verrucomicrobial methanotrophs in the environment.

4.3 Novel composite N-containing BHPs: potential biomarker to detect MOBs

Novel composite N-containing BHPs, which are connected to an ethenolamine functional group probably at position C-35 in the BHP side chain (Hopmans et al., 2021), were identified at both seep sites. At the Bissana seeps, all three ethenolamineBHPs (ethenolamine-BHT, ethenolamine-BHpentol and ethenolamine-BHhexol) reveal high relative abundances at the seeps and decrease with increasing distance from the seeps (Fig. 3c), whereas at the Censo seep the ethenolamineBHPs demonstrate high relative abundances at the seep and downhill from the seep (Fig. 2b). The distributions of the ethenolamineBHPs indicate a similar trend as the aminoBHPs (Figs. 2c versus 2a, 3a versus 3b), suggesting that they might derive from Type I MOBs and other MOBs that are present at both seep sites. Moreover, PCA revealed that the ethenolamineBHPs cluster closely with aminotetrol and aminopentol (Fig. 4a), further supporting that their source organism(s) are probably involved in the oxidation of methane

or from bacteria which occupy a similar environmental niche. Thus, they may represent new BHP based biomarkers for aerobic oxidation of methane. However, no source organism(s) are known yet as they have not been reported in any culture study to the best of our knowledge. For this, further research on pure MOB cultures is needed and our results suggest that probably Type I MOBs or groups of Gammaproteobacteria detected at Bissana might be good potential candidates to screen for the presence of ethenolamineBHPs.

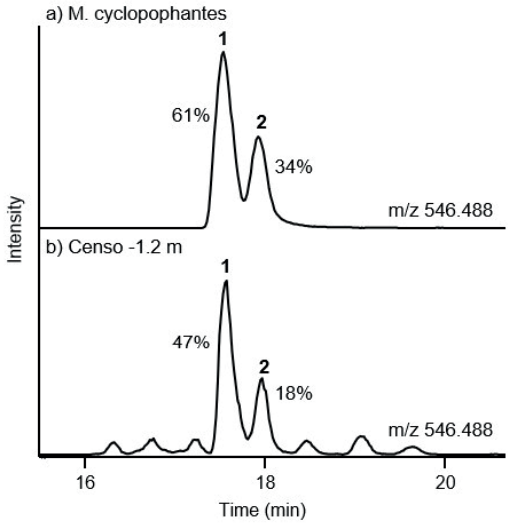


Figure 7. Partial mass chromatograms of aminotriol (1) and aminotriol isomer (2) and their relative abundances (in %) in a) *M. cyclopophantes* 3B (verrucomicrobial methanotroph) and b) Censo -1.2 m. The traces are labeled with the exact mass of *m/z* 546.488 $[M+H]^+$ for aminotriol used for searching.

Additionally, two composite aminoBHPs were found to be present in the Censo soils: $C_{16:0}$-N-acyl-aminotriol and $C_{11:0}$-N-acyl-aminotriol, of which only $C_{16:0}$-N-acyl-aminotriol could be identified also in the Bissana seep soils. At both sites, the $C_{16:0}$-N-acyl-aminotriol show high relative abundance at the seeps but also further away from the seeps (Figs. 2c and 3c). In the PCA, $C_{16:0}$-N-acyl-aminotriol clusters between 3-Me aminotriol and BHT and 3-Me BHT (Fig. 4a) suggesting a mixed origin. This is also supported by the finding of $C_{16:0}$ and $C_{16:1}$-N-acyl-aminotriol in *Nitrosomonas europaea*, an ammonia-oxidizing bacterium (Talbot et al., 2007a). The $C_{11:0}$-N-acyl-aminotriol reveals high relative abundance directly at the Censo seep and at –1.8 m (Fig. 2c) which is in accordance with higher relative abundance of Type I MOB *Methylococcaceae* species in these soils (Fig. 5b). Although, N-acyl-aminoBHPs were previously found in *N. europaea* and *R. vannieli* (Talbot et al., 2007a), no N-acyl-

aminoBHPs with such short fatty acid chains were previously reported. This suggests a potential production of this $C_{11:0}$-N-acyl-aminotriol by Type I MOB species. Since no source organism(s) of either N-acyl-aminotriols are known, further research is needed using Type I methanotroph cultures.

4.4 General and soil specific BHPs at both Sicilian gas seeps

The distribution of more ubiquitous BHPs (i.e. BHT) and the general soil-marker BHPs adenosylhopane and N1-methylinosylhopane (previously adenosylhopane type 1 and 2) was investigated at the two gas seeps (Figs. 2d, 3d and Tables S1, S2). In the Censo and Bissana seep soils, BHT show little variation in relative abundance over the seep transects and did not reveal a particular trend in respect to increasing distance from the seeps. BHT is one of the most ubiquitous BHPs found in terrestrial and marine environments and is synthesized by a wide variety of bacterial species (e.g. Cooke et al., 2008a; Talbot and Farrimond, 2007; Kusch et al., 2019). Therefore, BHT probably derives from numerous bacterial species present in both seep sites (Figs. 5 and 6) such as those from the phyla Actinobacteria, Acidobacteria, Planctomycetes or various Proteobacteria (e.g. Neunlist et al., 1985; Zundel and Rohmer, 1985; Sinninghe Damsté et al., 2004, 2017; Blumenberg et al., 2006). This interpretation is supported by the separate clustering of BHT from N-containing BHPs and soil-specific BHPs like adenosylhopane in the PCA plot (Fig. 4). In the Censo soils, BHT could also be produced by the highly abundant novel soil mycobacteria as some exploratory searches for the squalene-hopane-cyclase (shc) in closely related mycobacterial species (*M. simiae* complex) showed the potential to synthesize biohopanoids (Smit et al., 2021a).

Adenosylhopanes are precursors to all side-chain extended BHPs (Bradley et al., 2010), however, they have been reported to be abundant mainly in terrestrial environments and have thus been used as proxies for terrestrial input into marine systems (Cooke et al., 2008a, b; Zhu et al., 2011; Kusch et al., 2019). Adenosylhopane and N1-methylinosylhopane follow the opposite trend in their relative abundances with distance from the seeps compared to the N-containing BHPs, which is also demonstrated by the opposite behaviour of these two soil BHPs compared to the N-containing BHPs in the PCA plot (Figs. 2, 3 and 4). Likely, typical soil bacteria such as Acidobacteria and Actinobacteria present in the soils (Figs. 5 and 6) produce adenosylhopane and N1-methylinosylhopane (Alloisio et al., 2005; Sinninghe Damsté et al., 2017). Thus, MOBs and potential undescribed novel MOBs likely produce N-containing BHPs directly at and close to the seeps, whereas other general soil bacteria producing adenosylhopane

and N1-methylinosylhopane might dominate the BHP producing community further away from the seeps.

4.5 An index of N-containing BHPs to trace aerobic methane oxidation in the environment

The PCA analysis of BHP inventories over both seep transects show MOB-specific aminoBHPs (aminotriol, aminotetrol and aminopentol) plotting opposite of adenosylhopane and N1-methylinosylhopane derived from general soil bacteria. These results enable us to produce an index to trace the presence of MOBs, and thus aerobic oxidation of methane, in terrestrial environments. For this, we suggest the three aminoBHPs and their 3-methylated homologues as BHP-markers for aerobic oxidation of methane as they have well constrained MOB sources in contrast to the novel BHPs like the ethenolamineBHPs and N-acyl-aminotriols. In contrast, the soil-marker BHPs, adenosylhopane and N1-methylinosylhopane are fully unrelated to aerobic oxidation of methane and can be used as indicators for the general abundance of soil bacteria. This leads to the following definition of the so-called AminoBHP-index:

AminoBHP-index =

(**a**+ **a1**+ **b**+ **b1**+ **c**+ **c1**) / ((**j** + **k**) + (**a**+ **a1**+ **b**+ **b1**+ **c**+ **c1**)) **[1]**

with letters referring to the BHPs shown in Figure 1. This index theoretically ranges from 0 to 1, with 1 meaning BHPs only comprising methanotrophic biomarkers, and thus intense aerobic oxidation of methane, while 0 indicates the absence of the specific methanotrophic biomarkers. Applying this AminoBHP-index to the Sicilian seeps sites reveals values ranging from 0.44 to 0.84 at and very close to the Censo and Bissana seeps (Fig. 8 and Table S5). Between 1 m to 2 m distance from the seeps, most data points cluster around 0.2 with a few much higher indicating still a high input of MOB derived BHPs, in agreement with the high relative abundance of reads from MOB at the Censo site. Even further away, values are well below 0.2 suggests that there is no significant input of BHPs from methanotrophs. Therefore, we suggest 0.2 as a threshold, above which indicates the presence of MOBs, and thus aerobic oxidation of methane. It should be noted that aminotriol, included in this index, can also be produced by bacteria other than methanotrophs (e.g. Neunlist et al., 1985; Talbot and Farrimond, 2007; Blumenberg et al., 2012) although here they are likely partially sourced by methanotrophic Verrucomicrobia (see above). Nevertheless, it is

suggested to apply the index with caution when no 3-methylated homologue and no aminoBHPs other than aminotriols are abundant in the soils. In our study, the consistent decrease of the AminoBHP-index values with increasing distance from the main seep sites clearly demonstrates the potential to apply the index to modern and past environments where aerobic oxidation of methane is thought to be present. Indeed, some BHPs (BHT) have been dated back to 50.4–49.7 Ma ago in samples from marine Palaeogene cores from Tanzania (van Dongen et al., 2006) and back to ~125 ka in Mediterranean sapropels (Rush et al., 2019). Other studies showed that the aminoBHPs and especially aminopentol reflected periods of aerobic oxidation of methane in marine sediment records (1.2 Ma) (Talbot et al., 2014) and peat deposits which dated back until the PETM (56 Ma) implying a high preservation of aminoBHPs (van Winden et al., 2012; Talbot et al., 2016). Future work should investigate the long-term preservation potential of aminoBHPs and soil-specific BHPs as well as the impact of diagenesis on this index before it can be applied confidently to paleorecords.

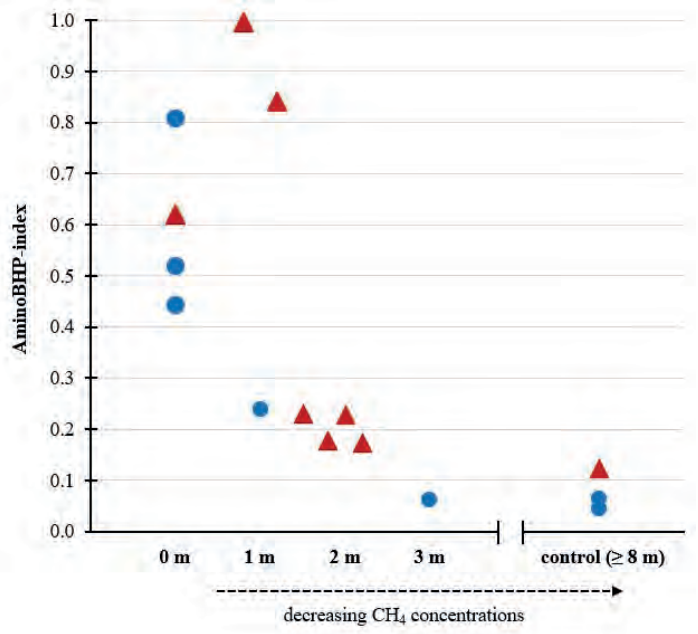


Figure 8. AminoBHP-index of soils from the Bissana and Censo seeps at increasing distance to the active seeps. The standard deviation for the AminoBHP-index from three triplicate samples ranges from 0.01 to 0.3. Key: Bissana seep = blue filled circles; Censo seep = red filled triangles.

Conclusions

The soils from the Censo and Bissana seeps show high relative abundances of aminoBHPs (aminotriol, aminotetrol and aminopentol) and novel

ethenolamineBHPs (ethenolamine-BHT, -pentol and -hexol) that generally decrease with increasing distance from the seeps. The Bissana seep contained 3-methylated aminoBHPs at and near the seeps while the Censo seep did not contain high abundances of 3-methylated BHPs but did show high abundances of novel $C_{11:0}$-N-acyl-aminotriol around the main seep. 16S rRNA gene analysis and clustering in PCA suggest that aminopentol and aminotetrol as well as the novel ethenolamineBHPs derive from Type I methanotrophs and possibly other methanotrophic bacteria. The Censo seep contains high abundances of aminotriol and a novel late eluting aminotriol isomer, based on its presence in the novel isolated verrucomicrobial strain *M. cyclopophantes* 3B and the abundance of specific 16S rRNA gene sequences in the Censo soils, which are attributed to the verrucomicrobia '*Candidatus* Methylacidiphilum'. The presence of aminotriol together with the late-eluting aminotriol isomer could be a new indicator for verrucomicrobial methanotrophs in the environment. The opposite behaviour of MOB-specific aminoBHPs compared to typical soil-marker BHPs (i.e. adenosylhopane) in PCA and the trends in their relative abundances, clearly suggest changes in the BHP-producing bacterial community away from the seeps related to the availability of methane. A new AminoBHP-index was developed based on aminoBHPs and their 3-metlylated homologues (as indicators for MOBs) and adenosylhopane and N1-methylinosylhopane (as indicators of typical soil bacteria). The application of this AminoBHP-index shows high values (≥ 0.4) close to the Bissana and Censo seeps and decreases with increasing distance from the seeps until it falls below 0.2 at more than 8 m distance from the seeps. The novel AminoBHP-index may offer a new biomarker tool to trace aerobic oxidation of methane in present and past terrestrial environments.

Author contributions

NTS, DR and SS planned the research. NTS and FG collected the samples. HJMoC and CH cultivated the verrucomicrobial *Methylacidimicrobium cyclopophantes* 3B and provided biomass. NTS performed lipid extraction and BHP data analysis, while DBC performed 16S rRNA analysis and LV analysed 16S rRNA gene sequencing data. NTS, ECH, DR, SS, JSSD and LV interpreted the data. NTS wrote the paper with help of DR, SS and JSSD. All authors approved the final text.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank Denise van der Slikke-Dorhout, Monique Verweij, Michel Koenen, Sanne Vreugdenhil and Maartje Brouwer for technical assistance. We also thank Caitlyn R. Witkowski for help with sample collection. NTS, SS and JSSD were supported by the Netherlands Earth System Science Centre (NESSC) through grant no. 024.002.001. LV, DR and JSSD were supported by the Soehngen Institute for Anaerobic Microbiology (SIAM) through Gravitation grant 024.002.002 from the Dutch Ministry for Education, Culture and Science. CH and HJMoC were supported by the European Research Council (ERC Advanced Grant project VOLCANO 669371). ECH received funding from the NWO middelgroot grant no. 834.13.004.

Supplemental Material (available upon request)

Table S1. Relative and total abundances of integrated BHPs with standard deviations of averaged triplicate samples for the Censo seep.
Table S2. Relative and total abundances of integrated BHPs with standard deviations of averaged triplicate samples for the Bissana seeps.

Table S3. Relative abundances of integrated BHPs for the verrucomicrobial methanotroph *M. cyclopophantes* 3B.

Table S4. Relative abundances of main bacterial phyla and known MOB species reported as a percentage of total bacterial 16S rRNA gene reads. Triplicate samples of each distance were averaged, and standard deviations noted.

Table S5. AminoBHP-index values for the Censo and Bissana seeps.

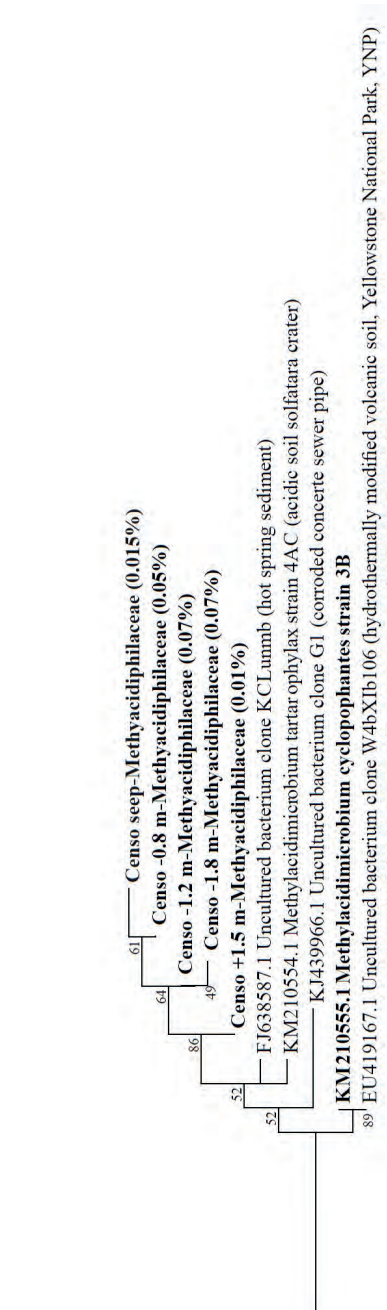


Fig S1. Maximum likelihood (ML) phylogenetic tree of the *Methylacidiphilaceae* 16S rRNA gene sequences (i.e., 292 bp; in bold, percentage between parentheses indicates percentage of the total reads attributed to the mentioned OTUs at the Censo seep 0 m) generated by amplicon sequencing and representative for the top five OTUs present in the soils from the Censo seep everlasting fire. Other *Candidatus* Methylacidiphilum species, including *Methylacidimicrobium cyclopophantes* 3B discussed in the text, are included in the tree. Bar indicates 0.01 nucleotide substitutions per site. Number in branches indicates percentage of support of a bootstrap analysis of 1,000 replicates.

Chapter 7

Novel hydrocarbon-utilizing soil mycobacteria synthesize unique mycocerosic acids at a Sicilian everlasting fire

Nadine T. Smit, Laura Villanueva, Darci Rush, Fausto Grassa, Caitlyn R. Witkowski, Mira Holzheimer, Adriaan J. Minnaard, Jaap S. Sinninghe Damsté, Stefan Schouten

Published in Biogeosciences

Fuoco di Censo everlasting fire in Sicily, Italy

Abstract

Soil bacteria rank among the most diverse groups of organisms on Earth and actively impact global processes of carbon cycling, especially in the emission of greenhouse gases like methane, CO_2 and higher gaseous hydrocarbons. An abundant group of soil bacteria are the mycobacteria, which colonize various terrestrial, marine and anthropogenic environments due to their impermeable cell envelope that contains remarkable lipids. These bacteria have been found to be highly abundant at petroleum and gas seep areas, where they might utilize the released hydrocarbons. However, the function and the lipid biomarker inventory of these soil mycobacteria are poorly studied. Here, soils from the Fuoco di Censo seep, an everlasting fire (gas seep) in Sicily, Italy, were investigated for the presence of mycobacteria via 16S rRNA gene sequencing and fatty acid profiling. The soils contained high relative abundances (up to 34 % of reads assigned) of mycobacteria, phylogenetically close to the *Mycobacterium simiae* complex and more distant to the well-studied *M. tuberculosis* and hydrocarbon-utilizing *M. paraffinicum*. The soils showed decreasing abundances of mycocerosic acids (MAs), fatty acids unique for mycobacteria, with increasing distance from the seep. The major MAs at this seep were tentatively identified as 2,4,6,8-tetramethyl tetracosanoic acid and 2,4,6,8,10-pentamethyl hexacosanoic acid. Unusual MAs with mid-chain methyl branches at positions C-12 and C-16 (i.e. 2,12-dimethyl eicosanoic acid and 2,4,6,8,16-pentamethyl tetracosanoic acid) were also present. The molecular structures of the Fuoco di Censo MAs are different from those of the well-studied mycobacteria like *M. tuberculosis* or *M. bovis* and have relatively $\delta^{13}C$-depleted values (-38 to -48 ‰), suggesting a direct or indirect utilization of the released seep gases like methane or ethane. The structurally unique MAs in combination with their depleted $\delta^{13}C$ values identified at the Fuoco di Censo seep offer a new tool to study the role of soil mycobacteria as hydrocarbon gas consumers in the carbon cycle.

1. Introduction

Soils harbor the largest diversity of microorganisms on our planet and have a large influence on the Earth's ecosystem as they actively impact nutrient and carbon cycling, plant production and the emissions of greenhouse gases (Tiedje et al., 1999; Bardgett and van der Putten, 2014; Delgado-Baquerizo et al., 2018). Soil bacteria rank among the most diverse and abundant groups of organisms on Earth. However, numerous studies suggest that most of their function and diversity in our ecosystems is still undescribed (Tiedje et al., 1999; Bardgett and van der Putten, 2014). The assessment of soil bacterial diversity has mainly

relied on 16S ribosomal RNA (rRNA) gene sequencing and has indicated that the most abundant bacterial phylotypes in global soils include Alphaproteobacteria, Gammaproteobacteria, Betaproteobacteria, Actinobacteria, Acidobacteria and Planctomycetes (Fierer et al., 2012; Delgado-Baquerizo et al., 2018). Besides the use of DNA-based techniques, lipid biomarkers offer an additional tool to investigate soil bacterial communities, such as branched glycerol dialky glycerol tetraether (brGDGTs) believed to derive from soil acidobacteria (Weijers et al., 2009; Peterse et al., 2010; Sinninghe Damsté et al., 2018) or lipids derived from methanotrophic bacteria like certain fatty acids (Bull et al., 2000; Bodelier et al., 2009), specific bacteriohopanepolyols (van Winden et al., 2012; Talbot et al., 2016) or ^{13}C-depleted hopanoids (Inglis et al., 2019; van Winden et al., 2020).

Mycobacteria of the genus *Mycobacterium* belonging to the phylum Actinobacteria form an abundant microbial group in global soils (Falkinham, 2015; Walsh et al., 2019). Some members of the genus *Mycobacterium* are obligate pathogens (e.g. *Mycobacterium tuberculosis* and *Mycobacterium leprae*) and are the cause of more than 1.5 million annual human deaths worldwide through the diseases tuberculosis and leprosy (World Health Organization, 2019). Thus, they have consequently been more frequently studied than opportunistic pathogenic and non-pathogenic environmental mycobacteria. Interestingly, early studies from the 1950s reported high abundances of non-pathogenic hydrocarbon-consuming mycobacteria (*M. paraffinicum*) in areas of oil and gas production, gas seeps, and in common garden soils (Davis et al., 1956; Dworkin and Foster, 1958; Davis et al., 1959). Cultivation and genomic studies show that mycobacteria can oxidize a range of greenhouse gases (ethane, propane, alkenes, carbon monoxide or hydrogen) and can degrade toxic polycyclic aromatic hydrocarbons (Miller et al., 2007; Hennessee et al., 2009; Coleman et al., 2012; Martin et al., 2014). Mycobacteria are able to colonize a wide variety of habitats from soils to aquatic and human-engineered environments (Brennan and Nikaido, 1995; Falkinham, 2009). Their impermeable cell envelope may play an important role in their ecological dominance. It consists of a peptidoglycan polymer that is surrounded by a thick hydrophobic lipid-rich outer membrane. This impermeability favors the formation of biofilms, thus enabling mycobacteria to be often the first colonizers at environmental interfaces like air-water or surface-water (Brennan and Nikaido, 1995). Additionally, the impermeable cell membrane allows a resistance to acidic conditions, anoxic survival and the possibility to metabolize recalcitrant carbon compounds. Mycobacteria feature an unusual lipid inventory

such as extremely long fatty acids with chains up to 90 carbon atoms long and with numerous methyl groups, hydroxylations and/or methoxylations produced by two fatty acid biosynthesis systems, i.e. FAS type I (eukaryotic type) and FAS type II (prokaryotic type) (Minnikin et al., 1985; Minnikin et al., 1993a; Donoghue et al., 2017; Daffé et al., 2019). Furthermore, mycobacteria are known to synthesize characteristic fatty acids like tuberculostearic acid (10-Me $C_{18:0}$) and the multi methyl-branched mycocerosic acids (MAs) that can contain 3 to 5 methyl branches at regularly spaced intervals such as at positions C-2, C-4, C-6 and C-8 (Minnikin et al., 1985; Minnikin et al., 1993a; Minnikin et al., 2002; Redman et al., 2009). MAs are synthesized by the mycocerosic acid synthase (encoded by the *mas* gene) through the FAS type pathway I using a methyl malonyl CoA instead of a malonyl CoA generating the unique methyl branching pattern of MAs (Brennan, 2003; Gago et al., 2011). These unusual fatty acids are bound to complex glycolipids like phthiocerol dimycocerosates (PDIMs), diacyl trehalose (DATs) or phenolic glycolipids (Minnikin et al., 2002; Jackson et al., 2007). However, in contrast to the pathogenic and opportunistic pathogenic mycobacteria, the lipid biomarker inventory of non-pathogenic mycobacteria in soils and other environments remains poorly described.

In this study, we investigated soils near a continuous gas seep named "Fuoco di Censo" ("Everlasting Fire") in Sicily, Italy to explore the presence of non-pathogenic, potentially hydrocarbon-utilizing, mycobacterial species using 16S rRNA gene amplicon sequencing and fatty acid profiling. It resulted in the identification of potential biomarkers for the presence of mycobacteria in terrestrial environments and hydrocarbon seeps. Furthermore, their stable carbon isotopic composition provided hints for their role in the carbon cycle in this gas seepage environment.

2. Material and Methods

2.1 Study area

The Fuoco di Censo seep (37°37'30.1''N, 13°23'15.0''E), in the following referred to as the Censo seep, is located at 800 m above sea level in the mountains of Southwestern Sicily, Italy (Etiope et al., 2002; Grassa et al., 2004). The area is part of the Alpine orogenic belt in the Mediterranean and located along the boundary of the African and European plates (Basilone, 2012). The Censo seep belongs to the Bivona area, which is characterized by a complex geological setting. The seep is located in an area with sandy clays, marls and evaporites from the Tortonian-Messinian that are covered by a thrusting limestone of Carnian-Rhetian age (Trincianti et al., 2015). The Censo seep is a typical

example of a natural 'Everlasting Fire', which is characterized by the absence of water and the temporal production of flames, which can be several meters high, by a continuous gas flux (Etiope et al., 2002). The Censo seep gas consists mainly of CH_4 (76-86 %) and N_2 (10-17%) as well as some other minor gases like CO_2, O_2, ethane, propane, He and H_2 (Etiope et al., 2002; Grassa et al., 2004). A diffuse soil degassing is detectable within an area of 80 m^2 with an average CH_4 flux of $7*10^6$ mg m^{-2} d^{-1} and a total CH_4 emission of $6.2*10^3$ kg yr^{-1} (Etiope et al., 2002; Etiope et al., 2007). The CH_4 is suggested to be generated by the thermal alteration of organic matter and is characterized by a stable carbon isotopic composition of $\delta^{13}C$ = -35 ‰ and δ^2H = -146 ‰ (Grassa et al., 2004). This thermogenic CH_4 possibly derives from mature marine source rocks (kerogen type II) with a thermal maturity beyond the oil window, resulting in a dry gas with $C_1/(C_2+C_3)$ ratios greater than 100 (Grassa et al., 2004).

2.2 Sample collection

Soil samples of the Censo seep were recovered during a field campaign in October 2017. The soil was collected from a horizon 5 to 10 cm below the surface and at three distances from the seep, i.e. 0 m (seep site), 1.8 m, and a control at 13.2 m distance from the main vent. The *in-situ* temperature of the soils at the time of collection was ca. 18 ºC. The soils were directly transferred into a clean geochemical sampling bag and stored frozen at -20 ºC until freeze drying and extraction.

2.3 Extraction and saponification

Freeze-dried Censo soils were extracted with a modified Bligh and Dyer extraction for various compound classes (Schouten et al., 2008; Bale et al., 2013). Soil samples (ca. 12 g) were ultrasonically extracted (10 min) with a solvent mixture containing methanol (MeOH), dichloromethane (DCM) and phosphate buffer (2: 1: 0.8, v: v: v). After centrifugation, the solvent was collected, combined and the residues re-extracted twice. A biphasic separation was achieved by adding additional DCM and phosphate buffer to a ratio of MeOH, DCM and phosphate buffer (1: 1: 0.9, v: v: v). The aqueous layer was washed two more times with DCM and the combined organic layers dried over a Na_2SO_4 column followed by drying under N_2.

Saponification (base hydrolysis) was conducted on aliquots (1-7 mg) of the Bligh Dyer extracts (BDEs) to release fatty acids from structurally complex intact polar lipids by the addition of 2 ml 1N KOH in MeOH solution and refluxing for 1 h at 130 °C. After cooling, the pH was adjusted to 5 by using a 2N HCL in MeOH

solution, separated with 2 ml bidistilled water and 2 ml DCM, and the organic bottom layer was collected. The aqueous layer was washed two more times with DCM and the combined organic layers dried over a Na_2SO_4 column followed by drying under N_2.

2.4 Derivatization of fatty acids

2.4.1 Preparation of fatty acid methyl esters using BF_3

Aliquots of the saponified Censo seep BDEs and aliquots of a mycocerosic acid standard (2,4,6-trimethyl-tetracosanoic acid; C_{27} MA standard) synthesized by hydrogenation with palladium and charcoal from mycolipenic acid (Holzheimer et al., 2020), were esterified with 0.5 ml of a boron trifluoride-methanol solution (BF_3 solution) for 10 min at 60 °C. After cooling, 0.5 ml bidistilled water and 0.5 ml DCM were added and shaken, and the DCM bottom layer pipetted off. The water layer was extracted twice with DCM and the combined DCM layers were dried over an $MgSO_4$ column. The soil extracts were eluted over a small silica gel column with ethyl acetate as an eluent to remove polar compounds. Extracts were subsequently separated using a small column packed with activated aluminum oxide into two fractions. The first fraction (fatty acid methyl ester fraction) was eluted with 4 column volumes of DCM followed by a second fraction (polar fraction) eluted with 3 column volumes of DCM/MeOH (1:1). The fatty acid methyl ester fractions were dried under a continuous flow of N_2 and analyzed using gas chromatography-mass spectrometry (GC-MS) and GC-isotope ratio mass spectrometry (IRMS).

2.4.2 Preparation of fatty acid "picolinyl esters" derivatives using 3-pyridylcarbinol

Aliquots of saponified Censo seep BDEs, as well as aliquots of the C_{27} MA standard, were derivatized into picolinyl esters. This technique enhances the abundance of diagnostic fragment ions in the mass spectrum, such as those of methyl branching points in fatty acids, enabling an improved structural identification (Christie, 1998; Harvey, 1998). Different 'picolinyl' derivatization protocols were tested on the C_{27} MA standard and the highest yields were achieved by the procedure in Harvey, 1998. In this procedure, 0.5 ml of thionyl chloride was added using a 1 ml disposable syringe to 1 mg aliquot of the dried saponified Censo seep BDEs in a pressure vial and left for ca. 2 min at room temperature. The vials were then dried by a continuous flow of N_2. 0.5 ml of a 1% 3-pyridylcarbinol in acetonitrile solution was added in the reaction vials and left at room temperature for 2 min. The volumes of reagents in this protocol were

reduced (0.1 ml) for 0.1 mg of the MA standard. The 'picolinyl' esters were transferred with acetonitrile to 2 ml analysis vials and the concentration was adjusted to 1 mg/ml with acetonitrile. The 'picolinyl' esters were analyzed using GC-MS with acetonitrile as injection solvent.

2.4.3 Preparation of fatty acid methyl sulfide esters using dimethyl disulfide (DMDS)

To determine the position of the double bonds in unsaturated fatty acids, dimethyl disulfide (DMDS) derivatization was used (Francis, 1981; Nichols et al., 1986). For this, 100 µl of hexane, 100 µl of DMDS solution (Merck ≥ 99%) and 20 µl of I_2/ether were added to the dry aliquot and heated overnight at 40 ºC. The mixture was left to room temperature and 400 µl of hexane and 200 µl of a 5% aqueous solution of $Na_2S_2O_3$ (for iodine deactivation) were added and mixed. The upper hexane layer was removed, and the aqueous layer washed twice with hexane. The three hexane layers were combined and dried over a Na_2SO_4 column before GC-MS analysis with hexane as injection solvent.

2.5 Instrumental analysis

2.5.1 Gas chromatography-mass spectrometry (GC-MS)

GC-MS was performed using an Agilent Technologies GC-MS Triple Quad 7000C in full scan mode. A CP-Sil5 CB column (25 m x 0.32 mm with a film of 0.12 µm, Agilent Technologies) was used for the chromatography with He as carrier gas (constant flow 2 ml min^{-1}). The samples (1 µl) were injected on column at 70 °C, the temperature was increased at 20 °C min^{-1} to 130 °C, raised further by 4 °C min^{-1} to 320 °C, at which it was held for 20 min. The mass spectrometer was operated over a mass range of *m/z* 50 to 850, the gain was set on 3, with a scan time of 700 ms.

2.5.2 Gas chromatography-isotope ratio mass spectrometry (GC-IRMS)

GC-IRMS was carried out with a Thermo Scientific Trace 1310 with a GC-Isolink II, a ConFlo IV and a Delta Advantage IRMS. The gas chromatography was performed on a CP-Sil5 CB column (25 m x 0.32 mm with a film thickness of 0.12 µm, Agilent) with He as carrier gas (constant flow 2 ml min^{-1}). The BF_3 methylated samples (dissolved in ethyl acetate) were on-column injected at 70 ºC and subsequently, the oven was programmed to 130 ºC at 20 ºC min^{-1}, and then at 4 ºC min^{-1} to 320 ºC, which was held for 10 min. Stable carbon isotope ratios are reported in delta-notation against Vienna Pee Dee Belemnite (VPDB)

^{13}C standard. Values were determined by two analysis and results averaged to a mean value.

2.6 DNA extraction, 16S rRNA gene amplification, analysis and phylogeny

DNA was extracted from sediments using the PowerMax soil DNA isolation kit (Qiagen). DNA extracts were stored at -80 °C until further analysis. The 16S rRNA gene amplicon sequencing and analysis was performed with the general 16S rRNA archaeal and bacterial primer pair 515F and 806RB targeting the V4 region (Caporaso et al., 2012; Besseling et al., 2018). Polymerase chain reaction (PCR) products were gel purified using the QIAquick Gel-Purification kit (Qiagen), pooled and diluted. Sequencing was performed at the Utrecht Sequencing Facility (Utrecht, the Netherlands) using an Illumina MiSeq sequencing platform. The 16S rRNA gene amplicon sequences were analyzed by an in-house pipeline (Abdala Asbun et al., 2020) that includes quality assessment by FastQC (Andrews, 2010), assembly of the paired-end reads with Pear (Zhang et al., 2013), and assignment of taxonomy (including picking representative set of sequences with 'longest' method) with blast by using the ARB Silva 128 database (https://www.arb-silva.de/, last access: 5 August 2018, release 128). Representative operational taxonomic unit (OTU) sequences (assigned with OTU picking method based on 97 % nucleotide similarity with Uclust) (Edgar, 2010), attributed to the family Mycobacteriaceae were aligned by using Muscle (Edgar, 2004) implemented in MEGA6, and then used to construct a phylogeny together with 16S rRNA gene sequences of characterized *Mycobacterium* species and closely related uncultured Mycobacteriaceae 16S rRNA gene sequences. The phylogenetic tree was inferred using the Maximum Likelihood method based on the General Time Reversible model (Nei and Kumar, 2000). The analysis involved 32 nucleotide sequences with 294 base pairs positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

3. Results and discussion

3.1 Microbial diversity in the Censo seep soils

Soils were sampled at the Censo seep and with increasing distance from the seep (Table 1). To investigate the microbial diversity, 16S rRNA gene libraries were generated from extracted DNA using 16S rRNA gene amplicon sequencing. This analysis showed a high relative abundance of 16S rRNA gene reads attributed to Mycobacteriaceae ranging from 0.7 to 34.1 % of assigned bacterial plus archaeal reads in the soils with relative abundances increasing with decreasing distance

to the seep (Table 1). Sequences assigned to known methanotrophs are Gammaproteobacteria (Methylococcales), Alphaproteobacteria (Methylocystaceae and Methylobacteriaceae) and Verrucomicrobia (‘*Candidatus* Methylacidiphilum’) but only accounted for 0.2 to 5.1 % of the total number of reads assigned (Table 1). Phylogenetic analysis indicated that there are two sequences representative for operational taxonomic units (OTUs) attributed to mycobacteria (i.e. sequences Censo seep 1 and Censo seep 2) present in the soils (Fig. 1). Both OTUs are phylogenetically most closely related to sequences of the *Mycobacterium simiae* complex (Tortoli, 2014) (Fig. 2; >98 % identical considering the 294 bp sequence fragment analyzed), which include *M. simiae*, *M. europaeum*, *M. kubicae* and *M. heidelbergense* (Hamieh et al., 2018). Previously described cultivated mycobacteria of the *M. simiae* complex are slow-growing mycobacterium species isolated from environmental niches but also associated to infections in humans as opportunistic pathogens (Lévy-Frébault et al., 1987; Heap, 1989; Bouam et al., 2018). The Censo seep sequences are more distantly related (94-95 % identical) to frequently studied pathogenic mycobacteria (such as *M. tuberculosis* and *M. leprae*) and other environmental mycobacteria like hydrocarbon-utilizers (e.g. *M. paraffinicum* and *M. vanbaalenii*) (Fig. 1). To the best of our knowledge the hydrocarbon-utilizing bacteria have not been isolated from humans or animals (e.g. *M. vanbaalenii*) and are mostly able to degrade aromatic hydrocarbons (Kweon et al., 2015). Our data reveal abundances of up to 34 % of uncultured mycobacteria (Censo 0 m) in the soils around the Censo seep. This is in line with previous reports of the occurrence of mycobacteria near petroleum seeps and gas fields (Davis et al., 1956; Davis et al., 1959).

3.2 Fatty acid composition of Censo seep soils

Analysis of the fatty acid fractions of the Censo seep soils reveal a distinct pattern that changes with increasing distance from the main seep (Fig. 2). Common fatty acids such as $C_{16:0}$, $C_{16:1\omega6}$, $C_{16:1\omega7}$, C_{18}, $C_{18:1\omega9}$ and $C_{18:1\omega7}$ as well as the longer chain C_{22} and C_{24} fatty acids occur in all three soils. C_{16} and C_{18} fatty acids are abundant lipids in soils and are synthesized by diverse bacteria and fungi, whereas the longer chain (C_{22}-C_{24}) fatty acids originate commonly from higher plants (Řezanka and Sigler, 2009; Frostegård et al., 2011). These fatty acids could also derive from mycobacteria which can produce fatty acids (C_{14} to C_{26}) with high amounts of C_{16} and C_{18} fatty acids and their unsaturated homologues (Chou et al., 1996; Torkko et al., 2003).

Table 1 Distribution of the main microbial groups (in percentage of assigned reads) based on 16S rRNA gene amplicon sequencing at three distances from the main gas seep in the Censo soils. The bold typeface annotates the relative abundances of mycobacteria in the Censo soils.

	0 m	**1.8 m**	**13.2 m**
Archaea; Euryarchaeota	53.5	0.0	0.0
Archaea; Thaumarchaeota	0.0	0.0	2.5
Bacteria; Acidobacteria	0.2	3.0	10.5
Bacteria; Actinobacteria	36.8	23.6	46.0
Acidimicrobiales, other	0.0	0.3	2.3
Corynebacteriales, Mycobacteriaceae, Mycobacterium	**34.1**	**8.5**	**0.7**
Frankiales, Geodermatophilaceae, Geodermatophilus	0.0	5.1	0.1
Micrococcales, Microbacteriaceae, Humibacter	0.1	3.0	0.0
Micromonosporales, Micromonosporaceae, Micromonospora	1.0	0.0	0.1
Pseudonocardiales, Pseudonocardiaceae, Pseudonocardia	0.0	0.0	1.6
Rubrobacterales, Rubrobacteriaceae, Rubrobacter	0.0	0.0	6.2
Gaiellales, Gaiellaceae, Gaiella	0.0	0.0	5.1
Solirubrobacterales, 288-2, uncultured bacterium	0.0	0.0	2.2
Solirubrobacterales, Elev-16S-1332, uncultured bacterium	0.0	0.6	4.3
Solirubrobacterales, Solirubrobacteraceae, Solirubrobacter	0.0	0.0	1.8
others	1.6	6.0	21.5
Bacteria; Armatimonadetes	0.0	0.0	0.2
Bacteria; Bacteroidetes	0.3	1.2	2.1
Bacteria; Chloroflexi	0.2	30.5	10.5
Anaerolineales, Anaerolineae, Anaerolineales	0.0	7.9	0.2
Ktedonobacteria	0.1	17.8	0.1
Bacteria; Firmicutes	5.2	15.7	0.6
Bacteria; Gemmatimonadetes	0.0	0.0	3.1

Bacteria; Latescibacteria	0.0	0.0	0.2
Bacteria; Nitrospirae	0.0	0.0	0.2
Bacteria; Planctomycetes	0.2	1.5	3.7
Bacteria; Proteobacteria	2.9	21.1	17.2
Alphaproteobacteria, Rhizobiales, Methylobacteriaceae	0.0	0.3	0.8
Alphaproteobacteria, Rhizobiales, Methylocystaceae	0.0	4.2	0.0
Gammaproteobacteria, Methylococcales	0.2	5.1	0.2
Bacteria; Saccharibacteria	0.1	2.1	0.2
Bacteria; Tectomicrobia	0.0	0.0	0.7
Bacteria; Verrucomicrobia	0.3	0.3	1.7
Verrucomicrobia Incertae Sedis, Candidatus Methylacidiphilum	0.2	0.0	0.0
Number of reads assigned	**140,206**	**63,916**	**259,714**

Besides mycobacteria which are abundant in the soils close to the main seepage (Table 1), the C_{16} fatty acids may also originate from Type I methanotrophs (Gammaproteobacteria), whereas C_{18} fatty acids could derive from type II methanotrophs (Alphaproteobacteria), present in these Censo seep soils (Fig. 2 and Table 1) (Bull et al., 2000; Bowman et al., 1993; Bodelier et al., 2009). However, the relative abundances of 16S rRNA gene reads of this Type I and II methanotrophs are only minor in the Censo soils (Table 1).

The Censo seep soils also feature C_{31}-C_{33} 17β,21β(H)-homohopanoic acids, the most abundant of which is the C_{32} 17β,21β(H)-hopanoic acid (bishomohopanoic acid) (Fig. 2). Hopanoic acids are common components in terrestrial environments (Ourisson et al., 1979; Rohmer et al., 1984; Ries-Kautt and Albrecht, 1989; Crossman et al., 2005; Inglis et al., 2018) and can be derived from a range of bacteria, including Alpha-and Gammaproteobacteria, Planctomycetes and Acidobacteria (Thiel et al., 2003; Sinninghe Damsté et al., 2004; Birgel and Peckmann, 2008; Sinninghe Damsté et al., 2017). Explorative searches of genomic databases for the biosynthetic gene encoding squalene-hopane-cyclase (shc) in mycobacteria from the *M. simiae* complex revealed a potential for biohopanoid production. In contrast, the more distantly related pathogenic mycobacteria, e.g. *M. tuberculosis*, are known to synthesize steroids instead of hopanoids (Lamb et al., 1998; Podust et al., 2001). Therefore,

mycobacteria from the *M. simiae* complex may be an additional source for hopanoic acids in the Censo seep soils.

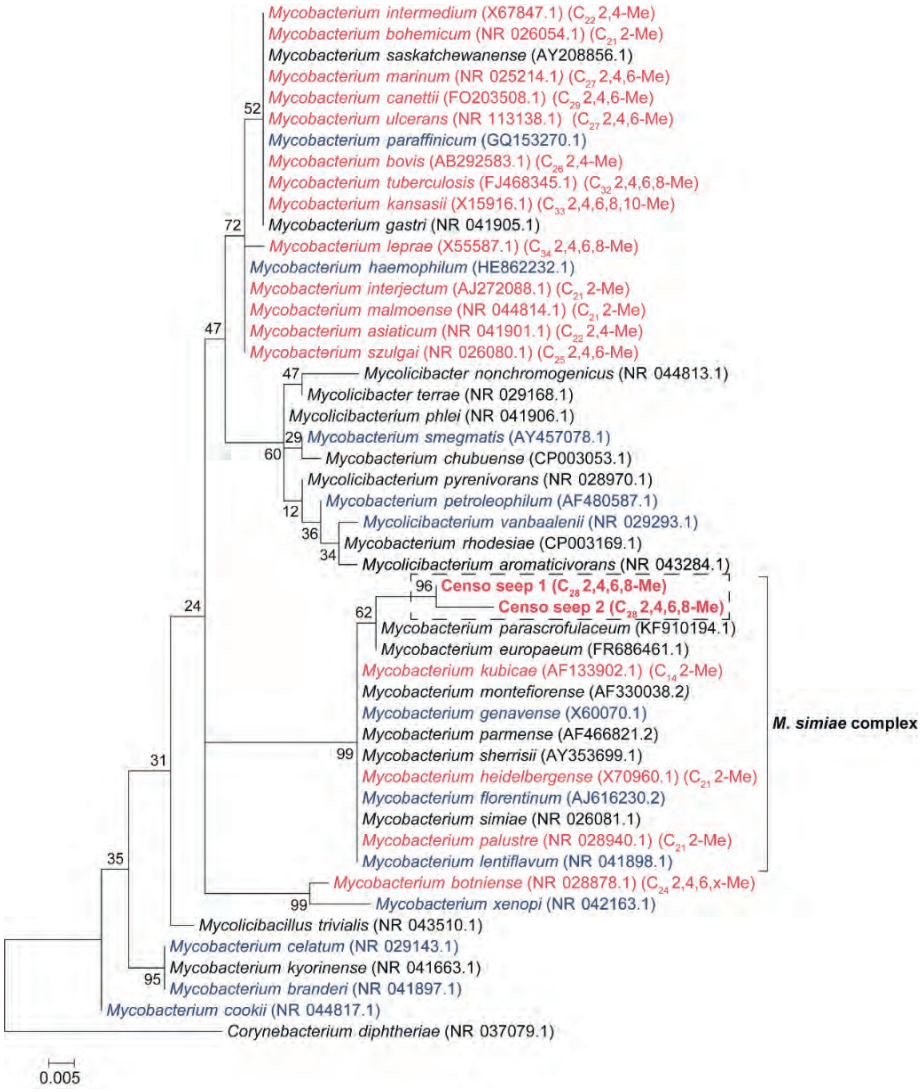


Figure 1. Maximum likelihood (ML) phylogenetic tree of the Mycobacterial 16S rRNA gene fragments (i.e. 294 bp; in bold) generated by amplicon sequencing and representative for the two OTUs present in the soils from the Censo seep "everlasting fire". The 16S rRNA gene sequence of *Corynebacterium diphtheriae* was used as an outgroup and other Mycobacterial 16S rRNA gene sequence are plotted for reference. The ML tree is based on the General Time Reversible model with gamma distribution plus invariable sites. Mycobacterial species biosynthesizing MAs are indicated in red font, species not containing MAs are shown in blue and species for which MAs have not been analyzed are shown in black. The mycobacterial species producing MAs (in red) are labelled with their dominant MA in brackets (total carbon number, Me= methyl, x = unidentified position of methyl group).

Interestingly, at the seep (0 m) the FA pattern is dominated by unusual FAs ranging from C_{19} to C_{31}, which are absent further away from the main seepage (Fig. 2). The mass spectra of the three most abundant representatives of these fatty acids are shown in Figure 3. Mass spectra of the methyl ester derivatives of these fatty acids show major fragment ions of *m/z* 88 and 101. These fragments result from "McLafferty" rearrangements associated with the presence of the carboxylic acid methyl ester group (Lough, 1975; Ran-Ressler et al., 2012). The presence of the even-numbered *m/z* 88 fragment ion, rather than the typical fragment ion at *m/z* 74 in the mass spectra of methyl esters of *n*-FAs, strongly suggests a methyl group at position C-2 (Fig. 3). One FA shows also high fragment ions at *m/z* 213 and *m/z* 241 (Fig. 3A). This difference of 28 Da hints at a second methyl group at position C-12 (Fig. 3A). Two of these fatty acids show a fragment ion at *m/z* 129 (Figs. 3C and E), suggesting the presence of an additional methyl at position C-4 of the fatty acids. The apparent methyl branches in these fatty acids are in agreement with the relatively early retention times of these FAs compared to the regular straight-chain counterparts (Fig. 2). Other fragment ions, including those potentially revealing the positions of additional methyl groups, were only present in low abundance, complicating further structural identification. Nevertheless, the presence of methyl branches at C-2 and C-4 in a number of these fatty acids does suggest that they may be related to mycobacteria-derived MAs, which share the same structural characteristics (Alugupalli et al., 1998; Nicoara et al., 2013). Indeed, the mass spectrum of the methyl ester of a synthetic C_{27} MA standard (2,4,6-trimethyl-tetracosanoic acid) (Holzheimer et al., 2020) shows identical mass spectral features (i.e. m/z 88 and 129; Fig. 4A). However, full structural interpretation of the mass spectrum of this authentic standard is also complicated by the low abundances of diagnostic fragment ions indicative for the position of the methyl branches in the alkyl chain.

To enhance the diagnostic fragmentation patterns of these potential MAs, the fatty acids were also transformed into a "picolinyl ester" (Harvey, 1998). The potential of this technique is revealed by the mass spectrum of the synthetic MA standard (2,4,6-trimethyl-tetracosanoic acid) "picolinyl ester" derivative (Fig. 4B), which shows fragment ions revealing all positions of methylation of the fatty acid *n*-alkyl chain. The high intensity of the fragment ion of *m/z* 165 indicates the presence of a methyl group at position C-2, while the presence of the fragment ions at *m/z* 178 and 206 combined with the absence of an *m/z* 192 fragment ion indicates the presence of a methyl group at C-4. Similarly, the presence of the third methyl group at position C-6 is revealed by

the fragment ions at *m/z* 220 and 248 and the low abundance of the fragment ion at *m/z* 234. Thus, the "picolinyl derivatization" technique substantially increases the confidence in the structural identification of multi-methyl-branched fatty acids using mass spectrometry. Therefore, this "picolinyl ester" derivatization technique was also applied to determine the methylation pattern of the potentially novel MAs in the Censo seep soils (Fig. 3).

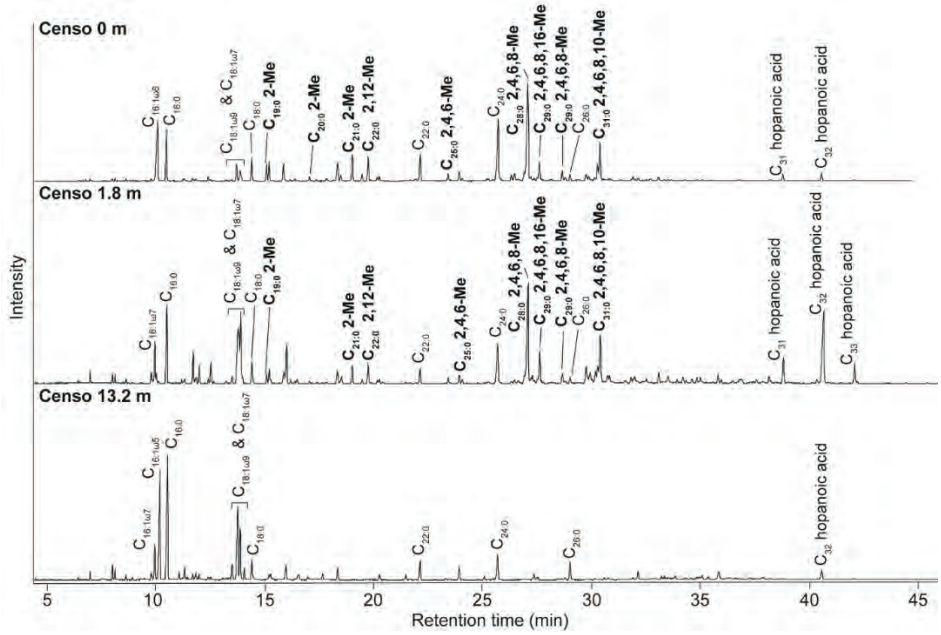


Figure 2. Total ion chromatograms of the saponified and derivatized (BF_3) fatty acid fractions from the Censo seep soils in increasing distance from the main seepage showing the distributions of FAs, hopanoic acids, and MAs. The black bold annotations show the tentatively identified MAs and the carbon position of their methyl groups (Me).

To illustrate this approach, we discuss the identification of the three major MAs. When analyzed as "picolinyl ester derivatives" (Fig. 3B, D and F), these MAs respectively showed molecular ions (M^+) 431, 515, and 557, indicating C_{22}, C_{28}, and C_{31} MAs, respectively. The mass spectrum of the "picolinyl ester derivative" C_{28} MA (Fig. 3D), the most abundant MA in the Censo seep soils, confirms the methylation at C-2 with the fragment ion of *m/z* 165. Furthermore, this spectrum also shows abundant fragment ions at *m/z* 178, 206, 220, 248, 262 and 290. Combined with the absence of the fragment ions at *m/z* 192, 234 and 276, this strongly suggests the presence of three additional methyl groups at position C-4, C-6 and C-8.

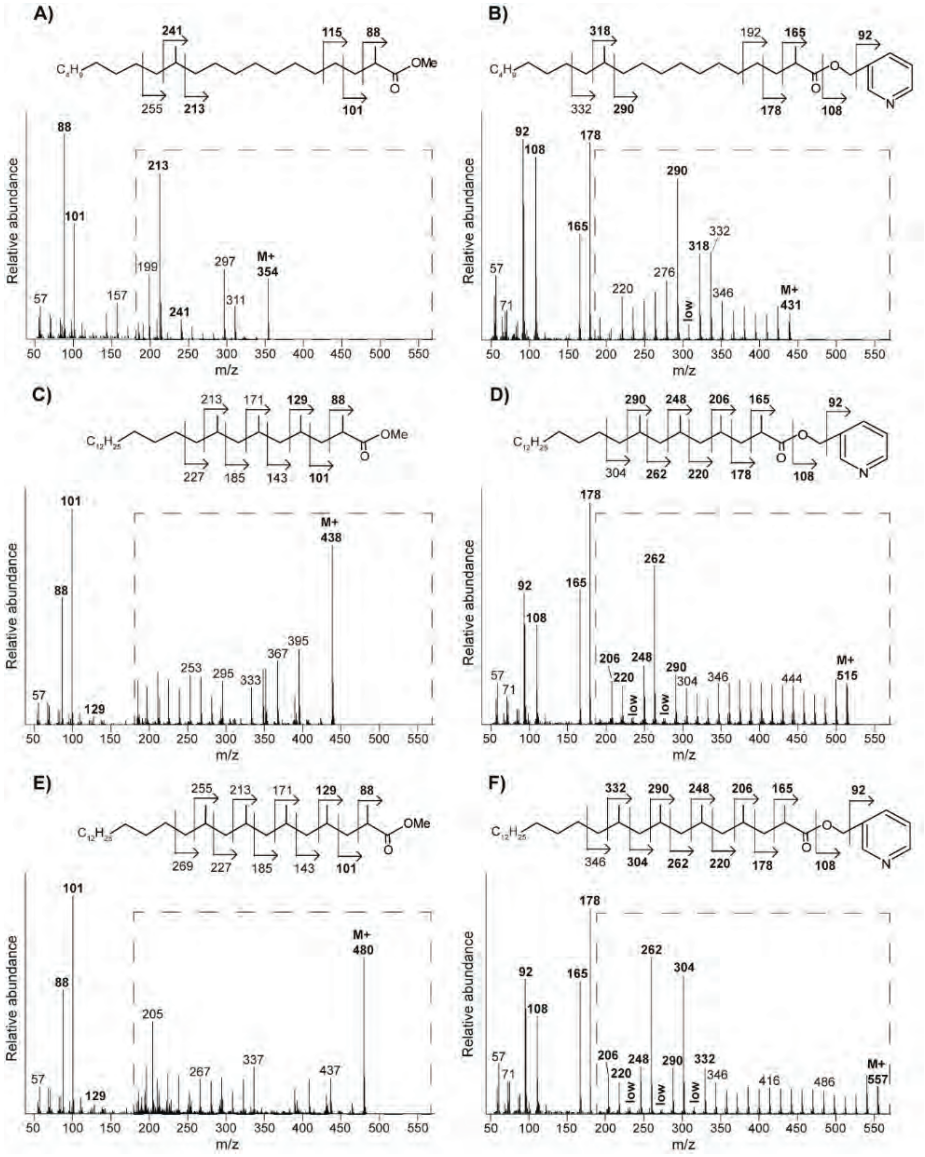


Figure 3. Mass spectra of the methyl ester (left panels) and 'picolinyl'-ester (right panels) derivatized MAs of the Censo 0 m soil sample, with proposed molecular structures and fragmentation patterns. 2,12-Dimethyl-eicosanoic acid (C_{22} 2,12-Me MA) (A and B), 2,4,6,8-tetramethyl-tetracosanoic acid (C_{28} 2,4,6,8-Me MA) (C and D), and 2,4,6,8,10-pentamethyl-hexacosanoic acid (C_{31} 2,4,6,8,10-Me MA) (E and F). The dashed boxes show a 10 times exaggerated view into the indicated area of the mass spectrum.

The mass spectrum of the C_{22} "picolinyl ester derivative" (Fig. 3B) also confirms the methyl branch at position C-2 through the mass ion *m/z* 165. Elevated fragment ions at *m/z* 290 and 318 in combination with the low intensity of the fragment ion at *m/z* 304 suggests a methyl group at position C-12. Further

mass spectral interpretations can be made for the C_{31} MA, with mass spectrum similar to that of the C_{28} MA but including an additional methyl group at position C-10, as indicated by the presence of fragment ions at *m/z* 304 and 332 and the absence of a fragment ion at *m/z* 318 (Fig. 3F). Thus, we tentatively identified these MAs as 2,12-dimethyl-eicosanoic acid (C_{22} 2,12-dimethyl MA), 2,4,6,8-tetramethyl-tetracosanoic acid (C_{28} 2,4,6,8-tetramethyl MA) and 2,4,6,8,10-pentamethyl-hexacosanoic acid (C_{31} 2,4,6,8,10-pentamethyl MA), respectively (Figs. 2, 3 and Table 2). Other abundant MAs tentatively identified include 2-methyl-octadecanoic acid (C_{19} 2-methyl MA), 2-methyl-nonadecanoic acid (C_{20} 2-methyl MA), 2-methyl-eicosanoic (C_{21} 2-methyl MA), 2,4,6-trimethyl-docosanoic acid (C_{25} 2,4,6-trimethyl MA), 2,4,6,8-tetramethyl-pentacosanoic acid (C_{29} 2,4,6,8-tetramethyl) and 2,4,6,8,16-pentamethyl-tetracosanoic acid (C_{29} 2,4,6,8,16-pentamethyl MA) MAs (Fig. 2 and Table 2).

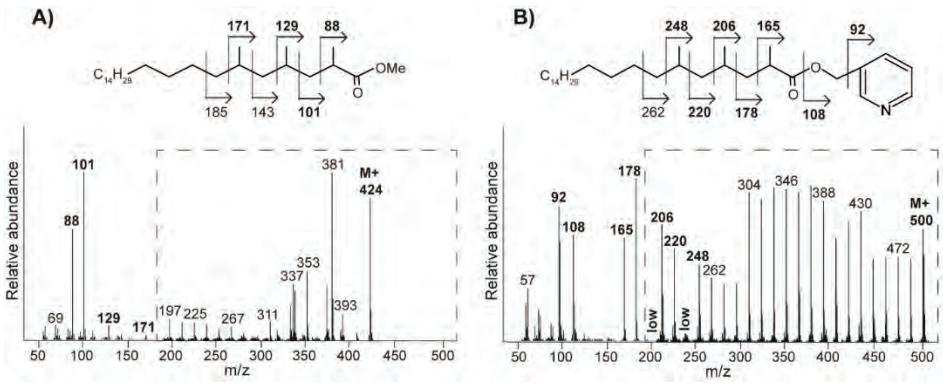


Figure 4. Mass spectra with fragmentation and annotated molecular structures of the A) methyl ester and B) 'picolinyl' ester synthetic 2,4,6-trimethyl-tetracosanoic acid (C_{27} 2,4,6-Me MA standard). The dashed boxes show a 10 times exaggerated view into the indicated area of the mass spectrum.

At the seep (0 m), the MAs have a high relative abundance, representing ca. 44% of the total FAs. Their abundance decreases to ca. 20% in the soil at 1.8 m from the seep, whereas MAs were not detected in the soil at 13.2 m distance from the seep (Fig. 2). These lipids show a similar distribution trend as the 16S rRNA gene sequencing results, which show high relative abundances of sequences from mycobacteria at the seep (ca. 34.1 % at 0 m), decrease to 8.5 % at 1.8 m, and are <1 % at 13.2 m (Table 1). Therefore, both the specific structure and the 16S rRNA gene data strongly suggest that the unusual FAs are derived from mycobacteria.

3.3 Mycocerosic acids as biomarkers for mycobacteria in the environment

Mycocerosic acids are thought to be only synthesized by mycobacteria and have been mainly studied as biomarkers for diseases from pathogenic (e.g. *M. tuberculosis* or *M. leprae*) or opportunistic pathogenic (e.g. the *M. simiae* complex) mycobacteria in the last decades (e.g. Minnikin et al., 1993a; Minnikin et al., 1993b; Torkko et al., 2003). These studies revealed a high structural variability of MAs with distribution patterns characteristic for different mycobacterial species. For example, the frequently studied *M. tuberculosis* shows a major C_{32} 2,4,6,8-tetramethyl MA and *M. leprae* a C_{34} 2,4,6,8-tetramethyl MA (Minnikin et al., 1993a; Minnikin et al., 1993b), whereas other mycobacteria feature shorter chain major MAs like C_{21} 2-methyl MA in *M. palustre* and C_{22} 2,4-dimethyl MA in *M. intermedium* (Chou et al., 1996; Torkko et al., 2002) (Fig. 1 and Table 2).

The Censo seep soils reveal a high number of tentatively identified MAs which have not been reported previously (Fig. 1 and Table 2), e.g. those biosynthesized by pathogenic mycobacteria like *M. tuberculosis* and *M. leprae* and by mycobacteria belonging to the more closely related *M. simiae* complex like *M. heidelbergense* (Minnikin et al., 1993a; Minnikin et al., 1993b; Torkko et al., 2003). The MA distribution of the Censo seep soils is characterized by a dominant C_{28} 2,4,6,8-tetramethyl MA, while the MA distribution of *M. heidelbergense* or *M. palustre* from the *M. simiae* complex is dominated by the C_{21} 2-methyl MA. Other more distantly related environmental opportunistic pathogens besides those of the *M. simiae* complex, like *M. marinum* or *M. intermedium*, produce a dominant 2,4,6- C_{27} trimethyl or C_{22} 2,4-dimethyl MA. As mentioned earlier, pathogenic mycobacteria like *M. tuberculosis* feature a major C_{32} 2,4,6,8-tetramethyl MA as well as *M. leprae* produces a dominant C_{34} 2,4,6,8-tetramethyl MA, clearly different from the major MA in the Censo soils (Fig. 1 and Table 2). Possibly, these unusual MAs could help to differentiate environmental Censo mycobacteria from opportunistic pathogenic and pathogenic mycobacteria in various modern and past environments.

Interestingly, the Censo mycobacteria show relatively high abundances of pentamethylated MAs (C_{29} 2,4,6,8,16-pentamethyl MA and C_{31} 2,4,6,8,10-pentamethyl MA) compared to other studied mycobacteria. *M. kansasii* has a dominant pentamethyl MA (C_{33} 2,4,6,8,10-pentamethyl MA; Table 2), which was also been reported in *M. tuberculosis* and *M. leprae* albeit in very low abundances, while *M. botniense* features a partially identified pentamethylated C_{27} (2,4,6,x,x) MA (Minnikin et al., 1985; Daffé and Laneelle, 1988; Torkko et al., 2003). Shorter chain MAs are also abundant in the Censo soils, some of

which have been identified in other mycobacterial species (Fig. 1 and Table 2): C_{20} 2-methyl MA (*M. palustre*), C_{21} 2-methyl MA (e.g. *M. palustre*, *M. heidelbergense* or *M. interjectum*) and C_{25} 2,4,6-trimethyl MA (*M. bohemicum*, *M. szulgai* and *M. intermedium*) (Torkko et al., 2001; Torkko et al., 2002; Torkko et al., 2003). The presence of C_{20} 2-methyl and C_{21} 2-methyl MAs in both Censo mycobacteria and mycobacteria from the closely related *M. simiae* complex indicate that these MAs might be a common feature in the *M. simiae* complex. However, these MAs have been also found in more distantly related mycobacterial species like *M. interjectum* and *M. malmoense*, while common pathogenic mycobacteria like *M. tuberculosis* or *M. bovis* do not produce these shorter chain MAs. These pathogenic mycobacteria contain a C_{27} 2,4,6-methyl MA (*M. tuberculosis*) and a C_{26} 2,4-methyl (*M. bovis*) as the shortest chain MAs (Minnikin et al., 1993a; Redman et al., 2009) which are not present in the Censo soils. Some more distantly related mycobacteria can even contain much shorter chain fatty acids like C_{11} 2-methyl MA (*M. interjectum* or *M. intermedium*), C_{15} 2-methyl MA (e.g. *M. kansaii* or *M. intermedium*) or C_{16} 2,4-dimethyl MA (*M. marinum*) (Torkko et al., 2003), but these are not found in the Censo MA inventory.

Table 2 Chemical variability and occurrence of MAs in the Censo seep soils and in the most relevant mycobacterial species. The underlined names of the mycobacterial species indicate the major MA configuration in the mycobacterial species. The x in the position of methylations in the *n*-alkyl chain features an unidentified position of the methyl group. MAs indicated in bold typeface are MAs identified in the Censo seep soils.

Chemical structure of MAs		**Occurrence**
Length of *n*-alkyl chain	**Position of methyl group(s)**	
C_{16}	2,4	*M. marinum* [a]
C_{18}	**2**	**Censo**
C_{19}	2	**Censo,** *M. palustre* [b]
C_{20}	**2**	**Censo**, *M. bohemicum*, *M. heidelbergense*, *M. malmoense*, *M. interjectum*, *M. palustre* [a,b,c,d,e]
C_{20}	2,4	*M. asiaticum*, *M. szulgai*, *M. intermedium*, *M. heidelbergense*, *M. malmoense* [a,b,d,e]
C_{20}	2,9	*M. palustre* [b]
C_{20}	**2,12**	**Censo**
C_{20}	2,4,6,x	*M. botniense* [d]

C_{22}	**2,4,6**	**Censo**, *M. bohemicum*, *M. szulgai*, *M. intermedium* [a,b,d]
C_{22}	2,4,6,x,x	*M. botniense* [d]
C_{24}	2,4	*M. bovis* [f,g]
C_{24}	2,4,6	*M. tuberculosis*, *M. bovis*, *M. kansasii*, *M. marinum*, *M. ulcerans*, *M. bohemicum*, *M. heidelbergense*, *M. malmoense*, *M. interjectum* [b,c,d,e,f,g,h]
C_{24}	**2,4,6,8**	**Censo**
C_{26}	2,4,6	*M. tuberculosis*, *M. leprae*, M. bovis, *M. kansasii*, *M. marinum*, *M. ulcerans* [f,g]
C_{24}	**2,4,6,8,16**	**Censo**
C_{25}	**2,4,6,8**	**Censo**
C_{26}	2,4,6,8	*M. tuberculosis*, *M.* leprae, *M. bovis*, *M. kansasii*, *M. marinum*, *M. ulcerans* [f,g,h,I,j]
C_{26}	**2,4,6,8,10**	**Censo**
C_{28}	2,4,6,8	*M. tuberculosis*, *M. leprae*, *M. bovis*, *M. kansasii*, *M. marinum* [f,g,h,k]
C_{28}	2,4,6,8,10	*M. tuberculosis*, *M. leprae*, *M. kansasii* [f,g]
C_{30}	2,4,6,8	*M. leprae* [f]

[a] Chou et al. (1996). [b] Torkko et al. (2002). [c] Torkko et al. (2001). [d] Torkko et al. (2003). [e] Valero-Guillén et al. (1988). [f] Minnikin et al. (1993a). [g] Minnikin et al. (1985). [h] Daffé and Laneelle (1988). [i] Donoghue et al. (2017). [j] Redman et al. (2009). [k] Minnikin et al. (2002).

The most unique feature that distinguishes the MAs of the Censo mycobacteria from cultivated mycobacterial species is the occurrence of methyl groups in the middle of the fatty acid chain at positions C-12 and C-16 in C_{22} 2,12-dimethyl and C_{29} 2,4,6,8,16-pentamethyl MAs, respectively. To the best of our knowledge, this mid-chain methyl branching has only been reported once before, in the mycobacterial species *M. palustre*, also from the *M. simiae* complex (Torkko et al., 2002), which is closely related to the species living in the Censo soil. However, the methyl branching in *M. palustre* is at position C-9 (C_{22} 2,9-dimethyl MA) (Torkko et al., 2002).

The fatty acid profile of the Censo soils shows longer chain MAs (e.g. C_{28} 2,4,6,8-tetramethyl and C_{31} 2,4,6,8,10-pentamethyl MAs) which are even more abundant than C_{24} and C_{26} long-chain *n*-alkyl fatty acids. This feature has not been previously reported in mycobacteria including mycobacteria from the closely-related *M. simiae* complex like *M. heidelbergense* and *M. palustre*,

which synthesize much higher amounts of regular fatty acids over MAs (Torkko et al., 2002; Torkko et al., 2003). Some mycobacterial species from the *M. simiae* complex (i.e. *M. lentiflavum, M. florentinum* and *M. genavense*) and other more distantly related mycobacteria (e.g. *M. paraffinicum* and *M. smegmatis*) (Torkko et al., 2002; Torkko et al., 2003; Fernandes and Kolattukudy, 1997; Chou et al., 1998) do not even contain MAs.

In conclusion, the MA patterns in the Censo soil mycobacteria are clearly different from those of previously cultivated mycobacterial species. This could be caused by environmental conditions near the Censo seep, which may have induced adaptions and regulation processes within the biosynthesis systems of MAs in the Censo mycobacteria or may just be a chemotaxonomic feature. Further studies of other soils that contain mycobacteria should reveal how unique the MAs detected in the Censo soils are.

3.4 Role of the mycobacteria at the Censo seep

The high relative abundances of mycobacteria and MAs based on both the relative 16S rRNA gene abundance and FA composition in the soil close to the main Censo seep (Table 1 and Fig. 2), combined with the decrease of these abundances in soils further away from the seep, hint to the potential involvement of mycobacteria in gas oxidation processes at the gas seep system. To further investigate this, the $\delta^{13}C$ values of the MAs, as well as regular fatty acids and hopanoic acids, were analyzed in the Censo seep soils (Fig. 5) and compared with that of the thermogenic derived methane (-30 to -35 ‰) and ethane (-25 ‰) at the Censo seep, as previously reported by Grassa et al. (2004).

At the seep site, regular and unsaturated C_{16}, C_{18}, C_{22} and C_{24} fatty acids showed no significant depletion in their carbon isotopic composition ($\delta^{13}C$ = -25 to -30 ‰), while at 1.8 m distance these FAs feature a bit more depleted $\delta^{13}C$ values ranging from -33 to -37 ‰ (Fig. 5). As mentioned before, the C_{16} and C_{18} FAs could originate from Type I and Type II methanotrophs (e.g. Bowman, 2006; Dedysh et al., 2007; Bodelier et al., 2009) although larger depletion of ca. 10 to 20 ‰ relative to the methane source is generally expected for fatty acids of aerobic methanotrophs (Jahnke et al., 1999; Blumenberg et al., 2007; Berndmeyer et al., 2013). Thus, a mixed bacterial community of heterotrophic and methanotrophic bacteria (e.g. Inglis et al., 2019) or other soil microbes using soil organic matter as a carbon source are likely to contribute to the pool of these fatty acids at Censo 0 m and 1.8 m. This agrees with the typical bulk $\delta^{13}C$ values of -25 to -30 ‰ in temperate soils (Balesdent et al., 1987; Huang et al., 1996)

and the presence of saturated and unsaturated C_{16} and C_{18} fatty acids even further away from the seep, at 13.2 m (Fig. 2).

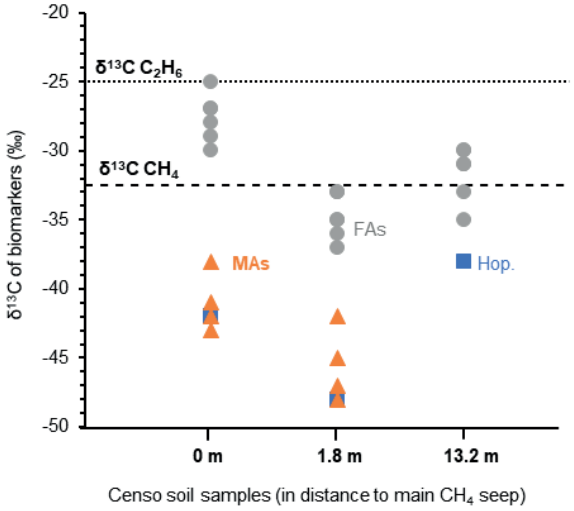


Figure 5. The stable carbon isotopic composition ($\delta^{13}C$) of biomarkers in the Censo soils at increasing distance from the main gas seepage. Biomarkers shown are fatty acids (FAs = grey circle), mycocerosic acids (MAs = orange triangle), and the C_{32} hopanoic acid (Hop. = blue square). Data points represent the mean average of two analysis. The $\delta^{13}C$ values of the released methane (CH_4 ~ -32.5 ‰) and ethane (C_2H_6 = -25 ‰) are indicated by dashed lines in the plot.

The C_{32} 17β,21β(H)-hopanoic acid shows more depleted $\delta^{13}C$ values ranging from -42 to -48 ‰, at Censo 0 m and 1.8 m, respectively (Fig. 5), suggesting an origin from bacteria involved in the cycling of a ^{13}C depleted carbon source like methane at this gas seep. The C_{32} hopanoic acid is a diagenetic product of bacteriohopanepolyols (Rohmer et al., 1984; Ries-Kautt and Albrecht, 1989; Farrimond et al., 2002), which could be produced by some of the aerobic methanotrophs (e.g. Methylocystaceae or Methylococcales) (Zundel and Rohmer, 1985; van Winden et al., 2012) identified in the Censo seep soils (Table 1). However, as discussed above, the mycobacteria in the soil, which are closely related to the *M. simiae* complex (Fig. 1), might also be able to synthesize hopanoids and therefore could be contributing to the hopanoid pool.

Depleted $\delta^{13}C$ values are observed in the MAs (-38 to -48 ‰) close to the Censo seep at 0 m and 1.8 m (Fig. 5), indicating that these are likely synthesized by organisms that use a ^{13}C-depleted carbon source rather than soil organic matter. The Censo seep releases high amounts of methane (76-86 % of total released gas) and minor amounts of higher gaseous hydrocarbons (ethane, propane etc.) as well as CO_2 and N_2 (Etiope et al., 2002; Grassa et al., 2004).

Thus, it would appear that the Censo mycobacteria are using ^{13}C-depleted methane as their carbon source as it is the major released gas at the Censo seep. This is in agreement with the decreasing relative abundance of mycobacteria and MAs away from the main seepage according to the decrease in the released gas. Furthermore, the δ^{13}C values of the MAs are more negative than the δ^{13}C value of the released methane, as expected for methanotrophs, and dissimilar to bulk soil organic matter and the simple FAs likely derived from heterotrophic bacteria.

However, these results are not completely in agreement with previous incubation and genetic studies, which showed that mycobacteria are not able to utilize methane but rather use other gaseous hydrocarbons like ethane and propane as well as alkenes, methanol, and carbon monoxide as their carbon source (Park et al., 2003; Coleman et al., 2011; Coleman et al., 2012; Martin et al., 2014). Studies from the 1950s reported high abundances of mycobacteria in soils from areas of oil and gas production and in areas of petroliferous gas seeps, hinting to their potential involvement in gas oxidation processes (Davis et al., 1956; Davis et al., 1959). Cultivation experiments of those soils confirmed that mycobacteria did not utilize methane but higher gaseous hydrocarbons (ethane and propane) (Davis et al., 1956; Dworkin and Foster, 1958). These results suggest that the mycobacteria in the Censo soils are perhaps not using methane, but possibly other gaseous hydrocarbons in the seep, like ethane or propane. However, it should be noted that two previous studies have described mycobacterial species *Mycobacterium flavum var. methanicum*, *Mycobacterium methanicum n. sp.* and *Mycobacterium* ID-Y that were able to oxidize methane (Nechaeva, 1954; Reed and Dugan, 1987). Alternatively, mycobacteria at the Censo seep could act as indirect methane utilizers by using secondary products of methane oxidation performed by other methanotrophs, like methanol. Indeed, some studies have shown that cultured pathogenic mycobacteria were able to utilize methanol (Reed and Dugan, 1987; Park et al., 2003; Park et al., 2010). However, this is difficult to reconcile with their very high abundances (up to 34.1 %) compared to the low abundance of typical methanotrophs like Methylococcales or Methylocystaceae (up to 5.1 %) near the seep (Table 1).

Overall, based on the clear abundance of mycobacterial 16S rRNA sequences in the Censo seep soils, the novel ^{13}C depleted MAs identified here may be useful biomarkers for the presence of hydrocarbon-oxidizing mycobacteria in soils. These unique MAs in combination with ^{13}C depletion could be used to trace mycobacteria in present and past environments, specifically those influenced by hydrocarbon seepage. Longer chain fatty acids

and branched fatty acids like MAs have been shown to be more resistant than other biomolecules (e.g. short-chain fatty acids) to diagenetic changes in diverse studies of fossil forests, sediment cores from the Gulf of California, and petroleum systems (Staccioli et al., 2002; Wenger et al., 2002; Camacho-Ibar et al., 2003). Under the right conditions, fatty acids may be preserved as bound compounds in ancient sediments through the Miocene (Ahmed et al., 2001). Indeed, studies have indicated the presence of MAs of *M. tuberculosis* on ancient bones from a 17,000 year old bison and from a ca. 200 year old human skeleton (Redman et al., 2009; Lee et al., 2012), suggesting a high preservation potential of these lipids.

Nevertheless, future research should investigate the presence and stable carbon isotope composition of MAs in other modern terrestrial and marine hydrocarbon seeps as well as in past environments where gas seepage might have played an important role. Additionally, further detailed incubation studies and genomic analysis of the Censo mycobacteria and mycobacteria at other terrestrial gas seeps are required to elucidate the exact role of the mycobacteria in gas oxidation processes at the Censo seep and in other gas rich environments.

Conclusion

Soils from the Fuoco di Censo Everlasting Fire show high relative abundances (up to 34 %) of uncultivated mycobacterial 16S rRNA gene sequences. These Censo mycobacteria are phylogenetically distant from the typical pathogenic mycobacteria *Mycobacterium tuberculosis* or *M. leprae*, and more closely related to the *M. simiae* complex like *M. heidelbergense* and *M. palustre*. At the main seep, Censo soils feature a unique MA pattern especially in the longer chain MAs. The most abundant MAs were tentatively identified as 2,4,6,8-tetramethyl-tetracosanoic acid (C_{28} 2,4,6,8-tetramethyl MA) and 2,4,6,8,10-pentamethyl-hexacosanoic acid (C_{31} 2,4,6,8,10-pentamethyl MA). The Censo soils also contained MAs with novel mid-chain methyl branching at positions C-12 and C-16 (C_{22} 2,12-dimethyl and C_{29} 2,4,6,8,16-pentamethyl MAs). The MA pattern in the Censo seep soils is clearly different from those reported for the well-studied mycobacteria like *M. tuberculosis* or *M. leprae* and from the closely related *M. simiae* complex. Only C_{20} 2-methyl, C_{21} 2-methyl and C_{25} 2,4,6-trimethyl MAs have been found previously in other mycobacteria from the *M. simiae* complex (e.g. *M. heidelbergense*) and three more distantly related mycobacteria (e.g. *M. interjectum*). These MAs have relatively low $\delta^{13}C$ values, suggesting that Censo mycobacteria use a carbon source depleted in ^{13}C, such as methane, higher gaseous hydrocarbons or secondary products of gas oxidation processes, like

methanol. The novel identified MAs in the Censo samples offer a new tool, besides DNA-based techniques, to investigate soils from present and past terrestrial environments for the presence of mycobacteria potentially involved in the cycling of gases.

Code availability

The 16S rRNA amplicon reads (raw data) have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject number PRJNA701386.

Data availability

Data will be made available on request to the corresponding author.

Author contribution

NTS, DR and SS planned research. NTS, FG and CRW collected samples. MH and AJM provided the synthetic C_{27} mycocerosic acid standard. NTS performed lipid analysis. LV analyzed 16S rRNA gene sequencing data. NTS, SS and LV interpreted the data. NTS wrote the paper with input from all authors.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank Marianne Baas, Monique Verweij, Jort Ossebaar, Ronald van Bommel, Sanne Vreugdenhil and Maartje Brouwer for technical assistance and Marcel van der Meer for discussion about isotopic values of the fatty acids. Sebastian Naeher, Rienk Smittenberg and Gordon Inglis are thanked for their useful comments which improved the manuscript.

Financial support

Stefan Schouten, Laura Villanueva and Jaap S. Sinninghe Damsté have been supported by the Netherlands Earth System Science Center (NESSC) and the Soehngen Institute for Anaerobic Microbiology (SIAM) through Gravitation grants (grant nos. 024.002.001 and 024.002.002) from the Dutch Ministry for Education, Culture and Science.

Review statement

This paper was edited by Sebastian Naeher and reviewed by Rienk Smittenberg and Gordon Inglis.

Chapter 8

The role of soil mycobacteria in the uptake of methane and ethane at a Sicilian everlasting fire

Nadine T. Smit, Darci Rush, Marcel T. J. van der Meer, Fausto Grassa, Jaap S. Sinninghe Damsté, Stefan Schouten

In preparation for Biogeosciences

Fuoco di Censo everlasting fire in Sicily, Italy

View from Fuoco di Censo downhill over the valleys of central Sicily, Italy

Abstract

Studies from the 1950s have shown that mycobacteria are ubiquitous near terrestrial gas seeps but their role in the cycling of the emitted gasses is still unclear. Recently, we found that novel mycobacterial species, closely phylogenetically related to the *Mycobacterium simiae* complex, as well as common aerobic methanotrophic bacteria were present in soils from a terrestrial Sicilian everlasting fire named Fuoco di Censo. The mycobacteria produced novel mycocerosic acids that were depleted in ^{13}C, suggesting a direct or indirect utilization of the gases like methane and ethane released at the Fuoco di Censo seep. Here, this hypothesis is tested by using polar lipid derived fatty acids–stable isotope probing (SIP) with ^{13}C-labeled methane or ethane on Fuoco di Censo seep soils, which showed rapid consumption of both gases. SIP incubations with ^{13}C-labeled methane for 5 days revealed incorporation of the ^{13}C label in regular and unsaturated C_{16} and C_{18} fatty acids, indicative of Type I and Type II methanotrophic activity, but mycocerosic acids did not incorporate ^{13}C. In contrast, the incubations with ^{13}C-labeled ethane for 9 days showed ^{13}C incorporation in C_{16} and C_{18} fatty acids as well as in mycocerosic acids, suggesting active ethane utilization by methanotrophs as well as mycobacteria. The ^{13}C incorporation of ^{13}C-labelled ethane but not methane in the structurally distinct mycocerosic acids at the Fuoco di Censo seep shows that soil mycobacteria might be an important group of microbes oxidizing gaseous higher hydrocarbons in global terrestrial gas seeps and soils. Therefore, the presence of unique mycocerosic acids, in combination with a ^{13}C-depleted signature, offers a new biomarker tool to detect hydrocarbon gas-consuming mycobacteria in present and past environments.

1. Introduction

Methane and higher gaseous hydrocarbons (ethane, propane and butane) are important greenhouse gases that directly affect global carbon cycles (Reeburgh, 1996; Etiope and Ciccioli, 2009). The reconstruction and understanding of the sources, sinks and chemical reaction pathways in greenhouse gas cycling is important to constrain the impact of methane and higher gaseous hydrocarbons on climate change. Worldwide, the aerobic oxidation of methane and of higher gaseous hydrocarbons (e.g., ethane, propane) in soils is one of their largest sinks (Solomon et al., 2007; Etiope and Ciccioli, 2009; Etiope et al., 2009; Dean et al., 2018). However, which microbes oxidize gaseous hydrocarbon in these terrestrial systems are not well constrained.

Soils harbour the most diverse and abundant groups of microbes on Earth, including aerobic methanotrophic bacteria and gaseous higher hydrocarbon-utilizing bacterial species (Fierer et al., 2012; Delgado-Baquerizo et al., 2018). Aerobic methanotrophic bacteria include three different phylogenetic groups: Type I methanotrophs (*Gammaproteobacteria*), Type II methanotrophs (*Alphaproteobacteria*) and *Verrucomicrobia* (e.g. Hanson and Hanson, 1996; Pol et al., 2007; Op den Camp et al., 2009; Bodelier et al., 2009). Some Type I methanotrophs are known to be able to utilize higher gaseous hydrocarbons, e.g. ethane (Kinnaman et al., 2007; Redmond et al., 2010). Other studies have shown that a group of gram-positive organisms in the Corynebacterium–Nocardia–Mycobacterium–Rhodococcus group (Dworkin and Foster, 1958; Ashraf et al., 1994) and some gram-negative *Pseudomonas* species are able to oxidize ethane (Takahashi, 1980; Hamamura et al., 1997). Genomic and cultivation studies revealed that mycobacteria can oxidize a range of greenhouse gases (ethane, propane, alkenes, carbon monoxide or hydrogen) and can degrade toxic polycyclic aromatic hydrocarbons (Hennessee et al., 2009; Coleman et al., 2012; Martin et al., 2014).

In the environment, bacteria involved in aerobic oxidation of methane can be traced via characteristic lipid biomarkers such as ^{13}C-depleted unsaturated, regular C_{16} and C_{18} fatty acids (FAs), hopanoids as well as bacteriohopanepolyols (e.g. Dedysh et al., 2007; Bodelier et al., 2009; van Winden et al., 2012; Talbot et al., 2014). In contrast, lipid biomarkers to detect microbes potentially involved in higher gaseous hydrocarbon oxidation are limited to ^{13}C-depleted C_{16} FAs derived from Type I methanotrophs like *Methylococcaceae* (Kinnaman et al., 2007; Redmond et al., 2010).

Several studies have shown that mycobacteria may be involved in gaseous hydrocarbon oxidation (Coleman et al., 2012; Martin et al., 2014). Indeed, studies from the 1950s found high abundances of non-pathogenic mycobacteria (*M. paraffinicum*) in areas of oil and gas production, gas seeps and common garden soils where mycobacteria have been shown to utilize ethane and propane (Davis et al., 1956, 1959; Dworkin and Foster, 1958). A recent study investigating soils from a terrestrial Sicilian everlasting fire named Fuoco di Censo found high relative abundances of 16S rRNA gene sequences attributed to novel environmental mycobacteria related to the *M. simiae* complex at and close to the gas seepage (Smit et al., 2021). Interestingly, unique environmental mycobacterial lipids namely multi-methyl-branched mycocerosic acids (MAs) with three to five methyl branches at regularly spaced intervals such as at positions C-2, C-4, C-6 and C-8 were detected as well as methyl branches at

positions C-12 and C-16 (Smit et al., 2021). The MAs identified at this seep were different from the known MAs in pathogenic and opportunistic pathogenic mycobacteria and showed relatively depleted $\delta^{13}C$ values, suggesting a potential involvement of their producers in the oxidation of the gases at the seep. However, the exact role and function of the mycobacteria and other methanotrophs in the cycling of methane and other gaseous hydrocarbons remained unclear.

Stable isotope probing (SIP)-incubations with $^{13}CH_4$ have been frequently used to study the activity and diversity of methane-oxidizing bacteria in terrestrial and marine environments (Bull et al., 2000; Crossman et al., 2005; Evershed et al., 2006; Maxfield et al., 2006; Redmond et al., 2010). The SIP technique allows the identification of organisms actively consuming ^{13}C-labelled substrates, based on the incorporation of ^{13}C into biomass, DNA, or lipid biomarkers. Therefore, we incubated soils directly taken from a terrestrial gas seep site in Sicily (Italy) with ^{13}C-labelled methane ($^{13}CH_4$) and ethane ($^{13}C_2H_6$) and analysed label uptake in fatty acid and mycocerosic acid biomarker inventories. These results shed new light on the role of mycobacteria in the cycling of methane and ethane at terrestrial gas seeps.

2. Material and methods

2.1 Sampling site and sample collection

The Fuoco di Censo seep (37°37’30.1’’N, 13°23’15.0’’E), in the following referred to as the Censo seep, is located in the mountains of Southwestern Sicily, Italy (Etiope et al., 2002, 2007; Grassa et al., 2004; Smit et al., 2021). The Censo seep is a typical example of a natural ‘everlasting fire’ at which the released thermogenic gas consists mainly of CH_4 (76–86%) and N_2 (10–17%) as well as some other minor gases like CO_2, O_2, ethane, propane, He and H_2 (Etiope et al., 2002; Grassa et al., 2004). The released CH_4 is characterized by a stable carbon isotopic composition of $\delta^{13}C = –35‰$, while the released ethane has a $\delta^{13}C$ values of –25‰ (Grassa et al., 2004).

Soil samples of the Censo seep were taken during a field campaign in October 2018 for incubation experiments. The soil was collected from a horizon 5–10 cm below surface directly at the main seep site with in-situ temperatures of ca. 18–20°C. The soils were directly transferred into a clean geochemical sampling bag and stored open to oxygen at 4°C for 7 days (soil for $^{13}CH_4$ incubations) and 14 days (soil for $^{13}C_2H_6$ incubations) until the start of the incubation experiments in the laboratory.

2.2 Incubation experiments

Two sets of incubation experiments with stable isotope probing (SIP), the first with 99% ^{13}C-labeled CH_4 (Sigma-Aldrich), and the second using fully ^{13}C-labeled 99% C_2H_6 (Sigma-Aldrich), were performed using the Censo soil samples. All soil slurries were prepared in clean autoclaved 160 ml glass bottles that were closed with butyl rubber stoppers and sealed with aluminium crimps. The bottles were inoculated with ca. 5 g aliquots of homogenized Censo soil and moistened with 0.5 ml Milli-Q water. The first incubation set comprised of three bottles amended with 5% $^{13}CH_4$ in the ca. 150 ml headspace and one heat-killed control bottle amended with 5% $^{13}CH_4$. The second incubation set comprised of three bottles amended with 5% of $^{13}C_2H_6$ in the ca. 150 ml headspace and one heat-killed control bottle amended with 5% $^{13}C_2H_6$. An unamended bottle (i.e. incubation control) was used to measure baseline levels of gases in the soil. The heat-killed control bottles from both set ups were autoclaved at 121 °C for 20 min and allowed to cool to ambient room temperature before $^{13}CH_4$ or $^{13}C_2H_6$ was added. Both incubation sets were incubated at 25°C in the dark and constantly shaken by a lab shaker. The oxygen levels in the gas phase of the incubation bottles were regularly measured by using a PICO-O_2 optical oxygen meter (*Pyroscience*) on oxygen sensor spots (OXSP5, *Pyroscience*), which were glued into the glass bottles before incubations started. During the incubations, oxygen concentration never fell below a threshold of 6.5%. The headspace methane and ethane concentrations were sampled at regular intervals with a gas-tight syringe (50 µl gas) and measured via gas chromatography with flame ionization detection (GC-FID; Thermo Scientific Focus). Methane and ethane concentrations were determined with reference to standard gas calibrations. $^{13}CH_4$ and $^{13}C_2H_6$ incubations continued until all $^{13}CH_4$ (5 days) and until about 95% of $^{13}C_2H_6$ (9 days) in the amended sets had been consumed (Fig. 1). Methane consumption rates were calculated based on a 150 ml headspace with 5 g soil (dry weight). After termination of both SIP incubation experiments, half of the incubated soils were freeze-dried and analysed for lipids.

2.3 Extraction, saponification, and derivatization

Freeze-dried Censo soils (ca. 0.5 g) from both SIP incubation set-ups were extracted with a modified Bligh and Dyer extraction for various compound classes (Schouten et al., 2008; Bale et al., 2013). Samples were ultrasonically extracted (10 min) with a solvent mixture containing methanol (MeOH), dichloromethane (DCM) and phosphate buffer (2 : 1 : 0.8; *v* : *v* : *v*). After centrifugation, the solvent was collected and combined, and the residues re-extracted twice. A biphasic separation was achieved by adding additional DCM

and phosphate buffer to a ratio of MeOH, DCM and phosphate buffer (1 : 1 : 0.9; *v* : *v* : *v*). The aqueous layer was washed two more times with DCM and the combined organic layers were eluted over a Na_2SO_4 column followed by drying under N_2.

Aliquots of the Bligh and Dyer extract (BDE) were saponified using 2 ml of a 1 N KOH MeOH solution and refluxed for 1 h at 130 °C. After cooling the pH was adjusted to 5 with a 2 N HCl MeOH solution. The layers were separated by 2 ml bidistilled water and 2 ml DCM, and the organic bottom layer was collected. The aqueous layer was washed two more times with DCM and the combined organic layers were dried first over a Na_2SO_4 column and then by N_2.

Aliquots of the saponified BDEs were esterified with 0.5 mL of a boron trifluoride–methanol solution (BF_3 solution) for 10 min. After cooling, 0.5 mL bidistilled water and 0.5 mL DCM were added and shaken, and the DCM bottom layer was pipetted off. The water layer was extracted twice with DCM, and the combined DCM layers were eluted over an $MgSO_4$ column. The derivatized extracts were eluted over a small silica gel column with ethyl acetate to remove polar compounds.

Subsequently, the derivatized extracts were separated using a small column packed with activated aluminium oxide into two fractions. The first fraction (fatty acid methyl ester fraction) was eluted with 4 column volumes of DCM and the second fraction (polar fraction) eluted with 3 column volumes of DCM/MeOH (1:1, v/v). The fatty acid methyl ester fractions were dried under a continuous flow of N_2 and afterwards analysed using gas chromatography-mass spectrometry (GC-MS) for compound identification and gas chromatography-isotope ratio mass spectrometry (GC-irMS) for compound specific isotope analysis with ethyl acetate as injection solvent.

2.4 Gas chromatography-mass spectrometry (GC-MS) and gas chromatography-isotope ratio mass spectrometry (GC-irMS)

GC-MS was carried out using an Agilent Technologies GC-MS Triple Quad 7000C in full scan mode with a scan time of 700 ms and a mass range of m/z 50 to 850. GC-irMS was performed with a 265 Thermo Scientific Trace 1310 with a GC-Isolink II, a ConFlo IV and a Delta Advantage irMS. GC for both GC-MS and GC-irMS was performed on a CP-Sil 5 column (25 m x 0.32 mm with a film thickness of 0.12 μm) with He as carrier gas (constant flow 2 ml min–1). The methylated samples (dissolved in ethyl acetate) were on-column injected at 70 ºC and subsequently, the oven was programmed to 130 ºC at 20 ºC min^{-1}, and

then at 4 ºC min^{-1} to 320 ºC, which was held for 10 min. Stable carbon isotope ratios are reported in atom% (AT%) ^{13}C. Values for fatty acids were corrected for the methyl groups introduced by the methylation and the results are reported as the average of the triplicate soil incubation bottles with standard deviation, the latter representing both analytical and experimental error.

3. Results

3.1 Methane and ethane consumption

Measurements of the $^{13}CH_4$ concentrations via GC-FID in the $^{13}CH_4$ incubation set up showed that ca. 40% of the amended $^{13}CH_4$ was consumed within 24 h (Fig. 1a). This corresponds to a consumption rate of 25 µmol CH_4 d^{-1} g $soil^{-1}$. The methane concentration decreased further in the following days until most of the $^{13}CH_4$ was oxidized after 5 days, with an average rate of $^{13}CH_4$ consumption of ca. 15 µmol CH_4 d^{-1} g $soil^{-1}$. The heat-killed control only showed slight changes in the CH_4 concentrations during the 5 days of the experiment.

In the first 24 h of the $^{13}C_2H_6$ incubations, ca. 9% of the amended $^{13}C_2H_6$ was utilized (Fig. 1b), corresponding to an ethane consumption rate of 8 µmol C_2H_6 d^{-1} g $soil^{-1}$. The ethane consumption increased after 1 day of incubation to ca. 20 µmol C_2H_6 d^{-1} g $soil^{-1}$. The average rate of $^{13}C_2H_6$ consumption was about 15 µmol C_2H_6 d^{-1} g $soil^{-1}$ during the whole 9 days of incubation, after which most of the ethane had been consumed. The ethane concentration of the heat-killed control of the $^{13}C_2H_6$ incubation remained relatively stable (Fig. 1b). The control incubation experiment where the Censo soil was not amended with either $^{13}CH_4$ or $^{13}C_2H_6$ did not reveal substantial production nor consumption of methane (Fig. 1a) or ethane.

3.2 Fatty acid composition

At the end of the incubations, the fatty acid (FA) fractions of the soils from both the $^{13}CH_4$ (5 days) and $^{13}C_2H_6$ Censo (9 days) incubations contained common FAs such as shorter chain FAs, including unsaturated and saturated C_{14}, C_{16} and C_{18} FAs, as well as longer chain FAs namely C_{22}, C_{24} and C_{26} FAs (Fig. 2a, b). The $^{13}CH_4$ incubations showed slightly higher relative abundance of C_{24} and C_{26} FAs than the $^{13}C_2H_6$ incubated Censo soils (Fig. 2a). Furthermore, the C_{31}, C_{32} and C_{33} 22R 17β,21β(H)-homohopanoic acids were also abundant in both Censo soil incubations with slightly higher relative abundance of the C_{32} hopanoic acid in the $^{13}CH_4$ Censo incubations. Additionally, both soil incubations showed high relative abundances of different MAs ranging from C_{21} to C_{31}. These MA had 2 to 5 methyl groups in the carbon chain mostly located close to the acid group in

a distribution, comparable to what was previously reported in these soils (Smit et al., 2021). In the $^{13}C_2H_6$ Censo incubations, the major C_{28} MA was relatively more abundant compared to common FAs like C_{22} and C_{24} (Fig. 2b).

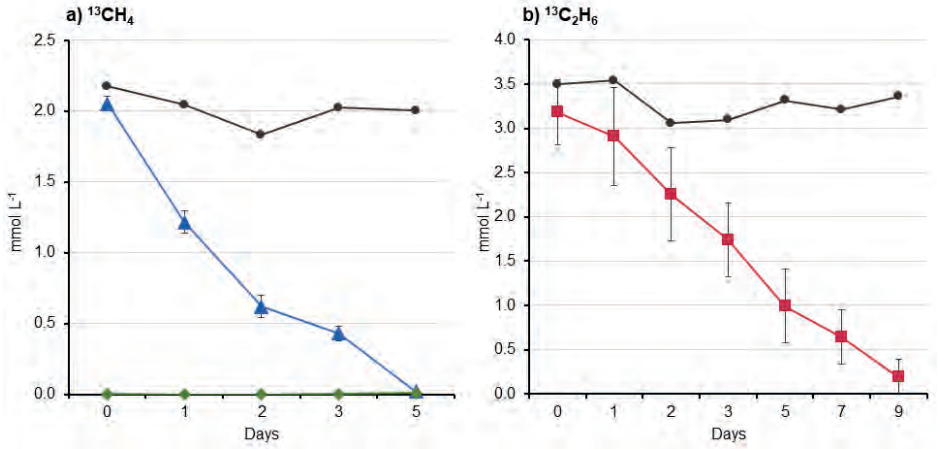


Figure 1. Concentration of methane and ethane (mmol L^{-1}) in the Censo soil incubations. a): data points show the $^{13}CH_4$ consumption (blue triangles) with standard deviation of three incubated soil samples, the heat killed control sample (HK; grey circles) and unamended soil sample (UA; green diamond) over 5 days. In b): data points show the $^{13}C_2H_6$ consumption (red squares) with standard deviation of three incubated soil samples and the heat killed control sample (HK; grey circles) over 9 days.

3.3 Labelling of FAs, MAs and hopanoic acids after incubation with $^{13}CH_4$ and $^{13}C_2H_6$

The ^{13}C labelling (expressed in AT% ^{13}C) of the main regular FAs, the hopanoic acids and several MAs was determined for the Censo soil $^{13}CH_4$ and $^{13}C_2H_6$ incubations (Fig. 3). The Censo soils incubated with $^{13}CH_4$ showed high levels of ^{13}C incorporation into regular and unsaturated C_{16} and C_{18} FAs (2 to 11 AT% ^{13}C) (Fig. 3a). Especially the unsaturated C_{16} and C_{18} FAs were heavily enriched with ^{13}C ranging from 8 to 11 AT% ^{13}C. In contrast, the C_{32} 22R 17β,21β(H)-hopanoic acid and all MAs (C_{21} to C_{31} MA) did not show a significant ^{13}C enrichment in the Censo soil $^{13}CH_4$ incubations. The heat-killed control showed no significant ^{13}C incorporation in any of the investigated FAs, C_{32} hopanoic acid or MAs, i.e. they had similar values as previously reported for the Censo soil (Smit et al., 2021).

The Censo soil incubations with $^{13}C_2H_6$ also revealed heavily ^{13}C-enriched regular and unsaturated C_{16} and C_{18} FAs with AT% ^{13}C values ranging from 6 to 13 % (Fig. 3b). The C_{32} 22R 17β,21β(H)-hopanoic acid showed no substantial incorporation of ^{13}C. In contrast to the methane incubation, all MAs had high AT% ^{13}C values, ranging from 1.4 to 2.0 AT% with the highest ^{13}C enrichment seen in the C_{21} MA. The heat killed control of the $^{13}C_2H_6$ Censo

incubations did not reveal any significant incorporation of ^{13}C label into the FAs, the C_{32} hopanoic acid or MAs.

4. Discussion

4.1 Methane and ethane consumption in the Censo seep soil incubation experiments

The consumption of $^{13}CH_4$ and $^{13}C_2H_6$ in the incubation bottles (Fig. 1) indicates that an active hydrocarbon gas-utilizing bacterial community is present in the Censo soil, i.e. hydrocarbon gas-utilizing bacteria which oxidized high amounts of CH_4 and C_2H_6 and incorporated the carbon derived from the gases into their biomass, including their lipids. The stable $^{13}CH_4$ and $^{13}C_2H_6$ concentrations, and thus lack of gas consumption, in the heat killed controls robustly indicate that the uptake of methane and ethane was mediated by microbes.

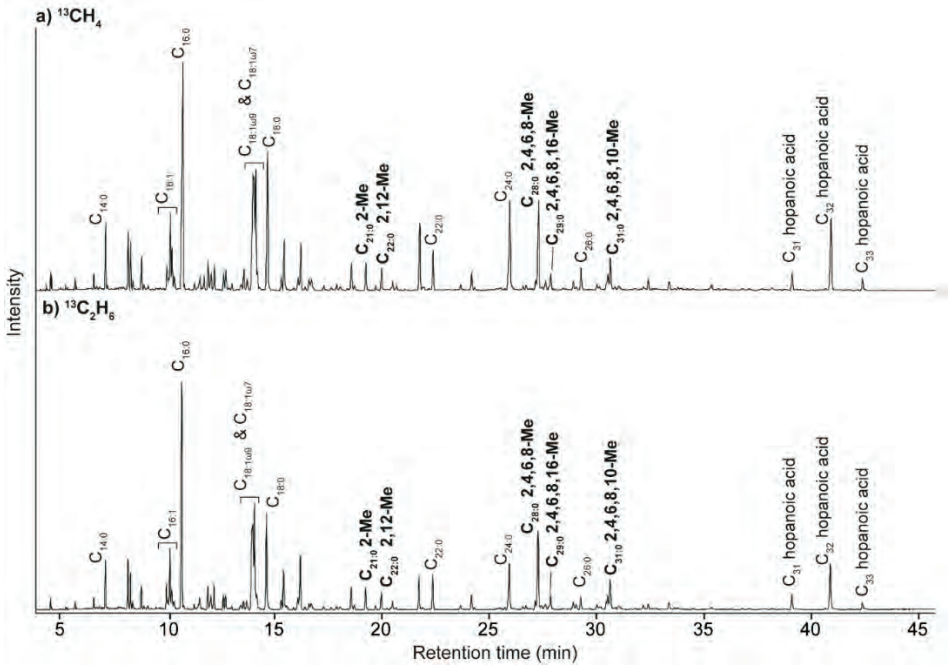


Figure 2. Total ion chromatograms of the fatty acid methyl ester fractions from the Censo seep soil incubations with a) $^{13}CH_4$ after 5 days and b) $^{13}C_2H_6$ after 9 days showing the distribution of FAs, hopanoic acids and mycocerosic acids (MAs). The black bold annotations show the previously identified MAs and the carbon position of their methyl groups (Me) (Smit et al., 2021).

The hydrocarbon gas consumption was fastest in the first 24 h in the $^{13}CH_4$ incubations (5 days) while consumption of $^{13}C_2H_6$ was about 3 times slower, although the overall average gas consumption rates were the same in both incubations set ups (Fig. 1). The differences in the early phase of the two

incubations might be related to the speed of adaptation and the initial microbial composition of the methane and ethane oxidizing community as well as to their ability to utilize high gas concentrations. A difference in initial gas consumption rates has also been observed in SIP incubations of sediments from marine hydrocarbon seeps at Coal Oil Point, California where the faster consumption of $^{13}CH_4$ than $^{13}C_2H_6$ was related to faster changes in the microbial communities, as revealed by DNA-based analyses (Kinnaman et al., 2007; Redmond et al., 2010).

4.2 Aerobic methane-oxidizing bacteria at the Censo seep

The regular and unsaturated C_{16} and C_{18} FAs produced during the $^{13}CH_4$ Censo soil incubations are most likely derived from Type I and Type II methanotrophs (e.g. *Methylococcales* or *Methylocystaceae*), which were identified previously in the Censo seep soil by 16S rRNA gene sequencing (Smit et al., 2021). The observation that methane oxidation in the Censo soil might be performed by these Type I and Type II methanotrophs confirms previous observations in fatty acid inventories of cultured methanotrophic bacteria (Dedysh et al., 2007; Bodelier et al., 2009) as well as $^{13}CH_4$ SIP studies of soils (Bull et al., 2000; Evershed et al., 2006). $^{13}CH_4$ SIP studies in temperate forest soils (Bull et al., 2000; Crossman et al., 2005; Evershed et al., 2006), peats (Chen et al., 2008) and marine gas seep sediments (Kinnaman et al., 2007; Redmond et al., 2010) also showed a high degree of ^{13}C incorporation in regular and unsaturated C_{16} and C_{18} FAs with moderate to high CH_4 concentrations. Furthermore, the lower ^{13}C incorporation in the saturated C_{16} and C_{18} FAs compared to their unsaturated homologues could be related to a mixed bacterial community of heterotrophic and methanotrophic bacteria being the source of the saturated compounds (Boschker et al., 1998), thereby diluting the ^{13}C signals of the latter bacteria. Interestingly, the natural ^{13}C abundance of these fatty acids was not depleted or only by 5–10‰ compared to the released CH_4 at the Censo seep (Smit et al., 2021).

No incorporation of ^{13}C was observed in the 22R 17β,21β(H)-bishomohopanoic acid in the $^{13}CH_4$ SIP incubations. Hopanoic acids are a diagenetic product of bacteriohopanepolyols, which can be produced by some methanotrophs but also from a variety of other bacterial species (Rohmer et al., 1984; Zundel and Rohmer, 1985; Talbot et al., 2001; Talbot and Farrimond, 2007). The lack of ^{13}C incorporation in the $^{13}CH_4$ SIP incubation suggests that the pool of C_{32} hopanoic acid was not produced by active methanotrophic bacteria during the short incubation time and may predominantly reflect fossil diagenetic compounds.

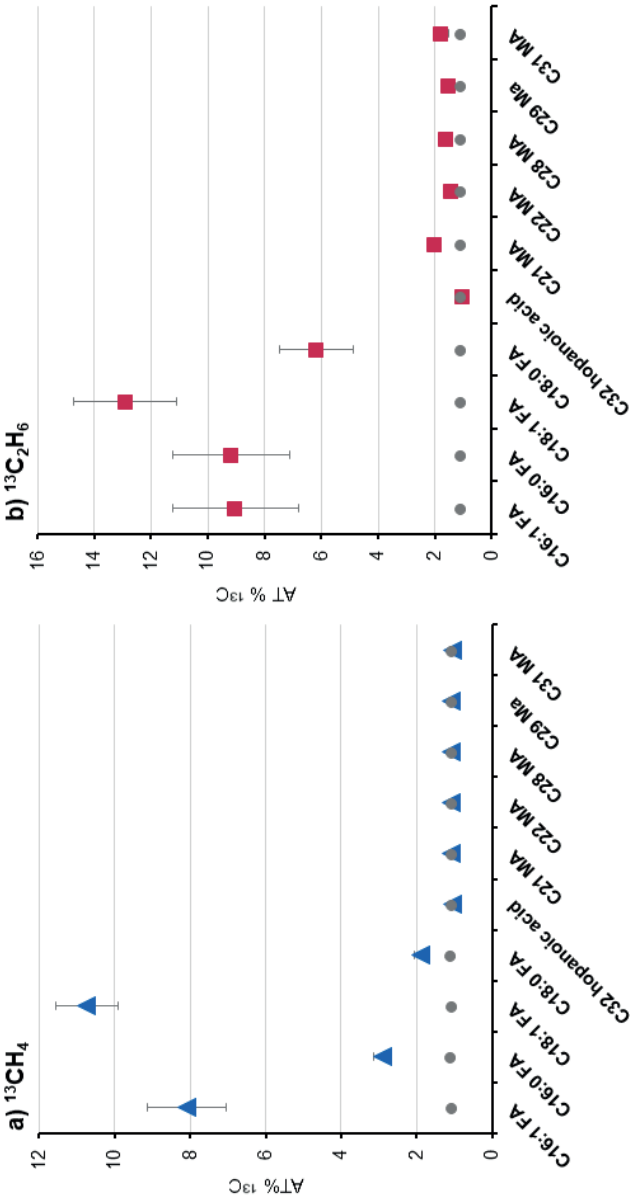


Figure 3. The stable carbon isotopic composition in atom% (AT%) ^{13}C of biomarkers in the Censo soil a) $^{13}CH_4$ after 5 days and b) $^{13}C_2H_6$ after 9 days incubations. Biomarkers shown are saturated and unsaturated fatty acids (FAs), the 22R 17β,21β(H) bishomohopanoic acid (C_{32} hopanoic acid) and mycocerosic acids (MAs). In a) $^{13}CH_4$ data points (blue triangles) and in b) $^{13}C_2H_6$ data points (red squares) represent the mean average of three incubated soil samples with standard deviations in light grey error bars. Some standard deviations are smaller than the symbol size and are therefore not visible. $\delta^{13}C$ values of heat killed controls for both incubations a) and b) are shown as grey circles.

Another possibility could be that the C_{32} hopanoic acid is synthesized by bacteria that do no utilise CH_4 as carbon source, such as heterotrophs or aerobic methanotrophs that do not directly incorporate CH_4. For example, Type II methanotrophs and methanotrophic *Verrucomicrobia* can assimilate both CH_4 and CO_2 as their carbon source thus only a portion of the fixed carbon may originate from depleted CH_4 or $^{13}CH_4$ in SIP incubations (Hanson and Hanson, 1996; Maxfield et al., 2006; Khadem et al., 2011).

The observation that MAs were relatively depleted in ^{13}C in natural soils at and close to the Censo seep, suggested a potential involvement of mycobacteria in CH_4 oxidation in the Censo soils (Smit et al., 2021). However, the lack of incorporation of ^{13}C in MAs in the $^{13}CH_4$ incubation experiments indicates that the Censo mycobacteria might not be able to utilize CH_4. This is supported by previous genome-based studies, which suggests that mycobacteria show the potential to oxidize higher gaseous hydrocarbons e.g. ethane and alkenes, methanol, and carbon monoxide but not methane (Park et al., 2003; Coleman et al., 2011, 2012; Martin et al., 2014). Although two studies have suggested that certain mycobacterial species are able to utilize methane directly (Nechaeva, 1949; Reed and Dugan, 1987), it is not clear if these mycobacteria use the CH_4 as their energy or carbon source and thus their biomass could remain unlabelled, like '*Ca*. Methylomirabilis oxyfera' which fixes carbon autotrophically (Kool et al., 2012; Rasigraf et al., 2014). Other possibilities include the incubation time being too short for the mycobacteria to grow or produce sufficient MAs, or that they were outcompeted by the other methanotrophs. Therefore, we cannot completely exclude that the Censo mycobacteria could oxidize CH_4, and this should be tested by future incubations experiments and metagenome analysis.

4.1 Ethane-utilizing bacteria in the Censo soils

The fatty acid fraction of the $^{13}C_2H_6$ SIP Censo incubations shows a similar distribution and inventory of FAs, hopanoic acids and MAs as the $^{13}CH_4$ SIP incubations (Fig. 2b). The oxidation of solely C_2H_6 has been rarely investigated in the past and the few studies done show that a group of gram-positive organisms in the Corynebacterium–Nocardia–Mycobacterium Rhodococcus group (Dworkin and Foster, 1958; Ashraf et al., 1994), along with some gram-negative *Pseudomonas* species are able to oxidize ethane (Takahashi, 1980; Hamamura et al., 1997). $^{13}C_2H_6$ SIP experiments of sediments from marine hydrocarbon seeps at Coal Oil Point offshore California showed that *Methylococcaceae* species can oxidize ethane in addition to methane and they

produce high abundances of unsaturated and saturated C_{16} FAs (Kinnaman et al., 2007; Redmond et al., 2010). In the Censo soils, *Methylococcaceae* species were identified previously (Smit et al., 2021) and might synthesize the ^{13}C-labelled saturated and unsaturated C_{16} FAs in the $^{13}CH_4$ as well as in the $^{13}C_2H_6$ incubations. Therefore, the ^{13}C label in C_{16} and C_{18} FAs in the $^{13}C_2H_6$ incubation experiments probably indicate that hydrocarbon gas-oxidizing bacteria are able to not only utilize CH_4 but also C_2H_6, in agreement with previous observations for some methanotrophic bacterial species (Hazeu and de Bruyn, 1980).

Intriguingly, the MAs in the $^{13}C_2H_6$ incubated Censo soil do show significant ^{13}C incorporation after nine days of incubation time, in contrast to the methane incubations (Fig. 3b). The ^{13}C-labelling of the MAs is in accordance with previous studies that showed mycobacteria are one of the rare groups of microbes that can utilize ethane (Park et al., 2003; Coleman et al., 2011, 2012). Furthermore, studies from the 1950s showed high abundances of mycobacteria in soils from areas of oil and gas production and in areas of petroliferous gas seeps, suggesting a potential involvement of mycobacteria in hydrocarbon gas oxidation processes (Davis et al., 1956, 1959). Subsequent cultivation experiments of those soils confirmed that mycobacteria did not utilize methane but higher gaseous hydrocarbons like ethane (Davis et al., 1956; Dworkin and Foster, 1958), which is in accordance with our Censo soil SIP incubation results. However, a slower rate of MA synthesis pathways, growth or abundance of mycobacteria could also influence the appearance of ^{13}C label in the ethane incubation experiments compared to the methane incubation experiments.

Our results of the $^{13}C_2H_6$ SIP incubations in tandem with the $^{13}CH_4$ incubations suggest that mycobacteria are involved in the oxidation of ethane rather than methane at the Censo seep. If confirmed in other settings, then MAs may be used as biomarkers for ethane oxidation by mycobacteria in recent and past environments. However, further research in the lipid inventory and genes responsible for hydrocarbon gas oxidation in environmental mycobacteria should be done in other terrestrial gas seeps and soils.

Conclusion

SIP incubations using $^{13}CH_4$ and $^{13}C_2H_6$ with soils from the terrestrial Censo gas seep showed a substantial ^{13}C label incorporation into regular and unsaturated C_{16} and C_{18} FAs characteristic for common Type I and II methanotrophs. These common methanotrophs seem to be able to not only utilize CH_4 but also C_2H_6. Furthermore, the ^{13}C incorporation in MAs was observed in the $^{13}C_2H_6$ incubations but not in $^{13}CH_4$ incubations, which indicates that Censo seep

mycobacteria oxidize C_2H_6 at the Censo seep and not CH_4. This is in accordance with previous genetic and cultivation studies of gas-utilization of mycobacteria. However, it should be noted that the experimental conditions, the time period of incubation or other factors could also influence the rate of synthesis of MAs by the Censo mycobacteria in the SIP incubations and thus the degree of the ^{13}C label incorporation. Our results do suggest, however, that ^{13}C depleted MAs in soils and terrestrial hydrocarbon seeps may offer new biomarkers to trace ethane oxidation in terrestrial gas seeps and soils.

Author contributions

NTS, DR and SS designed the study. NTS and FG collected the samples. NTS performed incubation experiments and lipid analysis. NTS, DR, MvdM, SS and JSSD interpreted the data. NTS wrote the manuscript with input from all authors.

Competing interests

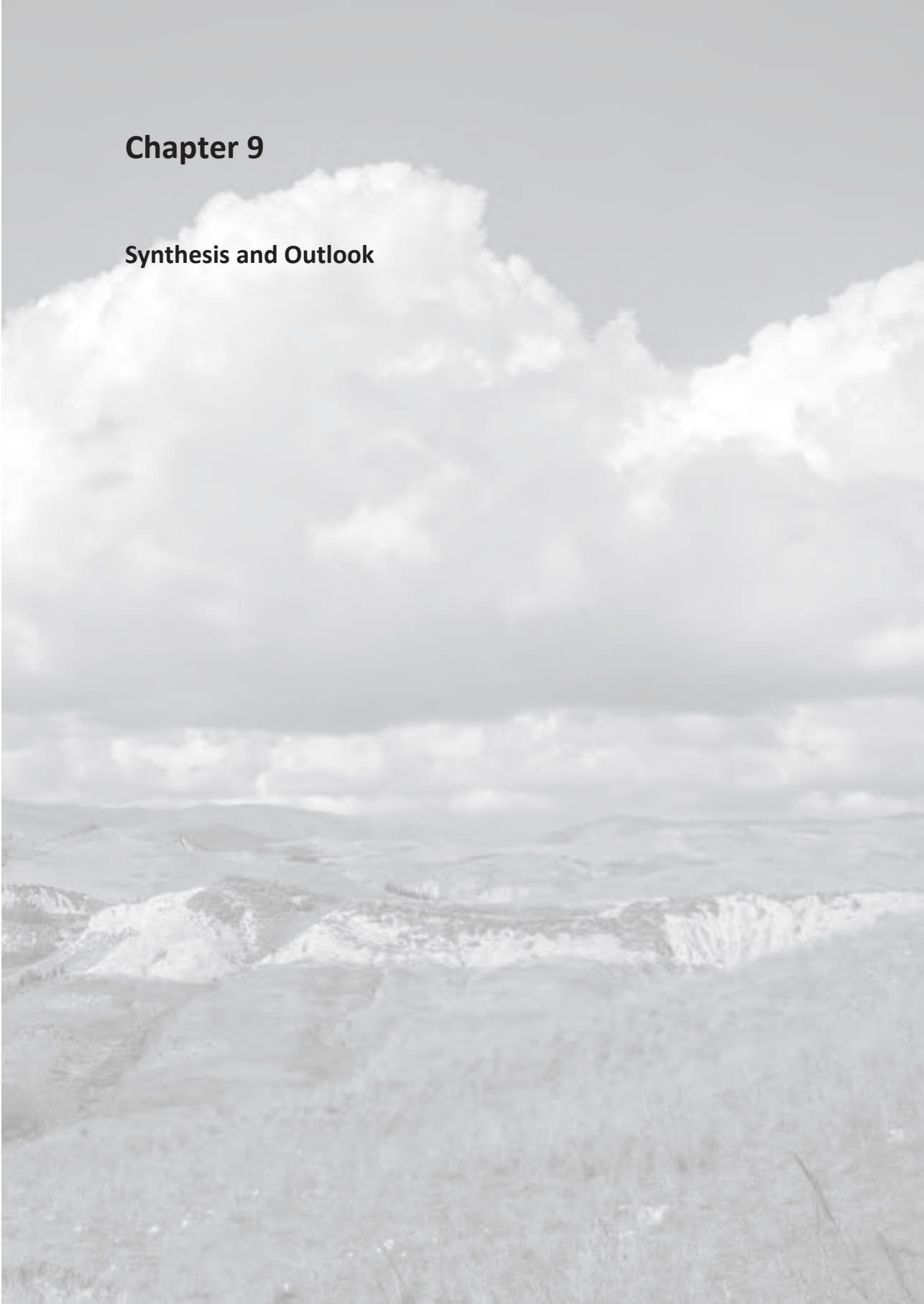
The authors declare that they have no conflict of interest.

Acknowledgements

We thank Marianne Baas, Monique Verweij, Jort Ossebaar and Ronald van Bommel for technical assistance. We also thank Helge Niemann and Maaike Goudriaan for providing the oxygen meter unit and assistance in SIP incubation set up. This study received funding from the Netherlands Earth System Science Centre (NESSC) through grant (024.002.001) to JSSD and SS from the Dutch Ministry for Education, Culture and Science.

Chapter 9

Synthesis and Outlook



View from the Bissana seep near Cattolica Eraclea in Sicily, Italy

Maccalube di Aragona mud volcano field (mud breccias from old eruptions) in Sicily, Italy

Synthesis and Outlook

In this thesis, novel characteristic lipid biomarkers are developed to detect microbes involved in the aerobic oxidation of methane but also of other gaseous hydrocarbons especially in terrestrial environments. The recognition of such lipid biomarkers is essential to unravel the cycling of methane and higher gaseous hydrocarbons, essential processes in global climate change in the present-day and past Earth system. Finding reliable lipid biomarkers to trace microbial methane oxidation has proven difficult in the past, highlighting the need for the research presented in this thesis.

Lipid biomarkers for nitrite/nitrate-dependent methane oxidation

Over the last decade the process of nitrite/nitrate-dependent methane oxidation was discovered and investigated (Raghoebarsing et al., 2006; Ettwig et al., 2010; Haroon et al., 2013) but distinct lipid biomarkers to detect this process have only rarely been investigated until now (Kool et al., 2012, 2014). Therefore, lipid biomarker inventories specific for bacteria and archaea performing nitrite/nitrate-dependent methane oxidation were investigated in **Chapters 2 and 3** using enrichment cultures of the intra-aerobic methanotroph '*Candidatus* Methylomirabilis oxyfera' ('*Ca*. M. oxyfera') and the ANME-2d archaeon '*Candidatus* Methanoperedens'. '*Ca*. M. oxyfera' synthesizes four unique demethylated hopanoids (22,29,30-trisnorhopan-21-ol, 3-methyl-22,29,30-trisnor hopan-21-one and 3-methyl-22,29,30-trisnorhopan-21-ol), of which only 22,29,30-trisnorhopan-21-one had been identified previously. A new multiple reaction monitoring (MRM) method was developed and used to successfully detect these trisnorhopanoids in a peatland. '*Ca.* Methanoperedens' produced archaeol and hydroxyarchaeol as well as isoprenoidal GDGTs and unusually high amounts of hydroxy-GDGTs. The novel demethylated hopanoids and high abundances of hydroxy-GDGTs may thus be used in future studies to trace nitrite/nitrate-dependent methane oxidation in various present and past environments. It is particularly promising to employ these new biomarkers for nitrite/nitrate-dependent methane oxidation in marine environments, e.g. oxygen minimum zones, since this process has only been investigated in freshwater environments and peatlands until now (Kool et al., 2012; Vaksmaa et al., 2017).

However, '*Ca*. M. oxyfera' was discovered in the offshore environment in northern Mexico and Costa Rica using molecular ecological techniques (Padilla et al., 2016) and these environments would form a promising testing ground for the new biomarkers. Furthermore, the preservation potential of especially the demethylated hopanoids should be determined so that it is clear

how far back in time nitrite-dependent methane oxidation in paleo-environments can be investigated.

Preliminary aerobic biodegradation experiments using '*Ca*. M. oxyfera' biomass for 90 days showed good preservation of the four demethylated hopanoids when typical fatty acids were already degraded (Smit et al., unpublished results). Interestingly, '*Ca*. M. oxyfera' (and Planctomycetes; Sinninghe Damsté et al., 2004) directly synthesize demethylated hopanoids, contrasting previous ideas that demethylated hopanoids represent diagenetic alteration products in the environment (Moldowan et al., 1984; Noble et al., 1985; Peters et al., 2005). This observation may impact the interpretation of the presence of demethylated hopanes in present and past environments significantly.

Methane oxidation in terrestrial mud volcanoes

Mud volcanoes (MVs) are a major source of methane, however, mainly marine MVs have been studied for aerobic and anaerobic oxidation of methane and their characteristic lipid biomarkers (Stadnitskaia et al., 2007; Elvert and Niemann, 2008; Lee et al., 2018). In **Chapter 4**, investigations of prominent terrestrial Sicilian MVs showed that freshly emitted mud breccias have high abundances of petroleum-derived hydrocarbons but low amounts of microbial lipid biomarkers such as fatty acids and hopanoic acids. These fatty and hopanoic acids were not depleted in ^{13}C, likely indicating an origin from bacteria other than methane-oxidizing bacteria (MOBs). In contrast, the isoprenoidal GDGTs present in one MV possess a distribution indicative for anaerobic oxidation of methane by ANME archaea, in contrast to studies in other terrestrial MVs which showed only presence of partly ^{13}C-depleted archaeol and dialkyl glycerol diethers (DAGE) specific for sulfate-dependent anaerobic oxidation of methane (Alain et al., 2006; Heller et al., 2012). The absence of strong indications for methane oxidation processes is surprising, in particular compared to marine mud volcanos (Stadnitskaia et al., 2005, 2008; Elvert and Niemann, 2008; Lee et al., 2018) and warrants further investigations. Future research should investigate the intact polar lipid inventory of these terrestrial MVs to more specifically examine living source organisms that actively synthesize GDGTs and fatty acids (Schubotz et al., 2009; Rossel et al., 2011) or study the distributions of BHPs, indicative for aerobic oxidation of methane (**Chapters 5 and 6)**.

Further confirmation of the importance of methane oxidation processes can come from molecular ecological techniques such as the sequencing of 16S rRNA or pmoA/mcrA genes (Hanson and Hanson, 1996; Hallam et al., 2003;

Knief, 2015), which may identify aerobic or anaerobic methanotrophs present in the mud breccias. The activity of these processes, as well as better constraints of fatty acid sources can come from SIP (stable isotope probing) or RIP (radioisotope probing), using ^{13}C or ^{14}C labeled methane and/or DIC or D_2O, respectively (e.g. Bull et al., 2000; Kellermann et al., 2012; Evans et al., 2019).

Using bacteriohopanepolyol derivatives to detect aerobic oxidation of methane
Bacteriohopanepolyols and their derivatives are known to be indicative for distinct bacterial species involved in methane oxidation (Neunlist and Rohmer, 1985; Zundel and Rohmer, 1985; Cvejic et al., 2000; Talbot et al., 2014; van Winden et al., 2012) and this was applied in **Chapters 5 and 6**. By analyzing underivatized bacteriohopanepolyol (BHP) derivatives using UHPLC-HRMS^2 numerous common and novel composite BHPs (e.g. ethenolamineBHPs or aminoBHPs acylated to fatty acid chains) were detected in bacterial cultures as well as in natural environments. Application of this methodology in two terrestrial methane seeps transects (Censo and Bissana seep) showed high relative abundance of aminoBHPs and the novel ethenolamineBHPs at and close to the seeps, which are likely synthesized by Type I MOBs (e.g. *Methylococcaceae*) and potentially other Proteobacteria species.

The detection of a new late eluting aminotriol isomer, which was also identified in a verrucomicrobial strain, *Methylacidimicrobium cyclopophantes* 3B, indicates for the first time that it may be used as a new biomarker for methanotrophic Verrucomicrobia. Verrucomicrobia have only been shown to produce common fatty acids (Op den Camp et al., 2009) and less distinct BHPs such as aminotriol (van Winden et al., 2012), both of which are also produced by other MOBs and bacterial species and are thus not useful for the identification of Verrucomicrobia in complex environmental samples. Therefore, the identification of the aminotriol isomer offers a new way to scan for Verrucomicrobia in modern and past environments besides DNA-based tools.

Soil-marker BHPs (adenosylhopane and N1-methylinosylhopane) showed contrasting behavior to N-containing BHPs (aminoBHPs and ethenolamineBHPs), implying general bacterial producer(s) such as Acidobacteria or Actinobacteria. Further analysis on the hopanoid synthesis (e.g. *shc* gene) potential (Pearson and Rusch, 2009; Sinninghe Damsté et al., 2017) in the detected bacterial reads is required to shed light on the potential of certain bacteria to synthesize bacteriohopanepolyols.

Based on the above work, the AminoBHP-index was developed as an index to trace aerobic oxidation of methane, showing high values close to the

seeps (≥ 0.4) and dropping to < 0.2 at distances >3 m from the active seeps. Future work should investigate the potential to apply this new AminoBHP-index to other terrestrial methane-rich environments such as other seeps, peatlands or stratified lakes where aerobic methane oxidation is thought to be present (e.g. Coolen et al., 2008; Talbot et al., 2016; Inglis et al., 2019). Furthermore, the long-term preservation potential of aminoBHPs and the soil-specific BHPs (adenosylhopane and N1-methylinosylhopane) should be investigated by diagenetic alteration experiments using for example hydrous pyrolysis. Once further established, this AminoBHP-index could help to reconstruct past aerobic oxidation of methane and thus past methane concentrations over longer timescales than the currently existing ice-core records, which only date back 800,000 years (Loulergue et al., 2008).

Unique mycocerosic acids trace novel hydrocarbon gas-oxidizing soil mycobacteria

Soils from a natural everlasting fire named Fuoco di Censo in Sicily, Italy showed high abundances of novel mycobacteria via 16S rRNA gene sequencing and mycocerosic acids (MAs), multi-methyl branched fatty acids, at and close to the seep (**Chapters 7 and 8)**. These novel soil mycobacteria are phylogenetically closely related to the *Mycobacterium simiae* complex but more remotely to well-studied pathogenic mycobacteria such as *M. tuberculosis*. A range of new MAs were identified and even very unusual ones with mid-chain methyl branches at positions C-12 and C-16 were present at and close to the seep. These MAs were naturally depleted in ^{13}C, suggesting a direct or indirect utilization of the released gases (methane and/or ethane). Confirmation of the use of these hydrocarbon gasses as a carbon source came from SIP incubations using ^{13}C labeled methane and ethane, which showed ^{13}C label incorporation of ethane, but not of methane, into the MAs. These results imply that mycobacteria at the Censo seep oxidize ethane but not methane, in agreement with studies from the 1950s in areas of petroleum production and gas seeps (Davis et al., 1956, 1959; Dworkin and Foster, 1958) as well as genomic studies (Coleman et al., 2012; Martin et al., 2014). Thus, soil mycobacteria might represent a major group of microbes involved in the cycling of ethane and maybe also higher gaseous hydrocarbons in terrestrial environments and possibly in marine environments, as they are present in a wide range of environmental niches (Brennan, 2003; Falkinham, 2015).

The novel structurally specific MAs in tandem with ^{13}C depleted isotopic signatures offer unique biomarkers to trace higher gaseous hydrocarbon

oxidation and gas-consuming soil mycobacteria in modern and past environments. Other terrestrial and marine gas seeps should be investigated for the presence of these new biomarkers to confirm their presence in other settings where also sequences of mycobacteria have been detected. Interestingly, mycobacteria are able to synthesize a number of other highly specific lipids due to the fact that they possess a double lipid membrane and two fatty acid biosynthesis systems (Brennan, 2003; Gago et al., 2011; Donoghue et al., 2017; Daffé et al., 2019). These lipids include extremely long fatty acids with chains up to C_{90} and with numerous methyl groups, hydroxylations and/or methoxylations bound to complex glycolipids, all known from pathogenic mycobacteria (Minnikin et al., 1993a, b; Donoghue et al., 2017). Potentially these lipids classes can be picked up by using UHPLC-HRMS^2 techniques. Finally, whole genome analysis of the mycobacteria in the Censo soils could reveal the pathways for MA synthesis (Brennan, 2003; Gago et al., 2011), and their relation with those of pathogenic mycobacteria, as well as shed more light on their carbon and energy assimilation pathways.

Overall the work presented in this thesis has substantially expanded our lipid biomarker tool set to detect past methane and ethane oxidation by the development of proxies such as mycocerosic acids and the AminoBHP-index. Together with more established biomarker proxies, these can potentially be used to study methane and higher gaseous hydrocarbon oxidation in present and past environments and thereby lead to more insight into past cycling of greenhouse gases.

References

References

Abdala Asbun, A., Besseling, M.A., Balzano, S., van Bleijswijk, J.D.L., Witte, H.J., Villanueva, L. and Engelmann, J.C. (2020) Cascabel: A Scalable and Versatile Amplicon Sequence Data Analysis Pipeline Delivering Reproducible and Documented Results. Frontiers in Genetics 11.

Ahmed, M., Schouten, S., Baas, M. and De Leeuw, J. (2001) Bound lipids in kerogens from the Monterey Formation, Naples Beach, California. The Monterey Formation: From Rock to Molecules. Columbia University Press, New York, 189-205.

Alain, K., Holler, T., Musat, F., Elvert, M., Treude, T. and Krüger, M. (2006) Microbiological investigation of methane-and hydrocarbon-discharging mud volcanoes in the Carpathian Mountains, Romania. Environmental Microbiology 8, 574-590.

Albaiges, J. and Albrecht, P. (1979) Fingerprinting marine pollutant hydrocarbons by computerized gas chromatography-mass spectrometry. International Journal of Environmental Analytical Chemistry 6, 171-190.

Alloisio, N., Marechal, J., Vanden Heuvel, B., Normand, P. and Berry, A.M. (2005) Characterization of a gene locus containing squalene-hopene cyclase (she) in Frankia alni ACN14a, and an she homolog in Aeidothermus eellulolytieus. Symbiosis 39, 83–90.

Alugupalli, S., Sikka, M. K., Larsson, L., and White, D. C. (1998) Gas chromatography–mass spectrometry methods for the analysis of mycocerosic acids present in Mycobacterium tuberculosis, Journal of Microbiological Methods, 31, 143-150.

Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R. and Stahl, D.A. (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Applied Environmental Microbiology 56, 1919-1925.

Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Arshad, A., Speth, D.R., de Graaf, R.M., Op den Camp, H.J.M., Jetten, M.S.M. and Welte, C.U. (2015) A Metagenomics-Based Metabolic Model of Nitrate-Dependent Anaerobic Oxidation of Methane by *Methanoperedens*-Like Archaea. Frontiers in Microbiology 6.

Ashraf, W., Mihdhir, A. and Murrell, J. C. (1994) Bacterial oxidation of propane. FEMS Microbiology Letters 122, 1-6.

Atlas, R. and Bartha, R. (1973) Inhibition by fatty acids of the biodegradation of petroleum. Antonie van Leeuwenhoek 39, 257-271.

Bale, N., Villanueva, L., Hopmans, E., Schouten, S. and Sinninghe Damsté, J. (2013) Different seasonality of pelagic and benthic Thaumarchaeota in the North Sea. Biogeosciences 10, 7195-7206.

Bale, N.J., Sorokin, D.Y., Hopmans, E.C., Koenen, M., Rijpstra, W.I.C., Villanueva, L., Wienk, H. and Sinninghe Damsté, J.S. (2019) New insights into the polar lipid composition of extremely halo(alkali)philic euryarchaea from hypersaline lakes. Frontiers in Microbiology 10, 377.

Balesdent, J., Mariotti, A., and Guillet, B. (1987) Natural ^{13}C abundance as a tracer for studies of soil organic matter dynamics, Soil Biology and Biochemistry, 19, 25-30.

Bard, M., Bruner, D., Pierson, C., Lees, N., Biermann, B., Frye, L., Koegel, C. and Barbuch, R. (1996) Cloning and characterization of ERG25, the Saccharomyces cerevisiae gene encoding C-4 sterol methyl oxidase. Proceedings of the National Academy of Sciences 93, 186-190.

Bardgett, R. D., and van der Putten, W. H. (2014) Belowground biodiversity and ecosystem functioning, Nature 515, 505, 2014.

Basilone, L. (2012) Litostratigrafia della Sicilia, Dipartimento di scienze della terra e del mare, Università degli studi, Arti Grafiche Palermitane s.r.l, Palermo, Italy, ISBN: 978-88-97559-09-2.

Beal, E.J., House, C.H., and Orphan, V.J. (2009) Manganese- and iron-dependent marine methane oxidation. Science 325, 184–187.

Bennett, B. and Abbott, G.D. (1999) A natural pyrolysis experiment-hopanes from hopanoic acids? Organic Geochemistry 30, 1509–1516.

Berg, I.A., Kockelkorn, D., Vera, W.H.R., Say, R.F., Zarzycki, J. and Hügler, M. (2010) Autotrophic carbon fixation in archaea. Nature Reviews Microbiology 8, 447–460.

Berger, S., Frank, J., Dalcin Martins, P., Jetten, M.S.M., and Welte, C.U. (2017) High-quality draft genome sequence of "*Candidatus* Methanoperedens sp." Strain BLZ2, a nitrate-reducing anaerobic methane-oxidizing archaeon enriched in an anoxic bioreactor. genome Announcements 5: e01159-17.

Bernard, B.B., Brooks, J.M. and Sackett, W.M. (1978) Light hydrocarbons in recent Texas continental shelf and slope sediments. Journal of Geophysical Research: Oceans 83, 4053-4061.

Berndmeyer, C., Thiel, V., Schmale, O. and Blumenberg, M. (2013) Biomarkers for aerobic methanotrophy in the water column of the stratified Gotland Deep (Baltic Sea). Organic Geochemistry 55, 103-111.

Besseling, M.A., Hopmans, E. C., Boschman, C.R., Sinninghe Damsté, J.S. and Villanueva, L. (2018) Benthic archaea as potential sources of tetraether membrane lipids in sediments across an oxygen minimum zone. Biogeosciences 15, 4047–4064.

Birgel, D. and Peckmann, J. (2008) Aerobic methanotrophy at ancient marine methane seeps: a synthesis. Organic Geochemistry 39, 1659-1667.

Bligh, E.G. and Dyer, W.J. (1959) A rapid method of total lipid extraction and purification. Canadian Journal Biochemistry Physiology 37, 911-917.

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

Blumenberg, M., Seifert, R., Nauhaus, K., and Pape, T. (2005) In vitro study of lipid biosynthesis in an anaerobically methane-oxidizing microbial mat. Applied Environmental Microbiology 71, 4345–4351.

Blumenberg, M., Krüger, M., Nauhaus, K., Talbot, H.M., Oppermann, B.I., Seifert, R., Pape, T. and Michaelis, W. (2006) Biosynthesis of hopanoids by sulfate-reducing bacteria (genus *Desulfovibrio*). Environmental Microbiology 8, 1220-1227.

Blumenberg, M., Seifert, R. and Michaelis, W. (2007) Aerobic methanothrophy in the oxic-anoxic transition zone of the Black Sea water column. Organic Geochemistry 38, 84-91.

Blumenberg, M., Hoppert, M., Krüger, M., Dreier, A. and Thiel, V. (2012) Novel findings on hopanoid occurrences among sulfate reducing bacteria: Is there a direct link to nitrogen fixation? Organic geochemistry 49, 1-5.

Bodelier, P.L., Gillisen, M.-J.B., Hordijk, K., Sinninghe Damsté, J.S., Rijpstra, W.I.C., Geenevasen, J.A. and Dunfield, P.F. (2009) A reanalysis of phospholipid fatty acids as ecological biomarkers for methanotrophic bacteria. The ISME Journal 3, 606.

Boere, A.C., Rijpstra, W.I.C., De Lange, G.J., Sinninghe Damsté, J.S. and Coolen, M.J.L. (2011) Preservation potential of ancient plankton DNA in Pleistocene marine sediments. Geobiology 9, 377-393.

Boetius, A., Ravenschlag, K., Schubert, C.J., Rickert, D., Widdel, F., Gieseke, A., Amann, R., Jørgensen, B.B., Witte, U. and Pfannkuche, O. (2000) A marine

microbial consortium apparently mediating anaerobic oxidation of methane. Nature 407, 623.

Bonini, M. (2009) Mud volcano eruptions and earthquakes in the Northern Apennines and Sicily, Italy. Tectonophysics 474, 723–735.

Boschker, H., Nold, S., Wellsbury, P., Bos, D., De Graaf, W., Pel, R., Parkes, R.J. and Cappenberg, T. (1998) Direct linking of microbial populations to specific biogeochemical processes by 13 C-labelling of biomarkers. Nature 392, 801-805.

Botz, R., Schmidt, M., Wehner, H., Hufnagel, H. and Stoffers, P. (2007) Organic-rich sediments in brine-filled Shaban-and Kebrit deeps, northern Red Sea. Chemical Geology 244, 520-553.

Bouam, A., Armstrong, N., Levasseur, A., and Drancourt, M. (2018) Mycobacterium terramassiliense, Mycobacterium rhizamassiliense and Mycobacterium numidiamassiliense sp. nov., three new Mycobacterium simiae complex species cultured from plant roots. Scientific Reports, 8, 1-13.

Bowman, J.P. (2011) Approaches for the characterization and description of novel methanotrophic bacteria. Methods in Enzymology 459, 45-62.

Bowman, J.P. (2006) The methanotrophs - the families Methylococcaceae and Methylocystaceae, The Prokaryotes, 5, 266-289.

Bowman, J.P., Sly, L.I., Nichols, P.D. and Hayward, A. (1993) Revised taxonomy of the methanotrophs: description of *Methylobacter* gen. nov., emendation of *Methylococcus*, validation of *Methylosinus* and *Methylocystis* species, and a proposal that the family *Methylococcaceae* includes only the group I methanotrophs. International Journal of Systematic and Evolutionary Microbiology 43, 735-753.

Bradley, A.S., Pearson, A., Sáenz, J.P. and Marx, C.J. (2010) Adenosylhopane: The first intermediate in hopanoid side chain biosynthesis. Organic Geochemistry 41, 1075-1081.

Brassell, S.C., Eglinton, G., Marlowe, I.T., Pflaumann, U. and Sarnthein, M. (1986) Molecular stratigraphy: a new tool for climatic assessment. Nature 320, 129-133.

Bray, E. and Evans, E. (1961) Distribution of *n*-paraffins as a clue to recognition of source beds. Geochimica et Cosmochimica Acta 22, 2–15.

Brennan, P.J., and Nikaido, H. (1995) The envelope of mycobacteria, Annual Review of Biochemistry 64, 29-63.

Brennan, P.J. (2003) Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. Tuberculosis 83, 91-97.

Brocks, J.J., Love, G.D., Summons, R.E., Knoll, A.H., Logan, G.A. and Bowden, S.A. (2005) Biomarker evidence for green and purple sulphur bacteria in a stratified Palaeoproterozoic sea. Nature 437, 866.

Brocks, J.J. and Pearson, A. (2005) Building the biomarker tree of life. Reviews in Mineralogy and Geochemistry 59, 233-258.

Bühring, S., Elvert, M. and Witte, U. (2005) The microbial community structure of different permeable sandy sediments characterized by the investigation of bacterial fatty acids and fluorescence in situ hybridization. Environmental Microbiology 7, 281–293.

Bull, I.D., Parekh, N.R., Hall, G.H., Ineson, P. and Evershed, R.P. (2000) Detection and classification of atmospheric methane oxidizing bacteria in soil. Nature 405, 175.

Burugupalli, S., Almeida, C.F., Smith, D.G.M, Shah, S., Patel, O., Rossjohn, J., Uldrich, A.P., Godfrey, D.I. and S. J. Williams (2020) α-Glucuronosyl and α-glucosyl diacylglycerides, natural killer T cell-activating lipids from bacteria and fungi. Chemical Science 11, 2161-2168.

Cadillo-Quiroz, H., Yashiro, E., Yavitt, J.B., and Zinder, S.H. (2008) Characterization of the archaeal community in a minerotrophic fen and terminal restriction fragment length polymorphism-directed isolation of a novel hydrogenotrophic methanogen. Applied Environmental Microbiology 74, 2059–2068.

Cai, C., Leu, A.O., Xie, G.-J., Guo, J., Feng, Y. and Zhao, J.-X. (2018) A methanotrophic archaeon couples anaerobic oxidation of methane to Fe (III) reduction. The ISME Journal 12, 1929–1939.

Camacho-Ibar, V. c. F., Aveytua-Alcázar, L., and Carriquiry, J. D. (2003) Fatty acid reactivities in sediment cores from the northern Gulf of California, Organic Geochemistry 34, 425-439.

Campbell, I.M. and Naworal, J. (1969) Composition of the saturated and monounsaturated fatty acids of *Mycobacterium phlei*. Journal of lipid research 10, 593–598.

Cangemi, M. and Madonia, P. (2014) Mud volcanoes in onshore Sicily: a short overview. Spongy, slimy, cosy & more, 123.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L. and Bauer, M. (2012) Ultra-high-

throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME Journal 6, 1621.

Catalano, S., De Guidi, G., Romagnoli, G., Torrisi, S., Tortorici, G. and Tortorici, L. (2008) The migration of plate boundaries in SE Sicily: Influence on the large-scale kinematic model of the African promontory in southern Italy. Tectonophysics 449, 41–62.

Chang, Y.H., Cheng, T.W., Lai, W.J., Tsai, W.Y., Sun, C.H., Lin, L.H. and Wang, P.L. (2012) Microbial methane cycling in a terrestrial mud volcano in eastern Taiwan. Environmental Microbiology 14, 895–908.

Chen, Y., Dumont, M.G., McNamara, N.P., Chamberlain, P.M., Bodrossy, L., Stralis-Pavese, N. and Murrell, J.C. (2008) Diversity of the active methanotrophic community in acidic peatlands as assessed by mRNA and SIP-PLFA analyses. Environmental Microbiology 10, 446-459.

Cheng, T.-W., Chang, Y.-H., Tang, S.-L., Tseng, C.-H., Chiang, P.-W., Chang, K.-T., Sun, C.-H., Chen, Y.-G., Kuo, H.-C. and Wang, C.-H. (2012) Metabolic stratification driven by surface and subsurface interactions in a terrestrial mud volcano. The ISME Journal 6, 2280–2290.

Chou, S., Chedore, P., Haddad, A., Paul, N. and Kasatiya, S. (1996) Direct identification of Mycobacterium species in Bactec 7H12B medium by gas-liquid chromatography. Journal of Clinical Microbiology 34, 1317-1320.

Chou, S., Chedore, P., and Kasatiya, S. (1998) Use of gas chromatographic fatty acid and mycolic acid cleavage product determination to differentiate among Mycobacterium genavense, Mycobacterium fortuitum, Mycobacterium simiae, and Mycobacterium tuberculosis. Journal of Clinical Microbiology 36, 577-579.

Christie, W.W. (1998) Gas chromatography-mass spectrometry methods for structural analysis of fatty acids. Lipids 33, 343-353.

Coleman, N.V., Yau, S., Wilson, N.L., Nolan, L.M., Migocki, M.D., Ly, M.a., Crossett, B. and Holmes, A.J. (2011) Untangling the multiple monooxygenases of Mycobacterium chubuense strain NBB4, a versatile hydrocarbon degrader. Environmental Microbiology Reports 3, 297-307.

Coleman, N.V., Le, N.B., Ly, M.A., Ogawa, H.E., McCarl, V., Wilson, N.L. and Holmes, A.J. (2012) Hydrocarbon monooxygenase in Mycobacterium: recombinant expression of a member of the ammonia monooxygenase superfamily. The ISME Journal 6, 171.

Collister, J.W., Summons, R.E., Lichtfouse, E. and Hayes, J.M. (1992) An isotopic biogeochemical study of the Green River oil shale. Organic Geochemistry 19, 265-276.

Connan, J., Bouroullec, J., Dessort, D. and Albrecht, P. (1986) The microbial input in carbonate-anhydrite facies of a sabkha palaeoenvironment from Guatemala: a molecular approach. Organic Geochemistry 10, 29–50.

Conrad, R. (2009) The global methane cycle: Recent advances in understanding the microbial processes involved. Environmental Microbiology Reports 1, 285–292.

Conte, M., Dickey, T., Weber, J., Johnson, R. and Knap, A. (2003) Transient physical forcing of pulsed export of bioreactive material to the deep Sargasso Sea. Deep Sea Research Part I: Oceanographic Research Papers 50, 1157-1187.

Conte, M.H., Weber, J. and Ralph, N. (1998) Episodic particle flux in the deep Sargasso Sea: an organic geochemical assessment. Deep Sea Research Part I: Oceanographic Research Papers 45, 1819-1841.

Cooke, M.P., Talbot, H.M. and Wagner, T. (2008a) Tracking soil organic carbon transport to continental margin sediments using soil-specific hopanoid biomarkers: A case study from the Congo fan (ODP site 1075). Organic Geochemistry 39, 965–971.

Cooke, M.P, Talbot, H.M. and Farrimond, P. (2008b) Bacterial populations recorded in bacteriohopanepolyol distributions in soils from Northern England. Organic Geochemistry 39, 1347–1358.

Cooke, M.P., van Dongen, B.E., Talbot, H.M., Semiletov, I., Shakhova, N., Guo, L. and Gustafsson, O. (2009) Bacteriohopanepolyol biomarker composition of organic matter exported to the Arctic Ocean by seven of the major Arctic rivers. Organic Geochemistry 40, 1151–1159.

Coolen, M.J.L., Talbot, H.M., Abbas, B.A., Ward, C., Schouten, S., Volkman, J.K. and Sinninghe Damsté, J.S. (2008) Sources for sedimentary bacteriohopanepolyols as revealed by 16S rDNA stratigraphy. Environmental Microbiology 10, 1783-1803.

Costantino, V., Fattorusso, E., Imperatore, C. and Mangoni, A. (2000) The first 12-methylhopanetetrol from the marine sponge *Plaktoris simplex*. Tetrahedron 56, 3781-3784.

Crossman, Z., Ineson, P. and Evershed, R. (2005) The use of 13 C labelling of bacterial lipids in the characterisation of ambient methane-oxidising bacteria in soils. Organic Geochemistry 36, 769-778.

Cvejic, J.H., Bodrossy, L., Kovács, K.L. and Rohmer, M. (2000a) Bacterial triterpenoids of the hopane series from the methanotrophic bacteria *Methylocaldum* spp.: phylogenetic implications and first evidence for an unsaturated aminobacteriohopanepolyol. FEMS Microbiology Letters 182, 361-365.

Cvejic, J.H., Putra, S.R., El-Beltagy, A., Hattori, R., Hattori, T. and Rohmer, M. (2000b) Bacterial triterpenoids of the hopane series as biomarkers for the chemotaxonomy of *Burkholderia*, *Pseudomonas* and *Ralstonia spp*.. FEMS Microbiology Letters 183, 295-299.

Daffé, M., and Laneelle, M. (1988) Distribution of phthiocerol diester, phenolic mycosides and related compounds in mycobacteria. Microbiology 134, 2049-2055.

Daffé, M., Quémard, A. and Marrakchi, H. (2019) Mycolic Acids: From Chemistry to Biology. Biogenesis of Fatty Acids, Lipids and Membranes, 181-216.

Daims, H., Brühl, A., Amann, R., Schleifer, K.-H. and Wagner, M. (1999) The Domain-specific Probe EUB338 is Insufficient for the Detection of all Bacteria: Development and Evaluation of a more Comprehensive Probe Set. Systematic and Applied Microbiology 22, 434-444.

Dalton, H. (1980) Oxidation of hydrocarbons by methane monooxygenases from a variety of microbes. In, *Advances in Applied Microbiology*. Elsevier, pp. 71–87.

D'Andréa, S., Canonge, M., Beopoulos, A., Jolivet, P., Hartmann, M., Miquel, M., Lepiniec, L. and Chardot, T. (2007) At5g50600 encodes a member of the short-chain dehydrogenase reductase superfamily with 11β-and 17β-hydroxysteroid dehydrogenase activities associated with Arabidopsis thaliana seed oil bodies. Biochimie 89, 222-229.

Davis, J., Chase, H. and Raymond, R. (1956) *Mycobacterium paraffinicum* n. sp., a bacterium isolated from soil. Applied Microbiology 4, 310-315.

Davis, J., Raymond, R. and Stanley, J. (1959) Areal contrasts in the abundance of hydrocarbon oxidizing microbes in soils. Applied Microbiology 7, 156-165.

Dean, J.F., Middelburg, J.J., Röckmann, T., Aerts, R., Blauw, L.G., Egger, M., Jetten, M.S.M., de Jong, A.E.E., Meisel, O.H., Rasigraf, O., Slomp, C.P., in't Zandt, M.H. and Dolman, A.J. (2018) Methane Feedbacks to the Global Climate System in a Warmer World. Reviews of Geophysics 56, 207-250.

Dedysh, S.N., Belova, S.E., Bodelier, P.L., Smirnova, K.V., Khmelenina, V.N., Chidthaisong, A., Trotsenko, Y.A., Liesack, W. and Dunfield, P.F. (2007)

Methylocystis heyeri sp. nov., a novel type II methanotrophic bacterium possessing 'signature'fatty acids of type I methanotrophs. International Journal of Systematic and Evolutionary Microbiology 57, 472-479.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-González, A., Eldridge, D.J., Bardgett, R.D., Maestre, F.T., Singh, B.K. and Fierer, N. (2018) A global atlas of the dominant bacteria found in soil. Science 359, 320-325.

Delmotte, M., Chappellaz, J., Brook, E., Yiou, P., Barnola, J.M., Goujon, C., Raynaud, D. and Lipenkov, V.I. (2004) Atmospheric methane during the last four glacial-interglacial cycles: Rapid changes and their link with Antarctic temperature. Journal of Geophysical Research: Atmospheres 109.

Deppenmeier, U., Müller, V. and Gottschalk, G. (1996) Pathways of energy conservation in methanogenic archaea. Archives of Microbiology 165, 149-163.

De Rosa, M. and Gambacorta, A. (1988) The lipids of archaebacteria. Progress in Lipid Research 27, 153–175.

De Rosa, M., Gambacorta, A., and Nicolaus, B. (1980) Regularity of isoprenoid biosynthesis in the ether lipids of archaebacteria. Phytochemistry 19, 791–793.

Didyk, B., Simoneit, B., Brassell, S.T. and Eglinton, G. (1978) Organic geochemical indicators of palaeoenvironmental conditions of sedimentation. Nature 272, 216–222.

Dimitrov, L.I. (2002) Mud volcanoes—the most important pathway for degassing deeply buried sediments. Earth-Science Reviews 59, 49–76.

Dimitrov, L.I. (2003) Mud volcanoes—a significant source of atmospheric methane. Geo-Marine Letters 23, 155–161.

Dlugokencky, E.J., Bruhwiler, L., White, J.W.C., Emmons, L.K., Novelli, P.C., Montzka, S.A., Masarie, K.A., Lang, P.M., Crotwell, A.M., Miller, J.B. and Gatti, L.V. (2009) Observational constraints on recent increases in the atmospheric CH_4 burden. Geophysical Research Letters 36.

Donoghue, H., Taylor, G., Stewart, G., Lee, O., Wu, H., Besra, G. and Minnikin, D. (2017) Positive diagnosis of ancient leprosy and tuberculosis using ancient DNA and lipid biomarkers. Diversity 9, 46.

Dowling, N.J., Widdel, F. and White, D.C. (1986) Phospholipid ester-linked fatty acid biomarkers of acetate-oxidizing sulphate-reducers and other sulphide-forming bacteria. Microbiology 132, 1815–1825.

Dworkin, M. and Foster, J. (1958) Experiments with some microorganisms which utilize ethane and hydrogen. Journal of Bacteriology 75, 592.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32, 1792-1797.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26, 2460-2461.

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

Eickhoff, M., Birgel, D., Talbot, H.M., Peckmann, J. and Kappler, A. (2013) Bacteriohopanoid inventory of Geobacter sulfurreducens and Geobacter metallireducens. Organic Geochemistry 58, 107-114.

Ekiel, I. and Sprott, G.D. (1992) Identification of degradation artifacts formed upon treatment of hydroxydiether lipids from methanogens with methanolic HCl. Canadian Journal of Microbiology 38, 764–768.

Elling, F.J., Könneke, M., Nicol, G.W., Stieglmeier, M., Bayer, B., Spieck, E., de la Torre, J.R., Becker, K.W., Thomm, M., Prosser, J.I., Herndl, G.J., Schleper, C. and Hinrichs, K.-U. (2017) Chemotaxonomic characterisation of the thaumarchaeal lipidome. Environmental Microbiology 19, 2681–2700.

Elvert, M., Suess, E., and Whiticar, M.J. (1999) Anaerobic methane oxidation associated with marine gas hydrates: superlight C-isotopes from saturated and unsaturated C_{20} and C_{25} irregular isoprenoids. Naturwissenschaften 86, 295–300.

Elvert, M., Boetius, A., Knittel, K. and Jørgensen, B.B. (2003) Characterization of specific membrane fatty acids as chemotaxonomic markers for sulfate-reducing bacteria involved in anaerobic oxidation of methane. Geomicrobiology Journal 20, 403-419.

Elvert, M., Hopmans, E., Treude, T., Boetius, A., and Suess, E. (2005) Spatial variations of methanotrophic consortia at cold methane seeps: implications from a high-resolution molecular and. Geobiology 3, 195–209.

Elvert, M. and Niemann, H. (2008) Occurrence of unusual steroids and hopanoids derived from aerobic methanotrophs at an active marine mud volcano. Organic Geochemistry 39, 167-177.

Etheridge, D.M., Steele, L.P., Francey, R.J. and Langenfelds, R.L. (1998) Atmospheric methane between 1000 A.D. and present: Evidence of anthropogenic emissions and climatic variability. Journal of Geophysical Research: Atmospheres 103, 15979-15993.

Etiope, G., Caracausi, A., Favara, R., Italiano, F. and Baciu, C. (2002) Methane emission from the mud volcanoes of Sicily (Italy). Geophysical Research Letters 29, 56-51-56-54.

Etiope, G., Martinelli, G., Caracausi, A. and Italiano, F. (2007) Methane seeps and mud volcanoes in Italy: gas origin, fractionation and emission to the atmosphere. Geophysical Research Letters 34.

Etiope, G., Feyzullayev, A. and Baciu, C.L. (2009) Terrestrial methane seeps and mud volcanoes: a global perspective of gas origin. Marine and Petroleum Geology 26, 333-344.

Etiope, G. and Ciccioli, P. (2009) Earth's Degassing: A Missing Ethane and Propane Source. Science 323, 478-478.

Ettwig, K.F., Zhu, B., Speth, D., Keltjens, J.T., Jetten, M.S.M., and Kartal, B. (2016) Archaea catalyze iron-dependent anaerobic oxidation of methane. Proceedings of the National Academy of Sciences 113, 127922–12796.

Ettwig, K.F., Butler, M.K., Le Paslier, D., Pelletier, E., Mangenot, S., Kuypers, M.M., Schreiber, F., Dutilh, B.E., Zedelius, J. and De Beer, D. (2010) Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature 464, 543-548.

Ettwig, K.F., Van Alen, T., van de Pas-Schoonen, K.T., Jetten, M.S. and Strous, M. (2009) Enrichment and molecular detection of denitrifying methanotrophic bacteria of the NC10 phylum. Applied and Environmental Microbiology 75, 3656-3662.

Ettwig, K.F., Shima, S., Van De Pas-Schoonen, K.T., Kahnt, J., Medema, M.H., Op Den Camp, H.J., Jetten, M.S. and Strous, M. (2008) Denitrifying bacteria anaerobically oxidize methane in the absence of Archaea. Environmental Microbiology 10, 3164-3173.

Evans, T.W., Coffinet, S., Könneke, M., Lipp, J.S., Becker, K.W., Elvert, M., Heuer, V. and Hinrichs, K.-U. (2019) Assessing the carbon assimilation and production of benthic archaeal lipid biomarkers using lipid-RIP. Geochimica et Cosmochimica Acta 265, 431-442.

Evershed, R.P., Crossman, Z.M., Bull, I.D., Mottram, H., Dungait, J.A.J., Maxfield, P.J. and Brennand, E.L. (2006) 13C-Labelling of lipids to investigate microbial communities in the environment. Current Opinion in Biotechnology 17, 72-82.

Falkinham, J.O. (2009) The biology of environmental mycobacteria. Environmental Microbiology Reports 1, 477-487.

Falkinham, J.O. (2015) Environmental sources of nontuberculous mycobacteria. Clinics in Chest Medicine 36, 35-41.

Farrimond, P., Griffiths, T. and Evdokiadis, E. (2002) Hopanoic acids in Mesozoic sedimentary rocks: their origin and relationship with hopanes. Organic Geochemistry 33, 965–977.

Farrimond, P., Talbot, H., Watson, D., Schulz, L. and Wilhelms, A. (2004) Methylhopanoids: molecular indicators of ancient bacteria and a petroleum correlation tool. Geochimica et Cosmochimica Acta 68, 3873-3882.

Farwell, C., Reddy, C.M., Peacock, E., Nelson, R.K., Washburn, L. and Valentine, D.L. (2009) Weathering and the fallout plume of heavy oil from strong petroleum seeps near Coal Oil Point, CA. Environmental Science & Technology 43, 3542–3548.

Fernandes, N.D., and Kolattukudy, P.E. (1997) Methylmalonyl coenzyme A selectivity of cloned and expressed acyltransferase and beta-ketoacyl synthase domains of mycocerosic acid synthase from Mycobacterium bovis BCG. Journal of Bacteriology 179, 7538-7543.

Ferretti, D.F., Miller, J.B., White, J.W.C., Etheridge, D.M., Lassey, K.R., Lowe, D.C., Meure, C.M.M., Dreier, M.F., Trudinger, C.M., van Ommen, T.D. and Langenfelds, R.L. (2005) Unexpected Changes to the Global Methane Budget over the Past 2000 Years. Science 309, 1714-1717.

Ferry, J.G. (1999) Enzymology of one-carbon metabolism in methanogenic pathways. FEMS Microbiology Reviews 23, 13-38.

Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H. and Caporaso, J.G. (2012) Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proceedings of the National Academy of Sciences 109, 21390-21395.

Fontaine, T., Lamarre, C., Simenel, C., Lambou, K., Coddeville, B., Delepierre, M. and Latgé, J-P (2009) Characterization of glucuronic acid containing glycolipid in *Aspergillus fumigatus* mycelium. Carbohydrate research 344, 1960-1967.

Francis, G. W. (1981) Alkylthiolation for the determination of double-bond position in unsaturated fatty acid esters. Chemistry and Physics of Lipids 29, 369-374.

Franzmann, P.D., Stackebrandt, E., Sanderson, K., Volkman, J.K., Cameron, D.E. and Stevenson, P.L. (1988) *Halobacterium lacusprofundi* sp. nov., a halophilic bacterium isolated from Deep Lake, Antarctica. Systematic and Applied Microbiology 11, 20–27.

Freeman, K.H., Hayes, J., Trendel, J.-M. and Albrecht, P. (1990) Evidence from carbon isotope measurements for diverse origins of sedimentary hydrocarbons. Nature 343, 254-256.

Frieling, J., Svensen, H.H., Planke, S., Cramwinckel, M.J., Selnes, H. and Sluijs, A. (2016) Thermogenic methane release as a cause for the long duration of the PETM. Proceedings of the National Academy of Sciences 113, 12059-12064.

Frostegård, Å., Tunlid, A., and Bååth, E. (2011) Use and misuse of PLFA measurements in soils. Soil Biology and Biochemistry, 43, 1621-1625.

Gachotte, D., Barbuch, R., Gaylor, J., Nickel, E. and Bard, M. (1998) Characterization of the Saccharomyces cerevisiae ERG26 gene encoding the C-3 sterol dehydrogenase (C-4 decarboxylase) involved in sterol biosynthesis. Proceedings of the National Academy of Sciences 95, 13794-13799.

Gachotte, D., Sen, S., Eckstein, J., Barbuch, R., Krieger, M., Ray, B. and Bard, M. (1999) Characterization of the Saccharomyces cerevisiae ERG27 gene encoding the 3-keto reductase involved in C-4 sterol demethylation. Proceedings of the National Academy of Sciences 96, 12655-12660.

Gago, G., Diacovich, L., Arabolaza, A., Tsai, S.-C. and Gramajo, H. (2011) Fatty acid biosynthesis in actinomycetes. FEMS Microbiology Reviews 35, 475-497.

Gibson, R.A., Talbot, H.M., Kaur, G., Pancost, R.D. and Mountain, B. (2008) Bacteriohopanepolyol signatures of cyanobacterial and methanotrophic bacterial populations recorded in a geothermal vent sinter. Organic Geochemistry 39, 1020-1023.

Gibson, J.A.E., Miller, M.R., Davies, N.W., Neill, G.P., Nichols, D.S., and Volkman, J.K. (2005) Unsaturated diether lipids in the psychrotrophic archaeon *Halorubrum lacusprofundi*. Systematic and Applied Microbiology 28, 19–26.

Granath, J. W., and P. Casero (2004) Tectonic setting of the petroleum systems of Sicily. In: R. Swennen, F. Roure, and J. W. Granath, eds., *Deformation, fluid flow, and reservoir appraisal in foreland fold and thrust belts*: AAPG Hedberg Series 1, 391–411.

Grassa, F., Capasso, G., Favara, R., Inguaggiato, S., Faber, E. and Valenza, M. (2004) Molecular and isotopic composition of free hydrocarbon gases from Sicily, Italy. Geophysical Research Letters 31, 6.

Grosjean, H. and Constantinesco, F. (1996) Enzymatic conversion of adenosine to inosine and to N1-methylinosine in transfer RNAs: A review. Biochimie 78, 488-501.

Guerrero-Cruz, S., Cremers, G., van Alen, T.A., Op den Camp, H.J.M., Jetten, M.S.M., Rasigraf, O., and Vaksmaa, A. (2018) Response of the anaerobic methanotroph "*Candidatus* Methanoperedens nitroreducens" to oxygen stress. Applied Environmental Microbiology 84, 1–17.

Hafenbradl, D., Keller, M., Thiericke, R., and Stetter, K.O. (1993) A novel unsaturated archaeal ether core lipid from the hyperthermophile *Methanopyrus kandleri*. Systematic and Applied Microbiology 16, 165–169.

Hallam, S.J., Putnam, N., Preston, C.M., Detter, J.C., Rokhsar, D., Richardson, P.M. and DeLong, E.F. (2004) Reverse Methanogenesis: Testing the Hypothesis with Environmental Genomics. Science 305, 1457-1462.

Hallam, S.J., Girguis, P.R., Preston, C.M., Richardson, P.M. and DeLong, E.F. (2003) Identification of methyl coenzyme M reductase A (mcrA) genes associated with methane-oxidizing archaea. Applied and Environmental Microbiology 69, 5483-5491.

Hamamura, N., Page, C., Long, T., Semprini, L. and Arp, D.J. (1997) Chloroform cometabolism by butane-grown CF8, *Pseudomonas butanovora*, and *Mycobacterium vaccae* JOB5 and methane-grown *Methylosinus trichosporium* OB3b. Applied and Environmental Microbiology 63, 3607-3613.

Hamieh, A., Tayyar, R., Tabaja, H., EL Zein, S., Bou Khalil, P., Kara, N., Kanafani, Z. A., Kanj, N., Bou Akl, I., and Araj, G. (2018) Emergence of Mycobacterium simiae: A retrospective study from a tertiary care center in Lebanon. PloS One 13, e0195390.

Hanson, R.S. and Hanson, T.E. (1996) Methanotrophic bacteria. Microbiological Reviews 60, 439-471.

Haroon, M.F., Hu, S., Shi, Y., Imelfort, M., Keller, J., Hugenholtz, P., Yuan, Z. and Tyson, G.W. (2013) Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. Nature 500, 567-570.

Harvey, D. J. (1998) Picolinyl esters for the structural determination of fatty acids by GC/MS, Molecular Biotechnology 10, 251-260.

Hayes, J.M. (1993) Factors controlling 13C contents of sedimentary organic compounds: principles and evidence. Marine Geology 113, 111-125.

Hayes, J.M. (2001) Fractionation of carbon and hydrogen isotopes in biosynthetic processes. Reviews in Mineralogy and Geochemistry 43, 225-277.

Hazeu, W. and de Bruyn, J.C. (1980) Ethane oxidation by methane-oxidizing bacteria. Antonie Van Leeuwenhoek 46, 443-455.

Heap, B. (1989) Mycobacterium simiae as a cause of intra-abdominal disease: a case report. Tubercle 70, 217-221.

Heller, C. (2011) Fluid venting structures of terrestrial mud volcanoes (Italy) and marine cold seeps (Black Sea) - Organo-geochemical and biological approaches. PhD-thesis, Georg-August-Universität Göttingen, Germany.

Heller, C., Blumenberg, M., Hoppert, M., Taviani, M. and Reitner, J. (2012) Terrestrial mud volcanoes of the Salse di Nirano (Italy) as a window into deeply buried organic-rich shales of Plio-Pleistocene age. Sedimentary Geology 263, 202-209.

Hemingway, J. D., Kusch, S., Walter, S. R. S., Polik, C. A., Elling, F. J. and Pearson, A. (2018) A novel method to measure the C-13 composition of intact bacteriohopanepolyols. Organic Geochemistry 123, 144–147.

Hennessee, C.T., Seo, J.-S., Alvarez, A.M. and Li, Q.X. (2009) Polycyclic aromatic hydrocarbon-degrading species isolated from Hawaiian soils: *Mycobacterium crocinum* sp. nov., *Mycobacterium pallens* sp. nov., *Mycobacterium rutilum* sp. nov., *Mycobacterium rufum* sp. nov. and *Mycobacterium aromaticivorans* sp. nov. International Journal of Systematic and Evolutionary Microbiology 59, 378-387.

Hinrichs, K.-U., Hayes, J.M., Sylva, S.P., Brewer, P.G. and DeLong, E.F. (1999) Methane-consuming archaebacteria in marine sediments. Nature 398, 802-805.

Hinrichs, K.-U., Pancost, R.D., Summons, R.E., Sprott, G.D., Sylva, S.P., Sinninghe Damsté, J.S., and Hayes, J.M. (2000) Mass spectra of sn-2-hydroxyarchaeol, a polar lipid biomarker for anaerobic methanotrophy. Geochemistry Geophysics Geosystems 1: 2000GC000042.

Hinrichs, K.-U. and Boetius, A. (2002) The anaerobic oxidation of methane: new insights in microbial ecology and biogeochemistry. In, Wefer, G., Billett, D., Hebbeln, D., Jørgensen, B., Schlüter, M., and van Weering, T. (eds), *Ocean margin systems*. Springer-Verlag, Berlin, pp. 457–477.

Hirayama, H., Sunamura, M., Takai, K., Nunoura. T., Noguchi, T., Oida, H., Furushima, Y., Yamamoto, H., Oomori, T. and Horikoshi, K. (2007) Culture-dependent and -independent characterization of microbial communities associated with a shallow submarine hydrothermal system occurring within a coral reef off Taketomi Island, Japan. Applied and Environmental Microbiology 73, 7642-7656.

Hirayama, H., Fuse, H., Abe, M., Miyazaki, M., Nakamura, T., Nunoura, T., Furushima, Y., Yamamoto, H. and Takai, K. (2013) *Methylomarinum vadi* gen.

nov., sp nov., a methanotroph isolated from two distinct marine environments. International Journal of Systematic and Evolutionary Microbioly 63, 1073-1082.

Hirayama, H., Abe, M., Miyazaki, M., Nunoura, T., Furushima, Y., Yamamoto, H. and Takai, K. (2014) *Methylomarinovum caldicuralii* gen. nov., sp. nov., a moderately thermophilic methanotroph isolated from a shallow submarine hydrothermal system, and proposal of the family *Methylothermaceae* fam. nov. International Journal of Systematic and Evolutionary Microbiology 64, 989-999.

Hoehler, T.M., Alperin, M.J., Albert, D.B., and Martens, C.S. (1994) Field and laboratory studies of methane oxidation in an anoxic marine sediment: Evidence for a methanogen-sulfate reducer consortium. Global Biogeochemical Cycles 8, 451–463.

Hoffmann, C., Mackenzie, A., Lewis, C., Maxwell, J., Oudin, J., Durand, B. and Vandenbroucke, M. (1984) A biological marker study of coals, shales and oils from the Mahakam Delta, Kalimantan, Indonesia. Chemical Geology 42, 1-23.

Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M. and Pääbo, S. (2001) Ancient DNA. Nature Reviews Genetics 2, 353-359.

Holmes, A.J., Roslev, P., McDonald, I.R., Iversen, N., Henriksen, K. and Murrell, J.C. (1999) Characterization of methanotrophic bacterial populations in soils showing atmospheric methane uptake. Applied and Environmental Microbiology 65, 3312-3318.

Holzheimer, M., Reijneveld, J. F., Ramnarine, A. K., Misiakos, G., Young, D. C., Ishikawa, E., Cheng, T.-Y., Yamasaki, S., Moody, D. B., Van Rhijn, I. and Minnaard A. J. (2020) Asymmetric Total Synthesis of Mycobacterial Diacyl Trehaloses Demonstrates a Role for Lipid Structure in Immunogenicity. ACS Chemical Biology 7, 1835-1841.

Hopmans, E.C., Schouten, S., Sinninghe Damsté, J.S., 2016. The effect of improved chromatography on GDGT-based palaeoproxies. Organic Geochemistry 93, 1–6.

Hopmans E.C., Smit N.T., Schwartz-Narbonne R., Sinninghe Damsté J.S. and Rush D. (2021) Analysis of non-derivatized bacteriohopanepolyols using UHPLC-HRMS reveals great structural diversity in environmental lipid assemblages. *Under review at Organic Geochemistry*.

Hu, S., Zeng, R.J., Burow, L.C., Lant, P., Keller, J., and Yuan, Z. (2009) Enrichment of denitrifying anaerobic methane oxidizing microorganisms. Environmental Microbiology Reports 1, 377–384.

Huang, Y., Bol, R., Harkness, D. D., Ineson, P., and 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. Organic Geochemistry 24, 273-287.

Inglis, G.N., Naafs, B.D.A., Zheng, Y., McClymont, E.L., Evershed, R.P. and Pancost, R.D., (2018) Distributions of geohopanoids in peat: Implications for the use of hopanoid-based proxies in natural archives. Geochimica et Cosmochimica Acta 224, 249–261.

Inglis, G.N., Naafs, B.D.A., Zheng, Y., Schellekens, J. and Pancost, R.D. (2019) δ13C values of bacterial hopanoids and leaf waxes as tracers for methanotrophy in peatlands. Geochimica et Cosmochimica Acta 260, 244-256.

Innes, H.E., Bishop, A.N., Head, I.M. and Farrimond, P. (1997) Preservation and diagenesis of hopanoids in recent lacustrine sediments of Priest Pot, England. Organic Geochemistry 26, 565-576.

Isaac, G., McDonald, S. and Astarita, G. (2011) Lipid separation using UPLC with charged surface hybrid technology. Waters Corporation, Milford, MA, USA, internal report.

Ivanov, M., Limonov, A. and Van Weering, T.C. (1996) Comparative characteristics of the Black Sea and Mediterranean Ridge mud volcanoes. Marine Geology 132, 253–271.

Jackson, M., Stadthagen, G., and Gicquel, B. (2007) Long-chain multiple methyl-branched fatty acid-containing lipids of Mycobacterium tuberculosis: biosynthesis, transport, regulation and biological activities, Tuberculosis 87, 78-86.

Jahnke, L.L., Summons, R.E., Hope, J.M. and Des Marais, D.J. (1999) Carbon isotopic fractionation in lipids from methanotrophic bacteria II: The effects of physiology and environmental parameters on the biosynthesis and isotopic signatures of biomarkers. Geochimica et Cosmochimica Acta 63, 79-93.

Judd, A.G. (2004) Natural seabed gas seeps as sources of atmospheric methane. Environmental Geology 46, 988-996.

Kannenberg, E.L. and Poralla, K. (1999) Hopanoid biosynthesis and function in bacteria. Naturwissenschaften 86, 168-176.

Kellermann, M.Y., Wegener, G., Elvert, M., Yoshinaga, M.Y., Lin, Y.-S., Holler, T., Mollar, X.P., Knittel, K. and Hinrichs, K.-U. (2012) Autotrophy as a predominant mode of carbon fixation in anaerobic methane-oxidizing microbial communities. Proceedings of the National Academy of Sciences 109, 19321-19326.

Kellermann, M.Y., Yoshinaga, M.Y., Wegener, G., Krukenberg, V., and Hinrichs, K.-U. (2016) Tracing the production and fate of individual archaeal

intact polar lipids using stable isotope probing. Organic Geochemistry 95, 13–20.

Khadem, A.F., Pol, A., Wieczorek, A., Mohammadi, S.S., Francoijs, K.-J., Stunnenberg, H.G., Jetten, M.S.M. and Op den Camp, H.J.M. (2011) Autotrophic Methanotrophy in Verrucomicrobia: *Methylacidiphilum fumariolicum*SolV Uses the Calvin-Benson-Bassham Cycle for Carbon Dioxide Fixation. Journal of Bacteriology 193, 4438-4446.

Kharbush, J.J., Ugalde, J.A., Hogle, S.L., Allen, E.E. and Aluwihare, L.I. (2013) Composite bacterial hopanoids and their microbial producers across oxygen gradients in the water column of the California Current. Applied and Environmental Microbiology 79, 7491-7501.

Killops, S.D. and Killops, V.J. (2013) Introduction to organic geochemistry. John Wiley & Sons.

Kinnaman, F.S., Valentine, D.L. and Tyler, S.C. (2007) Carbon and hydrogen isotope fractionation associated with the aerobic microbial oxidation of methane, ethane, propane and butane. Geochimica et Cosmochimica Acta 71, 271-283.

Kip, N., van Winden, J.F., Pan, Y., Bodrossy, L., Reichart, G.-J., Smolders, A.J.P., Jetten, M.S.M., Sinninghe Damsté, J.S. and Op den Camp, H.J.M. (2010) Global prevalence of methane oxidation by symbiotic bacteria in peat-moss ecosystems. Nature Geoscience 3, 617-621.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J.G., Dlugokencky, E.J., Bergamaschi, P., Bergmann, D., Blake, D.R. and Bruhwiler, L. (2013) Three decades of global methane sources and sinks. Nature Geoscience 6, 813-823.

Kits, K.D., Klotz, M.G. and Stein, L.Y. (2015) Methane oxidation coupled to nitrate reduction under hypoxia by the Gammaproteobacterium *Methylomonas denitrificans*, sp. nov. type strain FJG1. Environmental Microbiology 17, 3219-3232.

Knief, C. (2015) Diversity and Habitat Preferences of Cultivated and Uncultivated Aerobic Methanotrophic Bacteria Evaluated Based on pmoA as Molecular Marker. Frontiers in Microbiology 6.

Knittel, K. and Boetius, A. (2009) Anaerobic Oxidation of Methane: Progress with an Unknown Process. Annual Review of Microbiology 63, 311-334.

Koga, Y. and Morii, H. (2007) Biosynthesis of ether-type polar lipids in Archaea and evolutionary considerations. Microbiology and Molecular Biology Reviews 71, 97–120.

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

Koga, Y., Nishihara, M., Morii, H., and Akagawa-Matsushita, M. (1993) Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses. Microbiology Reviews 57, 164–182.

Kolaczkowska, E., Slougui, N.-E., Watt, D.S., Maruca, R.E. and Moldowan, J.M. (1990) Thermodynamic stability of various alkylated, dealkylated and rearranged 17α-and 17β-hopane isomers using molecular mechanics calculations. Organic Geochemistry 16, 1033–1038.

Kool, D.M., Talbot, H.M., Rush, D., Ettwig, K. and Sinninghe Damsté, J.S. (2014) Rare bacteriohopanepolyols as markers for an autotrophic, intra-aerobic methanotroph. Geochimica et Cosmochimica Acta 136, 114-125.

Kool, D.M., Zhu, B., Rijpstra, W.I.C., Jetten, M.S., Ettwig, K.F. and Sinninghe Damsté, J.S. (2012) Rare branched fatty acids characterize the lipid composition of the intra-aerobic methane oxidizer "*Candidatu*s Methylomirabilis oxyfera". Applied and Environmental Microbiology 78, 8650-8656.

Kopf, A., Klaeschen, D. and Mascle, J. (2001) Extreme efficiency of mud volcanism in dewatering accretionary prisms. Earth and Planetary Science Letters 189, 295–313.

Kopf, A. (2003) Global methane emission through mud volcanoes and its past and present impact on the Earth's climate. International Journal of Earth Sciences 92, 806–816.

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

Kuever, J., Könneke, M., Galushko, A. and Drzyzga, O. (2001) Reclassification of Desulfobacterium phenolicum as *Desulfobacula phenolica comb. nov*. and description of strain SaxT as *Desulfotignum balticum gen. nov., sp. nov*. International Journal of Systematic and Evolutionary Microbiology 51, 171–177.

Kusch, S., Walter, S., Hemingway, S.R. and Pearson, J.D.A. (2018) Improved chromatography reveals multiple new bacteriohopanepolyol isomers in marine sediments. Organic Geochemistry 124, 12–21.

Kusch, S., Sepúlveda, J. and Wakeham, S.G. (2019) Origin of Sedimentary BHPs Along a Mississippi River–Gulf of Mexico Export Transect: Insights From Spatial and Density Distributions. Frontiers in Marine Science 6.

Kweon, O., Kim, S.-J., Blom, J., Kim, S.-K., Kim, B.-S., Baek, D.-H., Park, S. I., Sutherland, J. B., and Cerniglia, C. E. (2015) Comparative functional pan-genome analyses to build connections between genomic dynamics and phenotypic evolution in polycyclic aromatic hydrocarbon metabolism in the genus Mycobacterium. BMC Evolutionary Biology 15, 21.

Lamb, D. C., Kelly, D. E., Manning, N. J., and Kelly, S. L. (1998) A sterol biosynthetic pathway in Mycobacterium. FEBS Letters 437, 142-144.

Lee, D.H., Kim, J.H., Lee, Y.M., Stadnitskaia, A., Jin, Y.K., Niemann, H., Kim, Y.G. and Shin, K.H. (2018) Biogeochemical evidence of anaerobic methane oxidation on active submarine mud volcanoes on the continental slope of the Canadian Beaufort Sea. Biogeosciences 15, 7419-7433.

Lee, A.K., Banta, A.B., Wei, J.H., Kiemle, D.J., Feng, J., Giner, J.-L. and Welander, P.V. (2018) C-4 sterol demethylation enzymes distinguish bacterial and eukaryotic sterol synthesis. Proceedings of the National Academy of Sciences 115, 5884-5889.

Lee, O.Y., Wu, H.H., Donoghue, H.D., Spigelman, M., Greenblatt, C.L., Bull, I.D., Rothschild, B.M., Martin, L.D., Minnikin, D.E., and Besra, G.S. (2012) Mycobacterium tuberculosis complex lipid virulence factors preserved in the 17,000-year-old skeleton of an extinct bison, Bison antiquus. PloS One 7, e41923.

Lee, S. and Poulter, C.D. (2008) Cloning, solubilization, and characterization of squalene synthase from Thermosynechococcus elongatus BP-1. Journal of bacteriology 190, 3808-3816.

Lévy-Frébault, V., Pangon, B., Buré, A., Katlama, C., Marche, C., and David, H. (1987) Mycobacterium simiae and Mycobacterium avium- M. intracellulare mixed infection in acquired immune deficiency syndrome. Journal of Clinical Microbiology 25, 154-157.

Lickorish, W.H., Grasso, M., Butler, R.W., Argnani, A. and Maniscalco, R. (1999) Structural styles and regional tectonic setting of the “Gela Nappe” and frontal part of the Maghrebian thrust belt in Sicily. Tectonics 18, 655–668.

Lipp, J.S., Morono, Y., Inagaki, F., and Hinrichs, K.-U. (2008) Significant contribution of Archaea to extant biomass in marine subsurface sediments. National Academy Science Letters 454, 991–994.

Liu, X.-L., Lipp, J.S., Simpson, J.H., Lin, Y.-S., Summons, R.E., and Hinrichs, K.-U. (2012) Mono- and dihydroxyl glycerol dibiphytanyl glycerol tetraethers in marine sediments: Identification of both core and intact polar lipid forms. Geochimica et Cosmochimica Acta 89, 102–115.

Liu, X., Leider, A., Gillespie, A., Gröger, J., Versteegh, G.J.M., and Hinrichs, K.-U. (2010) Organic Geochemistry Identification of polar lipid precursors of the ubiquitous branched GDGT orphan lipids in a peat bog in Northern Germany. Organic Geochemistry 41, 653–660.

Liu, Y. and Whitman, W.B. (2008) Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. Annals of the New York Academy of Sciences 1125, 171-189.

Llopiz, P., Neunlist, S. and Rohmer, M. (1992) Prokaryotic triterpenoids: O-α-D-glucuronopyranosyl bacteriohopanetetrol, a novel hopanoid from the bacterium Rhodospirillum rubrum. Biochemical Journal 287, 159-161.

López-Rodríguez, C., Stadnitskaia, A., De Lange, G.J., Martínez-Ruíz, F., Comas, M., Sinninghe Damsté, J.S. (2014) Origin of lipid biomarkers in mud volcanoes from the Alboran Sea, western Mediterranean. Biogeosciences 11, 3187–3204.

Lough, A. (1975) The chemistry and biochemistry of phytanic, pristanic and related acids. Progress in the Chemistry of Fats and other Lipids 14, 1-48.

Loulergue, L., Schilt, A., Spahni, R., Masson-Delmotte, V., Blunier, T., Lemieux, B., Barnola, J.-M., Raynaud, D., Stocker, T.F. and Chappellaz, J. (2008) Orbital and millennial-scale features of atmospheric CH_4 over the past 800,000 years. Nature 453, 383-386.

Lovley, D.R., Giovannoni, S.J., White, D.C., Champine, J.E., Phillips, E., Gorby, Y.A., Goodwin, S. (1993) *Geobacter metallireducens gen. nov. sp. nov.*, a microorganism capable of coupling the complete oxidation of organic compounds to the reduction of iron and other metals. Archives of Microbiology 159, 336–344.

MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., van Ommen, T., Smith, A. and Elkins, J. (2006) Law Dome CO_2, CH_4 and N_2O ice core records extended to 2000 years BP. Geophysical Research Letters 33.

Madigan, M.T., Martinko, J.M. and Parker, J. (1997) Brock biology of microorganisms. 8th Edition, Prentice Hall International, Inc., New York.

Madonia, P., Grassa, F., Cangemi, M. and Musumeci, C. (2011) Geomorphological and geochemical characterization of the 11 August 2008 mud volcano eruption at S. Barbara village (Sicily, Italy) and its possible relationship with seismic activity. Natural Hazards Earth System Sciences 11, 1545–1557.

Malott, R.J., Wu, C.-H., Lee, T.D., Hird, T.J., Dalleska, N.F., Zlosnik, J.E.A., Newman, D.K. and Speert, D.P. (2014) Fosmidomycin decreases membrane

hopanoids and potentiates the effects of colistin on *Burkholderia multivorans* clinical isolates. Antimicrobial Agents and Chemotherapy 58, 5211-5219.

Martin, K.E., Ozsvar, J. and Coleman, N.V. (2014) SmoXYB1C1Z of Mycobacterium sp. strain NBB4: a soluble methane monooxygenase (sMMO)-like enzyme, active on C2 to C4 alkanes and alkenes. Applied and Environmental Microbiology 80, 5801-5806.

Matsumi, R., Atomi, H., Driessen, A.J.M., and van der Oost, J. (2011) Isoprenoid biosynthesis in archaea - biochemical and evolutionary implications. Research in Microbiology 162, 39–52.

Mattavelli, L. and Novelli, L. (1990) Geochemistry and habitat of the oils in Italy. AAPG Bulletin 74, 1623–1639.

Maxfield, P.J., Hornibrook, E.R.C. and Evershed, R.P. (2006) Estimating High-Affinity Methanotrophic Bacterial Biomass, Growth, and Turnover in Soil by Phospholipid Fatty Acid ^{13}C Labeling. Applied and Environmental Microbiology 72, 3901-3907.

McAnulty, M.J., Poosarla, V.G., Kim, K.-Y., Jasso-Chávez, R., Logan, B.E., and Wood, T.K. (2017) Electricity from methane by reversing methanogenesis. Nature Communications 8, 15419.

McDonald, I.R., Bodrossy, L., Chen, Y. and Murrell, J.C. (2008) Molecular Ecology Techniques for the Study of Aerobic Methanotrophs. Applied and Environmental Microbiology 74, 1305-1315.

McKirdy, D., Aldridge, A. and Ypma, P. (1983) A geochemical comparison of some crude oils from pre-Ordovician carbonate rocks. Advances in Organic Geochemistry 1981. Wiley Chichester, 99–107.

McKirdy, D.M., Kantsler, A.J., Emmett, J.K. and Aldridge, A.K. (1984) Hydrocarbon genesis and organic facies in Cambrian carbonates of the Eastern Officer Basin, South Australia. AAPG Special Volumes 30, 13.

Mello, M.R., Gaglianone, P.C., Brassell, S.C. and Maxwell, J.R. (1988a) Geochemical and biological marker assessment of depositional environments using Brazilian offshore oils. Marine and Petroleum Geology 5, 205–223.

Mello, M.R., Telnaes, N., Gaglianone, P.C., Chicarelli, M.I., Brassell, S.C. and Maxwell, J.R. (1988b) Organic geochemical characterisation of depositional palaeoenvironments of source rocks and oils in Brazilian marginal basins, in: Mattavelli, L., Novelli, L. (Eds.), *Organic Geochemistry in Petroleum Exploration*. Pergamon, Amsterdam, pp. 31–45.

Menzel, P., Ng, K.L., and Krogh, A. (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nature Communications 7, 1–9.

Milkov A.V. (2005) Global Distribution of Mud Volcanoes and Their Significance in Petroleum Exploration as a Source of Methane in the Atmosphere and Hydrosphere and as a Geohazard. In: Martinelli G., Panahi B. (eds) *Mud Volcanoes, Geodynamics and Seismicity*. NATO Science Series (Series IV: Earth and Environmental Series), vol 51. Springer, Dordrecht.

Miller, C., Child, R., Hughes, J., Benscai, M., Der, J., Sims, R., and Anderson, A. (2007) Diversity of soil mycobacterium isolates from three sites that degrade polycyclic aromatic hydrocarbons. Journal of Applied Microbiology 102, 1612-1624.

Minnikin, D., Dobson, G., Goodfellow, M., Magnusson, M., and Ridell, M. (1985) Distribution of some mycobacterial waxes based on the phthiocerol family. Microbiology 131, 1375-1381.

Minnikin, D., Besra, G., Bolton, R., Datta, A., Mallet, A., Sharif, A., Stanford, J., Ridell, M. and Magnusson, M. (1993a) Identification of the leprosy bacillus and related mycobacteria by analysis of mycocerosate profiles, Annales-Societe Belge De Medecine Tropicale. Institute of Tropical Medicine, pp. 25-25.

Minnikin, D., Bolton, R., Hartmann, S., Besra, G., Jenkins, P., Mallet, A., Wilkins, E., Lawson, A. and Ridell, M. (1993b) An integrated procedure for the direct detection of characteristic lipids in tuberculous patients, Annales-Societe Belge De Medecine Tropicale. Institute of Tropical Medicine, pp. 13-13.

Minnikin, D. E., Kremer, L., Dover, L. G., and Besra, G. S. (2002) The methyl-branched fortifications of Mycobacterium tuberculosis. Chemistry & Biology 9, 545-553.

Moldowan, J.M. and McCaffrey, M.A. (1995) A novel microbial hydrocarbon degradation pathway revealed by hopane demethylation in a petroleum reservoir. Geochimica et Cosmochimica Acta 59, 1891-1894.

Moldowan, J.M., Sundararaman, P. and Schoell, M. (1986) Sensitivity of biomarker properties to depositional environment and/or source input in the Lower Toarcian of SW-Germany. Organic Geochemistry 10, 915–926.

Moldowan, J.M., Seifert, W.K., Arnold, E. and Clardy, J. (1984) Structure proof and significance of stereoisomeric 28, 30-bisnorhopanes in petroleum and petroleum source rocks. Geochimica et Cosmochimica Acta 48, 1651-1661.

Monaco, C. and Tortorici, L. (1996) Clay diapirs in Neogene-Quaternary sediments of central Sicily: evidence for accretionary processes. Journal of Structural Geology 18, 1265–1269.

Moore, E.K., Hopmans, E.C., Rijpstra, W.I.C., Villanueva, L. and Sinninghe Damsté, J.S. (2016) Elucidation and identification of amino acid containing

membrane lipids using liquid chromatography/high-resolution mass spectrometry. Rapid Communications in Mass Spectrometry 30, 739-750.

Murrell, J.C. and Dalton, H. (1983) Purification and properties of glutamine synthetase from *Methylococcus capsulatus* (Bath). Microbiology 129, 1187-1196.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., et al. (2014) Anthropogenic and Natural Radiative Forcing, in: Intergovernmental Panel on Climate, C. (Ed.), Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp. 659-740.

Nauhaus, K., Albrecht, M., Elvert, M., Boetius, A., and Widdel, F. (2007) In vitro cell growth of marine archaeal-bacterial consortia during anaerobic oxidation of methane with sulfate. Environmental Microbiology 9, 187–196.

NCBI Resource Coordinators (2016) Database resources of the National Center for Biotechnology Information. Nucleic Acids Research 44, 7–19.

Nechaeva, N. (1949) Two species of methane oxidizing mycobacteria. Mikrobiologiya 18, 310–317 (English Translation: Associated Technical Service, East Orange, NJ).

Nei, M., and Kumar, S. (2000) Molecular evolution and phylogenetics, Oxford University Press, ISBN: 0-19-513584-9.

Neunlist, S., Holst, O. and Rohmer, M. (1985) Prokaryotic triterpenoids: The hopanoids of the purple non-sulphur bacterium *Rhodomicrobium vannielii*: an aminotriol and its aminoacyl derivatives, N-tryptophanyl and N-ornithinyl aminotriol. European Journal of Biochemistry 147, 561-568.

Neunlist, S. and Rohmer, M. (1985) Novel hopanoids from the methylotrophic bacteria *Methylococcus capsulatus* and *Methylomonas methanica*. (22S)-35-aminobacteriohopane-30,31,32,33,34-pentol and (22S)-35-amino-3β-methylbacteriohopane-30,31,32,33,34-pentol. Biochemical Journal 231, 635-639.

Nichols, P.D. and Franzmann, P.D. (1992) Unsaturated diether phospholipids in the Antarctic methanogen *Methanococcoides burtonii*. FEMS Microbiology Letters 98, 205–208.

Nichols, P.D., Shaw, P.M., Mancuso, C.A., and Franzmann, P.D. (1993) Analysis of archaeal phospholipid-derived di- and tetraether lipids by high temperature capillary gas chromatography. Journal of Microbiological Methods 18, 1–9.

Nichols, P.D., Guckert, J.B., and White, D.C. (1986) Determination of monosaturated fatty acid double-bond position and geometry for microbial monocultures and complex consortia by capillary GC-MS of their dimethyl disulphide adducts. Journal of Microbiological Methods 5, 49-55.

Nicoara, S.C., Minnikin, D.E., Lee, O.C., O'Sullivan, D.M., McNerney, R., Pillinger, C.T., Wright, I.P., and Morgan, G.H. (2013) Development and optimization of a gas chromatography/mass spectrometry method for the analysis of thermochemolytic degradation products of phthiocerol dimycocerosate waxes found in Mycobacterium tuberculosis. Rapid Communications in Mass Spectrometry, 27, 2374-2382.

Niemann, H. and Elvert, M. (2008) Diagnostic lipid biomarker and stable carbon isotope signatures of microbial communities mediating the anaerobic oxidation of methane with sulphate. Organic Geochemistry 39, 1668-1677.

Niemann, H., Duarte, J., Hensen, C., Omoregie, E., Magalhaes, V., Elvert, M., Pinheiro, L., Kopf, A. and Boetius, A. (2006a) Microbial methane turnover at mud volcanoes of the Gulf of Cadiz. Geochimica et Cosmochimica Acta 70, 5336–5355.

Niemann, H., Lösekann, T., De Beer, D., Elvert, M., Nadalig, T., Knittel, K., Amann, R., Sauter, E.J., Schlüter, M. and Klages, M. (2006b) Novel microbial communities of the Haakon Mosby mud volcano and their role as a methane sink. Nature 443, 854-858.

Nisbet, E.G., Dlugokencky, E.J. and Bousquet, P. (2014) Methane on the Rise—Again. Science 343, 493-495.

Nishihara, M., Morii, H., and Koga, Y. (1989) Heptads of polar ether lipids of an archaebacterium, *Methanobacterium thermoautotrophicum*: structure and biosynthetic relationship. Biochemistry 28, 95–102.

Nishihara, M., Morii, H., Matsuno, K., Ohga, M., Stetter, O., and Koga, Y. (2002) Structural analysis by reductive cleavage with LiAlH4 of an allyl ether choline-phospholipid, archaetidylcholine, from the hyperthermophilic methanoarchaeon *Methanopyrus kandleri*. Archaea 1, 123–131.

Noble, R., Alexander, R. and Kagi, R.I. (1985) The occurrence of bisnorhopane, trisnorhopane and 25-norhopanes as free hydrocarbons in some Australian shales. Organic Geochemistry 8, 171-176.

Ochs, D., Tappe, C.H., Gaertner, P., Kellner, R. and Poralla, K. (1990) Properties of purified squalene-hopene cyclase from Bacillus acidocaldarius. European journal of biochemistry 194, 75-80.

O'Connor, F.M., Boucher, O., Gedney, N., Jones, C.D., Folberth, G.A., Coppell, R., Friedlingstein, P., Collins, W.J., Chappellaz, J., Ridley, J. and Johnson, C.E. (2010) Possible role of wetlands, permafrost, and methane hydrates in the methane cycle under future climate change: A review. Reviews of Geophysics 48, 4.

Ogniben, L. (1954) Le Argille Brecciate" Siciliane: con i rilievi di dettaglio di Grottacalda (Valguarnera, Enna), Passarello (Licata, Agrigento), Zubbi (S. Cataldo, Caltanissetta). Società Cooperativa Tipografica.

Ohtake, K., Saito, N., Shibuya, S., Kobayashi, W., Amano, R., Hirai, T., Sasaki, S., Nakano, C. and Hoshino, T. (2014) Biochemical characterization of the water-soluble squalene synthase from Methylococcus capsulatus and the functional analyses of its two DXXD (E) D motifs and the highly conserved aromatic amino acid residues. The FEBS journal 281, 5479-5497.

Op den Camp, H.J., Islam, T., Stott, M.B., Harhangi, H.R., Hynes, A., Schouten, S., Jetten, M.S., Birkeland, N.K., Pol, A. and Dunfield, P.F. (2009) Environmental, genomic and taxonomic perspectives on methanotrophic Verrucomicrobia. Environmental Microbiology Reports 1, 293-306.

Orphan, V., Hinrichs, K.-U., Ussler, W., Paull, C.K., Taylor, L., Sylva, S.P., Hayes, J.M. and DeLong, E.F. (2001) Comparative analysis of methane-oxidizing archaea and sulfate-reducing bacteria in anoxic marine sediments. Applied and Environmental Microbiology 67, 1922-1934.

Orphan, V.J., House, C.H., Hinrichs, K.-U., McKeegan, K.D., and DeLong, E.F. (2002) Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. Proceedings of the National Academy of Sciences USA 99, 7663–7668.

Ourisson, G. and Albrecht, P. (1992) Hopanoids. 1. Geohopanoids: the most abundant natural products on Earth? Accounts of Chemical Research 25, 398-402.

Ourisson, G., Albrecht, P. and Rohmer, M. (1979) The hopanoids: palaeochemistry and biochemistry of a group of natural products. Pure and Applied Chemistry 51, 709-729.

Padilla, C.C., Bristow, L.A., Sarode, N., Garcia-Robledo, E., Ramírez, E.G., Benson, C.R., Bourbonnais, A., Altabet, M.A., Girguis, P.R. and Thamdrup, B. (2016) NC10 bacteria in marine oxygen minimum zones. The ISME Journal 10, 2067.

Pagani, M., Caldeira, K., Archer, D. and Zachos, J.C. (2006) An ancient carbon mystery. Science 314, 1556.

Pan, J.-J., Solbiati, J.O., Ramamoorthy, G., Hillerich, B.S., Seidel, R.D., Cronan, J.E., Almo, S.C. and Poulter, C.D. (2015) Biosynthesis of squalene from farnesyl diphosphate in bacteria: three steps catalyzed by three enzymes. ACS central science 1, 77-82.

Pancost, R.D., Sinninghe Damsté, J.S., de Lint, S., van der Maarel, M.J. and Gottschal, J.C. (2000) Biomarker evidence for widespread anaerobic methane oxidation in Mediterranean sediments by a consortium of methanogenic archaea and bacteria. Applied and Environmental Microbiology 66, 1126-1132.

Pancost, R.D., van Geel, B., Baas, M. and Sinninghe Damsté, J.S. (2000b) δ13C values and radiocarbon dates of microbial biomarkers as tracers for carbon recycling in peat deposits. Geology 28, 663-666.

Pancost, R.D., Bouloubassi, I., Aloisi, G. and Sinninghe Damsté, J.S. (2001a) Three series of non-isoprenoidal dialkyl glycerol diethers in cold-seep carbonate crusts. Organic Geochemistry 32, 695-707.

Pancost, R.D., Hopmans, E.C. and Sinninghe Damsté, J.S. (2001b) Archaeal lipids in Mediterranean cold seeps: molecular proxies for anaerobic methane oxidation. Geochimica et Cosmochimica Acta 65, 1611-1627.

Pancost, R.D., Sinninghe Damsté, J.S. (2003) Carbon isotopic compositions of prokaryotic lipids as tracers of carbon cycling in diverse settings. Chemical Geology 195, 29-58.

Park, S.W., Park, S.T., Lee, J.E. and Kim, Y.M. (2008) *Pseudonocardia carboxydivorans sp. nov*., a carbon monoxide-oxidizing actinomycete, and an emended description of the genus Pseudonocardia. International Journal of Systematic and Evolutionary Microbiology 58, 2475–2478.

Park, H., Lee, H., Ro, Y. T., and Kim, Y. M. (2010) Identification and functional characterization of a gene for the methanol: N, N′-dimethyl-4-nitrosoaniline oxidoreductase from Mycobacterium sp. strain JC1 (DSM 3803). Microbiology 156, 463-471.

Park, S.W., Hwang, E.H., Park, H., Kim, J.A., Heo, J., Lee, K.H., Song, T., Kim, E., Ro, Y.T. and Kim, S.W. (2003) Growth of mycobacteria on carbon monoxide and methanol. Journal of Bacteriology 185, 142-147.

Parkes, J. (1999) Cracking anaerobic bacteria. Nature 401, 217–218.

Pearson, A. and Rusch, D.B. (2009) Distribution of microbial terpenoid lipid cyclases in the global ocean metagenome. The ISME Journal 3, 352-363.

Peisler, B. and Rohmer, M. (1992) Prokaryotic triterpenoids of the hopane series. bacteriohopanetetrols of new side-chain configuration from *Acetobacter* species. Journal of Chemical Research – Part S 9, 289-299.

Peters, K.E. and Moldowan, J.M. (1991) Effects of source, thermal maturity, and biodegradation on the distribution and isomerization of homohopanes in petroleum. Organic Geochemistry 17, 47–61.

Peters, K.E. and Moldowan, J.M. (1993) The Biomarker Guide, Prentice Hall, Englewood Cliffs, NJ, USA.

Peters, K.E., Peters, K.E., Walters, C.C. and Moldowan, J. (2005) The Biomarker Guide. Cambridge University Press.

Peterse, F., Hopmans, E.C., Schouten, S., Mets, A., Rijpstra, W.I.C., and Sinninghe Damsté, J.S. (2011) Identification and distribution of intact polar branched tetraether lipids in peat and soil. Organic Geochemistry 42, 1007–1015.

Peterse, F., Nicol, G. W., Schouten, S., and Sinninghe Damsté, J. S. (2010) Influence of soil pH on the abundance and distribution of core and intact polar lipid-derived branched GDGTs in soil. Organic Geochemistry 41, 1171-1175.

Piepoli, P. (1931) Studio geologico della zona Bivona-Cattolica Eraclea (Agrigento) con riferimento alle manifestazioni di idrocarburi, Riv. It. del Petrolio, n.7.

Pieri, M. and Mattavelli, L. (1986) Geologic framework of Italian petroleum resources. AAPG Bulletin 70, 103–130.

Podust, L. M., Poulos, T. L., and Waterman, M. R. (2001) Crystal structure of cytochrome P450 14α-sterol demethylase (CYP51) from Mycobacterium tuberculosis in complex with azole inhibitors. Proceedings of the National Academy of Sciences 98, 3068-3073.

Pol, A., Heijmans, K., Harhangi, H.R., Tedesco, D., Jetten, M.S.M. and Op den Camp, H.J.M. (2007) Methanotrophy below pH 1 by a new Verrucomicrobia species. Nature 450, 874-878.

Raghoebarsing, A.A., Pol, A., Van de Pas-Schoonen, K.T., Smolders, A.J., Ettwig, K.F., Rijpstra, W.I.C., Schouten, S., Sinninghe Damsté, J.S., Op den Camp, H.J. and Jetten, M.S. (2006) A microbial consortium couples anaerobic methane oxidation to denitrification. Nature 440, 918.

Ran-Ressler, R. R., Lawrence, P., and Brenna, J. T. (2012) Structural characterization of saturated branched chain fatty acid methyl esters by collisional dissociation of molecular ions generated by electron ionization. Journal of Lipid Research 53, 195-203.

Rasigraf, O., Kool, D.M., Jetten, M.S., Sinninghe Damsté, J.S. and Ettwig, K.F. (2014) Autotrophic carbon dioxide fixation via the Calvin-Benson-Bassham cycle by the denitrifying methanotroph Methylomirabilis oxyfera. Applied and Environmental Microbiology, AEM. 04199-04113.

Redman, J. E., Shaw, M. J., Mallet, A. I., Santos, A. L., Roberts, C. A., Gernaey, A. M., and Minnikin, D. E. (2009) Mycocerosic acid biomarkers for the diagnosis of tuberculosis in the Coimbra Skeletal Collection. Tuberculosis 89, 267-277.

Redmond, M.C., Valentine, D.L. and Sessions, A.L. (2010) Identification of novel methane-, ethane-, and propane-oxidizing bacteria at marine hydrocarbon seeps by stable isotope probing. Applied and Environmental Microbiology 76, 6412-6422.

Reeburgh, W.S. (1996) "Soft spots" in the global methane budget, Microbial growth on C1 compounds. Springer, pp. 334-342.

Reeburgh, W.S. (2007) Oceanic methane biogeochemistry. Chemical Reviews 107, 486-513.

Reed, W.M. and Dugan, P.R. (1987) Isolation and characterization of the facultative methylotroph Mycobacterium ID-Y. Microbiology 133, 1389-1395.

Rethemeyer, J., Schubotz, F., Talbot, H.M., Cooke, M.P., Hinrichs, K.-U., Mollenhauer, G., 2010. Distribution of polar membrane lipids in permafrost soils and sediments of a small high Arctic catchment. Organic Geochemistry 41, 1130–1145.

Rezanka, T., Siristova, L., Melzoch, K. and Sigler, K. (2010) Hopanoids in bacteria and cyanobacteria-their role in cellular biochemistry and physiology, analysis and occurrence. Mini-Reviews in Organic Chemistry 7, 300-313.

Řezanka, T., and Sigler, K. (2009) Odd-numbered very-long-chain fatty acids from the microbial, animal and plant kingdoms. Progress in Lipid Research 48, 206-238.

Ricci, J.N., Coleman, M.L., Welander, P.V., Sessions, A.L., Summons, R.E., Spear, J.R. and Newman, D.K. (2014) Diverse capacity for 2-methylhopanoid production correlates with a specific ecological niche. The ISME Journal 8, 675-684.

Ries-Kautt, M., and Albrecht, P. (1989) Hopane-derived triterpenoids in soils. Chemical Geology 76, 143-151.

Rohmer, M. and Ourisson, G. (1976) Structure des bactériohopanetétrols d'Acetobacter xylinum. Tetrahedron Letters 17, 3633-3636.

Rohmer, M., Dastillung, M. and Ourisson, G. (1980) Hopanoids from C30 to C35 in Recent muds - chemical markers for bacterial activity. Naturwissenschaften 67, 456-458.

Rohmer, M., Bouvier-Nave, P. and Ourisson, G. (1984) Distribution of hopanoid triterpenes in prokaryotes. Microbiology 130, 1137-1150.

Rohmer, M.and Ourisson, G. (1986) Unsaturated bacteriohopanepolyols from *Acetobacter aceti ssp. Xylinum*. Journal of Chemical Research. Synopses 10, 356 -357.

Rohmer, M. (1993) The biosynthesis of triterpenoids of the hopane series in the Eubacteria: A mine of new enzyme reactions. Pure and Applied Chemistry 65, 1293-1298.

Rossel, P.E., Lipp, J.S., Fredricks, H.F., Arnds, J., Boetius, A., Elvert, M., and Hinrichs, K.U. (2008) Intact polar lipids of anaerobic methanotrophic archaea and associated bacteria. Organic Geochemistry 39, 992–999.

Rossel, P.E., Elvert, M., Ramette, A., Boetius, A. and Hinrichs, K.-U. (2011) Factors controlling the distribution of anaerobic methanotrophic communities in marine environments: evidence from intact polar membrane lipids. Geochimica et Cosmochimica Acta 75, 164-184.

Ruddiman, W.F. (2003) The Anthropogenic Greenhouse Era Began Thousands of Years Ago. Climatic Change 61, 261-293.

Ruetters, H., Sass, H., Cypionka, H., and Rullkoetter, J. (2002) Phospholipid analysis as a tool to study complex microbial communities in marine sediments. Journal of Microbiological Methods 48, 149–160.

Rullkötter, J., Vonderdick, H. and Welte, D. (1982) Organic petrography and extractable hydrocarbons of sediment from the Gulf of California, Deep-Sea Drilling Project Leg-64. Initial reports of the deep sea drilling project 64, 837-853.

Rullkötter, J., Spiro, B. and Nissenbaum, A. (1985) Biological marker characteristics of oils and asphalts from carbonate source rocks in a rapidly subsiding graben, Dead Sea, Israel. Geochimica et Cosmochimica Acta 49, 1357–1370.

Rush, S., Sinninghe Damsté, J.S., Poulton, S.W., Thamdrup, B., Garside, A.L., González, J.A., Schouten, S., Jetten, M.S.M. and Talbot, H.M. (2014) Anaerobic ammonium-oxidising bacteria: A biological source of the bacteriohopanetetrol stereoisomer in marine sediments. Geochimica et Cosmochimica Acta 140, 50-64.

Rush, D., Osborne, K.A., Birgel, D., Kappler, A., Hirayama, H., Peckmann, J., Poulton, S.W., Nickel, J.C., Mangelsdorf, K., Kalyuzhnaya, M. Sidgwick, F.R. and Talbot, H.M. (2016) The bacteriohopanepolyol inventory of novel aerobic methane oxidizing bacteria reveals new biomarker signatures of aerobic methanothropy in marine systems. PLoS ONE 11, e0165635.

Rush, D., Talbot, H.M., van der Meer, M.T.J., Hopmans, E.C., Douglas, B. and Sinninghe Damsté, J.S. (2019) Biomarker evidence for the occurrence of anaerobic ammonium oxidation in the eastern Mediterranean Sea during Quaternary and Pliocene sapropel formation. Biogeosciences 16, 2467-2479.

Sáenz, J.P., Wakeham, S.G., Eglinton, T.I. and Summons, R.E. (2011) New constraints on the provenance of hopanoids in the marine geologic record: Bacteriohopanepolyols in marine suboxic and anoxic environments. Organic Geochemistry 42, 1351-1362.

Santos, V., Billett, D.S., Rice, A.L. and Wolff, G.A. (1994) Organic matter in deep-sea sediments from the Porcupine Abyssal Plain in the north-east Atlantic Ocean. I—Lipids. Deep Sea Research Part I: Oceanographic Research Papers 41, 787-819.

Sapart, C.J., Martinerie, P., Witrant, E., Chappellaz, J., van de Wal, R.S.W., Sperlich, P., van der Veen, C., Bernard, S., Sturges, W.T., Blunier, T., Schwander, J., Etheridge, D. and Röckmann, T. (2013) Can the carbon isotopic composition of methane be reconstructed from multi-site firn air measurements? Atmospheric Chemistry and Physics 13, 6993-7005.

Sassen, R., Roberts, H.H., Carney, R., Milkov, A.V., DeFreitas, D.A., Lanoil, B. and Zhang, C. (2004) Free hydrocarbon gas, gas hydrate, and authigenic minerals in chemosynthetic communities of the northern Gulf of Mexico continental slope: relation to microbial processes. Chemical Geology 205, 195–217.

Saunois, M., Bousquet, P., Poulter, B., Peregon, A., Ciais, P., Canadell, J. G., et al. (2016) The global methane budget 2000–2012. Earth System Science Data, 8(2), 697–751.

Scalan, E. and Smith, J. (1970) An improved measure of the odd-even predominance in the normal alkanes of sediment extracts and petroleum. Geochimica et Cosmochimica Acta 34, 611–620.

Schäffer, A.A., Aravind, L., Madden, T.L., Shavirin, S., Spouge, J.L., Wolf, Y.I., Koonin, E.V. and Altschul, S.F. (2001) Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. Nucleic acids research 29, 2994-3005.

Schmitz, R.A., Peeters, S.H., Versantvoort, W., Picone, N., Pol, A., Jetten, M.S.M. and Op den Camp, H.J.M. (2021) Verrucomicrobial methanotrophs:

ecophysiology of metabolically versatile acidophiles. FEMS Microbiology Reviews.

Schoell, M., Teschner, M., Wehner, H., Durand, B. and Oudin, J. (1983) Maturity related biomarker and stable isotope variations and their application to oil/source rock correlation in the Mahakam Delta, Kalimantan, Advances in Organic Geochemistry 1981. Wiley Chichester, pp. 156–163.

Schoell, M., McCaffrey, M., Fago, F. and Moldowan, J. (1992) Carbon isotopic compositions of 28, 30-bisnorhopanes and other biological markers in a Monterey crude oil. Geochimica et Cosmochimica Acta 56, 1391-1399.

Schouten, S., Bowman, J.P., Rijpstra, W.I.C. and Sinninghe Damsté, J.S. (2000) Sterols in a psychrophilic methanotroph, *Methylosphaera hansonii*. FEMS Microbiology Letters 186, 193-195.

Schouten, S., Hopmans, E.C., Pancost, R.D. and Sinninghe Damsté, J.S. (2000) Widespread occurrence of structurally diverse tetraether membrane lipids: Evidence for the ubiquitous presence of low-temperature relatives of hyperthermophiles. Proceedings of the National Academy of Sciences 97, 14421–14426.

Schouten, S., De Loureiro, M.R., Sinninghe Damsté, J.S. and de Leeuw, J.W. (2001a) Molecular biogeochemistry of Monterey sediments, Naples Beach, California: I. Distributions of hydrocarbons and organic sulfur compounds. The Monterey Formation: from Rocks to Molecules (eds. CM Isaacs and J. Rullkötter). Columbia University Press, New York, 150-174.

Schouten, S., Schoell, M., Sinninghe Damsté, J.S., Summons, R.E. and De Leeuw, J.W. (2001b) Molecular biogeochemistry of Monterey sediments, Naples Beach, California: II. Stable carbon isotopic compositions of free and sulphurbound carbon skeletons. The Monterey Formation: From Rocks to Molecules (eds. CM Isaacs and J. Rullkötter). Columbia University Press, New York, 175-188.

Schouten, S., Wakeham, S.G., Hopmans, E.C. and Sinninghe Damsté, J.S. (2003) Biogeochemical Evidence that Thermophilic Archaea Mediate the Anaerobic Oxidation of Methane. Applied and Environmental Microbiology 69, 1680-1686.

Schouten, S., Hopmans, E.C., Baas, M., Boumann, H., Standfest, S., Könneke, M., Stahl, D.A. and Sinninghe Damsté, J.S. (2008) Intact membrane lipids of "*Candidatus* Nitrosopumilus maritimus," a cultivated representative of the cosmopolitan mesophilic group I crenarchaeota. Applied and Environmental Microbiology 74, 2433-2440.

Schouten, S., Hopmans, E.C. and Sinninghe Damsté, J.S. (2013) The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: a review. Organic Geochemistry 54, 19-61.

Schubotz, F., Lipp, J.S., Elvert, M., Kasten, S., Mollar, X.P., Zabel, M., Bohrmann, G. and Hinrichs, K.-U. (2011) Petroleum degradation and associated microbial signatures at the Chapopote asphalt volcano, Southern Gulf of Mexico. Geochimica et Cosmochimica Acta 75, 4377–4398.

Schubotz, F., Wakeham, S.G., Lipp, J.S., Fredricks, H.F. and Hinrichs, K.-U. (2009) Detection of microbial biomass by intact polar membrane lipid analysis in the water column and surface sediments of the Black Sea. Environmental Microbiology 11, 2720-2734.

Schulenberg-Schell, H., Neuss, B. and Sahm, H. (1989) Quantitative determination of various hopanoids in microorganisms. Analytical Biochemistry 181, 120-124.

Schwartz-Narbonne, R., Schaeffer, P., Hopmans, E.C., Schenesse, M., Charlton, E. A., Jones, D. M., Sinninghe Damsté, J.S., Ul Haque, M.F., Jetten, M.S.M., Lengger, S.K., Murrell, J.C., Normand, P., Nuijten, G.H.L., Talbot, H.M. and Rush, D. 2020. A unique bacteriohopanetetrol stereoisomer of marine anammox. Organic Geochemistry 143, 103994.

Schwietzke, S., Sherwood, O.A., Bruhwiler, L.M.P., Miller, J.B., Etiope, G., Dlugokencky, E.J., Michel, S.E., Arling, V.A., Vaughn, B.H., White, J.W.C. and Tans, P.P. (2016) Upward revision of global fossil fuel methane emissions based on isotope database. Nature 538, 88-91.

Seifert, W.K. and Moldowan, J.M. (1978) Applications of steranes, terpanes and monoaromatics to the maturation, migration and source of crude oils. Geochimica et Cosmochimica Acta 42, 77-95.

Seifert, W.K. and Moldowan, J.M. (1980) The effect of thermal stress on source-rock quality as measured by hopane stereochemistry. Physics and Chemistry of the Earth 12, 229–237.

Sessions, A.L., Zhang, L., Welander, P.V., Doughty, D, Summons, R.E. and Newman, D.K. (2013) Identification and quantification of polyfunctionalized hopanoids by high temperature gas chromatography–mass spectrometry. Organic Geochemistry 56, 120-130.

Simoneit, B.R. (1977) The Black Sea, a sink for terrigenous lipids. Deep Sea Research 24, 813-830.

Simonin, P., Tindall, B. and Rohmer, M. (1994) Structure elucidation and biosynthesis of 31-methylhopanoids from *Acetobacter europaeus*. European Journal of Biochemistry 225, 765-771.

Sinninghe Damsté, J., Ten Haven, H., De Leeuw, J. and Schenck, P. (1986) Organic geochemical studies of a Messinian evaporitic basin, Northern Apennines (Italy) II. Isoprenoid and *n*-alkyl thiophenes and thiolanes. Organic Geochemistry 10, 791–805.

Sinninghe Damsté, J.S., Rijpstra, W.I.C., Schouten, S., Fuerst, J.A., Jetten, M.S. and Strous, M. (2004) The occurrence of hopanoids in planctomycetes: implications for the sedimentary biomarker record. Organic Geochemistry 35, 561-566.

Sinninghe Damsté, J.S., Rijpstra, W.I.C., Geenevasen, J.A., Strous, M. and Jetten, M.S.M. (2005) Structural identification of ladderane and other membrane lipids of planctomycetes capable of anaerobic ammonium oxidation (anammox). The FEBS Journal 272, 4270–4283.

Sinninghe Damsté, J.S., Rijpstra, W.I.C., Hopmans, E.C., Jung, M.-Y., Kim, J.-G., Rhee, S.-K., et al. (2012) Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids of group I.1a and I.1b Thaumarchaeota in soil. Applied and Environmental Microbiology 78, 6866–6874.

Sinninghe Damsté, J.S., Rijpstra, W.I.C., Dedysh, S.N., Foesel, B.U. and Villanueva, L. (2017) Pheno-and genotyping of hopanoid production in Acidobacteria. Frontiers in Microbiology 8, 968.

Sinninghe Damsté, J. S., Rijpstra, W. I. C., Foesel, B. U., Huber, K. J., Overmann, J., Nakagawa, S., Kim, J. J., Dunfield, P. F., Dedysh, S. N., and Villanueva, L. (2018) An overview of the occurrence of ether-and ester-linked iso-diabolic acid membrane lipids in microbial cultures of the Acidobacteria: Implications for brGDGT paleoproxies for temperature and pH. Organic Geochemistry 124, 63-76, 2018.

Smit, N.T., Rush, D., Sahonero-Canavesi, D.X., Verweij, M., Rasigraf, O., Cruz, S.G., Jetten, M.S., Sinninghe Damsté, J.S. and Schouten, S. (2019) Demethylated hopanoids in '*Ca*. Methylomirabilis oxyfera' as biomarkers for environmental nitrite-dependent methane oxidation. Organic Geochemistry 137, 103899.

Smit, N.T., Villanueva, L., Rush, D., Grassa, F., Witkowski, C.R., Holzheimer, M., Minnaard, A.J., Sinninghe Damsté, J.S. and Schouten, S. (2021a) Novel hydrocarbon-utilizing soil mycobacteria synthesize unique mycocerosic acids at a Sicilian everlasting fire. Biogeosciences 18, 1463-1479.

Smit, N. T., Rush, D., van der Meer, M. T. J., Grassa, F., Sinninghe Damsté, J. S. and Schouten, S. (2021b) The role of soil mycobacteria in the uptake of methane and ethane at a Sicilian everlasting fire. *In preparation for Biogeosciences.*

Smith, K.A., Dobbie, K.E., Ball, B.C., Bakken, L.R., Sitaula, B.K., Hansen, S., Brumme, R., Borken, W., Christensen, S., Priemé, A., Fowler, D., Macdonald, J.A., Skiba, U., Klemedtsson, L., Kasimir-Klemedtsson, A., Degórska, A. and Orlanski, P. (2000) Oxidation of atmospheric methane in Northern European soils, comparison with other ecosystems, and uncertainties in the global terrestrial sink. Global Change Biology 6, 791-803.

Sollai, M, Villanueava, L., Hopmans, E.C., Reichart, G.-J. and Sinninghe Damsté, J.S. (2019) A combined lipidomic and 16S rRNA gene amplicon sequencing approach reveals archaeal sources of intact polar lipids in the stratified Black Sea water column. Geobiology 17, 91–109.

Solomon, S., Qin, D., Manning, M., Averyt, K. and Marquis, M. (2007) Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC. Cambridge University Press.

Staccioli, G., McMillan, N., Meli, A., and Bartolini, G. (2002) Chemical characterisation of a 45-million-year bark from Geodetic Hills fossil forest, Axel Heiberg Island, Canada. Wood Science and Technology, 36, 419-427.

Stadnitskaia, A., Blinova, V., Ivanov, M., Baas, M., Hopmans, E., Van Weering, T. and Sinninghe Damsté, J.S. (2007) Lipid biomarkers in sediments of mud volcanoes from the Sorokin Trough, NE Black Sea: Probable source strata for the erupted material. Organic Geochemistry 38, 67-83.

Stadnitskaia, A., Muyzer, G., Abbas, B., Coolen, M., Hopmans, E., Baas, M., Van Weering, T., Ivanov, M., Poludetkina, E. and Sinninghe Damsté, J.S. (2005) Biomarker and 16S rDNA evidence for anaerobic oxidation of methane and related carbonate precipitation in deep-sea mud volcanoes of the Sorokin Trough, Black Sea. Marine Geology 217, 67-96.

Stadnitskaia, A., Nadezhkin, D., Abbas, B., Blinova, V., Ivanov, M.K. and Sinninghe Damsté, J.S. (2008) Carbonate formation by anaerobic oxidation of methane: Evidence from lipid biomarker and fossil 16S rDNA. Geochimica et Cosmochimica Acta 72, 1824-1836.

Sturt, H.F., Summons, R.E., Smith, K., Elvert, M., and Hinrichs, K. (2004) Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography / electrospray ionization multistage mass spectrometry — new biomarkers for biogeochemistry and microbial ecology. Rapid communications in mass spectrometry 18(6), 617–628.

Summons, R.E., Jahnke, L.L. and Roksandic, Z. (1994) Carbon isotopic fractionation in lipids from methanotrophic bacteria: relevance for interpretation of the geochemical record of biomarkers. Geochimica et Cosmochimica Acta 58, 2853-2863.

Summons, R.E. and Jahnke, L.L. (1992) Hopenes and hopanes methylated in ring-A: correlation of the hopanoids of extant methylotrophic bacteria with their fossil analogues. Biomarkers in sediments and petroleum. Prentice Hall, Englewood Cliffs, 182-200.

Summons, R.E. and Powell, T.G. (1986) Chlorobiaceae in Palaeozoic seas revealed by biological markers, isotopes and geology. Nature 319, 763-765.

Takahashi, J. (1980) Production of Intracellular and Extracellular Protein from n-Butane by *Pseudomonas butanovora* sp. nov. Advances in Applied Microbiology. Elsevier, pp. 117-127.

Takeuchi, M., Yoshioka, H., Seo, Y., Tanabe, S., Tamaki, H., Kamagata, Y., Takahashi, H.A., Igari, S., Mayumi, D. and Sakata, S. (2011) A distinct freshwater-adapted subgroup of ANME-1 dominates active archaeal communities in terrestrial subsurfaces in Japan. Environmental Microbiology 13, 3206-3218.

Talbot, H.M., Watson, D.F., Murrell, J.C., Carter, J.F. and Farrimond, P. (2001) Analysis of intact bacteriohopanepolyols from methanotrophic bacteria by reversed-phase high-performance liquid chromatography–atmospheric pressure chemical ionisation mass spectrometry. Journal of Chromatography A 921, 175-185.

Talbot, H.M., Squier, A.H., Keely, B.J. and Farrimond, P. (2003a) Atmospheric pressure chemical ionisation reversed-phase liquid chromatography/ion trap mass spectrometry of intact bacteriohopanepolyols. Rapid Communications in Mass Spectrometry 17, 728–737.

Talbot, H.M., Summons, R., Jahnke, L., Farrimond, P. (2003b) Characteristic fragmentation of bacteriohopanepolyols during atmospheric pressure chemical ionization liquid chromatography/ion trap mass spectrometry. Rapid Communications in Mass Spectrometry 17, 2788–2796.

Talbot, H.M., Watson, D.F., Pearson, E.J. and Farrimond, P. (2003c). Diverse biohopanoid compositions of non-marine sediments, Organic Geochemistry 34, 1353-1371.

Talbot, H.M. and Farrimond, P. (2007) Bacterial populations recorded in diverse sedimentary biohopanoid distributions. Organic Geochemistry 38, 1212-1225.

Talbot, H.M., Rohmer, M. and Farrimond, P. (2007a) Rapid structural elucidation of composite bacterial hopanoids by atmospheric pressure chemical ionization liquid chromatography/ion trap mass spectrometry. Rapid Communications in Mass Spectrometry 21, 880–892.

Talbot, H.M., Rohmer, M. and Farrimond, P. (2007b). Structural characterisation of unsaturated bacterial hopanoids by atmospheric pressure chemical ionisation

liquid chromatography/ion trap mass spectrometry. Rapid Communications in Mass Spectrometry 21, 1613-1622.

Talbot, H.M, Summons, R.E., Jahnke, L.L., Cockell, C.S., Rohmer, R. and Farrimond, P. (2008) Cyanobacterial bacteriohopanepolyol signatures from cultures and natural environmental settings. Organic Geochemistry 39, 232-263.

Talbot, H.M., Handley, L., Spencer-Jones, C.L., Dinga, B.J., Schefuß, E., Mann, P.J., Poulsen, J.R., Spencer, R.G.M., Wabakanghanzi, J.N. and Wagner, T. (2014) Variability in aerobic methane oxidation over the past 1.2 Myrs recorded in microbial biomarker signatures from Congo fan sediments. Geochimica et Cosmochimica Acta 133, 387-401.

Talbot, H.M., McClymont, E.L., Inglis, G.N., Evershed, R.P. and Pancost, R.D. (2016) Origin and preservation of bacteriohopanepolyol signatures in Sphagnum peat from Bissendorfer Moor (Germany). Organic Geochemistry 97, 95-110.

Talbot, H.M., Sidgwick, F.R., Bischoff, J., Osborn, K.A., Rush, D., Sherry, A and Spencer-Jones, C.L. (2016) Analysis of non-derivatized bacteriohopanepolyols by ultrahigh-performance liquid chromatography/tandem mass spectrometry. Rapid Communications in Mass Spectrometry 30, 2087-2098.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0, Molecular Biology and Evolution, 30, 2725-2729.

Ten Haven, H., De Leeuw, J. and Schenck, P. (1985) Organic geochemical studies of a Messinian evaporitic basin, northern Apennines (Italy) I: Hydrocarbon biological markers for a hypersaline environment. Geochimica et Cosmochimica Acta 49, 2181–2191.

Ten Haven, H., Baas, M., De Leeuw, J., Maassen, J., Schenck, P. (1987) Organic geochemical characteristics of sediments from the anoxic brine-filled Tyro basin (eastern Mediterranean). Organic geochemistry 11, 605-611.

Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W. and Hedderich, R. (2008) Methanogenic archaea: ecologically relevant differences in energy conservation. Nature Reviews Microbiology 6, 579-591.

Thiel, V., Blumenberg, M., Pape, T., Seifert, R. and Michaelis, W. (2003) Unexpected occurrence of hopanoids at gas seeps in the Black Sea. Organic Geochemistry 34, 81–87.

Tiedje, J. M., Asuming-Brempong, S., Nüsslein, K., Marsh, T. L., and Flynn, S. J. (1999) Opening the black box of soil microbial diversity. Applied Soil Ecology 13, 109-122.

Timmers, P.H.A., Welte, C.U., Koehorst, J.J., Plugge, C.M., Jetten, M.S.M., and Stams, A.J.M. (2017) Reverse methanogenesis and respiration in methanotrophic archaea. Archaea 2017: 1654237.

Torkko, P., Suomalainen, S., Iivanainen, E., Suutari, M., Paulin, L., Rudbäck, E., Tortoli, E., Vincent, V., Mattila, R., and Katila, M.-L. (2001) Characterization of Mycobacterium bohemicum isolated from human, veterinary, and environmental sources. Journal of Clinical Microbiology 39, 207-211.

Torkko, P., Suomalainen, S., Iivanainen, E., Tortoli, E., Suutari, M., Seppänen, J., Paulin, L., and Katila, M.-L. (2002) Mycobacterium palustre sp. nov., a potentially pathogenic, slowly growing mycobacterium isolated from clinical and veterinary specimens and from Finnish stream waters, International Journal of Systematic and Evolutionary Microbiology, 52, 1519-1525.

Torkko, P., Katila, M.-L., and Kontro, M. (2003) Gas-chromatographic lipid profiles in identification of currently known slowly growing environmental mycobacteria, Journal of Medical Microbiology, 52, 315-323.

Tornabene, T.G. and Langworthy, T.A. (1979) Diphytanyl and dibiphytanyl glycerol ether lipids of methanogenic archaebacteria. Science 203, 51-53.

Tortoli, E. (2014) Microbiological features and clinical relevance of new species of the genus Mycobacterium. Clinical Microbiology Reviews 27, 727-752.

Trincianti, E., Frixa, A. and Sartorio, D. (2015) Palynology and stratigraphic characterization of subsurface sedimentary successions in the Sicanian and Imerese Domains—Central Western Sicily. Review of Palaeobotany and Palynology 218, 48-66.

Trudinger, C.M., Etheridge, D.M., Rayner, P.J., Enting, I.G., Sturrock, G.A. and Langenfelds, R.L. (2002) Reconstructing atmospheric histories from measurements of air composition in firn. Journal of Geophysical Research 107 (D24), 4780.

Tsutsumi, Y., Iijima, A., Yoshida, K., Shoji, H. and Lee, J.T. (2009) HCCI combustion characteristics during operation on DME and methane fuels. International Journal of Automotive Technology 10, 645.

UniProt Consortium, 2018. UniProt: a worldwide hub of protein knowledge. Nucleic acids research 47, D506-D515.

Vaksmaa, A., Guerrero-Cruz, S., van Alen, T.A., Cremers, G., Ettwig, K.F., Lüke, C. and Jetten, M.S. (2017) Enrichment of anaerobic nitrate-dependent methanotrophic '*Candidatus* Methanoperedens nitroreducens' archaea from an

Italian paddy field soil. Applied Microbiology and Biotechnology 101, 7075-7084.

Vaksmaa, A., Jetten, M.S.M., Ettwig, K.F., and Lüke, C. (2017) *McrA* primers for the detection and quantification of the anaerobic archaeal methanotroph "*Candidatus* Methanoperedens nitroreducens." Applied Microbiology and Biotechnology 101, 1631–1641.

Valentine, D.L. and Reeburgh, W.S. (2000) New perspectives on anaerobic methane oxidation: minireview. Environmental Microbiology 2, 477-484.

Valero-Guillén, P., Martín-Luengo, F., Larsson, L., Jimenez, J., Juhlin, I., and Portaels, F. (1988) Fatty and mycolic acids of Mycobacterium malmoense. Journal of Clinical Microbiology 26, 153-154.

Van de Vossenberg, J., Rattray, J.E., Geerts, W., Kartal, B., van Niftrik, L., van Donselaar, E. G., Sinninghe Damsté, J.S., Strous, M. and Jetten, M.S.M. (2008) Enrichment and characterization of marine anammox bacteria associated with global nitrogen gas production. Environmental Microbiology 10, 3120–3129.

Van Dongen, B.E., Talbot, H.M., Schouten, S., Pearson, P.N. and Pancost, R.D. (2006) Well preserved Palaeogene and Cretaceous biomarkers from the Kilwa area, Tanzania. Organic Geochemistry 37, 539-557.

Van Duin, A.C.T., Sinninghe Damsté, J.S., Kodpmans, M.P., van de Graaf, B., de Leeuw, J.W., 1997. A kinetic calculation method of homohopanoid maturation: Applications in the reconstruction of burial histories of sedimentary basins. Geochimica et Cosmochimica Acta 61, 2409–2429.

Van Teeseling, M.C.F., Pol, A., Harhangi, H.R., van der Zwart, S., Jetten, M.S.M., Op den Camp, H.J.M. and van Niftrik, L. (2014) Expanding the Verrucomicrobial Methanotrophic World: Description of Three Novel Species of *Methylacidimicrobium* gen. nov.. Applied and Environmental Microbiology 80, 6782.

Van Winden, J.F., Kip, N., Reichart, G.-J., Jetten, M.S., Op den Camp, H.J. and Sinninghe Damsté, J.S. (2010). Lipids of symbiotic methane-oxidizing bacteria in peat moss studied using stable carbon isotopic labelling. Organic Geochemistry 41, 1040-1044.

Van Winden, J.F., Talbot, H.M., Kip, N., Reichart, G.-J., Pol, A., McNamara, N.P., Jetten, M.S., Op den Camp, H.J. and Sinninghe Damste, J.S. (2012) Bacteriohopanepolyol signatures as markers for methanotrophic bacteria in peat moss. Geochimica et Cosmochimica Acta 77, 52-61.

Van Winden, J.F., Talbot, H.M., Reichart, G.-J., McNamara, N.P., Benthien, A., and Sinninghe Damsté, J.S. (2020) Influence of temperature on the δ13C values

and distribution of methanotroph-related hopanoids in Sphagnum-dominated peat bogs. Geobiology, 18, 497-507.

Volkman, J.K., Alexander, R., Kagi, R.I. and Rullkötter, J. (1983) GC-MS characterisation of C27 and C28 triterpanes in sediments and petroleum. Geochimica et Cosmochimica Acta 47, 1033-1040.

Volkman, J.K., Holdsworth, D.G., Neill, G.P. and Bavor Jr, H. (1992) Identification of natural, anthropogenic and petroleum hydrocarbons in aquatic sediments. Science of the Total Environment 112, 203–219.

Volkman, J.K. (2003) Sterols in microorganisms. Applied Microbiology and Biotechnology 60, 495-506.

Wagner, T., Kallweit, W., Talbot, H.M., Mollenhauer, G., Boom, A. and Zabel, M. (2014) Microbial biomarkers support organic carbon transport from methane-rich Amazon wetlands to the shelf and deep sea fan during recent and glacial climate conditions. Organic Geochemistry 67, 85-98.

Wakeham, S.G., Lewis, C.M., Hopmans, E.C., Schouten, S. and Sinninghe Damsté, J.S. (2003) Archaea mediate anaerobic oxidation of methane in deep euxinic waters of the Black Sea. Geochimica et Cosmochimica Acta 67, 1359–1374.

Wakeham, S.G., Hopmans, E.C., Schouten, S. and Sinninghe Damsté, J.S. (2004) Archaeal lipids and anaerobic oxidation of methane in euxinic water columns: a comparative study of the Black Sea and Cariaco Basin. Chemical Geology 205, 427-442.

Wakeham, S.G., Amann, R., Freeman, K.H., Hopmans, E.C., Jørgensen, B.B., Putnam, I.F., Schouten, S., Sinninghe Damsté, J.S., Talbot, H.M. and Woebken, D. (2007) Microbial ecology of the stratified water column of the Black Sea as revealed by a comprehensive biomarker study. Organic Geochemistry 38, 2070-2097.

Walsh, C. M., Gebert, M. J., Delgado-Baquerizo, M., Maestre, F., and Fierer, N. (2019) A global survey of mycobacterial diversity in soil. Applied Environmental Microbiology, 85, e01180-19.

Wang, H-Y, Tatituri, R.V.V., Goldner, N.K., Dantas, G. and Hsu, F-F. (2020) Unveiling the biodiversity of lipid species in *Corynebacteria* - characterization of the uncommon lipid families in *C. glutamicum* and pathogen *C. striatum* by mass spectrometry. Biochimie 178, 158-169.

Wang, M., Tian, J., Bu, Z., Lamit, L.J., Chen, H., Zhu, Q., and Peng, C. (2019) Structural and functional differentiation of the microbial community in the

surface and subsurface peat of two minerotrophic fens in China. Plant and Soil 437, 1, 21-40.

Weber, H.S., Habicht, K.S., and Thamdrup, B. (2017) Anaerobic methanotrophic archaea of the ANME-2d cluster are active in a low-sulfate, iron-rich freshwater sediment. Frontiers in Microbiology 8, 1–13.

Wegener, G., Niemann, H., Elvert, M., Hinrichs, K.-U., and Boetius, A. (2008) Assimilation of methane and inorganic carbon by microbial communities mediating the anaerobic oxidation of methane. Environmental Microbiology 10, 2287–2298.

Wegener, G., Krukenberg, V., Ruff, S.E., and Kellermann, M.Y. (2016) Metabolic capabilities of microorganisms involved in and associated with the anaerobic oxidation of methane. Frontiers in Microbiology 7, 1–16.

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

Welander, P.V. and Summons, R.E. (2012) Discovery, taxonomic distribution, and phenotypic characterization of a gene required for 3-methylhopanoid production. Proceedings of the National Academy of Sciences 109, 12905-12910.

Wenger, L.M., Davis, C.L. and Isaksen, G.H. (2002) Multiple controls on petroleum biodegradation and impact on oil quality. SPE Reservoir Evaluation & Engineering 5, 375-383.

Welte, C. and Deppenmeier, U. (2014) Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1837, 1130-1147.

Whiticar, M.J. (1999) Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. Chemical Geology 161, 291-314.

Whitman, W.B. (1994) Autotrophic acetyl coenzyme A biosynthesis in methanogens. In, Drake,H.L. (ed), *Acetogenesis*. Chapman & Hall., New York, pp. 521–538.

Wolff, E. and Spahni, R. (2007) Methane and nitrous oxide in the ice core record. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365, 1775-1792.

World Health Organization (2019) World health statistics 2019: monitoring health for the SDGs, sustainable development goals, World Health Organization, 120 pp., ISBN: 9241565705.

Wörmer, L., Lipp, J.S., Schröder, J.M. and Hinrichs, K.-U. (2013) Application of two new LC–ESI–MS methods for improved detection of intact polar lipids (IPLs) in environmental samples. Organic Geochemistry 59, 10-21.

Wrede, C., Brady, S., Rockstroh, S., Dreier, A., Kokoschka, S., Heinzelmann, S., Heller, C., Reitner, J., Taviani, M. and Daniel, R. (2012) Aerobic and anaerobic methane oxidation in terrestrial mud volcanoes in the Northern Apennines. Sedimentary Geology 263, 210–219.

Wu, C.-H., Kong, L., Bialecka-Fornal, M., Park, S., Thompson, A.L., Kulkarni, G., Conway, S.J. and Newman, D.K. (2015) Quantitative hopanoid analysis enables robust pattern detection and comparison between laboratories. Geobiology 13, 391-407.

Zengler, K., Richnow, H.H., Rosselló-Mora, R., Michaelis, W. and Widdel, F. (1999) Methane formation from long-chain alkanes by anaerobic microorganisms. Nature 401, 266–269.

Zhang, G., Tian, J., Jiang, N., Guo, X., Wang, Y., and Dong, X. (2008) Methanogen community in Zoige wetland of Tibetan plateau and phenotypic characterization of a dominant uncultured methanogen cluster ZC-I. Environmental Microbiology 10, 1850–1860.

Zhang, J., Kobert, K., Flouri, T. and Stamatakis, A. (2013) PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30, 614-620.

Zhang, Y.G., Zhang, C.L., Liu, X.-L., Li, L., Hinrichs, K.-U. and Noakes, J.E. (2011) Methane Index: A tetraether archaeal lipid biomarker indicator for detecting the instability of marine gas hydrates. Earth and Planetary Science Letters 307, 525–534.

Zhu, B., van Dijk, G., Fritz, C., Smolders, A.J., Pol, A., Jetten, M.S. and Ettwig, K.F. (2012) Anaerobic oxidization of methane in a minerotrophic peatland: enrichment of nitrite-dependent methane-oxidizing bacteria. Applied and Environmental Microbiology 78, 8657-8665.

Zhu, C., Talbot, H.M., Wagner, T., Pan, J.M. and Pancost, R.D. (2010) Intense aerobic methane oxidation in the Yangtze Estuary: A record from 35-aminobacteriohopanepolyols in surface sediments. Organic Geochemistry 41, 1056-1059.

Zhu, C. Talbot, H.M., Wagner,T., Pan, J.-M. and Pancost, R.D. (2011) Distribution of hopanoids along a land to sea transect: Implications for microbial ecology and the use of hopanoids in environmental studies. Limnology and Oceanography 56, 1850–1865.

Zumberge, J.E. (1984) Source rocks of the La Luna Formation (Upper Cretaceous) in the Middle Magdalena Valley, Colombia. AAPG Special Volumes 30, 127.

Zundel, M. and Rohmer, M. (1985) Prokaryotic triterpenoids: 1. 3β-Methylhopanoids from Acetobacter species and Methylococcus capsulatus. European Journal of Biochemistry 150, 23-27.

Acknowledgements/ Danksagung

Acknowledgements / Danksagung

When I started my PhD about 4 years ago who would have thought that a global pandemic would get in the way of all our plans. Despite these difficult times shortly before and during Corona I, or better we, still managed to finish this thesis. There are many great people who I´m thankful to for their support and help.

First of all I would like to thank my three supervisors Stefan, Darci and Jaap for their help and support during my PhD. From the start on you believed in me, my decisions and my knowledge in Organic Geochemistry even if I doubted myself or was going through some difficult times. I´m very grateful that you taught me how to become an independent, critical thinking and confident scientist in the crazy world of academia.

Stefan: It is really hard to find the right words to say how thankful I am for all your support, knowledge and wisdom during the last 4 years. Your door was always open when I had questions or was puzzling with my data. Your undestroyable calmness helped me through a lot of difficult moments and taught me to stay calm before making a possibly stupid decision. One of my best memories are the endless afternoon meetings which we had together to puzzle all these weird new biomarkers (e.g. trisnorhopanoids or mycocerosic acids) together while drinking hot chocolate (Stefan) and coffee (I). Even when I wanted to go to Sicily to sample methane seeps, you supported my crazy ideas and who would have thought that an everlasting fire would give us so many exciting new data – Probably you are still trying to make sense out of these weird big glycolipids there…Thanks for being such an amazing and inspiring supervisor to me!

Darci: As you tend to say I was your test student since I was the first PhD student you supervised and it is time to let you know that your experiment worked out very well (at least I think so…). Thanks for all your help and support over the last years which have been always accompanied with lots of laughs and humor! You were always there for questions or just having a chat about some data, literature or experiments I wasn´t certain of. Your positive vibes and your humor were always there to cheer me up even when data didn´t turn out as we expected them to. Thanks for fascinating me about BHPs, helping me with presentations or setting up experiments. I will miss working with you and look forward to a lot of fun conferences together in the future where we can rock the dancefloor =).

Jaap: I´m so thankful for all your help and mentorship while interpreting data, writing manuscripts and getting this thesis finished. Unfortunately, we did not work together as much as we probably planned or thought in the beginning, however, I´m very grateful for all your input and help especially during the last year of my PhD. You taught me how to critically think about my data and write a perfect manuscript and I need to admit that you often drove me nuts with it. In the end, it turned out as a perfect manuscript which didn´t need many revisions. It was a pleasure working together with you and get inspired by your knowledge and never ending curiosity to understand our science even better than we already do.

I would like to thank the reading committee members and the sitting committee members for their assessments of my thesis and the public defense.

Before starting this PhD, my interest in the field of Organic Geochemistry was fueled by many other great people. I would like to thank the Organic Geochemistry group at the MARUM in Bremen to awaken my interest in biomarkers and isotopes and teaching me a lot of my first lab skills. Furthermore, there are three people who I would like to especially thank for their inspiration to become a scientist. First of all, **Florence**, you are an amazing and inspiring female scientist whose knowledge I will always appreciate. You were not only a fantastic mentor during my times as a master student, moreover, you (and your family) became very good friend to me who I can always ask for honest feedback and advice. I´m really looking forward to working together with you again at MARUM and having more nice dinners together. A very big thanks goes also to **Christian** and **Johan** who were great advisors and mentors during my time as a master student.

What would have been my thesis without all the great scientists, collaborators, technicians, friends and family at NIOZ and back home.

Ellen: Thank you so much for teaching me about crazy BHPs and so much more. I always enjoyed our meetings together with your humor and your endless knowledge about all kinds of intact lipids. I´m still amazed about many many new BHPs you identified in the cultures and the Censo soil (130 different BHPs). It was a pleasure working with you. I will miss our gezellig meetings and shitchats.

Laura: I´m thankful to you for all your help and input with the microbiological side of my data sets. Thanks for explaining a geochemist what 16S rRNA gene sequencing can do and what not. The whole mycobacteria story would only have been half as great as it became with your help and the DNA-data.

Marcel: Thanks a lot for having your door always open when I had some questions about my data or was just looking for some good advice. It was great to chat and discuss things with you and I always went back to my office with a more positive feeling.

Fausto: My Italian collaborator and friend. Grazie for your help with the two sampling trips in Sicily. It was great to see these fascinating places like the mud volcanoes, mud ponds and of course Fuoco di Censo. Thanks for your numerous explanations about these unique places and for your hospitality and the amazing food as well as the tasty fresh olive oil. I´m so happy that I did not only get a lot of nice research results but also found a very great person!

I would like to say a big thank you to all our amazing technicians at NIOZ, who do a great job in running the labs and instruments. Without all of you my research would have not been possible. Thanks to Monique (for all your support and help with the triplequad), Anchelique (extracting and analyzing a lot of my mud volcano samples), Denise (keeping the Orbitrap running), Jort and Ronald (running and maintaining all the IRMS), Marianne (doing a great job as labhead), Jessica, Irene, Michel, Caglar, Carsten, Alle, Maartje and Sanne.

Thanks to all the other PhDs and Postdocs for nice lunches, coffee breaks and drinks who accompanied me on my way at NIOZ and at MMB. Thank you, **Diana**, for all the nice exchanges over microbiology and geochemistry at the office next door and at coffee breaks together. Also thanks for the nice dinners together with you and **Ale**. I would also like to thank **Jessica and Philip** for nice evenings together with self-brewed beers and all the chats. Thank you, **Nicole**, for your humor, advice and chats when I sometimes popped into your office or at lunch. Thanks to **Alena** for the nice cooking events and long Texel walks together. Also I would like to thank **Marijke** and **Laura** for a great time at EGU, **Milou, Fons** and **Sigrid** for being always great people to chat to over a coffee.

My paranymphs **Dina and Maaike** (aka the MMB wannabe fit girls): I´m so grateful that I met you two during my time at NIOZ. We always had each other's backs and I feel we have been on a long rollercoaster ride regarding our PhD

projects. We went together through quite some good but also bad times and I think without your friendship I would have been definitely sometimes lost on Texel. I´m very happy that we became friends and I´m looking forward to many more exciting times in the future.

Cait: Thank you for being such a great friend to me over the last years. We always had a good time at conferences together, on Sicily or just at our numerous dinners and coffee breaks. I´m very happy that we still talk so often even if it is only via videochat and share our research experiences but also private matters with each other. You are always a critical and honest friend whose opinion I value very much.

Uli: Danke, dass du meine deutsche Freundin auf der niederländischen Insel warst. Unsere unzähligen Gespräche über das Leben und unsere PhDs auf Texel haben mir vieles versüßt. Danke für die vielen Barbecues und Partys zusammen. Ich bin so froh, dass du nun wieder so nah bei mir wohnst und wir beide als Postdocs am AWI und MARUM arbeiten werden. Hoffentlich ziehen du und deine Familie ganz bald nach Bremen. Ich freue mich auf gemeinsame Zeit zusammen hier als Dr. Uli und Dr. Nadine.

Emily: Thank you so much for our vacations and conferences together as well as for a bed to stayover in Utrecht. You are an amazing friend to me and we always have so much fun together. Can´t wait for the exciting times together which lie ahead of us.

I would also like to thank all my NESSC colleagues and conference buddies for great times at several events and conferences together. Especially I would like to thank the conference crew for lots of fun times at IMOG and GRC: **Tommy, Hendrik, Jerome, Allix, Cait and Emily** and many more.

Danke an alle unsere/meine Bremen Freunde, welche immer ein offenes Ohr haben und mit denen ich viele schöne Tage und Abende erlebt habe. Speziell danken möchte ich **Rebecca, Merle & Bene, Corinna & Lars, Daniel V. & Julia, Maja & Jonas und Anna & Till** für ihre Unterstützung während meines PhDs.

Jenny: Was wäre diese Danksagung, ohne dir zu danken für deine jahrelange großartige Freundschaft. Auch wenn du meine Forschung nicht immer verstehst, bin ich dir dankbar für deine Unterstützung und deinen Rat. Egal was wieder Komisches passiert, wir verstehen uns immer und haben zusammen viele schöne, aber auch nicht so positive Zeiten geteilt.

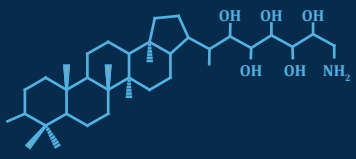
Als nächstes möchte ich meinem Partner und besten Freund **Stefan** von Herzen danken. Danke, dass du immer an mich glaubst auch wenn ich es manchmal selbst nicht tue. Danke für deine Unterstützung in den ganzen Jahren zusammen und besonders durch die letzten vier Jahre die uns vieles abverlangt haben. Nach PhD Stress plus Fernbeziehung Bremen-Texel kam auf uns Corona und vieles mehr zu trotz allem haben wir es geschafft zusammenzuziehen und vor allem immer zusammen zu bleiben. Ich kann mir keinen besseren Menschen vorstellen, mit dem ich zusammen sein möchte und der mich bedingungslos unterstützt und liebt. Danke für alles, auch wenn mir dafür die richtigen Worte fehlen. Ich liebe dich.

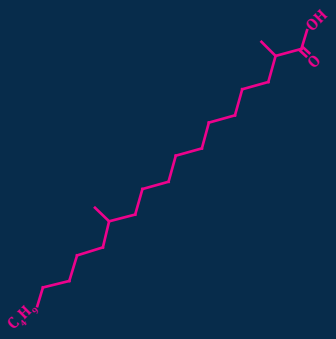
Danken möchte ich auch **Familie Wenau**, die mich herzlich in ihre Familie aufgenommen haben und die uns immer unterstützt. Danke für das großartige Flessenow, wo ich so manche Teile meiner Doktorarbeit geschrieben habe.

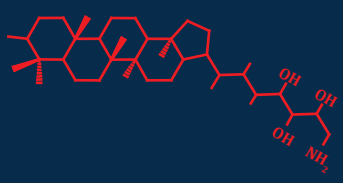
Zum Schluss möchte ich meiner Familie danken. Das größte Dankeschön geht an meine **Eltern**, die mich immer in meinen eigenen Entscheidungen unterstützt haben und den Grundstein für mein Wissen und meine Fähigkeiten mit ihrem Engagement während meiner Schulzeit gesetzt haben. **Mama**, danke für alles was du für mich und uns geleistet hast die ganzen Jahre. Du bist viel zu gut für diese Welt. **Papa**, danke für dein ganzes Wissen was du mir in den Naturwissenschaften vermittelt hast, wenn auch du nicht viel mit Chemie (meiner Leidenschaft) am Hut hast. Danke für deinen Rat und unsere unzählig lange Diskussion über die Welt. **Chantal**, danke fürs kleine Schwestersein und alles was damit zu tun hat ;). **Onkel Arno**, dankeschön für deine Unterstützung und Motivation auf allen meinen Lebenswegen. **Oma Lenchen**, danke für alles, deine Fürsorge als wir Kinder waren und dein Glaube an mich. Jetzt ist deine Enkeltochter sogar noch Dr. Nadine geworden! Schade, dass Opa viel zu früh gegangen ist, um dies mitzubekommen. Zuletzt auch ein großes Dankeschön an **Oma Edith** (auch wenn du ein paar Monate zu früh gegangen bist, um meine Promotion noch erleben zu können). Danke fürs Reden und manchmal einfach nur Zuhören. Viele Dank euch allen, ohne euch wäre ich nicht der Mensch der ich heute bin.

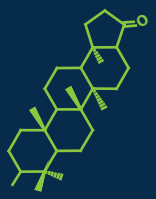
About the author

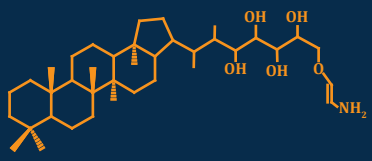
Nadine Talea Smit was born on the 7th of June 1990 in Meerbusch, Germany. From 2010 to 2013, she completed her bachelor's degree in Geoscience at the University of Bremen. During this time she worked as a student assistant in the Organic Geochemistry group of Kai-Uwe Hinrichs at MARUM, Bremen where she gained experience in laboratory skills and started to get fascinated by lipid biomarkers and isotopes. Her Bachelor thesis dealt with the calibration of the paleotemperature proxy TEX_{86} using pure archaeal cultures under different growth temperatures. Afterwards, Nadine attended the international MSc program of Marine Geoscience at the University of Bremen and at MARUM. During this time she worked together with Florence Schubotz and Christian Hallmann on several projects, gaining more insight into petroleum lipid biomarkers. Her thesis entitled "Geochemical characterization of asphalt deposits in the Campeche Bay (southern Gulf of Mexico) - Insights in the persistence of heavy oil in the marine environment" in cooperation with Shell Global Solutions in the Netherlands under the supervision of Florence Schubotz, Christian Hallmann and Johan Weijers was completed in 2016. Her keen interest in Organic Geochemistry led her to apply for a PhD position at the Royal Netherlands Institute for Sea Research (NIOZ) on Texel, the Netherlands. Nadine started her PhD studying novel lipid biomarkers for microbial methane oxidation in the environment within the NESSC program (Netherlands Earth System Science Center) in February 2017 under the supervision of Stefan Schouten, Darci Rush and Jaap Sinninghe Damsté. She will continue her scientific career as a postdoctoral researcher in a cluster project together with Victoria Orphan and Kai-Uwe Hinrichs at MARUM, University of Bremen, Germany and partly at California Institute of Technology, USA. In this project she will use diverse isotope tracing techniques to study e.g. methane oxidation and methanogenesis in aerobic and anaerobic marine sedimentary microbes.

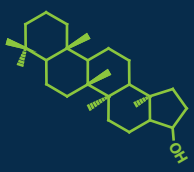


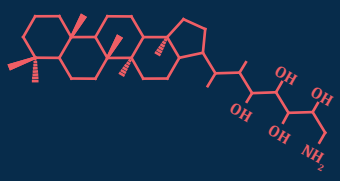


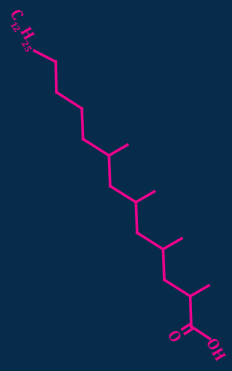


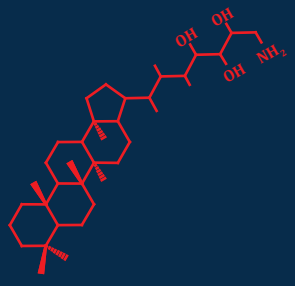


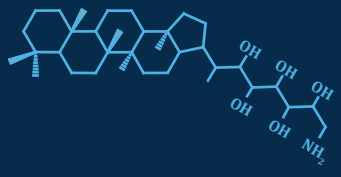


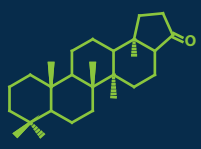


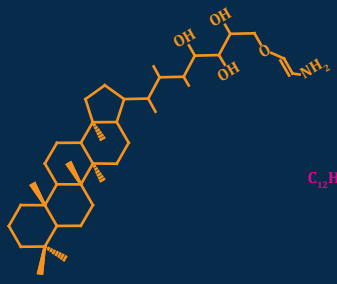


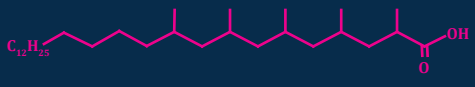


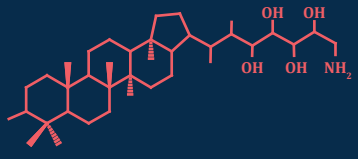












Royal Netherlands Institute for Sea Research & Utrecht University