**Non-invasive analysis of the soil microbiome: biomonitoring strategies using the volatilome, community analysis and environmental data**

Kelly R. Redeker1, Leda L. Cai1, Alex J. Dumbrell2, Alex Bardill1, James P. J. Chong1, Thorunn Helgason1

1- Department of Biology, University of York

2- School of Biological Sciences, University of Essex

**Contents**

1. Introduction 2

2. An overview of the soil volatilome 8

2.1 Microbes 10

2.2. Plants 11

2.3 Soils 12

3. Understanding Essex UK salt marsh sediments through the volatilome 14

3.1 Methods 15

3.1.1 The study sites 15

3.1.2 Sampling the volatilome and sediment in the field 16

3.1.3 Analysing volatilome and sediment samples 17

3.1.4 DNA extraction and sequencing 18

3.1.5 Microbial community sequence analysis 18

3.1.6 16S rRNA analysis 18

3.2 Relationships between environment, microbial community and the volatilome 19

3.2.1 Mud pan microbial communities across Essex salt marshes 19

3.2.2 Environmental conditions in sediments within Essex salt marshes 21

3.2.3 The volatilome across Essex salt marshes 22

3.2.4 Tying together the environment, volatilome, and microbial communities 23

4. Conclusion 27

References 31

**Abstract**

Within soils there are microorganisms that act to break down complex substrates (saprophytes), microorganisms that actively aid nutrient delivery (mycorrhizal fungi and nitrogen-fixing bacteria), and others that hijack the system to their own benefit (parasitic bacteria and fungi). The complex interaction between plants, these microbes and the soil determines how effectively nutrients will be recycled, with significant impact on regional productivity and biodiversity. Each microbe plays a role in overall soil function but, despite the critical role they play, soil microbial communities and their functions remain challenging to accurately quantify.

The functional behaviour of soils is difficult to quantify, in part due to the effects of disturbance when sampling. This suggests that non-invasive analytical tools are necessary to diagnose current soil function and to predict changes in soil behaviour with changing climate or land use. Microbial communities, the drivers of soil function, are diverse, and their individual metabolisms are often tightly coupled, such that the microbial community in aggregate may be considered to have a “net” metabolism. This net metabolism can be described by the volatile signatures that propagate from the soil in to the atmosphere and, by proxy, allowing a non-invasive analysis of the microbial community active in the subsurface.

Here, we detail here the complexities of the soil volatile metabolism, propose a “fingerprint” strategy to describe this complex community that uses trace gas fluxes combined with environmental data, and describe the promising outcomes from an initial foray using this method.

**1. Introduction**

Delivery of sustainable agriculture and effective conservation approaches rely on resilient soil functionality under a range of soil conditions (Kibblewhite *et al*., 2008). Understanding whether a soil is resilient and functional, often termed “soil health”, requires a broad understanding of physical, chemical, and biological soil properties (Allen *et al*., 2011). Many of the basic chemical and physical properties of soil are changing (driven by climate, land use, and agricultural pressures) with poorly understood impacts on soil biological communities and processes. Efficient and appropriate biomonitoring methods are needed to identify the ecological impacts of these changes and to manage ecosystems effectively before we lose functionality and critical ecoservice delivery.

Nearly all current methods used to explore these important subsurface aspects/processes require that the soil be collected and treated; thus modifying an ecosystem that is inherently sensitive to disturbance and producing potentially misleading conclusions (Choi *et al*., 2017). Furthermore, soil microbial communities are often monitored using DNA-based methods, which i) generally require substantial modification of the soil environment, ii) conflate the presence of a microbe with its activity, or functional impact and iii) tend to focus on community members with greatest abundance. In order to support sustainable management options and deliver better agricultural yields while maintaining soil capacity and function it is important to establish a non-invasive, biomonitoring methodology that allows characterization of soil microbial community and function without soil disturbance. We propose that the use of the volatile fraction of microbial metabolisms (the volabilome) allows us to by-pass many of these concerns. This non-invasive, surface sampling technique allows multiple analyses over time at the same location, reflects on-going microbial community function and activity, and therefore allows a more accurate assessment of the state of soils.

Soil volatile fluxes originate from microbial metabolic processes. Microbes in soil environments require constant metabolic activity; for cell maintenance, to grow and replicate, and to maintain functional capacity during periods of stasis. This metabolism either generates chemical outputs (as a metabolic endpoint, like CO2, or as a signalling compound, or for use in defence) or takes in chemicals for purpose (loose amino acids can be harvested for nitrogen or direct use, or sugars can be harvested for carbon/energy). These resources are obtained from, or are released into, the soil matrix (Box Figure 1).

The soil matrix consists of solid soil particles, soil particle aggregates, and open air- and/or water-filled volumes, called pore space. The soil pore space is important for root growth, water movement, water storage, and gas movement within the soil. Soils interact with the air above them through (primarily) diffusive transport, exchanging topsoil pore space air with the air lying over the soil surface. The topsoil pore spaces interact with the pore spaces in the layers below, leading to overall material movement throughout the soil column (Redeker *et al*., 2015). This air exchange allows organisms within the soil to interact with the overlying air, either through delivery of excess metabolic by-products from soil organisms to the atmosphere (e.g.- carbon dioxide from plant and microbial respiration) or through consumption of delivered chemical compounds via atmospheric air (e.g.- methane consumption in aerobic forest soils) (Conrad, 1996). This chemically complex combination of efflux (soil-to-atmosphere) and influx (atmosphere-to-soil) has been termed the “soil volatilome” and has been proposed as a tool to non-invasively study soil health and soil microbial communities (Insam & Seewald, 2010; Muller *et al*., 2013; Schmidt *et al*., 2015) (Box Figure 1).

The volatilome has already been used for some single-species diagnostics under controlled conditions. In particular, pathogens have been identified through headspace analysis of single organism cultures (headspace analysis is the practice of containing organisms in trapped air volumes and sampling that air volume at the beginning and end of a set period of time)(Box Figure 2). These analyses have led to the creation of profiles of specific volatiles and/or blends of volatiles by which contamination by these pathogenic organisms may be confirmed. Medicine, food, air and water have all been tested for contamination through analyses of these signature volatiles (Insam & Seewald, 2010; Garcia-Alcega et al., 2016 and references therein). Similarly, soil analyses have been performed to detect agricultural plant pathogens and moulds in soils through their identified volatile signatures (Penuelas *et al*., 2014 and references therein). It has been proposed that bacteria transformed with a plant-derived methyltransferase enzyme can be used to report on sub-surface microbial gene expression dynamics through volatilized methyl halides (Cheng *et al*., 2016).

Moving from single-species identification to determination of soil microbial communities through the volatilome has been challenging, although some environmental conditions allow simple functional assignments. Anoxic, saturated subsurface environments may support methanogens, which can be confirmed through methane release from soils (Le Mer & Roger, 2001), whereas methane, methanol or methyl halide (CH3X, where X = Cl, Br or I) uptake by soils indicates methylotrophs active within a (most likely) aerobic subsurface (Chistoserdova *et al*., 2009). In plant-free soils a substantial efflux of methyl halides appears to be indicative of aerobic conditions and fungal activity (Watling & Harper, 1998; Redeker *et al*., 2004; Redeker & Kalin, 2012).

The signal received at the soil surface is the net combination of all volatile metabolisms from all subsurface organisms, often acting counter to each other (e.g.- metabolites produced by one organism affect or are consumed by another)(Box Figure 1). The identification of soil microbial communities through the volatilome has been hampered by this complexity and requires advanced statistical/mathematical methods to disentangle the disparate signals received at the surface. To use these tools we must have a greater understanding of the emissions/consumption profiles for many of the organisms expected within the soil column.

Comprehensive databases that link microbes to defined metabolic uptake and outputs are crucial for effective biomonitoring, since they provide the metabolic profiles that allow specific identification of community (functional) members. Comprehensive volatile databases combined with non-invasive, surface-based sampling may be of more use in determining sub-surface microbial communities than DNA-based methods, analysed through extant 16S databases since non-invasive methods are linked to ongoing microbial activity as opposed to DNA methods which quantify the presence of organisms in the soil. Therefore, even if we accept the problems of disruption caused by sampling soils for DNA-based outcomes, DNA-based methods are still unlikely to provide data on which organisms remain active within the soil profile, which is often the specific information required. Unlike genomics (DNA-based), both transcriptomics (based on RNA) and proteomics (based on expressed proteins) approaches can probe microbial function. However, both still require soil disturbance for analysis, limiting their efficacy for understanding *in situ* soil community structure and function. They may be of greater use in creating effective databases linking microbial community members to specific metabolic outcomes. At least one database for the volatilome has been initiated. It incorporates published volatile emissions from over 10,000 microbial specimens (Lemfack *et al*., 2014).

The available databases for volatile profiles are still in the formative stage. Within the Lemfack *et al*. (2014) database the organisms reported are those organisms that are culturable, and the reported volatiles are those of sufficient concentration that they may be sampled, separable and identifiable through the methods chosen. This tends to bias the collected VOC profiles towards subsets of chemical compounds observable with commonly used methods (hydrocarbons, terpenes, alcohols, aldehydes and ketones) and the organisms to those that may be cultured within the lab, known to be less than 30% of the total microbial communities found within soils (Alessi et al., 2018). Microbes and plants have been shown to differ strongly in the total amount and the profile of sulfur-bearing, halogenated and metallic compounds volatilized (Gribble, 2003; Redeker *et al*., 2004, Bentley & Chasteen, 2004; Parks *et al*., 2013) suggesting that these types of compounds may have substantial predictive benefit, yet these are not fully represented in current databases since the relevant analytical methods are not as commonly used. It is also worth noting that the database focuses on produced volatiles and does not currently include consumed volatiles, a critical gap that needs to be addressed in order to understand complex metabolite interactions within soils. Despite current limitations, the database has confirmed that some species can be predictably identified from VOC profiles (Lemfack *et al*., 2014; Penuelas *et al*., 2014). The challenge remains, however, to identify microbial communities within the subsurface based entirely on the net sampled volatilome.

The challenge has been aided by the development of new soil community databases and paired isotopic analyses. These approaches are inappropriate for long term profiling of soil communities (either due to their high specificity or their disruptive nature) but they allow greater scope and confidence in assignments made through our proposed non-invasive volatile profiling. For instance, greater specificity of identification within the volatilome can be obtained from isotopic analyses of emissions, either singly (Redeker & Kalin, 2012) or paired (Oduro *et al*., 2013). The most common example of function-specific isotopic information comes from the methane cycle, where methanogens and methylotrophs generate very different isotopic profiles (von Fischer & Hedin, 2007) however species-specific isotopic signatures have been observed for methyl halide emissions from plants and fungi (Harper *et al*., 2001; Saito & Yokouchi, 2008). Paired isotopic analyses provide greater specificity than individual isotopic analyses, having been used successfully to characterise cycling of bacterial sulfur-volatiles (Oduro *et al*.,2013) and to analyse nitrous oxide emissions from nitrification and denitrification pathways (Zhang *et al*., 2016).

Furthermore, recent studies using soil-disruptive ‘omics technologies and databases derived from these data allow us to identify which members of the soil microbial community are more likely to co-exist, and under which environmental conditions (Fierer *et al*., 2009; Griffiths *et al*., 2011; Serna-Chavez *et al*., 2013). These data, in combination with local soil environmental data, may be used either i) to generate more precise microbial community predictions from non-invasive, non-destructive volatilome analyses or ii) to ground truth volatilome-derived community assignments.

Even with these additional technologies, successful characterization of soil microbial communities using the soil volatilome requires a fingerprint strategy. Criminal forensic fingerprinting (used to profile human culprits) relies on several dozen distinct minutiae, or patterns within a fingerprint, to accurately identify specific individuals and to avoid false positives. We expect that soil communities will also require multiple identification points to reduce the possible range of community assignments and provide greater accuracy in identification. At the moment only broad strokes descriptors of microbial communities can be depicted, with basic environment, functionality and family level organisms identified (e.g.- anoxic environments leading to methane generation by methanogenic archaea), but as databases incorporating functionality, favourable environments of specific organisms, likely co-existance of community members and species-specific volatilomes are developed more identification points may be added and genera or species level identification may be possible.

This paper will explore the utility of a volatile “fingerprint” approach to describing microbial communities. We begin by outlining the current state of knowledge regarding how major biological communities influence soil and sediment volatiles. We include a discussion of interactions driven purely through soil chemical and physical properties as well as an examination of complex environment/community interactions. We will then describe the methods and approaches used in a study that characterized UK salt marsh sediment microbial communities through analysis of environmental parameters, microbial community and a subset of volatile fluxes. The results of this study are presented to demonstrate the potential of the volatilome to coherently and self-consistently predict members of the soil community and the sediment environment.

**2. An overview of the soil volatilome**

Components of the volatilome can be loosely classed based on their function within the soil. The volatilome includes many basic metabolites and they interconnect with all major biogeochemical cycles (Falkowski *et al*., 2008). These metabolites include end products (e.g.- CO2, Yuste *et al*., 2007; CH4, Le Mer & Roger, 2001), intermediates (e.g.- N2O, Stein & Yung, 2003; acetic acid; Krzycki & Zeikus, 1984 and others; Penuelas *et al*., 2014) and consumed metabolites (e.g.- CH4, Le Mer & Roger, 2001; and other methyl donors; McAnulla *et al*., 2001). Metabolites can also fall into multiple categories, for instance methane and nitrous oxide are produced and consumed by different organisms within the same soil profile. The consumption of volatiles is often overlooked in discussions regarding the volatilome, despite the important cues that this provides for soil pH, anoxia, redox potential, and nutrient status.

VOCs produced by one group of organisms can promote or inhibit other organisms and their behaviours. For instance, some bacterial volatiles have been shown to promote growth in *Arabidopsis thaliana* (Ryu *et al*., 2003; Zhang *et al*., 2007) while volatiles from an Ascomycete (*Fusarium culmorum*) induce greater motility in some strains of bacteria (Schmidt *et al*., 2016). A number of plausible antifungal and antibacterial compounds have been identified from plant, fungal and bacterial emissions (Schmidt *et al*., 2015). These compounds do not consistently generate significant reduction in bacterial or fungal load, perhaps i) due to differences in soil chemistry/reactivity between study sites, ii) due to spatiotemporal variation between regional soil microbial communities or iii) due to greater efficacy of the antimicrobial compound when combined with other volatiles emitted from other members of the microbial community (Gallucci *et al*., 2009; Hemaiswarya & Doble, 2010). Although acute impacts of these produced antimicrobial compounds appears to be limited, long term, chronic effects from these compounds on microbial activity and composition may still represent an important component of ecosystem function by modifying local plant-microbe interactions.

Volatiles act to communicate information within soils, generating defence and/or stress tolerance responses which have been reported for plant-plant (Pierik *et al*., 2014), bacteria-plant (Toljander *et al*., 2007), bacteria-bacteria (Schmidt *et al*., 2015, Schulz-Bohm *et al*., 2015), and fungi-bacteria (Scmidt *et al*., 2015, Schulz-Bohm *et al*., 2015) interactions. More complex, multitrophic signalling has also been observed, as between plants signalling predatory nematodes to consume herbivorous insects (Ali *et al*., 2011), or between virally-infected plants signalling nematodes to provide transportation for the virus (Turlings *et al*., 2012). These signalling volatiles are targeted towards specific organismal interactions and are therefore some of the most likely components of the volatilome to provide species-, or genera-, specific identification when they have been characterized.

While the primary actors within soils that influence the volatilome are microbes and plants other biological actors can affect these primary actors through predation (Griffiths *et al*., 1999) and viral attack (Pan *et al*., 2014), and these may have significant impact on consumed and produced metabolites.

**2.1 Microbes**

The microbial community in soils consists of four separate (large in themselves) sub groups, including bacteria, fungi, archaea and protists. These organisms fulfil five primary functional roles as autotrophs (not fungi), decomposers, mutualists, predators/pathogens (not archaea) or lithotrophs (not fungi or protists) which ties them intimately to the biogeochemical cycling of chemical elements in soils. Many reviews and studies exist for bacterial volatiles (e.g.- Insam & Seewald, 2010 and references therein) but far fewer for fungi (and those are limited to pathogenic, ECM or saprophytic fungi- very little or no data exists for arbuscular mycorrhizal fungi despite their critical importance), and importantly there are no substantial reviews regarding non-greenhouse gas VOC emissions from archaea (but several reviews regarding their capacity to live in extreme environments, leading to potentially unique metabolisms with unique volatile signatures).

Attempts to classify functional group or microbial type by their volatile profile have recently been published (Muller *et al*., 2013; Penuelas *et al*., 2014). These studies have either performed analyses or collected reported fluxes for comparison across a range of organisms. Both studies demonstrate the challenge; most volatile metabolites are common across all microbial life. However, it is clear that fungal groups (ECM, pathogenic and saprophytic; Muller *et al*., 2013) or different classes of bacteria and fungi (Actinobacteria, Ascomycetes, Bacteroides, Cyanobacteria, Firmicutes, Alpha-, and Gammaproteobacterial; Penuelas *et al*., 2014) may be separable through their distinct volatile profiles. There exist differences in the reported signature volatiles between studies however, for instance sesquiterpene compounds were common across all fungal groups tested in the study performed by Muller *et al*. (2013) whereas in Penuelas *et al*. (2014) the fungal group was identified by their oxygenated metabolites (including alcohols, ketones and furans) and terpenes were not indicated. Also, Penuelas *et al*. (2014; and associated database) does not appear to include the broad range of halogenated compounds generated by fungi and bacteria as reported by Gribble, (2012) and Field, (2016).

It is probable that the differences in reported identifying compounds is due to one or a combination of i) context-specific responses, ii) different organisms studied within the same broad classification and iii) differing methods of analysing the volatiles released from these organisms. For example, the method used by Muller *et al*., (2013) preferentially selects for less volatile, more hydrophobic compounds leading to a probable underestimate of low molecular weight, more polar, halogenated volatiles and emphasizing the less polar, less volatile, higher molecular weight sesquiterpenes. This highlights one of the primary challenges facing the volatilome community. Multiple methods of analyses targeting different chemical families and functions are necessary in order to identify the most useful compounds for identification, which will then allow us to characterize the soil microbial community in the most versatile and robust manner.

**2.2 Plants**

On the whole, volatile metabolites generated by aboveground plant tissues have been intensively studied (e.g.- Loreto *et al*., 2006; Matsui *et al*., 2012; Niinemets *et al*., 2013; ul Hassan *et al*., 2015). Subsurface plant volatile interactions are less intensively studied than aboveground emissions but some aspects have been documented, including plant-microbe, plant-pest and multitrophic interactions (Massahla *et al*., 2017).

Root emissions are quite different from volatiles emitted by aboveground tissues and the environment in which they exist affects the manner in which these volatiles move or are metabolised (Read *et al*., 2003). Root exudates influence their local environment, producing acids and surfactants to aid nutrient acquisition (Jones, 1998; Mukherjee & Lal, 2013), toxins to defend against herbivory (Jassbi *et al*., 2010), and compounds that attract predators that consume herbivorous pests (Ali *et al*., 2011). Roots also influence their local environment by removing water through transpiration (Verstraeten *et al*., 2008; Nobrega *et al.*, 2017) and mobilization/translocation of water through root networks (Prieto et al., 2012), leading to modification of available pore air space within the rhizosphere which may affect volatile compounds diffusivity prior to soil microbial consumption, chemical degradation or emission to the atmosphere (Redeker *et al*., 2015).

Aerenchymous plant roots (aerenchymous plants have large air spaces in leaf, stem and root tissues allowing easier air passage between plant components) influence the local volatile environment through provision of a transit route into/out of the soil that is more rapid than diffusion through waterlogged soils (Laanbroek, 2010). The role of aerenchymous roots is to provide ready access of atmospheric oxygen to the subsurface plant components, and this allows subsurface volatiles to escape using the same transit route. Direct and indirect influences on soil hydrology/diffusivity from physical and chemical root impacts are varied and have been the subject of multiple studies and reviews (e.g.- Bodner *et al*., 2014 and references therein).

While root-based emissions are not as well profiled or quantified as aboveground emissions there are several families of plants that are known for specific trace gas emissions. The brassicas (Family Brassicaceae, the cabbages) in particular are known for their production of volatile, sulphurous compounds that are characteristic of the family (Virtanen, 1965). These sulphurous compounds are also produced by Brassicaceae root tissues and specifically in the form of an herbivory response that generates isothiocyanates (through enzymatic degradation of glucosinolates) (Doheny-Adams *et al*., 2017). Sulfur metabolisms have been linked to halogen-compound metabolisms through halide/thiol methyltransferases, and the Brassicaeae have been reported to be the plants that produce the most methyl halides per gram plant tissue, although this has yet to be tested substantially on root tissues (Saini *et al*., 1995).

**2.3 Soils**

Soils and sediments should be considered multi-layered compartments. In each compartment, biogenic and abiotic production may be occurring, along with metabolic consumption and abiotic degradation (although biological processes tend to dominate; Redeker & Kalin, 2012). Simultaneously each compound will be diffusing along a concentration gradient between its current location and the soil surface (Table 1; Box Figure 1). This diffusive rate determines the overall degradation and consumption of the compound of interest since it regulates the contact period with active sites and organisms within the soil matrix. It is estimated that as much as 90% of methane generated in the subsurface is consumed before it can be emitted at the surface whereas ebullition and/or transport through aerenchymous plant tissues, which lowers transit time, substantially reduces this metabolic loss (Zhu *et al*., 2012). The overall flux of compounds from the soil surface, the volatilome, will be determined by this balance between production, consumption, and rate of movement (Box Figure 1).

The soil and sediment matrix through which volatiles pass determines the residence time of the compound within the soil (Redeker & Kalin, 2012). Greater porosity leads to more rapid diffusion while greater pore water content reduces diffusive rates (Redeker *et al*., 2015). Surface winds may reduce residence time through advection or rotational pumping of pore air spaces (Redeker *et al*., 2015). Soils produce many volatiles through abiotic reactions and most biogenic compounds will react in the soil environment given sufficient time (Insam & Seewald, 2010).

The soil environment may directly influence the volatilome by affecting one or more of the above processes. Indirect effects from the soil environment are also possible. Abiotic stress tends to lead to changes in emissions, for instance drought and flooding have been shown to modify the type and extent of volatiles generated by bacteria (Asensio *et al*., 2007). Soil saturation and drought fill and empty soil pore spaces, leading to reduced and enhanced diffusivity respectively. Extreme drought can also lead to cracking at the surface, providing deep advective access to the subsurface, and creates open channels within the soil where root-soil connections are lost as the soil shrinks. Soil salinity, pH, nutrient status and temperature have been shown to be determining factors for soil microbial community composition (Lozupone & Knight, 2007) and therefore function (Nicol *et al*., 2008; Jones *et al*., 2014). Redox status also substantially influences the volatiles produced and released from the soil/sediment system (Devai & DeLaune, 1995). Greater surface exposure leads to enhanced winds over the soil surface, acting to remove volatiles from the soil rapidly and reducing the soil concentration gradient (Redeker *et al*., 2015).

**3. Understanding Essex UK salt marsh sediments through the volatilome**

Multiple challenges remain to the use of the volatilome as an effective, non-invasive tool for diagnosing soil microbial communities. The databases that are currently established need to be maintained, effectively curated, and expanded to include the volatilome of more individual organisms, across a broader range of volatiles, and to include consumption of volatiles as well as emission.

This database approach is currently limited by our capacity to culture soil microorganisms in the lab. Most soil organisms are not currently culturable and are unlikely to become so. It is therefore important to develop ‘omics techniques sufficiently that a full range of functional behaviour is ascribable, even when encountering unusual and/or rare organisms. During this development period, we will need to target functional behaviours of specific, poorly-understood organisms which lab-based culturing and ‘omics approaches are capable of resolving in order to provide maximum database coverage.

To develop an effective, volatilome-centred, non-invasive approach to monitoring microbial communities, in which specific volatile fluxes are linked to individual functions or organisms, it is important to place these micro-organisms in environmental and ecological context. Some organisms (e.g.- aerobes) will rarely co-habit in large numbers with other organisms (e.g. anaerobes) dependent upon both competitive and environmental pressures. By understanding organisms within a community and environmental context we can constrain other probable community inhabitants once a detected volatile signature is ascribed to a specific organism. Therefore, databases should also begin to incorporate environmental parameters and microbial community members with which the organisms (and their volatile fluxes) are associated. It remains an open question as to whether the fingerprint strategy can be applied across a range of soil types and conditions, or whether it will be ecosystem specific. To address this question, it is important to begin to apply this solution to a range of soils and sediments to observe the consistency of outcomes across a range of ecotypes.

We describe below the first effort to integrate many of these necessary components using data from Essex, UK salt marsh sediments. We have placed community members in context, both in terms of their preferred local environment and in terms of their common microbial (and plant) associates through genomics analysis of sediment paired with aboveground biomass and several important sediment characteristics. We have measured the net flux of a number of informative volatile compounds, taking advantage of both emission and consumption profiles. We discuss our preliminary findings in terms of volatile predictions of community members and environmental parameters.

**3.1 Methods**

*3.1.1 The study sites*

Three paired (natural and managed) salt marsh sites were selected along the Essex coastline within the Colne and Blackwater estuaries (Fig 1). The Colne Estuary provides a particularly rich resource for understanding how microbial communities have adjusted with historic environmental change (Nedwell *et al*., 2016). Sites included a location within the Fingringhoe Wick Nature Reserve (Fingringhoe Wick Range: 51° 49' 53.95'' N 0° 58' 11.54'' E), one near Mersea Island (51° 47' 50.29'' N 0° 55' 16.68'' E) and the third was located at the Essex Wildlife Trust Abbotts Hall Farm (51° 47' 11.95'' N 0° 51' 40.78'' E). Managed saltmarshes were paired with adjacent natural salt marshes and elevations of sampling sites were matched within 20cm among marshes.

[Insert Figure 1 here]

*3.1.2 Sampling the volatilome and sediment in the field*

Microbial communities, environmental parameters, and volatile fluxes were based on sequences and samples obtained from paired natural and managed salt marsh sediments (5-10 and 10-15cm depths) sampled once per season between July 2014 and June 2015. Four different vegetation-dominant communities were studied, including *Atriplex* spp, *Limonium* spp, *Puccinellia* spp., and unvegetated mud. Comparative volatile fluxes between natural and managed salt marshes were taken on the same date while sediment cores to describe local sediment characteristics were taken from within the volatile sampling “footprint” (Box Figure 2). Aboveground biomass and sediment microbial community cores were also taken from within the volatile sampling footprint directly after volatile sampling.

Seasonal samples (Sum 2014 – Spr 2015) for a limited suite of volatilome fluxes (CH4, CH3Cl, CH3Br, CH3I, dimethyl sulfide, isoprene, and CHCl3; Table 1) were taken using standard static chamber enclosures in each marsh for each of the four dominant vegetation covers with three replicates for each. Chamber bases were pinned to the sediment to avoid cutting effects on roots, but otherwise methodology followed the same protocol as described in Redeker & Kalin, 2012 (Box Figure 2). Isoprene, sulfur- and halogenated compound samples were stored in electropolished stainless steel canisters until analysis, within a two-week period in which the canisters have been demonstrated to be stable for reactive compounds (Low *et al*., 2003). Methane samples were stored in 12 ml exetainers until analysis.

[Insert Table 1 here]

Vegetation was harvested from within the volatilome chamber base after trace gas sampling. Once the chamber footprint was clear two sediment cores were taken from the central region of the chamber, one for environmental parameters (e.g.- pH, granulometry, nutrients, bulk density) and the other for microbial community analysis.

Sediment cores were stored on ice until arrival at the University of York where they were stored at 4°C and processed within 48 hours. Sections of the sediment cores between 5-10 cm and 10 -15 cm were selected to avoid root contamination and subsamples from within the interior of the cores at these depths were collected with aseptic technique, then frozen in liquid nitrogen immediately and stored at -80 °C until DNA extraction. Sediment parameter samples were collected from the same depths as the microbial community samples but were processed for granulometry, pH, bulk density, and nutrient measurements according to standard methodologies.

*3.1.3 Analysing volatilome and sediment samples*

The reactive volatilome samples (isoprene, halogen and sulfur compounds) were condensed onto a liquid nitrogen condensation trap and transferred to a Restek© PoraBond Q column (30m, 0.32 mm ID, 0.5 um thickness) within a HP 5972 MSD running in selective ion mode. Greenhouse gas samples were run on a PerkinElmer Autosystem XL GC ECD/FID using a hand-packed Restek© PoraPak Q column (2m, 1/8” OD, 50-80 mesh). Analyses and calibration of trace gases followed the methodology described in Redeker *et al.*, 2007 (Box Figure 2).

Bulk density, nutrient, and pH processing were performed on samples immediately after removal from the sediment cores. Samples were dried at 70oC for seven days and dry weight divided by known sample volume (corrected for compression during coring) was used to calculate bulk density. To normalize the ionic strength of the salt marsh sediment samples each was placed in a 1.0 M KCl solution prior to analysis on a Jenway Ltd 3310 pH meter. Sediment samples were treated with hydrogen peroxide (30% H2O2 concentration) to remove organic matter prior to grain size analysis on “Malvern Mastersizer 2000” (UK manufacturer, Malvern). Nutrient samples were prepared in accordance with Houba *et al*., 1995 with 1.0 M KCl, and stored at -20oC until analysis. Analysis for NH4+ and NO2-/NO3- was performed using a Seal Analytical AutoAnalyzer3.

*3.1.4 DNA extraction and sequencing*

NGS-based analysis of soils and sediments provides the current gold standard of microbial community (based on presence) and functional potential but does not allow repeat sampling for non-destructive applications (Jackson *et al*., 2016; Derocles *et al*., 2018). In this early stage of volatilome assessment we have aligned the observed volatilomics with this excellent data source to overcome incomplete volatile database issues. Total DNA metagenome sequencing was applied to avoid primer based artefacts in the data and to gain additional information on gene functions. Total DNA was extracted using the MoBio Powersoil kit (now Qiagen DNeasy Powersoil kit), using 250mg of soil. Samples were quantified by Qubit fluorometric quantitation (Thermo Fisher), diluted down to 0.2 ng µl-1 and requantified (again by Qubit). A total 1 ng DNA was used in library preparation using the Nextera XT DNA library prep kit v2, with set A indices (dual 8bp indices). The library preparation size selected the DNA for 300 to 600 bp in size. Final libraries were quantified by Qubit on HS DNA bioanalyzer chips prior to pooling, at approximately equimolar ratios for sequencing. All samples were run on a single lane on an Illumina HiSeq 3000 using 150bp paired end reads.

*3.1.5 Microbial community sequence analysis*

Sequencing produced a total of 328 million pairs of reads, ranging from 3.9-13.2 million read pairs per sample (mean 8.65 million). Reads were adaptor-trimmed using cutadapt (Martin, 2011) and uploaded to the OneCodex analysis platform (Minot et al., 2015). Taxa abundances were estimated from the number of matches to the OneCodex database. Taxa abundance tables provided by OneCodex were converted to BIOM format and correlations with environmental data were calculated using the observation\_metadata\_correlation.py script in QIIME (Caporaso et al., 2010).

*3.1.6 16S rRNA analysis*

Read files were separated by site and season before being assembled into contigs using Megahit (v 0.3.3-a, Li *et al*., 2015). Contigs greater than 30kb were removed using a custom Python script (see Supplementary Data)(<0.02% of contigs from each sample) before contigs containing 16S/23S and 18S/28S sequences were identified using SortMeRNA-2.1 (Kopylova *et al*., 2012) then classified using SSuMMo (Leach *et al*., 2012) and visualised using the Interactive Tree of Life (Letunic and Bork, 2016) (Fig 2, see Supplementary Material for a full scale image of Fig 2).

[Insert Figure 2 here]

**3.2 Relationships between environment, microbial community and the volatilome.**

Communities, environmental parameters, and volatile fluxes described below are based on sequences and samples obtained from natural and managed salt marsh sediments (5-10cm depth) during July/August 2014 (Summer) and February/March 2015 (Winter). To simplify we will not consider the sediment under the influence of plants and vegetation within the salt marshes and have focused entirely upon mud pan samples (See Box Figure 2). Figures 2, 3 and Tables 2, 3 demonstrate how the microbial communities, volatilome, and environmental parameters behave across space and time.

[insert Figure 3 here]

*3.2.1 Mud pan microbial communities across Essex salt marshes*

Even after aggregating sequences derived from three replicate mud pan samples per location and sampling date the sediment microbial communities show differences in space and time (Figs 2, 3). While individual species changes are too numerous to explicitly detail (Fig 2) there are several broad trends that can be easily observed. The Actinobacteria are much more represented at the managed Abbotts Hall site than in either managed Fingringhoe Wick or Mersea Island samples or than in any of the natural site samples (Fig 3). There are greater numbers of representative species from within the Rhodobacteraceae, Gammaproteobacteria and Epsilon/Deltaproteobacteria in the Fingringhoe Wick mud samples, particularly when compared to the samples from Abbotts Hall (Fig 3). Conversely, Fingringhoe Wick mud samples have poor representation in members of the Bacteroidetes relative to Abbotts Hall and Mersea Island (Fig 2).

The NGS sequence data was dominated by bacterial sequences, and more sediment samples will need to be analysed to obtain functional information regarding the rarer species. A greater number of samples is also needed to provide statistically useful volatilome/rare microbiome comparisons. For instance, the only identified, moderately abundant fungal community member was an uncultured Ascomycete, which was only present in realigned sites (Fig 2; AH Summer and FRW Winter). Archaeal community members were more commonly distributed, with notable differences including greater Halobacteriaceae presence in winter sampling periods and a reduced abundance of the Chrysiogenaceae in Fingringhoe Range Wick relative to their abundance in Abbotts Hall marshes (Fig 2).

Two of the microbial community organisms with greatest sequence representation are obligate anaerobes that lie within the Phylum Chloroflexi in the family Anaerolineaceae (Fig 2). However, in Mersea Island and Fingringhoe Wick the dominant Anaerolineaceae genera are *Levilinea* while in Abbotts Hall the dominant genera is *Longilinea*. The primary description of these organisms suggests that they have similar optimal temperatures (37oC) and pH (7.0) for growth but that *Levilinea* has a greater temperature range (25 -50oC) but a smaller pH range (6.0 – 7.2) relative to *Longilinea* (30 – 40oC and 5.0 -8.5 pH respectively) (Yamada *et al*., 2006; Yamada *et al*., 2007).

*3.2.2 Environmental conditions in sediments within Essex salt marshes*

Other broad trends reveal themselves when comparing environmental parameters obtained from the replicate mud pan samples (Table 2). Abbotts Hall samples from natural salt marshes tend to be more acidic than those from Mersea Island or Fingringhoe Wick Range. Managed sites tended to be more pH neutral than the natural marsh samples and overall there was a bimodal distribution of pH, with the majority of samples falling within a pH range of 6.5 to 8.0 (n = 20), however a substantial minority of sediment samples were acidic, falling in the pH range of 2.5 to 4.5 (n = 10) (Figure 4). The differences in pH within Essex sediments may have some explanatory capacity for the observed variance in Anaerolineaceae genera (phylum: Chloroflexi) since the lower pH sediments tend to harbour the *Longilinea*, which have been shown to tolerate a broader range of pH in lab cultures (Yamada *et al*., 2007).

Differences in bulk density within sediments also appears to have explanatory power (Table 2, Figure 3). Abbotts Hall sediments in managed salt marshes were consistently the densest of all sediments, with Fingringhoe Wick samples the least dense. This is opposite the trend observed in the natural marshes, where Abbotts Hall samples tended to be least dense and Fingringhoe Wick samples most dense. While the observed changes in sediment density in natural sediments does not appear to have substantial impacts, the high density managed sediments from Abbotts Hall clearly have greater Actinobacterial presence, suggesting that Actinobacteria outcompete other Classes in high density sediments.

Nitrogen availability and form varies between locations and season but is not obviously connected to microbial community (Table 2, Figures 2, 3). While there is great variation between samples taken from different locations in the same salt marsh on the same day there appears to be a trend in ammonia concentrations within natural salt marshes such that Abbotts Hall sediments have the greatest concentrations and Mersea Island sediments the least. Natural sediments demonstrate a strong seasonality, with greater concentrations of sediment-bound ammonia in the winter relative to the summer. In contrast there is no obvious seasonal trend in ammonia within managed salt marsh sediments.

There was significantly less NOx species when compared to ammonia in all sediments with the exception of the natural Fingringhoe Wick site. In all others oxidized nitrogen is less prevalent than reduced and there is no obvious seasonal trend in either natural or managed sediments.

*3.2.3 The volatilome across Essex salt marshes*

We measured fluxes of seven biologically informative trace gases (Table 1) from Essex salt marsh sediments, quantifying the net metabolism of these sediments over four seasons, three sites, and two management conditions. Fluxes of volatilome constituents from sediments are as variable as subsurface microbial communities and environmental parameters. There were significant differences in the fluxes between compounds, sites, and seasons (Fig 5; Table 3). Both net production of compounds and net consumption of compounds was observed. DMS and methyl iodide were both consistently emitted from salt marsh sediments across all sites and seasons, while methane tended towards efflux during winter and towards influx during the summer sampling period. Unsurprisingly the greatest effluxes and influxes are observed for methane (>10 mg CH4 m-2 day-1), but dimethyl sulphide follows directly with average and consistent efflux across all sites, at some locations exceeding 2 mg DMS m-2 day-1.

Methane behaves consistently between the natural and managed salt marshes; in both cases there is uptake of methane during the summer (-14±10 mg CH4 m-2 day-1) and efflux of methane during the winter (15±40 mg CH4 m-2 day-1). Methyl bromide is emitted by managed salt marshes (1.4±1.5 ug CH3Br m-2 day-1) while fluxes from natural salt marshes are, on average, not different from zero (0.5±3.3 ug CH3Br m-2 day-1).

[insert Figure 4 here]

While most other compounds were too variable between sites to generalize results from management style or seasonality there were some site-specific fluxes of interest. Abbotts Hall produced the greatest effluxes of methyl iodide across the study (managed: 11±9 ug CH3I m-2 day-1; natural: 5±9 ug CH3I m-2 day-1) as well as the most dimethyl sulphide (1.9±1.5 mg DMS m-2 day-1 both natural and managed), and the Abbotts Hall managed site was the only site to consistently generate and export chloroform from the sediment surface (43±10 ug CHCl3 m-2 day-1 during winter). Mersea Island sites generated the least DMS (0.2±0.1 mg DMS m-2 day-1) and Fingringhoe Wick sites produced the least methyl iodide (0.7±0.5 ug CH3I m-2 day-1).

At this gross, regional scale of analysis the links between volatile emissions and microbial community members are quite weak. Coastal sites rich in Bacteroidetes, known isoprene producers, do not generate more isoprene than other locations and sulfur-volatile generating bacteria from the Gammaproeobacteria, the Firmicutes and Actinobacteria are not strongly linked to lesser or greater dimethyl sulfide fluxes.

*3.2.4 Tying together the environment, volatilome, and microbial communities*

Comparison of microbial communities, environmental parameters and volatile fluxes is hampered by the substantial variability exhibited by natural and managed sediments. This variability across space and time does, however, provide a continuum of environmental conditions and volatile fluxes that can be correlated against microbial sequence abundance. When all sediment data is combined and analysed through Pearson correlations we can explore and confirm interactions suggested by the aggregate analysis. We focus here on the strongest correlations but, due to the often non-parametric nature of the data and limited replicate numbers, we were unable to completely correct for variable (environment, volatile, microbe) interactions. Therefore, we may underestimate some of the listed probabilities due to type I statistical errors.

It is important to note in the following discussion that we are discussing greater or lesser abundance of sequenced DNA as opposed to presence versus absence and correlations as described here are based on these relative abundances. As is clear from Figures 2 and 3 all major Classes are present in all sediment samples (Bacterially dominated, but inclusive of Archaea and Fungi), if not inclusive of every individual species observed. Therefore, positive correlations suggest greater abundance rather than sudden appearance and negative correlations suggest that a greater volatile flux or environmental condition is linked to a smaller representation of the microbial organism within the overall community. It should be noted that, negative correlations with volatile flux may indicate either i) fewer organisms exist where greater emissions occur, or ii) that larger microbial sequence abundances occur where greater influx occurs.

Environmental parameters appear to predict abundance of Actinobacteria (bulk density) and most abundant genera within the Chloroflexi (pH). In aggregate samples this may be true but when examined individually (environmental and microbial samples from sediment below 0.1 m2 volatile “footprints”) the relationship with pH is not as apparent. Overall the Phylum Chloroflexi is poorly correlated with pH (Pearson’s R2 [hereafter Pearson’s correlation coefficient will be annotated as PR2] = ~0.3, p = 0.15), with individual genera/species showing both significant negative correlations (*Chloroflexus spp* PR2 = -0.56, p < 0.001, *Dehalococcoides mccartyi spp* PR2 = -0.41, p < 0.001) and significantly positive correlations (SAR 202 cluster PR2 = 0.48, p < 0.001, *Herpetosiphon aurantiacus* PR2 = 0.48 p < 0.001). While the observed preference of *Levinelea* for more neutral pH sediments is not directly observed, there are strong correlations between the abundances of sequences derived from *Levilenea* and *Dehalococcoides mccartyi* (PR2 = 0.83, p < 1e-9) and *Levilenea* and *Chloroflexus* (PR2 = 0.83, p < 1e-9), both of which were the most significantly, negatively correlated genera/species with pH. This suggests that we may be witnessing a loose cohort effect, in which organisms that tend to associate may be driving some of the observed correlations between community and environmental parameters.

The relationship between Actinobacteria and bulk density is much more evident. The phylum Actinobacteria is strongly correlated with bulk density of salt marsh sediments (PR2 = 0.89, p < 1e-15) as is the class Actinobacteria (PR2 = 0.88, p < 1e-15). As well as the positively associated Actinobacteria there are several other microbial groupings that strongly negatively correlate with bulk density, including Chrysiogenetes (PR2 = -0.71, p < 1e-6), Deferribacteres (PR2 = -0.60, p < 1e-4) and Aquificae (PR2 = -0.55, p < 0.001) as well as groupings that are insensitive, including Gemmatimonadetes and Elusimicrobia (p > 0.4).

A selection of other reliable interactions between microbes and environmental parameters include pH (Deferribacteres PR2 = -0.50, p < 0.005), ammonia content (Gammaproteobacteria PR2 = 0.78, p = 1e-8), and NOx concentration (Halobacteria PR2 = 0.56, p < 0.001). Unsurprisingly, a number of fungal groupings held strong preferences for lower sediment water content (Hypocreomycetidae, Pezizomycotina, Leotiomyceta, Saccharomyceta < -0.80, p < 5e-11) while none showed a strong preference for saturated sediments. Other fungal preferences observed include a strong positive correlation with ammonia and bulk density for several Ascomycetes (ammonia- Cryphonectriaceae = 0.80, p = 2e-9; Diaporthales = 0.7, p = 1e-8) (bulk density- Sordariomycetes = 0.87, p = 1e-14; Leotiomyceta= 0.87, p = 3e-14).

The volatile fluxes from sediment surface are indicative of a different suite of microbial community members. Strong interactions between microbes and measured volatile compounds include DMS (Defluviimonas PR2 = 0.94, p < 1e-15), chloroform (Actinobacteria PR2 = 0.61, p < 1e-4 andChrysiogenetes PR2 = -0.39, p = 0.01), methyl iodide (Chrysiogenetes PR2 = -0.40, p = 0.01), isoprene (the Euryarchaeota Methanomicrobia PR2 = 0.53, p < 0.001 and the Cyanobacteria Microchaete and Mastigocoleus PR2 = 0.5, p < 0.001) and methane (Gemmatimonadetes PR2 = -0.49, p < 0.005). Fungal associations with trace gas fluxes include a strong positive correlation between the Ascomycetes Fusarium, Nectriaceae, Glomerellales and chloroform (PR2 = 0.74, p < 5e-8; PR2 = 0.70, p < 5e-7; PR2 = 0.58, p = 1e-4 respectively) and the Sordariaceae with methyl iodide (PR2 = 0.54, p = 5e-4).

These relationships/interactions are crucial for community assembly since these organisms can be reliably associated with certain sediment characteristics or functional outputs. Once the relative abundance of a number of organisms has been identified, their relationships with other microbial populations can fill in the broad picture of microbial community. In this way, if we explore the community associated with the Euryarchaeota we find that a) they associate with themselves consistently with the Chloroflexi (PR2 = 0.75, p < 1e-8), the Thermotogae (PR2 = 0.70, p < 1e-6), the Omnitrophica and the Latescibacteria (PR2 = 0.68, p < 1e-6). In contrast the Euryarchaeota are consistently negatively correlated with the members of the deltaproteobacteria order Burkholderiales (PR2 = -0.54, p < 5e-4).

Using this approach, a preliminary analysis of two important players in element cycling in marine sediments, the Desulfobulbaceae (cable bacteria involved in sulfur oxidation) and the Thaumarchaeota (ammonia oxidizing archaea) suggests that a coherent community picture can be generated (Fig 6). The Desulfobulbaceae tend to be negatively correlated with pH in sediments (PR2 ~ -0.3, p < 0.1), and are strongly negatively correlated with bulk density (PR2 = -0.7, p < 1e-5). The Thaumarchaeota on the other hand show a moderate negative correlation with ammonia in sediments (PR2 ~ -0.2, p < 0.1) and a strong positive correlation with pH (PR2 = 0.5, p < 0.005) and bulk density (PR2 = 0.7, p < 1e-12). This suggests in dense, more alkaline sediments we should expect to find a greater abundance of Thaumarchaeota and a reduced presence of Desulfobulbaceae and in less compact, more acidic sediments the reverse should be true.

[insert figure 5 here]

The volatiles associated with these organisms are also complementary, where the Desulfobulbaceae are negatively correlated with chloroform fluxes (PR2 = -0.4, p < 0.05) and methyl iodide (PR2 = -0.3, p < 0.05) but are weakly positively correlated with isoprene (PR2 = 0.4, p < 0.05). In contrast, the Thaumarchaeota are positively correlated with chloroform (PR2 = 0.7, p < 1e-5) and methyl iodide (PR2 = 0.7, p < 0.0005) and show negative correlations with methyl bromide (PR2 = -0.4, p < 0.05) and methane (PR2 = -0.4, p < 0.05). The volatile signals agree again, that where there are higher chloroform and methyl iodide fluxes we should expect greater numbers of Thaumarchaeota and reduced abundance for Desulfobulbaceae (Fig 6).

The microbial communities associated with these specific groups are also in agreement (Fig 6). Desulfobulbaceae are commonly found in association with Chloroflexi spp (PR2 = 0.5, p < 0.001), Euryarchaeota spp (PR2 = 0.4, p < 0.001) and Firmicutes spp (PR2 = 0.6, p < 1e-5) and do not tend to associate with Thaumarchaeota spp (PR2 = -0.4, p < 0.005) or Actinobacteria spp (PR2 = -0.7, p < 1e-7). The Thaumarchaeota are found in regions with high population representation of Alphaproteobacteria (Rhizobiales) spp (PR2 = 0.8, p < 1e-10) and reduced presence of Bacteriodetes spp (PR2 = -0.5, p < 0.0005) and Firmicutes spp (PR2 = -0.5, p < 0.0005). The microbial associations agree that where larger numbers of Firmicutes are found we can expect fewer Thaumarchaeota and greater numbers of Desulfobulbaceae.

**4. Conclusion**

Within the context of variable and complex salt marsh sediments we have demonstrated that there are a number of organisms whose relative abundance can be predicted by volatile emissions, and these inhabit a limited range of environmental conditions and associate with a limited subset of other organisms. These correlations and interactions lead to basic predictions of individual phyla abundance derived from measured volatilome fluxes. Importantly, these interactions, based on separate measurements of sediment characteristics, volatile trace gas fluxes, and sediment microbial communities generate self-consistent outcomes. This preliminary success, based here on limited numbers of samples and volatiles, suggests that analysis of a small number of volatiles will allow us to constrain the sub-surface microbial community to the phylum level once community builder algorithms have been trained in common relationships. That this has been possible despite the limited number of sediment samples and the reduced range of observed volatiles suggests that this method has promise for more accurate non-invasive diagnosis of soil microbial communities and function.

It is important to note that we do not claim that the volatiles chosen in this analysis are the most useful for all future approaches. The compounds chosen in this analysis are not often used in current diagnostic approaches and may provide an insight into a different set of biological functions than more commonly reported compounds. Indeed, the process of selecting the most informative compounds for use in a “fingerprint” volatile analysis of sub-surface microbial communities is one of the greater challenges facing the application of this approach. Future work will need to explore single-species emissions and metabolic uptake (to identify “known unknowns”) and these studies will need to be combined with scanning of net volatilomes from soils and sediments (to identify “unknown unknowns”). Once a suite of informative compounds has been identified they may be individually targeted for maximum sensitivity in analysis (scanning methods tend to be less sensitive than targeted methods) and the selected list of compounds may be refined after comparison to soil and sediment communities.

We focus in our analysis on community structure as described through individual species abundance. However, our sampling and analytical approach, particularly WGS sequencing of sediments paired with trace gas fluxes, also allow us to study the community structure through the lens of functionality. This is often done in a targeted manner, by looking for the total abundance of specific genetic sequences that allow for individual metabolisms or reactions to be carried out. For instance, conserved genetic sequences that allow for nitrification or denitrification are carried across a range of microorganisms and these sequences can be targeted in sediment samples to identify the overall nitrification/denitrification (functional) potential of the microbial community, regardless of which specific organisms carry the genetic sequence. A caveat to this approach; specific, genetically encoded enzymes may play different or opposing roles in different organisms and under different environmental pressures. For example, the enzyme Dsr catalyses the reduction of sulfur in some sulfur-cycling bacteria and the opposite, oxidation reaction in others (Wasmund et al., 2017). Likewise, nitrogen reductase has been shown to oxidize as well as reduce nitrogen compounds while the ammonia-oxidizing enzyme, ammonia monooxygenase, also oxidizes methane and other organics (Prosser, 2015). Despite these concerns there is substantial value in exploring the metagenome using this functional approach, although clearly better definition and characterization of multifunctional genes is required for optimal outcomes.

We propose that analysis of the net volatilome should become an integral part of future biomonitoring strategies. We have shown here that this technique avoids many of the common issues surrounding other soil sampling and community and analysis methods, it i) provides a non-destructive sampling route, allowing us to understand sub-surface environment and community without disruption, ii) specifically focuses on net functional outcomes (the net volatilome) rather than the present microbial community and iii) this allows us to focus on community members with greatest impact rather than abundance.

A targeted volatilomics approach is already in use to detect and deter specific parasites, pests and diseases across a range of applications. However, the net volatilome fingerprint can also be used to observe changes in the sub-surface that are driven by changing climate (often linked to enhanced pest and disease incidence), providing a prior warning diagnostic rather than an after-the-fact signal. The volatilome fingerprint can be used to better understand local soil cycling of nutrients and within agroecosystems may provide information on maximally efficient routes for nutrient delivery to crops. Transformed and managed landscapes can be monitored through this non-invasive technology at the same locations over time to observe the impacts of land use on soil function.

For this method to reach its full potential there remain a number of important further steps however, including i) maintenance and expansion of the current volatilome databases (to include both efflux and influx, and to grow to include AMF, archaea and unculturable bacteria), ii) expansion of analytical methodologies to provide a more comprehensive range of volatiles per organism, iii) a greater number of replicates across a range of soil and sediment environments, iv) attention to spatial and temporal variability and v) robust statistical approaches to link complex volatile fluxes to individual species within complex microbial communities. Assuming that the listed challenges are met, assembling sub-surface microbial communities through analysis of the volatilome will require attention to the chemical and physical properties of the soil, as well as an appreciation for the impact that non-microbial organisms may have on microbial community behaviour and function.

**Acknowledgements**

J.P.J.C. is a Royal Society Industry Fellow. The authors thank the staff of the Essex Wildlife Trust, particularly those at Abbotts Hall and Fingringhoe Wick, for access, support and advice. We would also like to thank the Colchester Wildfowling & Conservation Club for access to the Mersea Island site. The authors thank Phil Brailey and Emma Rand for figure creation and assistance. This work used the York Biosciences Technology Facility and the authors thank Peter Ashton and Sally James for their assistance. K.R., A.D., and T.H. acknowledge Natural Environment Research Council award no. NE/K01546X/1.

**References**

Ali J.G., Alborn H.T., Stelinski L.L. (2011) Constitutive and induced subterranean plant volatiles attract both entomopathogenic and plant parasitic nematodes. *J. Ecol*., **99** (1), 26-35. http://www.jstor.org/stable/41058834

Allen D.E., Singh B.P., Dalal R.C. (2011) Soil Health Indicators Under Climate Change: A Review of Current Knowledge. In: Singh B., Cowie A., Chan K. (eds) Soil Health and Climate Change. Soil Biology, vol 29. Springer, Berlin, Heidelberg

Alessi A., Redeker K.R, Chong J.P.J. (2018) “A practical introduction to microbial molecular ecology through the use of iChips”, *Ecol & Evol*, in review

Asensio D., Penuelas J., Filella I., Llusià J. (2007) On-line screening of soil VOCs exchange responses to moisture, temperature and root presence. *Plant Soil,* **291**, 294–261.

Bentley, R. & Chasteen, T. G. (2004) Environmental VOSCs—formation and degradation of dimethyl sulfide, methanethiol and related materials. *Chemosphere,* **55**, 291–317.

Bodner, G., Leitner, D. & Kaul, HP. (2014) *Plant Soil*, **380**: 133. <https://doi.org/10.1007/s11104-014-2079-8>

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7, 335–336. doi:10.1038/nmeth.f.303

Cheng, H.-. Y, Masiello, CM, Bennett, GN, Silberg, JJ. (2016). Volatile gas production by methyl halide transferase: An in-situ reporter of microbial gene expression in soil. *Env Sci & Tech,* **50**, 8750-8759.

Chistoserdova, L., Kalyuzhnaya, M. G., Lidstrom, M. E. (2009). The Expanding World of Methylotrophic Metabolism. *Ann Rev Microbiol*, **63**, 477–499. http://doi.org/10.1146/annurev.micro.091208.073600

Choi S., Song H., Tripathi BM, Kerfahi D, Kim H, Adams JM (2017) Effect of experimental soil disturbance and recovery on structure and function of soil community: a metagenomic and metagenetic approach. *Scientific Reports,* **7**, Article number: 2260 doi:10.1038/s41598-017-02262-6

Conrad, R. (1996). Soil Microorganisms as Controllers of Atmospheric Trace Gases (H2, CO, CH4, OCS, N2O, and NO). *Microbiol Rev*, **60**, 609-620.

Devai, I. & DeLaune, R.D. (1995). Formation of volatile sulfur compounds in salt marsh sediment as

influenced by soil redox condition, *Org. Geochem,* **23** (4), 283-287.

Doheny‑Adams, T., Redeker, K., Kittipol, V., Bancroft, I., Hartley, S. E. (2017) Development of an efficient glucosinolate extraction method, *Plant Methods*, **13** (17), doi:10.1186/s13007‑017‑0164‑8.

Falkowski, P.G., Fenchel, T., Delong, E.F. (2008) The Microbial Engines That Drive Earth’s Biogeochemical Cycles, *Science*, **320**, 1034, doi:10.1126/science.1153213

Field, J.A. (2016) Natural Production of Organohalide Compounds in the Environment. In: L. Adrian and F.E. Löffler (eds.), Organohalide-Respiring Bacteria, doi:10.1007/978-3-662-49875-0\_2 Springer-Verlag Berlin Heidelberg 2016

Fierer, N., Strickland, M. S., Liptzin, D., Bradford, M. A. and Cleveland, C. C. (2009) Global patterns in belowground communities. *Ecol Lett*, **12**, 1238–1249. doi:10.1111/j.1461-0248.2009.01360.x

Gallucci, M.N., Oliva, M., Casero, C., Dambolena, J., Luna, A., Zygadlo, J. *et al*. (2009). Antimicrobial combined action of terpenes against the food-borne microorganisms *Escherichia coli*, *Staphylococcus aureus* and *Bacillus cereus*. *Flavour Fragr J*, **24**, 348–354.

Garcia-Alcega, S., Nasir, Z.A., Ferguson, R., Withby, C., Dumbrell, A.J., Colbeck, I., Gomes, D., Tyrrel, S., Coulon, F. (2016). Fingerprinting outdoor air environments using microbial volatile organic compounds (MVOCs). *Trends in Analytical Chemistry*, **86**, 75-83.

Gribble, G. W. (2003) The diversity of naturally produced organohalogens. *Chemosphere* 52, 289–297.

Gribble G.W. (2012) Recently discovered naturally occurring heterocyclic organohalogen compounds. *Heterocycles,* **84** (1), 157–207. doi:10.3987/rev-11-sr(p)5

Griffiths, B.S., Bonkowski, M., Dobson, G., Caul, S. (1999) Changes in soil microbial community structure in the presence of microbial-feeding nematodes and protozoa. *Pedobiologia*, **43**, 297-304.

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M. and Whiteley, A. S. (2011) The bacterial biogeography of British soils. *Environ Microbiol*, **13**, 1642–1654. doi:10.1111/j.1462-2920.2011.02480.x

Harper, D. B., Kalin, R. M., Hamilton, J. T. G. and Lamb, C. (2001) Carbon Isotope Ratios for Chloromethane of Biological Origin:  Potential Tool in Determining Biological Emissions. *Environ. Sci. Technol.* **35**, 18, 3616-3619.

Hemaiswarya, S., Doble, M. (2010). Synergistic interaction of phenylpropanoids with antibiotics against bacteria. *J Med Microbiol,* **59**, 1469–1476.

Insam, H. & Seewald, M.S.A. (2010) Volatile organic compounds (VOCs) in soils, *Biol Fertil Soils* 46: 199. https://doi.org/10.1007/s00374-010-0442-3

Jassbi, A.R., Zamanizadehnajari, S., Baldwin, I.T. (2010) Phytotoxic Volatiles in the Roots and Shoots of *Artemisia tridentata* as Detected by Headspace Solid-phase Microextraction and Gas Chromatographic-mass Spectrometry Analysis. *J Chem Ecol*, **36**, 1398–1407, doi:10.1007/s10886-010-9885-0

Jones, D. (1998). Organic acids in the rhizosphere – a critical review. *Plant Soil*, **205** (1), 25-44. <http://www.jstor.org/stable/42949904>

Jones, C.M., Spor, A., Brennan, F.P., Breuil, M-C., Bru, D., Lemanceau, P., Griffiths, B., Hallin, S., Philippot, L. (2014) Recently identified microbial guild mediates soil N2O sink capacity. *Nature Climate Change*, **4**, 801-805. doi: 10.1038/NCLIMATE2301

Kibblewhite M.G, Ritz K, Swift M.J (2008) Soil health in agricultural systems. *Phil. Trans. R. Soc. B*. 363, 685–701. doi:10.1098/rstb.2007.2178.

Kopylova E, Noé L, Touzet H. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*. 28:3211-7.

Krzycki, J.A., Zeikus, J.G. (1984) Acetate catabolism by *Methanosarcina barkeri*: Hydrogen-dependent methane production from acetate by a soluble cell protein fraction, *FEMS Microbiol Lett*, **25** (1), 27–32, <https://doi.org/10.1111/j.1574-6968.1984.tb01369.x>

Laanbroek, H.J. (2010) Methane emission from natural wetlands: interplay between emergent

macrophytes and soil microbial processes. A mini-review, *Ann Bot,* **105**, 141–153. doi:10.1093/aob/mcp201

Leach AL, Chong JP, Redeker KR. (2012) SSuMMo: rapid analysis, comparison and visualization of microbial communities. *Bioinformatics*. 28:679-86.

LeMer J, Roger P (2001) Production, oxidation, emission and consumption of methane by soils: a review. *Eur J Soil Biol* 37:25–50

Lemfack, M. C., Nickel, J., Dunkel, M., Preissner, R., & Piechulla, B. (2014). mVOC: a database of microbial volatiles. *Nucleic Acids Res*, 42(Database issue), D744–D748. http://doi.org/10.1093/nar/gkt1250

Letunic I, Bork P. (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res*. 44:W242-5.

Li D, Liu CM, Luo R, Sadakane K, Lam TW. (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 31:1674-6.

Loreto, F., Barta, C., Brilli, F., Nogues, I. (2006) On the induction of volatile organic compound emissions by plants as consequence of wounding or fluctuations of light and temperature. *Plant, Cell Environ*, **29**, 1820–1828. doi: 10.1111/j.1365-3040.2006.01561.x

Low, J.C., Wang, N.-Y., Williams, J., Cicerone, R.J. (2003) Measurements of ambient atmospheric C2H5Cl and other ethyl and methyl halides at coastal California sites and over the Pacific Ocean. *J Geophys Res*, **108** (D19), 4608. doi:10.1029/2003JD003620

Lozupone CA. & Knight R. (2007) Global patterns of bacterial diversity., *Proc Natl Acad Sci.* 104 (27), 11436-40.

Massalha, H., Korenblum, E., Tholl, D., Aharoni, A. (2017) Small molecules below-ground: the role of specialized metabolites in the rhizosphere. *The Plant Journal*, **90**, 788–807. doi: 10.1111/tpj.13543

Martin, M., (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal,* 17, 10–12.

Matsui, K., Sugimoto, K., Mano, J., Ozawa, R., Takabayashi, J. (2012) Differential metabolisms of Green Leaf Volatiles in injured and intact parts of a wounded leaf meet distinct ecophysiological requirements. *PLoS ONE,* **7**(4), e36433. doi:10.1371/journal.pone.0036433

McAnulla, C., McDonald, I.R., Murrell, J.C. (2001) Methyl chloride utilising bacteria are ubiquitous in the natural environment. *FEMS Microbiol Lett*, **201**: 151–155. doi:10.1111/j.1574-6968.2001.tb10749.x

Minot, S.S., Krumm, N., Greenfield, N.B., (2015). One Codex: A Sensitive and Accurate Data Platform for Genomic Microbial Identification. *bioRxiv* 027607. doi:10.1101/027607

Mukherjee, A., Lal, R. (2013) Biochar Impacts on Soil Physical Properties and Greenhouse Gas Emissions, *Agronomy*, **3**, 313-339. doi:10.3390/agronomy3020313

Muller A., Faubert P., Hagen M., zu Castell W., Polle A., Schnitzler J. P., *et al*. (2013). Volatile profiles of fungi–chemotyping of species and ecological functions. *Fungal Genet. Biol*. **54** 25–33. doi:10.1016/j.fgb.2013.02.005

Nedwell, D.B., Underwood, G.J.C., McGenity, T.J., Whitby, C., Dumbrell, A.J. (2016). The Colne Estuary: A Long-Term Microbial Ecology Observatory. *Adv. Ecol. Res*., **55**, 227-281.

Nicol, G.W., Leininger, S., Schleper, C., Prosser, J.I. (2008) The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Env Microbiol*, **10** (11), 2966–2978. doi:10.1111/j.1462-2920.2008.01701.x

Niinemets, Ü., Kännaste, A., Copolovici, L. (2013) Quantitative patterns between plant volatile emissions induced by biotic stresses and the degree of damage. *Front. Plant Sci.*, **4** (262). doi: 10.3389/fpls.2013.00262

Nobrega, R.L.B., Guzha, A.C., Torres, G.N., Kovacs, K., Lamparter, G., Amorim, R.S.S., et al. (2017) Effects of conversion of native cerrado vegetation to pasture on soil hydro-physical properties, evapotranspiration and streamflow on the Amazonian agricultural frontier. *PLoS ONE,* **12** (6),

e0179414. https://doi.org/10.1371/journal.pone.0179414

Oduro H, Kamyshny A, Zerkle A, Li Y, Farquhar J. (2013) Quadruple sulfur isotope constraints on the origin and cycling of volatile organic sulfur compounds in a stratified sulfidic lake. *Geochimica et Cosmochimica Acta*. 120. 251-262. 10.1016/j.gca.2013.06.039.

Pan, D., Watson, R., Wang, D., Tan, Z.H., Snow, D.D., Weber, K.A. (2014) Correlation between viral production and carbon mineralization under nitrate-reducing conditions in aquifer sediment. *ISME* *Journal*, **8**, 1691–1703

Parks, J.M., Johs, A., Podar, M., *et al* (2013) The Genetic Basis for Bacterial Mercury Methylation, *Science,* 339, 1332-1335.

Peñuelas, J., Asensio, D., Tholl, D., Wenke, K., Rosenkranz, M., Piechulla, B., Schnitzler, J.P. (2014) Biogenic volatile emissions from the soil. *Plant, Cell Environ*, **37**, 1866–1891. doi:10.1111/pce.12340

Pierik, R., Ballaré, C.L., Dicke, M. (2014) Ecology of plant volatiles: taking a plant community perspective. *Plant, Cell Environ*, **37**, 1845–1853. doi: 10.1111/pce.12330

Prieto, I., Armas, C., Pugnaire, F.I. (2012) Water release through plant roots: new insights into its consequences at the plant and ecosystem level. *New Phytol*, **193**, 830–841. doi: 10.1111/j.1469-8137.2011.04039.x

Prosser, J. I. (2015). Dispersing misconceptions and identifying opportunities for the use of 'omics' in soil microbial ecology. *Nature Reviews Microbiology*, **13**(7), 439-446. DOI: 10.1038/nrmicro3468

Read, D.B., Bengough, A.G., Gregory, P.J., Crawford, J.W., Robinson, D., Scrimgeour, C.M*. et al*. (2003) Plant roots release phospholipid surfactants that modify the physical and chemical properties of soil. *New Phytol*, **157**, 315–326.

Redeker, KR, Treseder, K, Allen, MF (2004) Ectomycorrhizal Fungi: A New Source of Atmospheric Methyl Halides?, *Global Change Biology* **10** (6), 1009-1016, 2004.

Redeker, K.R., Davis, S., Kalin, R.M. (2007) Isotope values of atmospheric halocarbons and hydrocarbons from Irish urban, rural, and marine locations*. J Geophys Res*, **112**, D16307, doi:10.1029/2006JD007784

Redeker KR, Kalin, R (2012) Methyl chloride isotopic signatures from Irish forest soils and a comparison between abiotic and biogenic methyl halide soil fluxes, *Global Change Biology*, **18**, 1453–1467. doi: 10.1111/j.1365-2486.2011.02600.x.

Redeker, K.R., Baird, A.J., The, Y.A. (2015) Quantifying wind and pressure effects on trace gas fluxes across the soil-atmosphere interface. *Biogeosciences*; **12** (24), 7423-7434.

Ryu, C.-M., Farag, M.A., Hu, C.-H., Reddy, M.S., Wei, H.-X., Paré, P.W. Kloepper, J.W. (2003) Bacterial volatiles promote growth in Arabidopsis. *PNAS*, **100** (8) 4927-4932. doi:10.1073/pnas.0730845100

Saini, H.S., Attieh, J.M., Hanson, A.D. (1995) Biosynthesis of halomethanes and methanethiol by higher plants via a novel methyltransferase reaction, *Plant Cell Environ*., **18**, 1027–1033.

Saito, T. & Yokouchi, Y. (2008) Stable carbon isotope ratio of methyl chloride emitted from glasshouse-grown tropical plants and its implication for the global methyl chloride budget, *Geophys. Res. Lett*., **35**, L08807, doi:10.1029/2007GL032736

Schmidt, R., Etalo, D. W., de Jager, V., Gerards, S., Zweers, H., de Boer, W., & Garbeva, P. (2015). Microbial Small Talk: Volatiles in Fungal–Bacterial Interactions. *Front. Microbiol*, **6**, 1495. <http://doi.org/10.3389/fmicb.2015.01495>

Schmidt, R., Etalo, D.W., de Jager, V., Gerards, S., Zweers, H., de Boer, W., Garbeva, P. (2016) Microbial Small Talk: Volatiles in Fungal–Bacterial Interactions. *Front. Microbiol*. **6**, 1495. doi: 10.3389/fmicb.2015.01495

Schulz-Bohm, K., Zweers, H., de Boer, W., Garbeva, P. (2015) A fragrant neighborhood: volatile mediated bacterial interactions in soil. *Front. Microbiol*. **6**, 1212. doi: 10.3389/fmicb.2015.01212

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

Sorek, R., & Cossart, P. (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet*. **11**, 9–16. doi: 10.1038/nrg2695

Stein, L.Y., Yung, Y.L. (2003) Production, Isotopic Composition, and Atmospheric Fate of Biologically Produced Nitrous Oxide. *Annu. Rev. Earth Planet. Sci.,* **31**, 329–356. doi:10.1146/annurev.earth.31.110502.080901

Toljander, J.F., Lindahl, B.D., Paul, L.R., Elfstrand, M., Finlay, R.D. (2007) Influence of arbuscular mycorrhizal mycelial exudates on soil bacterial growth and community structure. *FEMS Microbiol Ecol,* **61**, 295–304. doi:10.1111/j.1574-6941.2007.00337.x

Turlings, T.C.J., Hiltpold, I., Rasmann, S. (2012) The importance of root-produced volatiles as foraging cues for entomopathogenic nematodes. *Plant Soil*, **358**, 51–60. doi:10.1007/s11104-012-1295-3

ul Hassan, M.N., Zainal, Z., Ismail, I. (2015) Green leaf volatiles: biosynthesis, biological functions and

their applications in biotechnology. *Plant Biotech J*., **13**, 727–739. doi: 10.1111/pbi.12368

Verstraeten, W.W., Veroustraete, F., Feyen, J. (2008) Assessment of Evapotranspiration and Soil Moisture Content Across Different Scales of Observation. *Sensors*, **8**, 70-117.

Virtanen, A.I. (1965) Studies on organic sulphur compounds and other labile substances in plants. *Phytochem*., **4**, 207-228.

von Fischer, J. C. & Hedin, L. O. (2007) Controls on soil methane fluxes: Tests of biophysical mechanisms using stable isotope tracers, *Global Biogeochem. Cycles*, **21**, GB2007, doi:10.1029/2006GB002687

Wahman, D. G., Katz, L. E., Speitel Jr., G. E. (2005) Cometabolism of Trihalomethanes by *Nitrosomonas europaea*, *Appl. Environ. Microbiol*., **71**, 12, 7980-7986.

Wasmund, K., Mußmann, M., Loy, A. (2017) The life sulfuric: microbial ecology of sulfur cycling in marine sediments, *Environ. Microbiol. Rep*., **9**, 323-344. doi:10.1111/1758-2229.12538

Watling, R., & Harper, D. (1998). Chloromethane production by wood-rotting fungi and an estimate of the global flux to the atmosphere. *Mycological Research*, **102**(7), 769-787.

Yamada, T., Sekiguchi, Y., Hanada, S., Imachi, H., Ohashi, A., Harada, H., Kamagata, Y. (2006) *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum *Chloroflexi*. *Int J Syst* *Evol Microbiol*., **56** (6), 1331-1340. doi: 10.1099/ijs.0.64169-0

Yamada, T., Imachi, H., Ohashi, A., Harada, H., Hanada, S., Kamagata, Y., Sekiguchi, Y. (2007) *Bellilinea caldifistulae* gen. nov., sp. nov. and *Longilinea arvoryzae* gen. nov., sp. nov., strictly anaerobic, filamentous bacteria of the phylum *Chloroflexi* isolated from methanogenic

propionate-degrading consortia. *Int J Syst* *Evol Microbiol*., **57**, 2299-2306. doi:10.1099/ijs.0.65098-0

Yuste, J.C., Baldocchi, D.D., Gershenson, A., Goldstein, A., Misson, L., Wong, S. (2007) Microbial soil respiration and its dependency on carbon inputs, soil temperature and moisture. *Global Change Biology,* **13**, 1–18, doi: 10.1111/j.1365-2486.2007.01415.x

Zhang H., Kim M.S., Krishnamachari V., Payton P., Sun Y., Grimson M., et al. (2007). Rhizobacterial volatile emissions regulate auxin homeostasis and cell expansion in *Arabidopsis*. *Planta* **226**, 839–851. doi:10.1007/s00425-007-0530-2

Zhang, W., Li, Y., Xu, C., Li, Q., Lin, W. (2016) Isotope signatures of N2O emitted from vegetable soil: Ammonia oxidation drives N2O production in NH4+-fertilized soil of North China, *Scientific Reports,* 6, Article number: 29257

Zhu, B., van Dijk, G., Fritz, C., Smolders, A.J.P., Pol, A., Jetten, M.S.M., Ettwig, K.F. (2012) Anaerobic Oxidization of Methane in a Minerotrophic Peatland: Enrichment of Nitrite-Dependent Methane-Oxidizing Bacteria. *App Env Microbio*., **78** (24), 8657–8665. doi:10.1128/AEM.02102-12

**Table 1:** Biological and chemical information for analysed volatiles

**Table 2:** Averaged selected environmental parameters.

**Table 3:** Averaged measured fluxes of selected components of the volatilome. Negative fluxes indicate that sediments degraded or consumed the compound from the chamber headspace while positive fluxes denote movement of the material from the soil to the air. All fluxes are in g m-2 day-1.

**Box Figure 2**: Volatilome sampling. Volatiles are sampled within a trapped airspace (A), first at an initial time point and then 9 minutes later. This allows the analyst to ascribe changes in the sampled air mass to processes occurring in the trapped airspace over the sampling period while avoiding excessive stress on the living system. Samples are drawn from the chamber into the sampling canister (C) via an Ascarite trap (B), which reduces water and carbon dioxide content of the sampled air. Once returned to the lab the canisters are attached to the liquid nitrogen condensation trap (D) which cools the air under vacuum, removing nitrogen, oxygen, and argon from the sample and leaving only compounds normally present in the part-per-billion range. This condensed, concentrated sample is delivered to the GC-MS (F) which separates the sample (gas chromatography) and analyses each individual compound (G, H, I, J, K, L, and M) based on their mass (mass spectrometry).

**Figure 1**.- Essex UK salt marsh sampling sites. The furthest site westward is the Abbotts Hall (AH) site, while the furthest site Eastward is the Fingringhoe Wick Range (FWR) site. The Mersea Island (MI) sampling site is centrally located, but sheltered, relative to the other sites. The sites were paired, both natural and managed marshes were adjacent to each other and are represented by single demarcations.

**Figure 2**.- SSuMMo derived microbial community comparison between sites and seasons assigned by Class. Bars indicate positive presence of the closest microbial match and size of the bar is indicative of the abundance of the organism within the sample. Samples are as indicated, from the inside ring out: ordered as Abbots Hall, Fingringhoe Wick and Mersea Island, summer natural and realigned marshes then winter natural and realigned marshes. Archaea were found to be between 2 and 12% of the identified community and generally more abundant in winter at realigned sites.

**Figure 3**.- Class level microbial community comparison between sites and seasons. Only the ten most abundant classes are shown.

**Figure 4**. Distribution of pH across sediment samples from Essex salt marshes

**Figure 5:** Averaged trace gas fluxes for managed (grey columns) and natural (white columns) salt marsh sediments across summer (stippled, left) and winter (clear, right) seasons. Note that the fluxes are shown at a range of scales and include both efflux from (positive) and influx to (negative) the sediment. Fluxes shown are all g m-2 day-1 ± stderr.

**Figure 6**: The combined interactions between microbes, their environment and the net volatilome observed. Each microbe is shown individually (satellite images) in its preferred environmental state (high or low pH versus greater or lesser bulk density) and, where appropriate, with associated volatile fluxes (filled symbols indicates greater flux, open symbols indicates lesser-to-negative flux). The central image shows the combined, coherent community with microbial assemblages within environmental context and with net volatilome.