UNIVERSITÀ DEGLI STUDI DI NAPOLI "FEDERICO II"

TESI DI DOTTORATO DI RICERCA IN BIOLOGIA

XXXV CICLO

Microbial influences on subduction zone volatiles cycling

Dottorato in Biologia

Ph.D. Course in Biology

Coordinators:

Chiar.mo Prof. Sergio Esposito

Tutor: ***Candidate:***

Chiar.mo Prof. Donato Giovannelli **Matteo Selci**

2022/2023

DON’T PANIC

Neil Gaiman

This thesis is devoted to all the people who have always supported me.
To my parents and my brother, who never stopped believing in me.
To my friends for being there when I need them.
A special mention goes to my tutor Donato Giovannelli, for the trust, the mentorship, and the "paccheri" on my back.
To Angela Cordone, for the help and the infinite trust.
To all the GiovannelliLab crew, past and present, too many to mention one by one. I feel so lucky for the time and the adventures we shared. I learn so much from every one of you. I will never forget it.

Open-source statement

The work presented in this thesis was realized with open-source software, in a convinced effort to push the **open-source philosophy** forward. Where possible all scientific software used was open-source. The entire work has been carried out on the **Ubuntu** operating system, with **R** for statistical computing, **LibreOffice** and **Google sheets** as office suites, and **Inkscape** for graphics and charts.

This PhD thesis is one of the many examples that working entirely with open-source software is possible and should be done to support informatics and cultural freedom.

Table of Contents

Chapter 1

Geobiochemical cycling of volatiles at convergent margins

Convergent margins: gateway between Earth's surface and deep mantle

Tectonic plays a critical role in shaping the Earth's surface through several processes such as the recycling of the crust, the creation of new oceanic crust at ocean spreading centers, and the progressive accretion of new continental crust (Parnell, 2004). At the same time, these processes carry the continuous cycling of volatile species such as carbon, water, hydrogen, and sulfur between Earth's surface and subsurface, influencing the atmospheric composition and contributing to keep our planet habitable for at least the last 3.8 billion years (Figure 1; Moore et al., 2017; Bekaert et al., 2021). Convergent plate margins, where denser oceanic plates subduct beneath the continental crust, are key interfaces between the oxidized exospheric reservoirs (atmosphere, hydrosphere, and biosphere) and the reduced interior mantle where volatile species are cycled (Figure 2; Kelemen and Manning, 2015). The physical and chemical properties of the mantle, including its convection and the formation of continental crust, are greatly influenced by the presence of volatile elements such as H_2O and CO_2. These elements reduce the melting point of mantle materials, influencing the geochemical features of the lava that is eventually released (Bekaert et al., 2021). Additionally, elements like sulfur, nitrogen, and carbon influence the redox state of the magma influencing the mobility of elements, the possible mineralogy of the melts, and the overall behavior of the volatiles (Frost and McCammon, 2008).

Despite the terrestrial mantle's relatively low volatile element abundances, it represents the majority of the bulk silicate Earth by mass (Schaefer et al., 2012), which makes it a key player in shaping the amount and the quality of volatile elements on our planet's surface. Volatiles additionally control the flux of slab components into the mantle wedge and are responsible for melt generation and crystallization of phase assemblages in the overriding crust (Zellmer et al., 2015). The depth at which magma accumulates can be partially influenced by its volatile content, as well as its degassing rate and extent during storage and decompression. These volatiles also are important in determining the rheological properties of magma and the resulting style of eruption, which in turn influences the release of volatiles (Zellmer et al., 2015). The lithosphere, composed of the crust and the upper layer of the mantle, is constantly enriched by significant amounts of volatiles and trace elements through hydrothermal and seawater circulations which pervade for kilometers within the crust (Ranero et al., 2003). During the subduction, the host lithology as well as the pressure and temperature define the metamorphic dehydration reactions that control what fraction of volatiles is recycled back to the surface and what fraction is returned to the mantle (Zellmer et al., 2015). A diverse suite of parameters linked with the subduction geometry, such as dip angle and convergence rate, as well as secondary processes such as forearc erosion influence volatile cycling (Bekaert et al., 2021).

The quality and quantity of volatile recycled have profound effects on Earth's evolution and climate. For example, the quantity and speciation of carbon subducted and recycled to the surface affect the redox potential of the Earth's surface and have influenced the atmospheric composition over geological timescales influencing climate (Hayes and Waldbauer, 2006;

Galvez and Pubellier, 2019). On the other hand, biological processes such as oxygenic photosynthesis, the burial and degradation of organic carbon, actively contribute to the production and maintenance of the current Earth's breathable atmosphere (Bergin et al., 2015). The extremely high abundance of microorganisms on Earth (Fullerton et al., 2021) highlights their quantitative importance in influencing the biogeochemical cycling of elements. Microorganisms are responsible for up to 45 % of carbon fixation (and oxygen inputs to the atmosphere (Field et al., 1998)) and about 95 % of its consumption during the remineralization of organic matter (King et al., 2006). They are also responsible for about 50 % of nitrous oxide (Matthews, 1994; Fowler et al., 2013) and between 26 and 30 % of methane (Knittel and Boetius, 2009) emissions.

Rock and plate geochemistry composition at convergent margins and parameters variation

Subduction zones have traditionally been considered as made up of arc crust, which is composed of sediments, igneous rocks (whether volcanic or intrusive) mixed with their metamorphic equivalents and a thick layer of lithospheric mantle. Its geochemical composition can span from mostly felsic to intermediate, reflecting a high degree of heterogeneity (Miller and Christensen, 1994; Christensen and Mooney, 1995; Taira et al., 1998; Jagoutz et al., 2009; Rioux et al., 2010; Bekaert et al., 2021). Element recycling in subduction zones can take place both in the trench and in back-arc basins, where elements can come back to the surface. Back-arc basins form in oceanic environments where the subducting slab is undergoing rollback, allowing the extension and formation of new oceanic crust behind the arc front. Under this condition, magmas formed are predominately mafic. However, this does not mean that there isn't a diversity in the type of basalts present. Most back-arc basins include Mid-Oceanic Ridge Basalts (MORB's), alkaline basalts, island arc tholeiites, and transitional compositions (Saunders and Tarney, 1984), and have a characteristic slab signature such as high H_2O, large ion lithophile elements (LILE) contents, and a relative depletion in Nb and Ta.

In subduction zones, fluids play a very important role, since they act as the modifying agent for the lithologies that make up the slab, and also as the carrier of the elements that can contribute to the formation of specific rocks like serpentinites (Spandler and Pirard, 2013). The alteration of the subducting crust by sea water or hydrothermal fluids near the spreading center increases the content of volatiles (H_2O, Cl, S, and SO_2), and trace elements such as K, Rb, Cs, U, Pb, Sr, As and B, with H_2O being the main fluid released during plate subduction, thanks to slab dehydration. These dehydration fluids tend to be impoverished in trace elements, except for B, As, Cs, and Li, and might provide habitats that could be colonized by microbial communities capable to thrive under different temperatures, pH, Eh, pressure, salinity and elemental compositions (Vitale Brovarone et al., 2020; Fullerton et al., 2021).

Subducting slabs can also influence the inventories and contributions of volatiles to crustal and atmospheric reservoirs, with this contribution depending on slab relative thickness, slab composition, slab age, slab subduction rate and angle (Bekaert et al., 2021). Dip angle is also thought to control the composition of both fluids and volatiles released during subduction (Dumitru, 1991; Hu and Gurnis, 2020). A shallow angle of subduction can promote the preservation of volatiles-rich minerals, whereas a slab with a steeper angle is more prone to degassing and the release of volatiles. The age of the subducting slab in thought has an impact on volatiles release as well. Older slabs are more hydrated than the younger ones and typically enriched with volatile-bearing minerals, leading to a higher volatile release. The age of the slab can be also correlated to its thickness, the amount of sediments present, and thus its composition, eventually impacting the overall volatile release efficiency.

All of the above highlights how depending on the nature, age and the setting of the convergent margin, differences in the amount and the elements released will take place. As an example of the influence of relative thickness over the geochemical composition of erupted lavas, we can consider two different volcanic zones in the Andes, the central volcanic zone (CVZ) and the southern volcanic zone (SVZ). The CVZ represents a thicker (50-70 km) continental crust of Paleozoic to Mesozoic age, lies above a Benioff zone with an easterly dip between 25° to 30° and has a mantle wedge of nearly 70 km, with andesite-dacite and dacite-rhyolite ignimbrites as main eruptive products. On the other hand, in the SVZ the mantle wedge is less developed (nearly 50 km), and is located 90 km above a slightly shallow dipping (<25°) Benioff zone with Al-rich basalts, basaltic andesites and andesites as the main lithologies (Hilton et al., 1993). The subduction of the Cocos plate below the Caribbean plate at nearly 83 mm/year (DeMets et al., 2010) shows again another example of how different subduction zones can be. The Nicoya and Osa peninsula are closer to the trench than any other coastline in subduction zones, placing them directly above the seismogenic zone (van Rijsingen, 2019). In the case of the Central America Volcanic Front (CAVF), the geochemical composition of the eruptive products from Platanar, Poás, Barva, Turrialba, and Irazu volcanoes shows an oceanic island basalt (OIB) signature, different from the rest of the CAVF where rocks with a geochemical composition typically associated with convergent margins are observed (Di Piazza et al., 2015) that with a more or less andesitic in composition (Spandler and Pirard, 2013).

Thus, convergent margins have a severe impact on the geochemical characteristics of Earth's deep mantle and shallow crustal reservoirs, as well as on the size of the different carbon sinks and reservoirs (Barry et al., 2019b, 2019a). The Central American Volcanic Arc has been the object of several studies focused on evaluating the links between biology and subduction and its effects on biogeochemical cycles (Barry et al., 2019a), in particular in the area between Costa Rica and Panama. The interaction between the Caribbean and the Nazca-Coco plates with the Panama microplate has allowed the development of a complex tectonic setting where the Nazca-Coco plate subducts under the Caribbean and the Panama microplate (Avard et al., 2021). This phenomenon creates a very active seismic zone, from the trench up until the hinterland of Costa Rica, and releases a high amount of volatiles that can be used by microorganisms in the deep biosphere.

Volatile budgets at convergent margins

Subduction zones are considered the most important tectonic driver of geochemical cycles since this process allows the recycling of elements from the subducting slab back to the overlying crust, atmosphere, and oceans (Spandler and Pirard, 2013; Zheng, 2019). This is especially true for carbon, where its fate and fluxes are controlled mainly by volcanic emissions and the transformation of subducted carbon-bearing sediments and carbonates (Burton et al., 2013; Fischer et al., 2019; Plank and Manning, 2019; Werner et al., 2019; Bekaert et al., 2021). In fact, subduction zones can provide a link between the long-term (abiotic/sedimentary) and short-term (biotic) carbon cycle, where subducting slabs bring carbon back into Earth's interior and control how much carbon can re-exit the system through volcanic arc emissions (Galvez and Pubellier, 2019). During its ascent from the mantle to the surface, carbon is altered by complex geological and biological processes, capable of transforming its redox state and altering its mobility and residence time in the crust (Barry et al., 2019b; Vitale Brovarone et al., 2020; Bekaert et al., 2021), and depending on the characteristics of the convergent margins, the efficiency carbon is brought back to the surface with changes (Plank and Manning, 2019). Recent estimates of carbon inputs from subduction zones consider nearly 84 ± 14 Mt of C per year, with different properties such as isotopic signature and attributes depending on the subduction zone. For instance, the Central America subduction slab is thought of being composed mainly of carbonate sediments, with high biological productivity, organic carbon content and a mean $\delta^{13}C$

of -3.0 ‰ (Plank and Manning, 2019). Emissions from terrestrial volcanoes contribute to nearly 300 Mt yr-1 of CO_2 (Mörner and Etiope, 2002; Chiodini et al., 2006), with other sources of CO_2 like diffuse degassing and submarine volcanism at spreading centers being responsible for releasing CO_2 fluxes that can rival the amount released by eruptive activity (between 83 and 93 Mt yr-1 and 58 and 96.8 Mt yr-1 respectively) based on estimates conducted by several studies (Marty and Tolstikhin, 1998; Fischer et al., 2019; Werner et al., 2019; Fischer and Aiuppa, 2020). From a biological standpoint, subduction zones can also act as a shuttle that connects the deep Earth with its surface having an impact on the evolution of Earth's interior and surface (Barry et al., 2019b; Zheng, 2019; Vitale Brovarone et al., 2020; Fullerton et al., 2021). Recent studies led by Barry et al., (2019a) have provided evidence that even if nearly 91±4.0 % of carbon released during the subduction process is sequestered in the form of calcite, while 3.3 ± 1.3 % is incorporated into biomass highlighting the fact that subduction zones can be areas where microbial life can develop, and that convergent margins are tectonic settings where transformations at a molecular level of the organic carbon can take place (Galvez et al., 2020) linking the organic and the inorganic components of the carbon cycle.

Biological overprints on convergent margin's volatile cycling

Elements and compounds are cycled through the biosphere, atmosphere, hydrosphere, and geosphere by a number of biotic and abiotic processes. The cycling of elements, including carbon, nitrogen, phosphorus, and many others, is crucially influenced by microbial metabolism, as microorganisms have the ability to break down and transform various compounds. For example, in the nitrogen cycle, through the action of nitrogen-fixing bacteria, nitrogen gas is reduced to ammonia, which is in turn oxidized by nitrogen-oxidizing bacteria, and thus converted back to nitrogen gas (Schlesinger et al., 2011).

These cycles are driven by redox reactions, in which electrons are transferred between different chemical species (donor/acceptor), and are largely mediated by microorganisms that can catalyze these reactions to support their life and energy yields. For instance, in the process of anaerobic respiration, microorganisms can transfer electrons to an inorganic electron acceptor, such as sulfur or nitrogen compounds (Richardson, 2000). In contrast, in the process of photosynthesis, microorganisms accept electrons from donors such as water or carbon dioxide and use the energy to produce organic compounds. The cycling of elements continues in the Earth's subsurface, where chemical elements and compounds move through the deep soil and sediments, as well as subsurface waters. These cycles are driven by a complex network of physical, chemical, and biological processes, where microbes-driven redox reactions play a critical role (Falkowski et al., 2008). For instance, subsurface microbes can oxidize and reduce compounds such as methane and sulfur, having a significant impact on the global cycles of these two elements. Deep subsurface microbial communities are influenced by parameters like temperature, redox state, pH, pressure, and energy sources alternative to sunlight (Fullerton et al., 2021). Scientists are currently investigating the subsurface microbial diversity, the microbial co-occurrence networks, and the most abundant metabolisms. Heterotrophic anaerobic microbial communities exist in porous sandstone or sandy sediments, located adjacent to organic-rich deposits representing the main source of electron donors for the in situ microbial communities (Beeman and Suflita, 1987; Balkwill et al., 1989). Additionally, highly diverse microbial communities represented by iron-reducing, sulfate-reducing, methanogenic, acetogenic, and fermentative bacteria (Stevens et al., 1993) can occur in regions where a source of organic matter is not apparent but where igneous rock is present. In these circumstances, microbial communities can be sustained by the geochemical energy obtained from reduced minerals. It's also known that the water interaction with iron-containing minerals in silica-rich and basaltic bedrock, causes the release of hydrogen gas (Dunham et al., 2021). Hydrogen can be used as an energy source for many types of microorganisms and specifically in the subsurface ecosystem,

the primary producers appear to be the hydrogenotrophic bacteria (Krumholz, 2000). These organisms may provide organic materials to support a variety of other microbial chemosynthetic metabolisms, such as methanogenesis.

Potential of convergent margin to host a diverse subsurface biosphere

In this context, convergent margins connect the large reservoir of carbon present in the deep Earth with the planet surface. In detail, when an oceanic plate collides with a continental plate, the heavier one dives under the other and sinks into the mantle, dragging carbon and other volatile compounds with it, which are thus transferred from the earth's surface to the interior. Contextually, part of these compounds are recycled back on the surface from arc volcanoes and other secondary geothermal processes (Kelemen and Manning, 2015). This dynamic geochemical equilibrium induced by the subduction of the oceanic plate under the continental crust, promotes the cycling and mobilization of carbon compounds and other volatiles through a crust layer potentially below the isotherm of 135-150 ˚C (Figure 3; Harris and Wang, 2002; Rogers et al., 2023a), thus providing microbial communities with a wide range of habitats and energy sources. In these environments, subsurface microorganisms perform carbon fixation, and participate in nitrogen, sulfur and hydrogen cycling. Moving along the volcanic arc, changes in subduction parameters will impact the variety and amount of volatiles released, fueling chemolithoautotrophs-based microbial communities and causing variations in their metabolic diversity . Additionally, serpentinization-sourced hydrogen with low dissolved oxygen in convergent margins create the right conditions for metabolisms like as reported by Fullerton et al. (2021), deep subsurface microbial communities investigated along the like reductive tricarboxylic acid and Wood-Ljungdahl pathways, hydrogen oxidation and sulfate reduction. Costa Rican convergent margins are not only influenced by parameters like pH and temperature, but also correlate with major cation and anions of subsurface fluids, directly driven by tectonic processes. This suggests that the microbial community composition and distribution is related to subsurface geochemical parameters.

Carbon in convergent margin

Carbon is the main element in terms of contribution to the wellbeing and sustainability of life on Earth (Hazen and Schiffries, 2013). Carbon is generally stored in both organic and inorganic forms, in a variety of reservoirs like rocks, sediments, oceans, living organisms, and in the atmosphere (Falkowski et al., 2000). Carbon is continuously exchanged and recycled between these reservoirs through complex interactions between geosphere and biosphere (Ussiri and Lal, 2017; Isson et al., 2020). The flux between the different reservoirs occurs at different times scales, from the shortest, where carbon is subtracted and released in the atmosphere by photosynthesis and respiration, to the longest, where over geologic time scale organic and inorganic carbon is buried in deep sediments and recycle into the atmosphere through plate tectonics (Dean and Gorham, 1998; Falkowski et al., 2000; Zellmer et al., 2015; Plank and Manning, 2019). The balance between these processes is essential for planet habitability and climate stability (Kwon et al., 2009; Isson et al., 2020). Unlike surface environments where carbon cycling is well constrained (Crisp et al., 2022), its role in the subsurface environment remains still not well understood. To date, numerous studies have focused on the fate of carbon through the deep layers of our planet, looking at the subsurface microbiota as a key player in carbon sequestration (Emerson et al., 2016; Barry et al., 2019c). In subduction zones, the two main forms of carbon available for microorganisms are CO_2 and CH_4 (Figure 4), while organic carbon and CO are less abundant. Here, a number of chemolithoautotrophic microorganisms work as carbon sinks, utilizing CO_2 and CO as electron acceptors, donors and as carbon sources. Between the others, the reductive tricarboxylic acid cycle (rTCA) and the Wood-Ljungdahl pathway are recognised as the most ancient microbial metabolisms involved in carbon fixation.

As shown by (Rogers et al., 2023b), in Costa Rican convergent margins, these two pathways are well represented, with the Wood-Ljungdahl pathway abundant in all the subsurface area investigated, and the rTCA cycle mostly concentrated in the arc and forearc area. The rTCA cycle is responsible for producing carbon dioxide and water, using electron donors that can be easily found in subduction zones, such as hydrogen or sulfide. The Wood-Ljungdahl pathway is used by a numerous anaerobic microorganisms for energy generation and carbon fixation, producing acetate or methane depending on the presence of bacteria (acetogens) or archaea (hydrogenotrophs), respectively (Bender et al., 2011; Borrel et al., 2016). Only a limited number of archaeal groups, Euryarchaeota, TACK (Thaumarchaeota, Aigarchaeota, Crenarchaeota, and Korarchaeota), DPANN (Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, and Nanohaloarchaeota), and the Asgard superphylum, are able to perform methanogenesis, using geothermal substrates like CO_2 and H_2, as well as other by-products of organic matter fermentation such as acetate, and methylated compounds (Cord-Ruwisch et al., 1988; Evans, 2012; Sieber et al., 2012). At convergent margins, methane generated deep by thermogenic or abiotic processes and shallower in the crust by biological activity is then transported to the surface by fluids and gases (Vitale Brovarone et al., 2020). During ascent, methane can be used by methanotrophic communities which perform methane oxidation in presence of oxygen or anaerobically using different electron acceptors like sulfate, iron, manganese, nitrite and nitrate (Guerrero-Cruz et al., 2021). However, large spatial-scale studies about biogeological processes as well as constraints that clarify what feedbacks are present between deep and surface Earth in controlling the methane degassing across convergent margins have not been performed yet.

Nitrogen and sulfur

Nitrogen is one of the fundamental elements required by living systems, as it is incorporated into the structure of amino acids, nucleic acids, and proteins (Zerkle and Mikhail, 2017). However, the cycling of nitrogen between Earth's reservoirs is still debated, given that different parameters, such as background lithologies, and the structure of the subducting slab, influence the rate at which nitrogen is returned to the surface through arc volcanism (Füri et al., 2021; Labidi et al., 2021). Recently, it has been estimated that between 11 % to 16 % of Earth's total nitrogen budget resides in both the continental, as well as oceanic crust (Bebout et al., 2016), while the vast majority of nitrogen remains in the atmosphere in the form of N_2 gas (78 % of the total atmospheric gases). Even though nitrogen can be delivered to the biosphere by inorganic processes, such as rock weathering, the vast majority of atmospheric nitrogen is fixed by living systems. Nitrogen-fixing prokaryotes, by the means of the nitrogenase complex, can convert the inert N_2 gas into biomass (Pi et al., 2022). Fixed nitrogen in the form of ammonium (NH_4^+), can be readily oxidized to nitrite (NO_2^-) and nitrate (NO_3^-), and reduced back to ammonium by nitrate and nitrite oxidation. The fixed nitrogen can return to the atmosphere in the form of N_2 by denitrification and anaerobic ammonium oxidation (anammox) (Zerkle and Mikhail, 2017). Biologically fixed nitrogen, however, can be transferred into low-temperature mineral phases, such as the ones found in clay minerals (Boyd, 2001). Additionally, ammonium-bound nitrogen in micas and feldspars moves from surface reservoirs into the deep Earth by means of subduction (Bebout et al., 2016). As the slab subducts (Figure 4), nitrogen bearing minerals can return a large fraction of nitrogen to the surface through metamorphism (Zerkle and Mikhail, 2017). For instance, 15-40 % of the subducted nitrogen returns to the surface at the Central American margin, and 4-17 % at the Izu-Bonin arc (Mitchell et al., 2010; Lee et al., 2017). These differences were suggested to be related to the thermal regime of the subduction zone, controlling the recycling efficiency of nitrogen, and likely other volatiles (Bekaert et al., 2021; Halama and Bebout, 2021). However, the amount of degassed is recycled back from the subsurface biosphere, still remains unconstrained. For example, Fullerton et al., (2021) found that the deep biosphere was responsible for the additional removal of 1–13 % of the carbon inputs to the subduction zone, which suggests that the same mechanisms can be responsible for

trapping nitrogen. Furthermore, a recent study found the presence of a high abundance of genes related to nitrogen metabolism in the arc and forearc of the Costa Rica subduction zone (Rogers et al., 2023). The presence of these genes can hint on the possibility of these communities to rely on degassed nitrogen from the subsurface.

Sulfur is one of the key elements capable of connecting the four main spheres of our planet e.g., atmosphere, lithosphere, hydrosphere and biosphere, and therefore plays a fundamental role in many geobiological processes. It has a significant impact on the partition coefficients between silicates and metal minerals in Earth's accretion and core formation, as evidenced by the fact that sulfur is a key element during early metal-silicate partitioning on Earth (Wood et al., 2014). Additionally, it is one of the most important volatile components in magmas, surface degassing processes and in the catalysis of atmospheric ozone depletion, as a result of anthropogenic emissions, biological processes and oxidative weathering of continental sulfides (Farquhar et al., 2000). Sulfur occurs in several isotopically distinct reservoirs, probably due to the effect of higher surface UV resulting from a reduced ozone column depth of 2090 Ma in the past (Zahnle and Walker, 1982). In fact, SO_3 and H_2SO_4 compounds are formed from the photochemical oxidation of SO_2 in contact with water (Wayne, 1979). In mineralization processes, sulfur is one of the main players in the formation of ore deposits e.g., arsenopyrite, pyrite ore, chalcopyrite, and more generally, sulfide minerals (Simon and Ripley, 2011). Sulfur can transition from different oxidation states into different compounds ranging from 6+ to 2- with strong redox ability in geological and biological reservoirs e.g., S^0, S^{2-} in H_2S, S^{4+} in SO_2 and S^{6+} in SO_4^{2-}. These properties made sulfur one of the fundamental players in understanding the co-evolution of Earth systems. For instance, at the surface of many bio-geological processes, such as the mobilization of mineral sulfides in crushed industrial ores, and in biomineralization pathways (Li et al., 2021). From a compositional perspective, marine sediments, serpentinized peridotite and altered oceanic in subduction zones are rich in sulfur (Canfield, 2004), however, the determination of the subduction input of sulfur in these areas (Figure 4) is subject to large uncertainties due to the large variation in its concentration in the oceanic lithosphere. Only 6.3 % to 20 % of the total sulfur budget is released from the descending plate (the fluid-driven release pulse occurs at 90 km depth in subduction zones (Walters et al., 2020). The 25 % to 40 % of the total sulfur subducts in the altered oceanic crust is lost to dehydrating plate fluids, but little is known about the mobilization and redistribution of sulfur during fluid-rock interactions caused by fluid flows within the plate (Li et al., 2021). The redox state of sulfur during metamorphism and its redistribution during the above processes can be measured through in situ measurements of sulfur isotopes. For instance, the studies from Li et al., (2021) showed that the mobilization and redistribution during intra-plate fluid-rock interactions play an important role in desulfurization and sulfur recycling in subduction zones. The biological recycling of sulfur in subduction zones is still uncertain. The potential activity of sulfate-reducing microorganisms, in the Mariana forearc, has been suggested in the South Chamorro Serpentinite Seamounts (serpentinite mud volcanoes in non-accretionary forearcs on convergent margins provide a unique opportunity to study geochemical and microbial activity in subduction zones) as a basis of the sulfate depletion and sulfide enrichment in pore water (Mottl et al., 2003). However, the vertical distribution of sulfate-reducing microorganisms and the origin of the sulfate are still uncertain. Mottl et al, 2003, shows that microorganisms capable of producing carbonate, bisulfide, and ammonia are present in the forearc. Archeal biomass is exceptionally high in subsurface samples (from phospholipid-derived diphytanoyl diether biomarkers), and it has a bimodal distribution with the highest peaks at the same depths as the bisulfide maxima. This suggests that archaea oxidize methane and reduce sulfate, in a process known as anaerobic oxidation of methane, through a novel alkaliphilic-lithotrophic pathway (Valentine, 2002). There is no clear evidence of sulfate-reducing bacteria at different key depths. (Mottl et al., 2003).

Conclusions

In summary, tectonic processes have been affecting the planetary redox potential, conditioning atmospheric composition, climate stability, and planet habitability over the geological time scale. Convergent margins are directly involved in driving the continuous cycling of volatile species between the xenosphere and the interior of the planet. Host lithology, pressure, and temperature lead the metamorphic dehydration reactions influencing the fraction of volatiles that is recycled back to the surface. Quality and quantity of volatiles have deeply affected the Earth's subsurface geochemistry. Where the thermal gradients are not constraining for life, the chemical energy supplied by different elements have been supporting the wide subsurface microbial communities which mediate redox reactions and control the biogeochemical cycles. In subduction zones, microbial communities have been found to have a pivotal role in leading the sequestration and the releasing of several compounds back to the atmosphere. However, most of the knowledge present in literature is strictly linked to the carbon cycle rather than nitrogen and sulfur. Although these elements are equally important to understand the intrinsic dynamics that govern the subsurface biosphere, most of the investigations have been focused on the input and output of carbon compounds during the subduction process. To date, the number of studies about the geobiological processes in convergent margins is still poor and it is focused on small areas. For this reason, new large spatial-scale investigations about the interactions between geosphere and biosphere processes are needed to clarify what feedbacks are present between deep and surface Earth.

Figures

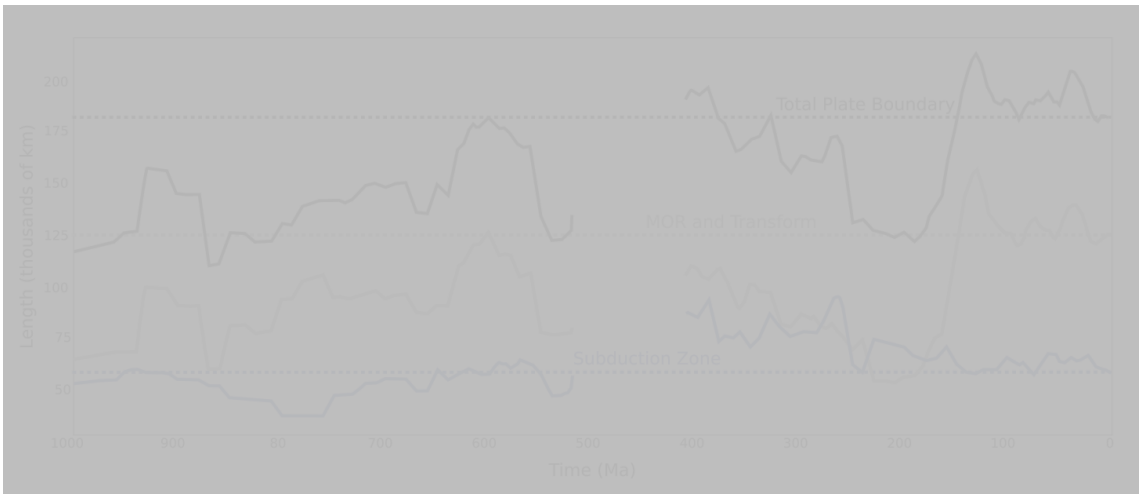


Figure 1. Length (in thousands of km) of the total plate boundaries (dark gray), Mid Ocean Ridges (MOR) and transform boundaries (light gray) and subduction zone length (light blue) through the last 1 Ga (adapted from Merdith et al., 2017) calculated as 10 Myr rolling average. The present-day length of plate boundaries are plotted as straight dashed lines.

Figure 2. Map of the extension of present day subduction zones globally.

Figure 3. Schematic representation of a convergent margin (adapted from Giovannelli et al., 2020) showing the subducting slab, the overriding plate, the major provinces (outer forearc, forearc, volcanic arc and backarc) and some of the key features. The area highlighted in cyano represents the portion of the subducting slab and crust potentially below 135-150 ˚C.

Figure 4. Pathway and major species of Carbon (dark gray), Sulfur (dark orange) and Nitrogen (blue) recycling at convergent margins. Dashed lines represent a lack of knowledge regarding the source and circulation pathway for the specific element). Major species for each element are reported. Question marks indicate lack of information regarding the abundance and origin of the specific molecule in each province, while the molecules in light gray in the outer forearc and forearc represent molecules derived from biological transformation at depth.

References

Avard, G., Mora, M. M., Bakkar, H., Angarita, M., and Cascante, M. (2021). Peligro y vigilancia volcánica en Costa Rica.

Balkwill, D. L., Fredrickson, J. K., and Thomas, J. M. (1989). Vertical and horizontal variations in the physiological diversity of the aerobic chemoheterotrophic bacterial microflora in deep southeast coastal plain subsurface sediments. *Applied and Environmental Microbiology* 55, 1058–1065.

Barry, P. H., de Moor, J. M., Giovannelli, D., Schrenk, M., Hummer, D. R., Lopez, T., et al. (2019a). Forearc carbon sink reduces long-term volatile recycling into the mantle. *Nature* 568, 487–492. doi: 10.1038/s41586-019-1131-5.

Barry, P. H., Nakagawa, M., Giovannelli, D., Maarten de Moor, J., Schrenk, M., Seltzer, A. M., et al. (2019b). Helium, inorganic and organic carbon isotopes of fluids and gases across the Costa Rica convergent margin. *Sci Data* 6, 284. doi: 10.1038/s41597-019-0302-4.

Bebout, G. E., Lazzeri, K. E., and Geiger, C. A. (2016). Pathways for Nitrogen Cycling in Earth's Crust and Upper Mantle: A Review and New Results for Microporous Beryl and Cordierite - Invited Centennial Article.

Beeman, R. E., and Suflita, J. M. (1987). Microbial Ecology of a Shallow Unconfined Ground Water Aquifer Polluted by Municipal Landfill Leachate. *Microbial Ecology* 14, 39–54.

Bekaert, D. V., Turner, S. J., Broadley, M. W., Barnes, J. D., Halldórsson, S. A., Labidi, J., et al. (2021). Subduction-Driven Volatile Recycling: A Global Mass Balance. *Annual Review of Earth and Planetary Sciences* 49, 37–70. doi: 10.1146/annurev-earth-071620-055024.

Bergin, E. A., Blake, G. A., Ciesla, F., Hirschmann, M. M., and Li, J. (2015a). Tracing the ingredients for a habitable earth from interstellar space through planet formation. *Proceedings of the National Academy of Sciences* 112, 8965–8970. doi: 10.1073/pnas.1500954112.

Bergin, E. A., Blake, G. A., Ciesla, F., Hirschmann, M. M., and Li, J. (2015b). Tracing the ingredients for a habitable earth from interstellar space through planet formation. *Proceedings of the National Academy of Sciences* 112, 8965–8970. doi: 10.1073/pnas.1500954112.

Borrel, G., Adam, P. S., and Gribaldo, S. (2016). Methanogenesis and the Wood–Ljungdahl Pathway: An Ancient, Versatile, and Fragile Association. *Genome Biology and Evolution* 8, 1706–1711. doi: 10.1093/gbe/evw114.

Boyd, S. R. (2001). Nitrogen in future biosphere studies. *Chemical Geology* 176, 1–30. doi: 10.1016/S0009-2541(00)00405-8.

Canfield, D. E. (2004). The evolution of the Earth surface sulfur reservoir. *American Journal of Science* 304, 839–861. doi: 10.2475/ajs.304.10.839.

Cord-Ruwisch, R., Seitz, H.-J., and Conrad, R. (1988). The capacity of hydrogenotrophic anaerobic bacteria to compete for traces of hydrogen depends on the redox potential of the terminal electron acceptor. *Archives of Microbiology* 149, 350–357.

Crisp, D., Dolman, H., Tanhua, T., McKinley, G. A., Hauck, J., Bastos, A., et al. (2022). How Well Do We Understand the Land-Ocean-Atmosphere Carbon Cycle? *Reviews of Geophysics* 60, e2021RG000736. doi: 10.1029/2021RG000736.

Dean, W. E., and Gorham, E. (1998). Magnitude and significance of carbon burial in lakes, reservoirs, and peatlands. *Geology* 26, 535–538.

Di Piazza, A., Rizzo, A. L., Barberi, F., Carapezza, M. L., De Astis, G., Romano, C., et al. (2015). Geochemistry of the mantle source and magma feeding system beneath Turrialba volcano, Costa Rica. *Lithos* 232, 319–335. doi: 10.1016/j.lithos.2015.07.012.

Dumitru, T. A. (1991). Effects of subduction parameters on geothermal gradients in forearcs, with an application to Franciscan Subduction in California. *Journal of Geophysical Research: Solid Earth* 96, 621–641. doi: 10.1029/90JB01913.

Dunham, E. C., Dore, J. E., Skidmore, M. L., Roden, E. E., and Boyd, E. S. (2021). Lithogenic

hydrogen supports microbial primary production in subglacial and proglacial environments. *Proceedings of the National Academy of Sciences* 118, e2007051117. doi: 10.1073/pnas.2007051117.

Evans, K. A. (2012). The redox budget of subduction zones. *Earth-Science Reviews* 113, 11–32. doi: 10.1016/j.earscirev.2012.03.003.

Falkowski, P., Scholes, R. J., Boyle, E., Canadell, J., Canfield, D., Elser, J., et al. (2000). The Global Carbon Cycle: A Test of Our Knowledge of Earth as a System. *Science* 290, 291–296. doi: 10.1126/science.290.5490.291.

Farquhar, J., Bao, H., and Thiemens, M. (2000). Atmospheric Influence of Earth's Earliest Sulfur Cycle. *Science* 289, 756–758. doi: 10.1126/science.289.5480.756.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *science* 281, 237–240.

Fowler, D., Coyle, M., Skiba, U., Sutton, M. A., Cape, J. N., Reis, S., et al. (2013). The global nitrogen cycle in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368, 20130164. doi: 10.1098/rstb.2013.0164.

Frost, D. J., and McCammon, C. A. (2008). The redox state of Earth's mantle. *Annu. Rev. Earth Planet. Sci.* 36, 389–420.

Fullerton, K. M., Schrenk, M. O., Yücel, M., Manini, E., Basili, M., Rogers, T. J., et al. (2021). Effect of tectonic processes on biosphere–geosphere feedbacks across a convergent margin. *Nat. Geosci.* 14, 301–306. doi: 10.1038/s41561-021-00725-0.

Galvez, M. E., Fischer, W. W., Jaccard, S. L., and Eglinton, T. I. (2020). Materials and pathways of the organic carbon cycle through time. *Nat. Geosci.* 13, 535–546. doi: 10.1038/s41561-020-0563-8.

Galvez, M. E., and Pubellier, M. (2019). "How do subduction zones regulate the carbon cycle?," in *Deep carbon: Past to present* (Cambridge University Press), 276–312.

Guerrero-Cruz, S., Vaksmaa, A., Horn, M. A., Niemann, H., Pijuan, M., and Ho, A. (2021). Methanotrophs: Discoveries, Environmental Relevance, and a Perspective on Current and Future Applications. *Frontiers in Microbiology* 12. Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.678057 [Accessed January 31, 2023].

Halama, R., and Bebout, G. (2021). Earth's Nitrogen and Carbon Cycles. *Space Sci Rev* 217, 45. doi: 10.1007/s11214-021-00826-7.

Hayes, J. M., and Waldbauer, J. R. (2006). The carbon cycle and associated redox processes through time. *Philosophical Transactions of the Royal Society B: Biological Sciences* 361, 931–950. doi: 10.1098/rstb.2006.1840.

Hazen, R. M., and Schiffries, C. M. (2013). Why deep carbon? *Reviews in Mineralogy and Geochemistry* 75, 1–6.

Hu, J., and Gurnis, M. (2020). Subduction Duration and Slab Dip. *Geochemistry, Geophysics, Geosystems* 21, e2019GC008862. doi: 10.1029/2019GC008862.

Isson, T. T., Planavsky, N. J., Coogan, L. A., Stewart, E. M., Ague, J. J., Bolton, E. W., et al. (2020). Evolution of the Global Carbon Cycle and Climate Regulation on Earth. *Global Biogeochemical Cycles* 34, e2018GB006061. doi: 10.1029/2018GB006061.

Kelemen, P. B., and Manning, C. E. (2015). Reevaluating carbon fluxes in subduction zones, what goes down, mostly comes up. *Proceedings of the National Academy of Sciences* 112, E3997–E4006.

King, G. M., Kirchman, D., Salyers, A. A., Schlesinger, W., and Tiedje, J. M. (2006). Global environmental change: microbial contributions, microbial solutions 2010: 15. *American Society for Microbiology, Washington, DC*.

Knittel, K., and Boetius, A. (2009). Anaerobic Oxidation of Methane: Progress with an Unknown Process. *Annual Review of Microbiology* 63, 311–334. doi: 10.1146/annurev.micro.61.080706.093130.

Kwon, E. Y., Primeau, F., and Sarmiento, J. L. (2009). The impact of remineralization depth on the air–sea carbon balance. *Nature Geosci* 2, 630–635. doi: 10.1038/ngeo612.

Lee, H., Fischer, T. P., de Moor, J. M., Sharp, Z. D., Takahata, N., and Sano, Y. (2017). Nitrogen recycling at the Costa Rican subduction zone: The role of incoming plate structure. *Sci Rep* 7, 13933. doi: 10.1038/s41598-017-14287-y.

Li, J.-L., Schwarzenbach, E. M., John, T., Ague, J. J., Tassara, S., Gao, J., et al. (2021). Subduction zone sulfur mobilization and redistribution by intraslab fluid–rock interaction. *Geochimica et Cosmochimica Acta* 297, 40–64. doi: 10.1016/j.gca.2021.01.011.

Li, L.-G., Xia, Y., and Zhang, T. (2017). Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection. *The ISME journal* 11, 651–662.

Matthews, E. (1994). Nitrogenous fertilizers: Global distribution of consumption and associated emissions of nitrous oxide and ammonia. *Global Biogeochemical Cycles* 8, 411–439. doi: 10.1029/94GB01906.

Mitchell, E. C., Fischer, T. P., Hilton, D. R., Hauri, E. H., Shaw, A. M., de Moor, J. M., et al. (2010). Nitrogen sources and recycling at subduction zones: Insights from the Izu-Bonin-Mariana arc: IBM NITROGEN SYSTEMATICS AND RECYCLING. *Geochem. Geophys. Geosyst.* 11, n/a-n/a. doi: 10.1029/2009GC002783.

Moore, E. K., Jelen, B. I., Giovannelli, D., Raanan, H., and Falkowski, P. G. (2017). Metal availability and the expanding network of microbial metabolisms in the Archaean eon. *Nature Geosci* 10, 629–636. doi: 10.1038/ngeo3006.

Parnell, J. (2004). Plate tectonics, surface mineralogy, and the early evolution of life. *International Journal of Astrobiology* 3, 131–137. doi: 10.1017/S1473550404002101.

Pi, H.-W. (n.d.). Origin and Evolution of Nitrogen Fixation in Prokaryotes.

Plank, T., and Manning, C. E. (2019). Subducting carbon. *Nature* 574, 343–352. doi: 10.1038/s41586-019-1643-z.

Ranero, C. R., Phipps Morgan, J., McIntosh, K., and Reichert, C. (2003). Bending-related faulting and mantle serpentinization at the Middle America trench. *Nature* 425, 367–373.

Richardson, D. J. (2000). Bacterial respiration : a flexible process for a changing environment.

Rogers, T. J., Buongiorno, J., Jessen, G. L., Schrenk, M. O., Fordyce, J. A., de Moor, J. M., et al. (2022). Chemolithoautotroph distributions across the subsurface of a convergent margin. *The ISME Journal*. doi: 10.1038/s41396-022-01331-7.

Schaefer, L., Lodders, K., and Fegley, B. (2012). Vaporization of the Earth: Application to exoplanet atmospheres. *The Astrophysical Journal* 755, 41.

Sieber, J. R., McInerney, M. J., and Gunsalus, R. P. (2012). Genomic insights into syntrophy: the paradigm for anaerobic metabolic cooperation. *Annual review of microbiology* 66, 429–452.

Simon, A. C., and Ripley, E. M. (2011). The Role of Magmatic Sulfur in the Formation of Ore Deposits. *Reviews in Mineralogy and Geochemistry* 73, 513–578. doi: 10.2138/rmg.2011.73.16.

Spandler, C., and Pirard, C. (2013). Element recycling from subducting slabs to arc crust: A review. *Lithos* 170–171, 208–223. doi: 10.1016/j.lithos.2013.02.016.

Stevens, T. O., McKinley, J. P., and Fredrickson, J. K. (1993). Bacteria associated with deep, alkaline, anaerobic groundwaters in Southeast Washington. *Microb Ecol* 25, 35–50. doi: 10.1007/BF00182128.

Ussiri, D. A. N., and Lal, R. (2017). "Introduction to Global Carbon Cycling: An Overview of the Global Carbon Cycle," in *Carbon Sequestration for Climate Change Mitigation and Adaptation*, eds. D. A. N. Ussiri and R. Lal (Cham: Springer International Publishing), 61–76. doi: 10.1007/978-3-319-53845-7_3.

Valentine, D. L. (2002). Biogeochemistry and microbial ecology of methane oxidation in anoxic environments: a review. *Antonie Van Leeuwenhoek* 81, 271–282. doi: 10.1023/A:1020587206351.

Vitale Brovarone, A., Sverjensky, D. A., Piccoli, F., Ressico, F., Giovannelli, D., and Daniel, I. (2020). Subduction hides high-pressure sources of energy that may feed the deep

subsurface biosphere. *Nat Commun* 11, 3880. doi: 10.1038/s41467-020-17342-x.

Walters, J. B., Cruz-Uribe, A. M., and Marschall, H. R. (2020). Sulfur loss from subducted altered oceanic crust and implications for mantle oxidation. *Geochem. Persp. Let.*, 36–41. doi: 10.7185/geochemlet.2011.

Wood, B. J., Kiseeva, E. S., and Mirolo, F. J. (2014). Accretion and core formation: The effects of sulfur on metal–silicate partition coefficients. *Geochimica et Cosmochimica Acta* 145, 248–267. doi: 10.1016/j.gca.2014.09.002.

Zahnle, K. J., and Walker, J. C. G. (1982). The evolution of solar ultraviolet luminosity. *Reviews of Geophysics* 20, 280–292. doi: 10.1029/RG020i002p00280.

Zellmer, G. F., Edmonds, M., and Straub, S. M. (2015). Volatiles in subduction zone magmatism. *Geological Society, London, Special Publications* 410, 1–17. doi: 10.1144/SP410.13.

Zerkle, A. L., and Mikhail, S. (2017). The geobiological nitrogen cycle: From microbes to the mantle. *Geobiology* 15, 343–352. doi: 10.1111/gbi.12228.

Zheng, Y.-F. (2019). Subduction zone geochemistry. *Geoscience Frontiers* 10, 1223–1254. doi: 10.1016/j.gsf.2019.02.003.

Figures

Chapter 2

Aims and Objectives

This thesis aims to increase our knowledge of the interactions between biosphere and geosphere with a first focus on the microbial overprint on the volatiles that are cycled at convergent margins. Hot springs and other secondary geothermal manifestations, broadly referred to as deeply-sourced seeps, are used throughout this thesis to access the deep subsurface from the surface, an approach known as "window into the subsurface" originally described by (Deeming and Baross, 1993) and recently adapted by Giovannelli et al. (2022) to the large scale approach linking geology and biology underlying this thesis. Besides their use as a window to the subsurface, hot springs have been used for centuries for leisure as well as religious rituals by human populations. Given the large number of samples collected in this work, one of the secondary aims has been to investigate the anthropic use of the springs and its link to the potential presence of pathogens.

Specific questions addressed in this work are:

- What biological processes influence volatile cycling in convergent margins? To what extent biology can influence volatile cycling? Are there specific links with geological processes?
- In what way geosphere and biosphere interact to alter volatiles in secondary geothermal manifestations? What types of feedback between geology and biology are present? What kind of consequences may these produce?
- Can human presence and activities affect the hot springs microbial communities? Are there implications for public health?

The entire work presented in this thesis deals with microbial composition and fundamental functions from numerous deeply sourced seeps and secondary geothermal emissions, providing new information between microbial communities and volatiles that are cycled through the convergent margins. The used approach is a mix of comparative genomic and geochemistry techniques combined with molecular assays, integrating laboratory and field survey. Deeply sourced seeps and secondary geothermal emissions were sampled and analyzed, looking at the prokaryotic community structure and at specific subgroups present *in situ*, the driving factors shaping their functional diversity, and their role in specific biogeochemical cycles.

The samples analyzed in this work were obtained from three different scientific campaigns conducted across Costa Rica and Panama, in 2017, 2018, and 2019, spanning a 700 km segment of the Central American Convergent Margin. Given the multidisciplinary nature of this work, in addition to the University of Naples Federico II, a large number of international partners from

diverse institutions have contributed to data analysis and interpretations. This includes researchers from the University of Tennessee, Knoxville (USA), Woods Hole Oceanographic Institution (USA), the Observatorio Volcanológico y Sismológico de Costa Rica (OVSICORI), Universidad Nacional (Costa Rica) and many others.

The thesis opens with a mini-review on the geobiological cycling of volatiles in convergent margins (Chapter 1). The relationship between carbon cycling, with a focus on methane, microbial diversity is investigated using a combination of advanced geochemical techniques and metagenomic approaches (Chapter 3). A meromictic lake in Costa Rica is selected as a case study to investigate the feedback between geological and biological processes affecting the carbon cycle and the potential effects of microbial metabolism on limnic eruptions (Chapter 4). The role of recreational hot springs as potential reservoirs for multi-resistant pathogens is investigated (Chapter 5), linking the anthropic use of these locations with potential public health hazards. Finally, the information obtained during the doctorate project is combined together in a short overview linking the main findings along a common thread (Chapter 6).

References

Deming, J. W., and Baross, J. A. (1993). Deep-sea smokers: Windows to a subsurface biosphere? Geochimica et Cosmochimica Acta 57, 3219–3230. doi: 10.1016/0016-7037(93)90535-5.

Giovannelli, D., Barry, P. H., de Moor, J. M., Jessen, G. L., Schrenk, M. O., and Lloyd, K. G. (2022). Sampling across large-scale geological gradients to study geosphere–biosphere interactions. Frontiers in Microbiology 13. doi: 10.3389/fmicb.2022.998133.

Chapter 3

Origin and fate of methane in the Central American convergent margin

Abstract

Convergent margins are gateways to Earth's interior where volatile species such as carbon, water, hydrogen, and sulfur are cycled between the surface and the interior of the planet. At these locations, carbon is recycled from deep reservoirs in two main forms: oxidized carbon in the form of carbon dioxide, and reduced carbon in the form of methane. While the former is quantitatively more important and its volcanic fluxes have been better constrained, the latter represents the most reduced possible form of carbon and greatly contributes to greenhouse effects and climate stability. Understanding the geological and biological processes underpinning the origin and fate of methane in convergent margins is thus pivotal to constrain carbon cycling and redox balance in convergent margins. Here we present coupled data from 47 geothermal deeply-sourced seeps spanning the Costa Rica and Panama convergent margin. By analyzing the presence and diversity of methane-cycling microorganisms and using clumped isotope data, we are able to provide an unprecedented snapshot of the geobiological processes controlling methane cycling in convergent margins.

Keywords: Convergent margins, methane cycle, clumped isotopes, deep biosphere, carbon cycle

Introduction

The Earth's history has been shaped by important ecological shifts driven by metabolic advancements, which significantly altered the composition of the oceans and atmosphere. Despite hundreds of millions of years passing before the evolution of photosynthesis, a broad knowledge about these early transitions is still lacking. Earth's carbon cycle and the continuous cycling of volatile species such as carbon, water, hydrogen, and sulfur between Earth's surface and deep reservoirs likely created and maintained temperate climatic conditions, contributing to keep our planet habitable for at least the last 3.8 billion years (1, 2). Plate tectonics controls volatile recycling mainly through plate recycling, mantle convection and volcanisms (1). These conditions soon hosted the first ecosystems of the early Earth in the late Hadean and early Archean eras, most likely involving chemolithotrophs releasing methane as a byproduct (3). The development of methanogenic metabolisms had a profound impact on atmospheric composition, leading to a warmer climate.

Convergent margins are gateways between the oxidized Earth's surface and the reduced interior where volatile species are cycled (4). The quantity and speciation of carbon subducted and recycled to the surface affect the redox potential of the Earth's surface and alter the atmospheric

composition over geological timescales influencing climate (5, 6). During its ascent from the mantle to the surface, carbon is altered by complex geological and biological processes, capable of transforming its redox state and altering its mobility and residence time in the crust (1, 7, 8). For example, carbon dioxide released from the subducting slab can precipitate as calcite during its ascent in the forearc region of convergent margins (7). At convergent margins carbon can be released through volcanism and secondary geothermal manifestation in two main forms: oxidized carbon as carbon dioxide (CO_2) and reduced carbon as methane (CH_4) (5). Other quantitatively less important forms of carbon with intermediate redox oxidation states are released to the surface, such as carbon monoxide (CO) or organic carbon. While carbon dioxide represents the vast majority of the carbon released through volcanism (9) and its origin, global flux and contribution to Earth's atmospheric composition and climate stability through time has been somehow constrained, methane's origin and fate at convergent margin has received comparatively less attention (8), and studies have been generally focused on methane released underwater near plate boundaries (10, 11).

Methane is a powerful greenhouse gas that plays a significant role in regulating Earth's climate. Its global warming potential is 28 times greater than that of carbon dioxide over a 100-year time frame (12). Methane is also important for understanding the balance of reducing equivalents in Earth's interior (8), and the formation of methane, regardless of the processes, consumes electrons removing reducing equivalents from the mantle (5). Recently, convergent margins have been shown to hide large natural fluxes of methane to the atmosphere (8). As such, understanding the sources and sinks of methane at plate boundaries is crucial to constrain carbon cycling and Earth redox balance.

Methane can be produced through a variety of natural processes, including microbial methanogenesis (13), thermogenic generation in sedimentary basins (14), abiotic synthesis following serpentinization of subducted oceanic crust and mantle wedge (15) and metamorphism of subducted organic-rich sediments (16). Methane generated deep by subduction processes and shallower in the crust by biological activity is then transported to the surface by fluids and gases (8). During ascent, biological and geological processes alter the quantity and isotopic signature of the released methane, confounding the primary source with secondary processes. Previous methods based on the utilization of carbon ($^{13}C/^{12}C$) and hydrogen (D/H) isotope ratio (51) were unsuccessful because of the substantial overlap in isotopic signatures associated with microbial, thermogenic, and abiogenic gases (27). Recent advances in doubly substituted (clumped) methane isotopologue analysis provides a key tool to estimate the temperature at which a sample of methane was formed or thermally equilibrated, unlocking critical information on the origin of methane in geological reservoirs (17).

The abundance and isotopic composition of CH_4 has been reported for hot springs and volcanic hosted geothermal systems in recent years (17–22) revealing a high variability in the isotopic signature of methane in volcanic hosted systems. Multiple hypotheses have been proposed to explain the obtained results, including abiotic generation from mantle derived CO_2, pyrolysis of organic matter in high temperature geothermal systems, thermogenic origin from surface derived or ancient organic matter, microbial production from mantle derived CO_2 or surface-derived sources of carbon, or a combination of the above (17–22). The large variability in isotopic signatures observed in geothermal systems has not been systematically linked to any measured variable, and a clear picture of the origin of methane and secondary geological and

biological processes altering its quantity and isotopic composition at convergent margins is still lacking.

Here, we report results on the gas geochemistry, methane isotopologues, aqueous geochemistry and microbiology of 47 deeply-sourced seeps (*sensu* (23) spanning the Costa Rica and Panama convergent margins. We link the subsurface and near-surface microbial community involved in methane cycling to gas, aqueous and solid phase geochemistry, different lithologies and to methane isotopologues across diverse rock reservoirs of the Central American convergent margins (Figure 1). In this area, the Cocos and the Nazca plates subduct beneath the Caribbean plate, affecting the geologic evolution of the upper plate and influencing latitudinal gradients in the the degree of distention and compression, the nature of volcanism, the presence of sedimentary basins, and the development of a backarc thrust belt (24). Our results show that the shift in the methane isotopic composition is consistent with a overlapping geological and biological processes controlled by the geological settings and dominant rock type present in the area, ultimately showing that the origin and fate of methane in the Central American convergent margin is controlled by complex feedbacks between deep and surface Earth processes.

Materials and Methods

Sampling approach

We collected 47 samples from diverse geothermal features including geothermal springs, steam heated waters, mud pools, fumaroles, acid-sulfate springs, sodium chloride springs, bicarbonate springs and alkaline springs, collectively referred to as deeply-sourced seeps by Giovannelli et al., (2022) (23), across a 700 km section of the Central American convergent margin during two field campaigns in 2017 and 2018. We used field observation, pH and temperature measurements to identify the main inlet in each seep and collected samples in close proximity to minimize surface contamination. Gas phase samples and water samples were collected in pre-evacuated 250 mL Giggenbach bottles containing 50 mL of 4 M NaOH and samples for methane clumped isotopologue determination were collected into 1L pre-evacuated glass bottles containing ~ 200 mL 4 M NaOH. Hydrothermal fluids (up to 2 L) were collected using a 0.22 μm sterivex filter and where present, about 30 g of sediments were collected into a sterile falcon tube. All samples were immediately frozen at −196 °C for transport and analysis. Trace metal and major ion measurements were performed on filtered subsamples and samples for trace elements in the solid-sediment fraction were collected in acid washed 15 ml falcon tubes and stored at 4 °C. Fluids were also collected in gas tight acid washed vials for carbon isotope analysis of dissolved inorganic carbon (DIC).

Gas geochemistry

Giggenbach bottle samples for gas composition were analyzed as previously described (25). Briefly, the bottles were connected to a GC via a vacuum line and turbo pump. The headspace gas was introduced to the GC through two parallel columns using two 250 μL loops and two 6-way valves switching simultaneously. One column operated with Ar carrier gas and the other with H_2 carrier gas. Methane was analyzed on the H_2 carrier gas column with a flame ionization detector and other gasses were analyzed on thermal conductivity detectors. The CH_4

concentration in water samples was calculated from the mass of water collected and the partial pressure of CH_4 measured in the headspace. The CO_2/CH_4 ratio in gas samples was determined by calculating the total moles of CO_2 and CH_4 collected in the Giggenbach bottle. The error for CH_4 concentrations and CO_2/CH_4 ratios was estimated at less than 5 % and less than 10 % respectively. Carbon isotope compositions of gas samples were analyzed at UNA on a Picarro G2201-I by acidifying NaOH solutions extracted from Giggenbach bottle samples. The CO_2-N_2 mixture was transferred to the Picarro G2201-I and analyzed for 8 minutes at a flow rate of ~25 ml/min. The $\delta^{13}C$-PDB values were calibrated against a set of 8 standards with uncertainties of <0.1 ‰ based on repeat analyses of standards and samples.

Methane clumped isotopes

Methane isotopologues were analyzed as previously described (26). Briefly, methane was extracted and purified from the gas phase of the flask samples by repeatedly vacuuming and flushing with a He carrier gas through a cold trap filled with activated charcoal submerged in liquid nitrogen, trapping other gasses such as CH_4, N_2, CO, and traces of CO_2. The trap was then heated to separate adsorbed gasses through a gas chromatography column packed with Carboxen-1000 (MilliporeSigma, St Louis, MI) held at 30 °C, and eluted methane was trapped again in a U-trap containing silica gel at liquid nitrogen temperature. The purified CH_4 samples were measured for $^{12}CH_4$, $^{13}CH_4$, $^{12}CH_3D$ and $^{13}CH_3D$ isotopologue composition using a tunable infrared laser direct absorption spectroscopy (TILDAS) instrument (26, 27). Samples were bracketed by measurements of methane standard gasses of known isotopic composition (δD and $\delta^{13}C$) spanning a large δD range (~200 ‰), which were heated to 250 °C for at least two weeks with a Pt catalyst to ensure isotopologue equilibrium. Isotope values are reported using standard delta notation against VPDB and VSMOW for the ratios $^{13}C/^{12}C$ and D/H, respectively. This isotope scale was calibrated by the measurements of NGS-1 and NGS-3, using reference $\delta^{13}C$ values of −29.0 and −72.8 ‰, and δD of −138 and −176 ‰, for NGS-1 and NGS-3, respectively.

Aqueous and sediment geochemistry

Anions and cations were determined via ion chromatography. For anions, a Dionex AS4A-SC separation column, sodium hydroxide eluent, and ASRS-I suppressor were used. For cations, a Dionex CS12-SC separation column, methane sulfonic acid eluent, and CSRS-I suppressor were used. Trace metal concentrations were determined in aqueous and acid-digested solid samples with a PerkinElmer NexIon 350X inductively coupled plasma mass spectrometry instrument using multi-element calibration standard solution of metals as described in our SOP.

Genomic DNA extraction and sequencing

DNA extractions from Sterivex filters were performed using a modified phenol-chloroform extraction optimized for low biomass samples based on previously published methods (28), with additional modifications for use with Sterivex filters (29). Briefly, extractions were performed via chemical lysis with lysozyme, proteinase-K and SDS treatment, then purified with phenol-chloroform extractions and precipitated with sodium acetate and isopropyl alcohol. DNA extractions from sediment samples were performed using the Qiagen DNeasy PowerSoil HTP 96 Kit. Difficult to extract samples were extracted using a modified phenol-chloroform

extraction followed by concentration using the Zymo Genomic DNA Clean and Concentrator Kit (Zymo). Extracted DNA was quantified using a NanoDrop 2000c with additional polymerase chain reaction screening performed using universal bacterial primers. High quality DNA was extracted from the fluids and sediments of the majority of the 46 sites; 31 sites worked for both fluids and sediments, 5 were successful only in sediments and 10 only in fluids. Amplicon sequencing and shotgun metagenome sequencing were performed as part of the Census of Deep Life initiative with the Deep Carbon Observatory on an Illumina MiSeq platform at the Marine Biological Laboratory sequencing facility (https://www.mbl.edu/). Amplicon sequencing is carried out after amplifying the bacteria-specific V4-V5 region of the 16S rRNA gene using primers 518F (AATTGGANTCAACGCCGG) and 926R (CCGYCAATTYMTTTRAGTTT). The same extracted DNA was used for shotgun metagenomic sequencing without amplification. The Illumina Nextera Flex kit for MiSeq+NextSeq was used and obtained shotgun metagenomes varied from 25 to 150 million base pairs. The amplicon sequencing data is available from the NCBI SRA archive with accession numbers PRJNA579365, while the shotgun metagenomes are deposited in PRJNA627197.

Bioinformatic and statistical approaches

Raw reads received from the sequencing center were processed using the DADA2 package (30). Primers and adapters were trimmed, followed by a quality profile step, where only sequences with a call quality for each base between 20 and 40, were kept for further analysis. Amplicon sequence variants (ASVs) were estimated through error profile, and assigned taxonomy with the SILVA database - release 138.1 (31). The resulting taxonomic assignments, in combination with variant abundance tables, were used with the phyloseq package (32). Sequences related to Chloroplasts, Mitochondria, and Eukaryotes, as well as groups related to human pathogens and common DNA extraction contaminants (33), were removed from the dataset. All downstream statistical analyses, data processing and plotting were carried out in the R statistical software version 4.2.2. Shotgun Metagenomic short reads were quality checked and trimmed using Trimmomatic (34), and then functionally annotated with Mifaser (35) using the provided Gold Standard Plus (GS21 ALL) database, including gene sequences from biogeochemically relevant pathways. Post QC raw reads were used to assess the taxonomic composition with a sensitive profiling tool called Kaiju (36). The obtained count tables were imported in R. To account for the influence of different gene lengths on the read recruitments, and for the different sequencing depth associated with each metagenome, read counts were normalized to the average length of the corresponding protein in the GS21-ALL database and multiplied by the median library size across the dataset. To account for the compositionality of sequencing data (9), we further divided each EC abundance by the abundance of the DNA directed RNA polymerase subunit β (rpoB gene), and scaled to 1,000,000. This procedure returns the relative abundance of each EC for 1,000,000 reads mapped to rpoB, assuming equal sequencing depth. This allows within-sample and between-sample comparison, effectively opening the data and removing the problem of compositionality. The obtained normalized counts (mifaser_nrpo) were used for downstream analyses. A complete R script containing all the steps to reproduce our analysis is available at https://github.com/giovannellilab/Selci_et_al_Origin_and_fate_of_methane and is released as a permanent version (version 1.0) using Zenodo with doi: XXXXXXXX.

Results

To understand the origin, the fate, and the distribution of methane seeps, amplicon libraries for methane microbes and metagenomic DNA sequences of key genes for methane cycling were compared to the different methane isotopic signatures observed in 47 seeps sampled along the Central America convergent margin (Figure 1A). The sampled deeply sourced seeps are located along a 700 km section of the Central American convergent margin, and positioned to cover the major provinces across the convergent margin: the outer forearc, the forearc, the volcanic arc and the backarc. Seeps positioned at each location show differences in the origin of the volatiles, with near arc sites showing higher Rc/Ra ratio (Supplementary Figure S1). The sampled seeps are also located in different geologic units of diverse composition, including ocean floor basalts, quaternary sedimentary deposits, sedimentary rocks and volcanic rocks (Figure 1B). The distribution of the different lithologies is partially related to the location with respect to the across-subduction axis, with ocean seafloor basalt present in the outer forearc, volcanic rocks in the main volcanic arc and deep in the forearc province, sedimentary rocks associated with uplifting and orogenic events and quaternary sedimentary deposits filling the basins at the margin of the forearc and backarc created by the local topography (Figure 1B).

Methane was detected in 39 of the 47 sampled sites Table 1), with concentrations ranging from 0.00009 to 864.2 mmol/mol of total dry gas. Carbon isotopic ratios for the methane ($\delta^{13}C_{CH4}$) were obtained only for 11 sites (Table 2), corresponding to the locations with the highest methane concentrations. The obtained values for $\delta^{13}C_{CH4}$ showed values ranging from -22.5 to -76.3 ‰. Methane cycling microorganisms, recognized based on their taxonomic annotation, were identified in nearly all sites that yielded amplifiable DNA (n=43 sites) with the exception of BQ, HN and SR. Sequences associated with known methane cycling microorganisms constitute in total 5.2 % of the identified ASVs, and ranged between 0.01 % (RRS) and 36.2 % (LWF) of the total sequences (Supplementary Figure S2), and were identified in 69 samples, of which 34 were fluids. The community was highly diverse, and represented by members of the major Bacteria and Archaea orders known to be involved in methane cycles, either through methanogenesis or methane oxidation (both aerobic and anaerobic).

To provide records on formation mechanism and condition molecules, we have used higher sensitivity methods which included the utilization of doubly substituted isotopologues (clumped isotopes; Figure 2A). All the investigated sites, with the exception of CI and BR1 fall in areas of the plot grouping samples with known formation mechanisms, in broad agreement with their placement in a Whiticar plot (Supplementary Figure S3). Samples like PS, LW, and CW fall within the “bacterial carbonate reduction” field (hydrogenotrophic methanogenesis) resulting in a light ^{13}C, as found in pore waters and gas hydrates of marine sediments (26). SI and SR are placed within the “bacterial methyl type and fermentation” field (methylotrophic and acetoclastic methanogenesis) which is D-depleted compared to the other. As expected, samples like BQ, HN, QN, BR1, and CI are enriched in ^{13}C and D, falling within the geothermal and thermogenic fields. Although the average value of δD_{CH4}, MT falls within the “mix and transition” field, since the lighter $\delta^{13}C_{CH4}$, suggesting mixing process during the methane formation.

Between the samples related to hydrogenotrophic methanogenesis (HM; Figure 2A), only PS ($\Delta^{13}CH_3D$=5.65 ‰) yields a $\Delta^{13}CH_3D$ similar to the microbial methane measured in pore waters

and gas hydrates from northern Cascadia margins (26), being also in line with the expected low formation temperatures in marine sediment basins. In PS sediments, anaerobic methane microbes have the highest abundance (Supplementary Figure S4C, S4D) with Methanosarcinales and ANME-1 as main groups. LW and CW samples are lightly depleted in $^{13}CH_3D$ ($\Delta^{13}CH_3D$=4.68 ‰ and $\Delta^{13}CH_3D$=4.48 ‰) showing a $\Delta^{13}CH_3D$ based-temperature not congruent with their expected formation temperatures.

In the LW fluids, aerobic methane microbes are dominant against the anaerobic groups (Figure 2B,C), highlighted by the higher abundance of Rhizobiales compared to the Methanosarcinales. In LW sediments, instead, anaerobic methane microbes are the dominant groups (Supplementary Figure S4C, D) with Methanosarcinales and Methanomassiliicoccales as most abundant taxa (Supplementary Figure S2). As LW, also CW sediments are characterized by anaerobic methane microbes (Figure 2B, C) like Methanosarcinales, Methanomassiliicoccales, and Methanobacteriales (Supplementary Figure S3C, D). Likewise, CW fluids are dominated by anaerobic methanogens such as Methanosarcinales and Methanobacteriales, with the addition of the aerobic group of the Rhizobiales.

Samples related to methylotrophic and acetoclastic methanogenesis (SI, and SR) as well as to "mix and transition" field (MT; Supplementary Figure S3) are highly depleted in $^{13}CH_3D$ ($\Delta^{13}CH_3D$=-0.59 ‰, $\Delta^{13}CH_3D$=0.28 ‰, and $\Delta^{13}CH_3D$=1.58 ‰, respectively) similarly to swamp and cow rumen samples (26). These low $\Delta^{13}CH_3D$ values correspond to improper high $\Delta^{13}CH_3D$ temperatures which indicate the absence of a thermodynamic equilibrium (26, 37). In SI, kinetic isotopic effects may be caused by the major presence of anaerobic methane microbes (Supplementary Figure S4D) which are dominated by Methanosarcinales, Methanobacteriales, and Methanomassiliicoccales. In a similar way, in MT, anaerobic methane microbes are dominant compared to the aerobic ones (Figure 2B,C) with Methanosarcinales as main taxa detected.

Samples associated to thermogenic methanogenesis *i.e.,* BQ, QN, and HN, yielded consistently low $\Delta^{13}CH_3D$ ($\Delta^{13}CH_3D$=1.65 ‰, $\Delta^{13}CH_3D$=1.76 ‰, and $\Delta^{13}CH_3D$=1.46 ‰, respectively) like similar geothermal systems (38) which corresponded to high $\Delta^{13}CH_3D$ based-temperature, remaining consistent with the expected formation temperatures. Among them, only in QN methane related microbes are detected, with a major abundance of anaerobic groups (Figure 2C). These are mainly composed of ANME-1 and Methanosarcinales in the fluids against the aerobic methane microbes which are dominant in the sediments (Supplementary Figure S4C) and represented by Methylococcales and Betaproteobacteriales. Although BR1 displayed a thermogenic $\delta^{13}C_{CH4}$ fingerprint (-26.7 ‰) similar to BQ, HN and QN (-29.9 ‰, -22.5 ‰, and -28.5 ‰), it showed an interesting enrichment in $^{13}CH_3D$ (3.48 ‰), suggested as the effect of anaerobic methane oxidizers which are able to isotopically equilibrate CH_4 with dissolved CO_2 when small concentration of sulfate (> 0.5 mM) are present (39, 40). In BR1, a concentration of 0.44 mM of sulfate is detected and a higher abundance of aerobic methane microbes is found (Figure 2B, C) with Methylomirabiales and Methylococcales as dominant groups in both fluid and sediment samples (Supplementary Figure S4). CI shows the lightest $\delta^{13}C_{CH4}$ fingerprint (-41.6 ‰) compared to the other thermogenic samples (described above) but the highest $\Delta^{13}CH_3D$ (8.32 ‰), suggesting kinetic isotopologue fractionation (non-equilibrium clumped value) as potential reason. This may be related by the prevalence of anaerobic methane microbial groups (Figure 2C) in fluid samples which are dominated by Methanosarcinales

unlike the sediment sample where Methanobacteriales is the highest in abundance (Supplementary Figure S2).

Looking at the methane cycle key genes (methane formation and oxidation), we have considered the sequences translating for the methyl-coenzyme M reductase (*mcrA*), the enzyme responsible for the last step of the biological methane formation by the reduction of a methyl group (41, 42) and the formylmethanofuran dehydrogenase (*fwd*), a key enzyme at beginning of the chain for the CH_4 formation from H_2/CO_2 (43) (Supplementary Figure S5). For the methane oxidation, the two evolutionarily related enzymes methane monooxygenase (*mmo*) and the methane/ammonia monooxygenase (*pmo-amo*) (44) are included. The methanol dehydrogenase (*mdh*) is also considered since it is a key intermediate in methylotrophic metabolism in which one-carbon (C1) compounds are reduced (45). While these genes were identified in the majority of the investigated metagenomes (n= 43 sites), their abundance is higher in sites with high methane concentrations measured in the gas phase. In PS sediments, the methanogenesis genes *mcrA* and *fwd* displayed the highest abundance compared to those (*i.e.*, *mmo*, *pmo-amo*, and *mdh*) involved in the methane oxidation and post methane oxidation pathways, suggesting a potential biological production of methane against its biological oxidation. Gene sequences related to the methane utilization like the *mmo* and the *pmo-amo* are found in highest abundance in both LW and CW fluids, as well as for the gene of the *mdh*. The highest abundance of sequences that translate for the *mcrA* is found, instead, within the LW and CW sediments, followed by the sequences for the *fwd*.

MT fluids are characterized by a higher number of methanogenesis gene sequences compared to those involved in methane oxidation. In SI, high content of sequences related to *mcrA* and *fwd* are mainly found in sediments compared to the fluids which showed a high amount of sequences related to methanol oxidation (*mdh*). In BQ, the higher number of sequences are associated with *fwd* and *mdh*, in both fluid and sediments, where in this latter, *mcrA* sequences and smaller amounts of *pmo-amo* and *mmo* are revealed. In a similar way, also QN fluids and sediments displayed both a bigger amount of sequences related to methane oxidation (high abundance in *mmo*, *pmo-amo*, and *mdh*) rather than methane formation processes. In BR1, fluids and sediments are both characterized by high quantities of *mmo* and *pmo-amo* sequences. Sequences for the *mcrA*, instead, are found only in smaller amounts within the fluids against their complete absence within the sediments. Similarly, CI fluids are dominated by sequences for the methane oxidation (*mmo* and *pmo-amo*) and a lower abundance of sequences related to *mcrA* and *fwd*. CI sediments displayed instead a main presence of sequences related to CO_2 utilization and methanol oxidation (*fwd* and *mdh*). The distribution of the key methane cycling functional genes on the clumped isotope $\Delta^{13}CH_3D$ values against $\delta^{13}C_{CH4}$ (Figure 3; Supplementary Figure S6) is in broad agreement with the distribution of the taxonomic sequences presented in Figure 2C and 2D.

When looking at the distribution of functional genes annotated at each sites using jaccard based PCoA multivariate ordinations (Supplementary Figure S7), a clear partition between fluids and sediments is present (adonis p.value<0.0001; Supplementary Figure S7C and S7D) both in the unweighted and weighted ordination, highlighting the dual distribution between the subsurface and surface community metabolic profiles along the Central America convergent margins. The different bedrock associated with each sampling site were also statistically significant (adonis, p<0.001; Supplementary Figure S7A and S7B). The percentage of variance explained was 16.2

% and 38.1 % for the weighted and unweighted Jaccard based PCoA, respectively. When considering only the enzymes involved in methane cycling assigned through KEGG modules instead, the percentage of explained variance increased to 29.3 % and 71.2 % for the weighted and unweighted Jaccard based PCoA, respectively. The difference between fluid and sediment was still present when considering methane cycling genes (adonis p-value<0.005), suggesting a distinct role of methane cycling in both the matrix types, but in this case the different bedrock composition was also significative (Figure 4A; adonis, p.value<0.0001), for both the weighted and unweighted ordination (Supplementary Figure S8A and S8B). The deeply sourced seeps distributed in quaternary deposits (QD), where methane gas relayed a strong biogenic signal (Figure 4B), showed a similar functional composition with the seeps located in ocean floor basalts (OFB), which also displayed a similar lighter $\delta^{13}C_{CH4}$. A different trend is observed with the seeps distributed along the central regions of Costa Rica and Panama, characterized by sedimentary and volcanic rocks (SR and VR) where the methane emissions had a thermogenic/abiotic signature (Figure 4B, Figure 2A).

Further information on the potential origin of the methane and CO_2 can be furnished by the relative compositions of the He isotopes and CO_2 or CH_4. In the ternary diagram of CO_2, ^3He and ^4He (Figure 4C) most of the samples located in volcanic rock reservoirs are characterized by higher mantle contributions ($^3He/^4He$ ratio above air) and only in two samples (BC and PF), located in Panama and in the Costa Rica backarc respectively, an enrichment in crustal CO_2 above average arc samples is observed (Figure 4C). Samples from sites hosted in sedimentary rocks and quaternary deposits are instead on average enriched in radiogenic helium compared to volcanic rock hosted samples. He isotopes and CH_4 ternary (Figure 4D) show an increase in CH_4 concentrations in volcanic rock hosted samples from amantle-like source, while a clear enrichment in CH_4 is visible in some of the sites hosted in quaternary deposits.

Discussion

Convergent margins are key locations for the recycling of volatiles between the Earth surface and interior (4). While carbon dioxide represents the vast majority of the carbon released through volcanism (9) and its origin, global flux and contribution to Earth's atmospheric composition and climate stability through time has been somehow constrained, methane's origin and fate at convergent margin has received comparatively less attention (8), and studies have been generally focused on methane released underwater near plate boundaries (10, 11). Additionally, although the carbon budget of volcanic arcs has been investigated in detail in the last two decades, information regarding the origin and fate of carbon recycled in the outer forearc, forearc and backarc regions are scarce (7). Understanding the relative contribution of the different carbon species recycled at convergent margins as well as their relative origins is a key step in constraining global carbon cycling.

While several studies have investigated the origin of methane on diverse tectonic settings (17, 20), detailed information on the origin of methane at convergent margin is limited to a few data points (17–22) that revealed a high variability in the isotopic signature of methane in volcanic hosted systems. Additionally, the biological overprint imposed by the large subsurface biosphere present in the crust (46) has been neglected so far. Biology has been shown to play a key role in altering the quality and quantity of CO_2 recycled in the forearc region of convergent margins

(29, 47) and backarc regions (48), suggesting a potential role in methane cycling as well. Here we present gas geochemistry and clumped methane isotope data coupled to metagenomic-based microbial diversity data from deeply-sourced seeps of the Central America convergent margin.

Our results show that methane is present and detectable in the majority of the sampled seeps distributed across the 700 km segment of the Central America convergent margin (Figure 1A and Table 1). While the concentrations of methane varied by 6 orders of magnitude in the sampled seeps, the highest concentrations of methane were linked to sites hosted in sedimentary quaternary deposits (Figure 1B and Figure 4). Microbial diversity and functional data suggest that the communities present in the sites hosted in quaternary deposits have the potential to produce the observed methane (Figure 3, 4 and Supplementary Figure S2), similarly to sites hosted in ocean floor basalts. This observation is consistent with the predicted biological origin of the methane using clumped isotope data (Figure 2), that clearly places these sites within known areas ascribed to hydrogenotrophic or methylotrophic and acetoclastic methanogens. Conversely, sites present in volcanic and sedimentary rocks show a higher potential for aerobic and anaerobic utilization of methane (Figure 3, 4 and Supplementary Figure S2), especially considering that the highest abundances sequences related to known methane cycling microorganisms is present in these sites in the surface sediments rather than the deep-derived fluids (Supplementary Figure S2). Taken together our observations suggest that the sampled sites belong to four distinct categories with regard to the methane origin broadly consistent with the bedrock each site population is hosted in (Figure 4).

The first group is composed of outer forearc sites hosted in ocean floor basalt and other mafic rocks where the presence of highly reduced fluids with high pH suggests the potential for the upflow of serpentinizing fluids supporting biological methanogenesis mainly linked to hydrogen oxidation and the use of deeply derived CO_2. This observation is broadly in agreement with methane and microbiological data reported from other mafic and ultramafic ecosystems (49, 50), The second group is represented by sites hosted in sedimentary quaternary deposits where the redox conditions and concentration of organic carbons likely support methylotrophic and acetoclastic methanogenesis, as supported by our clumped isotope and metagenomic results. In this group the source of carbon for methanogenesis is likely surface-derived and linked with paleo-deposition, although Rc/Ra suggests the presence of volatiles of deeper origins as well (Supplementary Figure S1), potentially suggesting that in the case of hydrogenotrophic methanogenesis, which appears to be minor in these sites, deeply-derived CO_2 could directly contribute to the starting carbon pool (7). In these two groups the role of biology in controlling the origin and fate of the methane is evident.

The third group is composed of sites hosted in sedimentary rocks (Figure 1 and Figure 4). The microbial community identified in the fluids of these sites is composed by sequences related to diverse methanogens and occasionally, albright less prevalent, some anaerobic methanotrophs (like sequences related to the order Methylomirabilales), suggesting that methane in these sites might be biological in nature, potentially derived from a mix of carbon present in the sedimentary rocks as well as the conversion of deeply-derived CO_2. The clumped isotopes, available for a single site in this category, CI, suggest that this site might be influenced by anaerobic methane oxidation, while the Withicar plot suggests a thermogenic origin for the methane (Supplementary Figure S3). In these sites our data suggests the potential for cryptic methane cycling, with unknown effect on clumped isotopes signature.

The last group of sites is represented by seeps hosted in volcanic rocks (Figure 1 and Figure 4). These sites show a stronger signal of deeply-derived volatiles with varying degrees of mantle influence (Figure 4C and Supplementary Figure S1), suggesting a deep origin of the methane (8) or that methane is derived abiotically or thermogenically from deep CO_2. The clumped isotopes of methane in these sites suggest a thermogenic/abiotic origin, with possible deviation from the expected signal likely due to methane utilization near or at the surface (Figure 2 and Figure 3). In these sites biology is represented mainly by sequences related to known methane oxidizers, and observation supported also by the abundance of key genes involved in methane oxidation in the metagenomic data, suggesting the potential of overprinting in part the geological signal.

Overall our data show a strong agreement between the origin and processes affecting methane at convergent margin predicted using clumped isotopes and the microbiological data collected at the same sites. Moreover, our results suggest that the hosting rock type as well as the position within a specific province in the convergent margin (*i.e.*, outer forearc, forearc, arc and backarc) exerts a strong control over the quantity and signature of the methane cycled to the surface. Previous studies on the origin and fate of CO_2 in this segment of the Central American convergent margin showed that, together with the (possibly biologically) mediated precipitation of calcite, the nature of the subducting plate was one of the major control on the measured carbon at the surface (7, 29, 47). Our results instead suggest that the overriding plate might bear a larger influence on the origin of the methane racing the surface. This does not exclude that the carbon utilized for methane synthesis by the diverse processes identified might be directly recycled through the subduction processes, at least in some of the main groups identified.

Taken together the data presented show that the geological setting is one of the main factors in controlling the last steps of the carbon path through the Earth's crust before the release within the atmosphere, and that geological and biologically mediated secondary processes can in part overprint the deeper signal and influence the fate of methane at convergent margins. Future studies are needed in order to assess the quantitative contribution of the different processes to the convergent margins methane flux. We expect that the diversity of rock types and geological settings in the overriding plate might exert a similar control on the origin and fate of methane across all subduction zones, with potentially large implications for global methane cycling as well as atmospheric contribution through deep time.

Acknowledgments

This work was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program Grant Agreement No. 948972—COEVOLVE—ERC-2020-STG to DG.

Competing Interest Statement

The authors declare no competing interests.

Classification

Earth, Atmospheric, and Planetary Sciences

References

1. D. V. Bekaert, *et al.*, Subduction-Driven Volatile Recycling: A Global Mass Balance. *Annu. Rev. Earth Planet. Sci.* **49**, null (2021).

2. E. K. Moore, B. I. Jelen, D. Giovannelli, H. Raanan, P. G. Falkowski, Metal availability and the expanding network of microbial metabolisms in the Archaean eon. *Nat. Geosci.* **10**, 629–636 (2017).

3. F. U. Battistuzzi, A. Feijao, S. B. Hedges, A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* **4**, 1–14 (2004).

4. P. B. Kelemen, C. E. Manning, Reevaluating carbon fluxes in subduction zones, what goes down, mostly comes up. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3997-4006 (2015).

5. J. M. Hayes, J. R. Waldbauer, The carbon cycle and associated redox processes through time. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 931–950 (2006).

6. T. Plank, C. E. Manning, Subducting carbon. *Nature* **574**, 343–352 (2019).

7. P. H. Barry, *et al.*, Forearc carbon sink reduces long-term volatile recycling into the mantle. *Nature* **568**, 487 (2019).

8. A. Vitale Brovarone, *et al.*, Subduction hides high-pressure sources of energy that may feed the deep subsurface biosphere. *Nat. Commun.* **11**, 3880 (2020).

9. A. Aiuppa, T. P. Fischer, T. Plank, P. Bani, CO2 flux emissions from the Earth's most actively degassing volcanoes, 2005–2015. *Sci. Rep.* **9**, 5442 (2019).

10. A. G. Judd, The global importance and context of methane escape from the seabed. *Geo-Mar. Lett.* **23**, 147–154 (2003).

11. A. H. Jahren, C. P. Conrad, N. C. Arens, G. Mora, C. Lithgow-Bertelloni, A plate tectonic mechanism for methane hydrate release along subduction zones. *Earth Planet. Sci. Lett.* **236**, 691–704 (2005).

12. I. C. Change, The physical science basis. *Contrib. Work. Group Fifth Assess. Rep. Intergov. Panel Clim. Change* **1535**, 2013 (2013).

13. M. A. Lever, A New Era of Methanogenesis Research. *Trends Microbiol.* **24**, 84–86 (2016).

14. D. A. Stolper, *et al.*, Formation temperatures of thermogenic and biogenic methane. *Science* **344**, 1500–1503 (2014).

15. M. O. Schrenk, W. J. Brazelton, S. Q. Lang, Serpentinization, Carbon, and Deep Life. *Rev. Mineral. Geochem.* **75**, 575–606 (2013).

16. G. Etiope, B. Sherwood Lollar, Abiotic Methane on Earth. *Rev. Geophys.* **51**, 276–299 (2013).

17. P. M. J. Douglas, *et al.*, Methane clumped isotopes: Progress and potential for a new isotopic tracer. *Org. Geochem.* **113**, 262–282 (2017).

18. J. J. Moran, *et al.*, Dual stable isotopes of CH4 from Yellowstone hot-springs suggest hydrothermal processes involving magmatic CO2. *J. Volcanol. Geotherm. Res.* **341**, 187–192 (2017).

19. D. T. Wang, E. P. Reeves, J. M. McDermott, J. S. Seewald, S. Ono, Clumped isotopologue constraints on the origin of methane at seafloor hot springs. *Geochim. Cosmochim. Acta* **223**, 141–158 (2018).

20. D. A. Stolper, *et al.*, The utility of methane clumped isotopes to constrain the origins of methane in natural gas accumulations. *Geol. Soc. Lond. Spec. Publ.* **468**, 23–52 (2018).

21. K. Suda, *et al.*, The origin of methane in serpentinite-hosted hyperalkaline hot spring at Hakuba Happo, Japan: Radiocarbon, methane isotopologue and noble gas isotope approaches. *Earth Planet. Sci. Lett.* **585**, 117510 (2022).

22. P. Beaudry, A. Stefánsson, J. Fiebig, J. H. Rhim, S. Ono, High temperature generation and equilibration of methane in terrestrial geothermal systems: Evidence from clumped isotopologues. *Geochim. Cosmochim. Acta* **309**, 209–234 (2021).

23. D. Giovannelli, *et al.*, Sampling across large-scale geological gradients to study geosphere–biosphere interactions. *Front. Microbiol.* **13** (2022).

24. K. D. Morell, Late Miocene to recent plate tectonic history of the southern Central America convergent margin. *Geochem. Geophys. Geosystems* **16**, 3362–3382 (2015).

25. P. H. Barry, *et al.*, Helium, inorganic and organic carbon isotopes of fluids and gases across the Costa Rica convergent margin. *Sci. Data* **6**, 1–8 (2019).

26. D. T. Wang, *et al.*, Nonequilibrium clumped isotope signals in microbial methane. *Science* **348**, 428–431 (2015).

27. S. Ono, *et al.*, Measurement of a Doubly Substituted Methane Isotopologue, 13CH3D, by Tunable Infrared Laser Direct Absorption Spectroscopy. *Anal. Chem.* **86**, 6487–6494 (2014).

28. D. Giovannelli, G. d'Errico, E. Manini, M. M. Yakimov, C. Vetriani, Diversity and phylogenetic analyses of bacteria from a shallow-water hydrothermal vent in Milos island (Greece). *Front. Extreme Microbiol.* **4**, 184 (2013).

29. K. M. Fullerton, *et al.*, Effect of tectonic processes on biosphere–geosphere feedbacks across a convergent margin. *Nat. Geosci.* **14**, 301–306 (2021).

30. B. J. Callahan, *et al.*, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

31. C. Quast, *et al.*, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).

32. P. J. McMurdie, S. Holmes, phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217 (2013).

33. C. S. Sheik, *et al.*, Identification and Removal of Contaminant Sequences From Ribosomal Gene Databases: Lessons From the Census of Deep Life. *Front. Microbiol.* **9** (2018).

34. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

35. C. Zhu, *et al.*, Functional sequencing read annotation for high precision microbiome analysis. *Nucleic Acids Res.* **46**, e23–e23 (2018).

36. P. Menzel, K. L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* **7**, 1–9 (2016).

37. Y. Shuai, H. Xie, S. Zhang, Y. Zhang, J. M. Eiler, Recognizing the pathways of microbial methanogenesis through methane isotopologues in the subsurface biosphere. *Earth Planet. Sci. Lett.* **566**, 116960 (2021).

38. P. Beaudry, A. Stefánsson, J. Fiebig, J. H. Rhim, S. Ono, High temperature generation and equilibration of methane in terrestrial geothermal systems: Evidence from clumped isotopologues. *Geochim. Cosmochim. Acta* **309**, 209–234 (2021).

39. D. A. Stolper, *et al.*, Distinguishing and understanding thermogenic and biogenic sources of methane using multiply substituted isotopologues. *Geochim. Cosmochim. Acta* **161**, 219–247 (2015).

40. M. Y. Yoshinaga, *et al.*, Carbon isotope equilibration during sulphate-limited anaerobic oxidation of methane. *Nat. Geosci.* **7**, 190–194 (2014).

41. R. S. Wolfe, My kind of biology. *Annu. Rev. Microbiol.* **45**, 1–36 (1991).

42. R. Thauer, Biodiversity and unity in biochemistry. *Antonie Van Leeuwenhoek* **71**, 21–32 (1997).

43. P. A. Bertram, R. K. Thauer, Thermodynamics of the formylmethanofuran dehydrogenase reaction in Methanobacterium thermoautotrophicum. *Eur. J. Biochem.* **226**, 811–818 (1994).

44. A. J. Holmes, A. Costello, M. E. Lidstrom, J. C. Murrell, Evidence that participate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiol. Lett.* **132**, 203–208 (1995).

45. S. Heux, T. Brautaset, J. A. Vorholt, V. F. Wendisch, J. C. Portais, "Synthetic Methylotrophy: Past, Present, and Future" in *Methane Biocatalysis: Paving the Way to Sustainability*, M. G. Kalyuzhnaya, X.-H. Xing, Eds. (Springer International Publishing, 2018), pp. 133–151.

46. C. Magnabosco, *et al.*, The biomass and biodiversity of the continental subsurface. *Nat. Geosci.*, 1 (2018).

47. T. J. Rogers, *et al.*, Chemolithoautotroph distributions across the subsurface of a convergent margin. *ISME J.*, 1–11 (2022).

48. M. Cascone, *et al.*, Microbial diversity in the backarc hot springs of Argentina and its role in biogeochemical cycles in (GOLDSCHMIDT, 2021) (May 31, 2022).

49. W. J. Brazelton, *et al.*, Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy. *PeerJ* **5**, e2945 (2017).

50. E. M. Fones, *et al.*, Physiological adaptations to serpentinization in the Samail Ophiolite, Oman. *ISME J.* **13**, 1750–1762 (2019).

51. M. J. Whiticar, Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. *Chem. Geol.* **161**, 291–314 (1999).

Figures

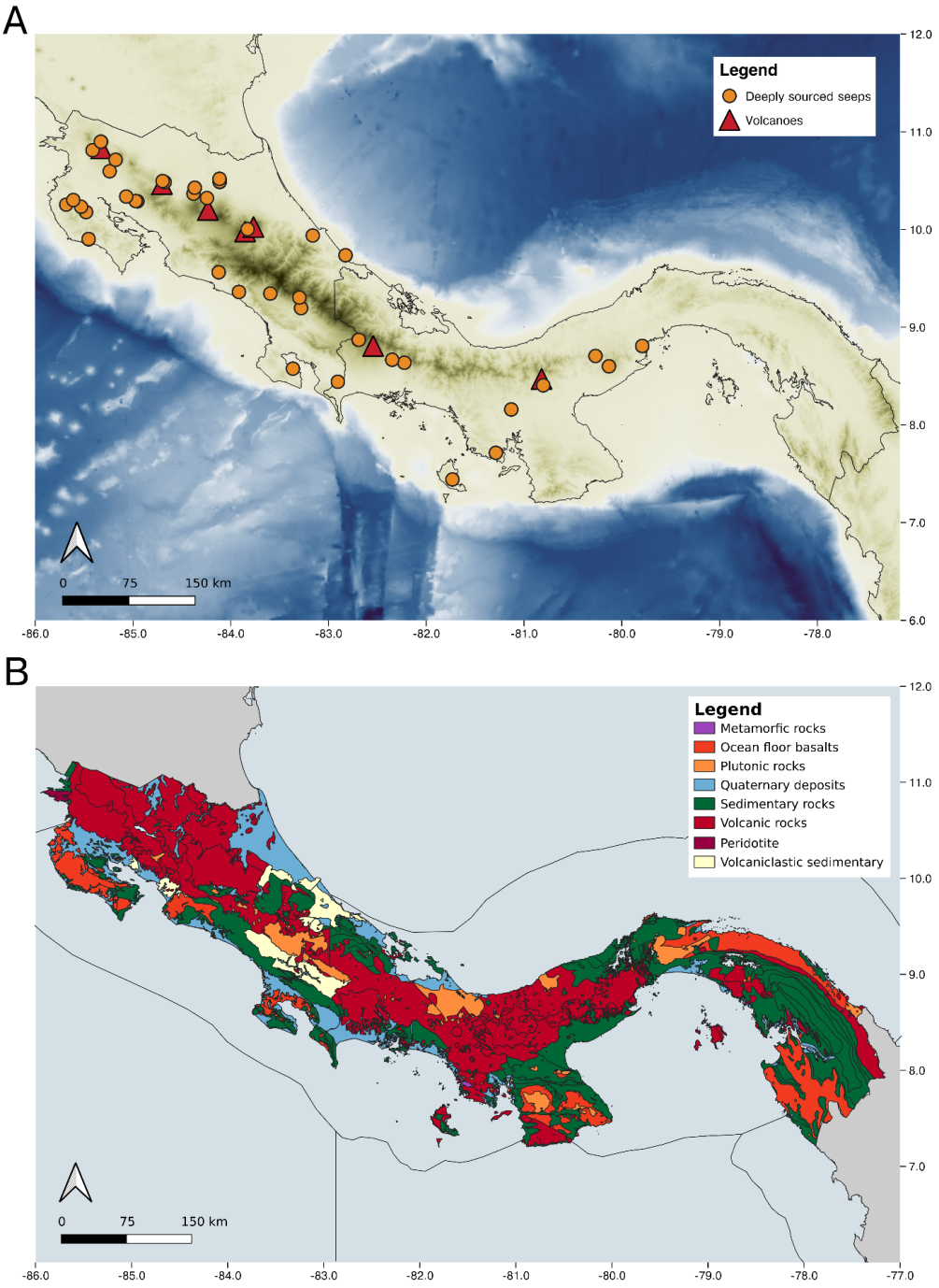


Figure 1. Central America convergent margin maps. A. Topographic map where the sampled sites are indicated as orange circles while volcanoes are indicated as red triangles, from North to South: Rincón de la Vieja, Arenal, Poás, Turrialba, Irazú, Barú, and La Yeguada, respectively. B. Geologic map that includes deposits and different rock types along Costa Rica and Panama.

Figure 2. A. Clumped isotopes $\Delta^{13}CH_3D$ values against $\delta^{13}C_{CH4}$ are shown for 11 deeply sourced springs investigated in Central America (red circles) and compared to samples from different environments (purple triangle: Cow rumen; light blue triangle: Northern Cascadia Margin; green triangle: Swamp; inverted purple triangle: Hydrothermal vents; teal diamond: Santa Monica Basin; yellow square: Volcanic and geothermal samples); B. Relative abundance of aerobic methane cycling related groups (blue) and, C. relative abundance of anaerobic methane cycling related groups (orange), both from 16S rRNA libraries of fluid samples. HM indicates the area of hydrogenotrophic methanogenesis, MAM indicates the area of methylotrophic and acetoclastic methanogenesis, and TM indicates the area of thermogenic methane formation, based on the Whiticar plot (Supplementary Figure S2).

Figure 3. Clumped isotopes $\Delta^{13}CH_3D$ values against $\delta^{13}C_{CH4}$ are shown for the normalized abundance of the methanogen cycle key genes found in fluid samples which are reported in orange for the mcrA gene (A), and in blue for the mmo (B), the pmo-amo (C), and the mdh (D) genes.

Figure 4. (A) Functional based Principal Coordinates Analysis (PCoA) based on weighted Jaccard similarity of the shotgun metagenome functional read assignment, subsetted for methane cycling; (B) $\delta^{13}C_{CH4}$ distribution across different geologic units (QD: quaternary deposits; OFB: oceanic floor basalts; SR: sedimentary rocks; VR: volcanic rocks) with the size proportional to the CH_4 concentration (mg/kg); (C) Ternary plot of CO_2, ^3He, and ^4He and (D) CH_4, ^3He, and ^4He relative compositions. Mantel, Air, and crust gas samples are also reported for comparison. All the figures are colored by geologic units.

Supplementary Information

Supplementary figure 1. Rc/Ra plotted against the distance from the trench, coloured by the belonging rock type.

Supplementary Figure S2. Relative abundance of 16S rRNA libraries of methane cycling related taxa at the order level (above is shown the real relative abundance before the subsetting).

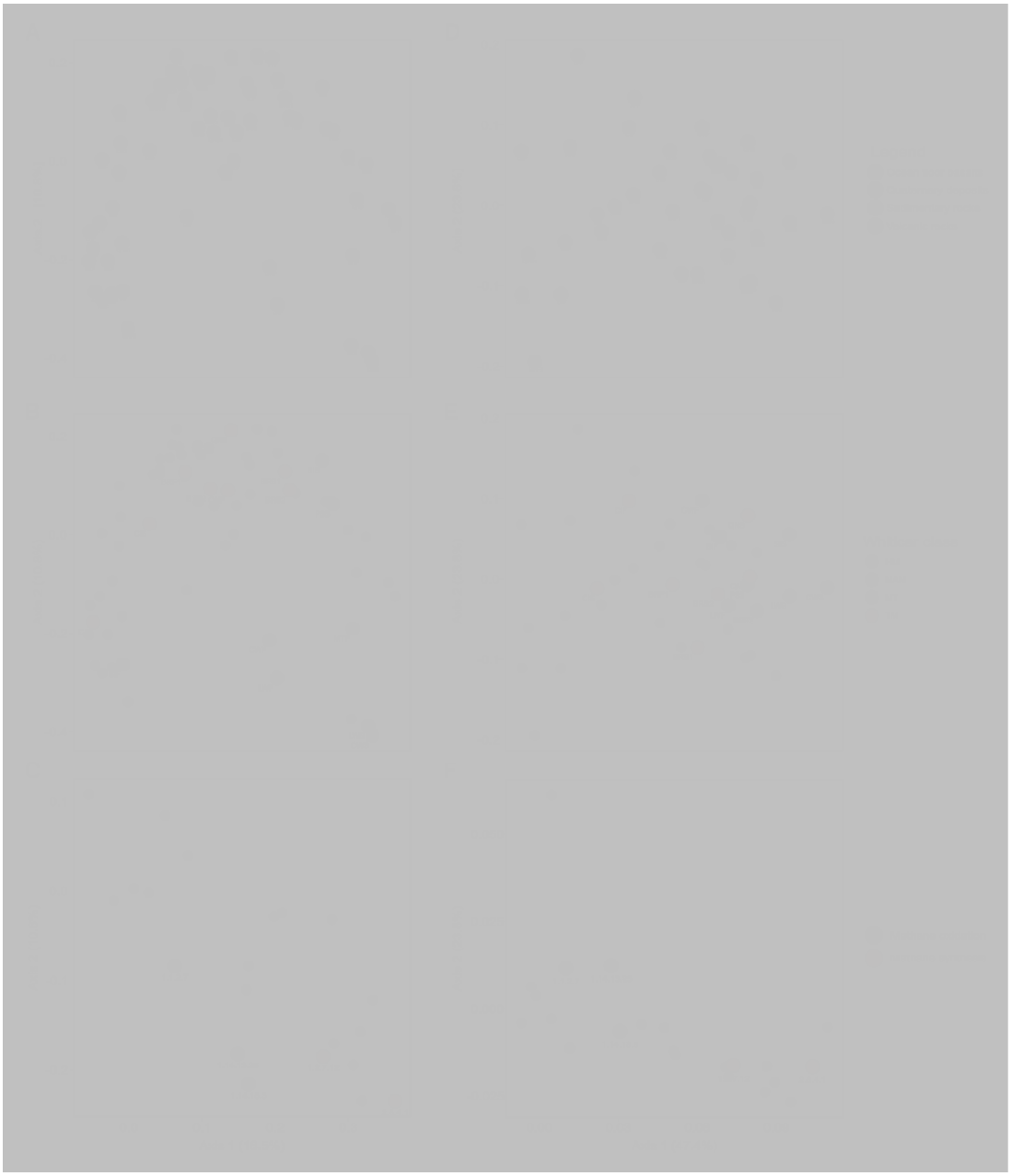
Supplementary Figure S3. Whiticar plot for the Identification of the origin of methane based on the relationship between δD_{CH4} and $\delta^{13}C_{CH4}$ which allow the separation between biogenic and thermogenic gasses from one another. The boundaries that define the common composition space are given from Whiticar, 1999 (51). From this work, 11 deeply sourced springs, coloured by the belonging geologic unit, are compared to samples from different environments (purple triangle: Cow rumen; light blue triangle: Northern Cascadia Margin; green triangle: Swamp; inverted purple triangle: Hydrothermal vents; teal diamond: Santa Monica Basin; yellow square: Volcanic and geothermal samples).

Supplementary Figure S4. Clumped isotopes $\Delta^{13}CH_3D$ values against $\delta^{13}C_{CH4}$ are shown for fluid (A, B) and sediment (C, D) samples where aerobic (blue) and anaerobic (orange) groups, involved in methane cycling, are present. The relative abundance is obtained from the 16S rRNA libraries of the respective samples and is proportional to the size of the circles.

Supplementary Figure S5. Shotgun metagenome functional read assignment for methanogenesis (orange) and methane oxidation (blue). The normalized abundance is indicated as the size of the circles. Sites in which a biogenic signature for the methane was found are highlighted in black.

Supplementary Figure S6. Clumped isotopes $\Delta^{13}CH_3D$ values against $\delta^{13}C_{CH4}$ are shown for the normalized abundance of the methane cycle key genes found in sediment samples. The *mcrA* gene is reported in orange (A), while the *mmo* (B), the *pmo-amo* (C), and the *mdh* (D) genes are reported in blue.

Supplementary Figure S7. Functional based Principal Coordinate Analysis (PCoA) based on (A, C) weighted Jaccard similarity and (B, D) unweighted Jaccard similarity among the shotgun metagenome functional read assignment, coloured by geologic units, and (C, D) by sample type.



Supplementary Figure S8. PCoA based on weighted Jaccard similarity index on the shotgun metagenome functional read assignment subsetted for methane cycling, coloured by geologic units (A) and whiticar class (B). PCoA based on weighted (C) Jaccard similarity index on the shotgun metagenome functional read assignment subsetted for methane cycling, with main functions for methane synthesis in orange, and main functions for methane oxidation in blue. PCoA based on unweighted (E) Jaccard similarity index on the shotgun metagenome functional read assignment subsetted for methane cycling, coloured by Whiticar class (Supplementary figure 3). PCoA based on unweighted Jaccard similarity index on the shotgun metagenome functional read assignment subsetted for methane cycling, coloured by geologic units (D) and whiticar class (E). PCoA based on unweighted (F) Jaccard similarity index on the shotgun metagenome functional read assignment subsetted for methane cycling, with main functions for methane synthesis in orange, and main functions for methane oxidation in blue.

Table 1. Sampled sites descriptors and environmental factors

ExpID	SiteID	Province	Nation	Geologic unit	Lat	Long	Alt	Temp (°C)	pH	Salinity
CR17	BQ									
CR17	BR1									
CR17	BR2									
CR17	CY									
CR17	EP									
CR17	ES									
CR17	ET									
CR17	FA									
CR17	MT									
CR17	PF									
CR17	QH1									
CR17	QH2									
CR17	QN									
CR17	RS									
CR17	RV									
CR17	SI									
CR17	SL									
CR17	ST									
CR17	TC									
CR17	VC									
CR18	BC									
CR18	BS									
CR18	BW									
CR18	CH									
CR18	CI									
CR18	CL									
CR18	CV									
CR18	CW									
CR18	CZ									
CR18	ER									
CR18	GE									
CR18	HA									
CR18	LB									
CR18	LE									
CR18	LH									
CR18	LP									
CR18	LW									
CR18	MC									
CR18	PS									
CR18	PX									
CR18	RC									
CR18	RR									
CR18	SC									
CR18	XF									
CR17	HN									
CR17	SR									
CR18	YR									

Table 2.

ExpID	SiteID	CH_4 (ppm)	$\delta^{13}C_{CH4}$	δD_{CH4}	$\Delta^{13}CH_3D$	Calc. Temp Clumped (°C)
CR17	BR1					
CR17	QN					
CR17	BQ					
CR17	MT					
CR17	HN					
CR17	SR					
CR17	SI					
CR18	CI					
CR18	PS					
CR18	LW					
CR18	CW					

Chapter 4

"*The straw that broke the camel's back*": microbial controls on a meromictic lake carbon cycle

Abstract

Volcanic lakes are considered "blue windows" into the depths of a volcano whose activity can vary from a status of quiescence to strong degassing and eruptions depending on chemistry and volatiles releasing. After two limnic eruptions documented in Cameroon during the second half of the 1980s, low activity volcanic lakes as well as meromictic lakes began to be investigated worldwide. The cause of the lake overturning process and consecutive degassing are still not clear although the general hypothesis suggests that multiple factors affect the steady state of the lakes. To date, few studies considered the interactions between geosphere and biosphere in these events. Microbial communities play a key role in biogeochemical cycles regulation, by affecting not only the chemical features of the entire water column but also its entire carbon reservoir. Generally, meromictic lakes show large amounts of CO_2 and CH_4 dissolved within the lake fluids, representing the main carbon source together with autochthonous and allochthonous forms of organic carbon that sink to the bottom. All of them sustain different microbial metabolisms which may increase CO_2 concentration overtime and work as a trigger for rolling events. Lake Rio Cuarto, in Costa Rica, is a low activity meromictic lake characterized by magmatic/hydrothermal input from the bottom, producing high content of dissolved gasses like CO_2 and CH_4. In our study, we used a multidisciplinary approach that included genomic and gas and fluid geochemistry to better understand the role of microbes in carbon cycle regulation in these stratified volcanic lakes, using Lake Rio Cuarto as a model.

Keywords: limnic eruptions, meromictic lake, hydrothermal lake, microbial diversity, carbon cycle

Introduction

Volcanic lakes are dynamic environments where hydrothermal-magmatic systems and Earth's surface interact with one another (Rouwet *et al.*, 2015). Unlike volcanic lakes characterized by high activity, where a broad range of volatiles is released (de Moor *et al.*, 2016; Vaselli *et al.*, 2019), lakes with lower activity are generally dominated by a minor rate of CO_2 injection. This favors a permanent stratification, distinctive of meromictic lakes where seasonal water movements do not include the lowest layers, generating a divergence between the physico chemical conditions of the shallower epilimnion and the deeper hypolimnion layers (Rouwet *et al.*, 2015; Fuchs *et al.*, 2022). The absence of recirculation prevents the vertical dissolved gas transferring into the lakes upper layers and to the atmosphere (Boehrer *et al.*, 2017; Fuchs *et al.*, 2022), until multiple factors affect the steady state of the lakes (Mortimer, 1953; Shy & Breidenthal, 1990; Cotel, 1999; Woods & Phillips, 1999; Rice, 2000; Zhang & Kling, 2006; Mott & Woods, 2010; Noguchi & Niino, 2010; Schmid *et al.*, 2010; Tassi *et al.*, 2018; Fuchs *et*

al., 2022). Such destabilization can induce an overturning process and the consequent releasing of huge amounts of CO_2 in the atmosphere as the catastrophic limnic eruptions that took place in Cameroon during the second half of the 1980s (Sabroux *et al.*, 1987; Nayar, 2009; Jones, 2010). The capacity of meromictic lakes to harbor significant amounts of CO_2 make them important carbon reservoirs (Juutinen *et al.*, 2009; Boehrer *et al.*, 2019; Fuchs *et al.*, 2022) in which biotic functions play a significant role in influencing the local carbon budget (Cabassi *et al.*, 2014; Rouwet *et al.*, 2014; Fuchs *et al.*, 2022). Unlike the epilimnion, where phototrophic communities are sustained by sunlight and atmospheric CO_2 (Dokulil & Teubner, 2000), the deeper hypolimnion layers are dominated by chemotrophic microbial communities which are supported by different carbon sources (Pasche *et al.*, 2009). Meromictic lakes with anoxic bottom waters have been noted for the active production of biogenic CH_4 as result of geogenic CO_2 reduction (common in volcanic lakes with degassing activity) and as the final steps of organic matter degradation after which a variety of potential substrates such as methyl containing compounds, acetate, formate, CO_2, and H_2 are released (Pasche *et al.*, 2011; Lyu *et al.*, 2018). Diffusion and accumulation of CH_4 along the water column is in function of the electron acceptors (e.g. O_2, NO_3^-, SO_4^{2-}, and Fe^{3+}) that may be rapidly consumed in high productivity areas for the degradation of the organic matter (Lyu *et al.*, 2018). The availability of electron acceptors influence the aerobic and anaerobic methanotrophic bacteria which mitigate the vertical fluxes of CH_4, coupling this latter consumption with: i. oxygenic photosynthesis (Milucka *et al.*, 2015; Oswald *et al.*, 2015; Brand *et al.*, 2016), ii. denitrification under oxygen limitation performed by Fe(II)/CH_4-oxidizing and nitrate/nitrite-reducing bacteria (Walter, 2011; Kopf *et al.*, 2013; Kits *et al.*, 2015a, 2015b; Michiels *et al.*, 2017; Tischer *et al.*, 2022), and iii. metal cycling (Fe- and Mn-hydroxide) as suggested by Crowe et al. (2011). Vertical fluxes of CO_2 in meromictic lakes are in function of methane cycling as well as on heterotrophic activity (Torres *et al.*, 2011) that similarly to other aquatic environments, exploits sinking particles that supply “hot spots” for organic matter decomposition and remineralization to CO_2 (Azam, 1998; Azam & Malfatti, 2007; Torres *et al.*, 2011). In place and external carbon inputs affect the entire carbon pool along the water column (Toming *et al.*, 2020), triggering a positive feedback between the enhancement of biological activity and the production rate of CO_2 and CH_4 (Cabassi *et al.*, 2014; Rouwet *et al.*, 2014; Fuchs *et al.*, 2022). To date, meromictic lakes have been investigated worldwide, focusing on different aspects that span from the iron-rich lakes as model for studying Archean oceans (Walter *et al.*, 2014) to the present effects of greenhouse gases accumulation (Juutinen *et al.*, 2009; Fuchs *et al.*, 2022). However, only a poor attention has given to more complex systems where volcanic settings and microbial network interact with each other affecting the biogeochemical features of a meromictic lakes (Aeschbach-Hertig *et al.*, 1999; Caliro *et al.*, 2008; Tassi *et al.*, 2009, 2018; Cabassi *et al.*, 2013, 2014; Rouwet *et al.*, 2015). We use a multidisciplinary approach in which microbiology and geochemistry are combined to better understand the biogeochemical processes that affect the physical-chemical features and the carbon cycling along the meromictic Lake Rio Cuarto vertical profile, investigating whether the new biogenic carbon input in the hypolimnion could trigger a lake overturning.

Material and Methods

Sampling method

In 2019, fluid and gas samples were collected along the water column of the Lake Rio Cuarto, a crater lake present along the Cordillera Central of Costa Rica (Fig. 1; Table 1). Fluid samples were filtered using sterivex membranes for molecular analysis and frozen at -20 °C, while samples for fluid geochemistry (*i.e.,* trace elements and major ions; Table 3) were filtered through 0.22 µm membranes and stored at +4 °C. Additionally, gas samples were collected in evacuated glass flasks using standard collection procedures (Giggenbach & Goguel, 1989), whereby precautions were taken to minimize any possible air contamination (Barry *et al.*, 2013). Physico-chemical parameters, such as temperature, pH, and salinity were measured using a multi-parameter sensor probe (HANNA, HI98194; Table 1).

Geochemistry

The concentrations of major cations and anions were measured using ion chromatography (ECO, IC metrohm) equipped with conductivity detectors. Calibration curves were run in the range of 0.1 and 10 ppm with $R \geq 0.999$. All dilutions and black subtractions were made using 18.2 MΩ/cm type I water. Anions were run using a 3.2 mM Na_2CO_3 + 1 mM $NaHCO_3$ mobile phase on a Metrosep A Supp 5 column equipped with a 0.15 M ortho-Phosphoric acid suppressor. The flow of the anionic eluent was 0.7 ml min-1 for 30 min, effective for good resolution of all anions in the standard solution. Cations were run using a 2.5 mM HNO_3 + 0.5 mM $(COOH)_2 \times 2H_2O$ mobile phase on a Metrosep C4 column. The flow of the cationic eluent was 0.9 ml min-1 with a total separation of 35 min, effective for good resolution of all cations in the standard solution. Data analysis was carried out through MagIC Net 3.3 software. calibration curves were carried out using certified (CPAchem) external standards for each of the anions and cations analyzed.

The concentrations of trace elements were measured using an inductively coupled plasma mass spectrometer (ICP-MS, Agilent 7900). Calibration curves were run in the range of 10 to 400 ppb with $R \geq 0.999$. All sample dilutions, standards and blanks were made with 1 % HNO_3 . Samples were introduced into ICP-MS through a peristaltic pump, and once arrived at the nebulizer they were converted to aerosol. The aerosol passed through a spray chamber, where the larger droplets were removed. The fine droplets were transported to the plasma torch based on Argon gas (high purity 99.999 %) at a rate of around 15 l/min, where they were ionized. All samples were analyzed with helium collision mode with a flow of 4.7 ml/min, that is effective in case of polyatomic interference. Finally, each element passed through the detector in order to be counted. An internal standards mix was used to cover the analytic mass range. Data acquisition and analysis was carried out through MassHunter 4.6 software.

Total alkalinity was measured by adding 0.5 mL of fluid into a falcon tube, with the addition of 0.2 mL of bromophenol blue. Subsequently, 0.5 mL of HCl 0.02 N was added in increments, with constant mixing, until a color change was observed (from blue to yellow). The volume of acid added to the sample was then used to calculate total alkalinity with the following equation: $Total\ alkalinity,\ mg\ CaCO_3/L = (A \times N \times 50000)/(mL\ of\ sample)$ where A represents the mL of acid added to each sample, and N the normality of the acid used.

$\delta^{13}C$ was evaluated with an automated continuous flow preparation system (isoFLOW, Elementar) coupled with an isotope ratio mass spectrometer (isoprime precisION, Elementar). Approximately 200 µl of samples were transferred into septum sealed vials and placed in the autosampler. An addle flush helium gas was used to remove atmospheric gasses, followed by the addition of phosphoric acid to digest the carbonate samples with an incubation time of 3 h at 70 °C to convert it into CO_2. The CO_2 was taken from the hand space and sent to the mass spectrometer. Data acquisition and analysis were carried out with lyticOS Software Suite. Calculated isotope ratios were reported in the δ-notation relative to the Vienna Peedee Belemnite (VPDB).

Carbon isotopes on gas samples from the 2019 campaign were analyzed at the OVSICORI on a Picarro G2201-I. $\delta^{13}C$-VPDB values were calibrated against a set of eight standards with values ranging from +2.42‰ to −37.21‰, including internationally accepted standards NBS19 and Carrara Marble. Reported $\delta^{13}C$ values have uncertainties of <0.1‰ based on repeated analyses of standards and samples.

DNA extraction and sequencing

DNA extraction was performed on three samples using a modified phenol-chloroform extraction method adapted for geothermal fluids (Giovannelli et al., 2013, 2016). Briefly, the Sterivex membrane filters are fragmented and incubated at 37 °C shaking for 1 hour with 6 mL of extraction buffer solution of extraction buffer 1X (100 mM Sodium Phosphate Buffer pH 8.0 + 100 mM EDTA pH 8.0 + 100 mM Tris-HCl pH 8.0 + 1.5 M NaCl + 1 % CTAB) and further supplemented with 50 µl of Lysozyme (100 mg/ml), followed by two incubation periods at 37 °C for 30 minutes, the latter of them following the addition of 50 µl of Proteinase K (20 mg/ml). Subsequently, 1 mL of 20 % SDS is added to the samples, which undergo 4 steps of freezing/thawing in dry ice/65 °C, respectively; then, samples are incubated for 2 hours at 65 °C, and vortexed every 15 minutes. The supernatant is recovered and 1 volume of a phenol:chloroform:isoamyl alcohol (25:24:1/24:1) solution is added. After inverting 100 times the solution, it is centrifuged at 7000 rpm for 10 minutes. The aqueous phase is recovered and 1 volume of Chloroform:Isoamyl alcohol (24:1) is added. After inverting 100 times the solution, it is centrifuged at 7000 rpm for 10 minutes. The aqueous phase is recovered and supplemented with Na-acetate (0.1 vol) and isopropanol 100 %, and incubated at room temperature overnight. The precipitated DNA is then washed with 70 % ice-cold ethanol, and re-suspended in 50 µl of Tris-HCl (50 mM). DNA was visualized with agarose gel electrophoresis and quantified spectrophotometrically and spectrofluorimetrically (NanoDrop and Qubit). Obtained DNA was sequenced at the Integrated Microbiome Resource (IMR, https://imr.bio) using primers targeting the V4-V5 of the 16S rRNA (515FB = GTGYCAGCMGCCGCGGTAA 926R = CCGYCAATTYMTTTRAGTTT 515FB – 926R) and using Illumina MiSeq technology. The extracted DNA was also used for shotgun metagenomic sequencing. Microbial community metagenomes were prepared using the Illumina Nextera Flex Kit for MiSeq + NextSeq, which requires a minimal amount of starting material (1 ng), as it is a PCR-based library preparation procedure. Samples are "tagged" (enzymatically cut and tagged with adapters), amplified by PCR when adding barcodes, purified by columns or beads, normalized by Illumina beads or manually, then pooled for upload to NextSeq, where 2X sequencing depth was achieved (~8 M PE reads = 16 M single reads & 2.4 Gb/sample).

Bioinformatic analysis

Raw 16S rRNA sequence reads were processed using the DADA2 package (Callahan *et al.*, 2017). After a trimming step, only sequences with an average quality call for each base between 20 and 40 remained for downstream analysis. The obtained amplicon sequence variants (ASV) were assigned taxonomy with the SILVA database - release 138 (Quast *et al.*, 2013). Abundance tables and taxonomic assignments were used to create a phyloseq object using the phyloseq package (McMurdie and Holmes, 2013). ASVs related to Mitochondria, Chloroplast, and Eukaryotes, as well as groups related to human pathogens and common DNA extraction contaminants (Sheik *et al.*, 2018), were removed from the dataset. The remaining reads represented 55 % of the original raw reads, with 245 assigned to 41 AVS. Alpha diversity was investigated using Shannon and Simpson diversity index. All statistical analysis, data processing and plotting were carried out with the R statistical software version 4.1.2 (Core, 2021), using the vegan (Oksanen *et al.*, 2018) and ggplot2 (Wickham, 2011) packages. A complete R script containing all the steps to reproduce our analysis is available at https://github.com/giovannellilab/Selci_et_al_Meromictic_lake_carbon_cycle and is released as a permanent version (version 1.0) using Zenodo with doi: XXXXXXXX.

Raw metagenomic reads were trimmed and mate-paired by Trimmomatic, version 0.36 (Bolger *et al.*, 2014), using default settings. Reads were then assembled *de novo* in metaSPAdes (Nurk *et al.*, 2017) to create site-specific assemblies with a minimum contig length of 2.0 kb. Reads that passed the trimming step were used for functional annotation through the mi-faser software and the GS-21-ALL database, a manually curated database of enzymes with experimentally validated functions representing 2,716 individual Enzyme Commission (EC) numbers, using the quality control pipeline in mi-faser (Zhu *et al.*, 2018). Obtained raw counts for each function were combined in a single dataset (mifaser_counts_raw). To account for the influence of different gene lengths on the read recruitments, and for the different sequencing depth associated with each metagenome, read counts were normalized to the average length of the corresponding protein in the GS21-ALL database and multiplied by the median library size across the dataset. To account for the compositionality of sequencing data, we further divided each EC abundance by the abundance of the DNA directed RNA polymerase subunit β (rpoB gene), and scaled to 1,000,000. This procedure returns the relative abundance of each EC for 1,000,000 reads mapped to rpoB, assuming equal sequencing depth. This allows within-sample and between-sample comparison, effectively opening the data and removing the problem of compositionality. The obtained normalized counts (mifaser_nrpo) were used for downstream analyses with the package 'phyloseq' (McMurdie & Holmes, 2013) in R 4.2.2 (Core, 2021). Taxonomy assignment on reads was carried out on Kaiju (Menzel *et al.*, 2016) which uses microbial and viral proteins as a reference database.

Assembled contigs were converted in amino acid sequences with the software Prodigal, v. 2.6.1 (Hyatt *et al.*, 2010) and then searched against the MEROPS database which classifies proteolytic enzymes using hierarchical classification by homologous sequences (Rawlings *et al.*, 2018). The software DIAMOND (Buchfink *et al.*, 2021) and the BLASTP implementation were used for this analysis. At the end of the analysis, only the hits with almost 70 % of coverage and 70 % of similarity were retained. All the sequences analyzed in the present study are available through the European Nucleotide Archive (ENA) with the project accession XXXX, under the ENA Umbrella project CoEvolve PRJEB55081.

Mesocosm experiments

The singular physical-chemical conditions of the lake, i.e. magmatic/hydrothermal CO_2 input from the bottom and high methane production along the water column, make Lake Rio Cuarto an opensky laboratory. To investigate the CH_4 production rate of the lake, fluid samples were collected from the deeper layer (68 m), where the highest CH_4 concentration was measured in a previous study (Cabassi *et al.*, 2014). Fluids were split in three groups of 120 mL pre-vacuumed vials (A, B, and C), in three replicates for each group, respectively. A mix of 16 mL of CO_2 and 64 mL of H_2 (1:4 ratio) was added to the group A and B at time zero (Zeikus, 1977), while the group C was used as control. Headspace gas was sampled on different days for each group, for a total of 21 days. To analyze the CH_4 and CO_2 concentration and the respective $\delta^{13}C$, 1 mL of headspace gas was diluted with 500 mL of N_2 and measured with the Picarro Picarro G2201-I. CH_4 and CO_2 concentration and respective $\delta^{13}C$ are reported in Table 4. At the end of the experiment, the group A fluids were filtered using sterivex membranes for molecular analysis and frozen at -20 °C, waiting for DNA extraction and sequencing as mentioned above.

Results

Geochemistry

The physico-chemical parameters measured at the different samples depths, as well as all the geochemical parameters, are listed in table 1. Along the water column, the temperature ranged from 28.5 °C at the highest depth sampled (RC_68), to 32.7 °C at 30 meter depth (RC_30) (average water temperature of 30.6 °C between all sampled points). The highest pH value (6.92) was recorded at 20 meter depth, at the site RC_20, and the lowest pH value (6.48) at 40 meter depth, at the RC_40 site (average pH of 6.59 °C between all sampled points). In all the sampled sites, no significant variations in salinity were recorded (1.5 psu).

The concentration of the major elements in water samples retrieved at different depths was measured in the forms of major anions and cations (Table 2). Chlorine concentrations ranged from 1.61 ppm at RC_50 to 3.09 ppm at RC_30 (average concentration of 1.9 ppm). Sodium concentrations ranged from 5.17 ppm at RC_20 to 8.2 ppm at RC_68 (average concentration of 8.2 ppm), while potassium ranged from 2.38 ppm at RC_20 to 6.1 ppm at RC_30 (average concentration of 3.7 ppm). Ammonia concentrations ranged from 4.68 ppm at RC_30 to 14.97 ppm at RC_68 (average concentration of 9.5 ppm), however, it was not possible to measure it in RC_20 (values below detection limit). Magnesium concentrations ranged from 4.95 ppm at RC_20 to 6.37 ppm at RC_50 (average concentration of 5.9 ppm), while calcium concentrations ranged from 13.87 ppm at RC_20 to 17.54 ppm at RC_68 (average concentration of 16.3 ppm). Both sulfate and nitrate concentrations were below the detection limit in all sampled sites.

Trace elements concentrations were measured for the trace metals nickel, iron, copper, and magnesium, which are known to play key biological roles (Falkowski *et al.*, 2008; Raanan *et al.*, 2018), since they act as cofactors in many important redox-reactions. Nickel, iron, and copper could only be measured in the samples RC_20 and RC_68, since in the other sampled locations the concentrations for these metals were below the detection limit of our instrument. Nickel concentrations ranged from 0.28 ppm in RC_20 to 0.3 ppm in RC_68, while iron concentrations

ranged from 5.59 ppm in RC_20 to 226.39 ppm in RC_68. Copper concentrations ranged from 0 ppm in RC_20 to 0.05 ppm in RC_68. Magnesium concentrations ranged from 26.14 ppm in RC_20, to 584.37 ppm in RC_68. Total alkalinity in the form of bicarbonate (HCO_3^-) was 300 ppm in RC_20, RC_50, and RC_68, and below the detection rate in the other samples (Table 2).

Analysis of dissolved gases like CH_4 showed a concentration of 234.8 ppm in the shallower layer (RC_20) and 968.9 ppm in the deepest layer (RC_68), with the maximum concentration of 1252.8 ppm, at 40 m of depth (RC_40). The concentration of CO_2 showed a low content equal to 64.6 ppm at 20 m of depth (RC_20) against higher concentrations in the deeper layers (60 and 68 m of depth), 227.6 ppm and 203.4, respectively. $\delta^{13}C$ signatures for CH_4 showed values lower than -60 ‰ along the entire water column, while $\delta^{13}C_{CO2}$ isotopic signatures ranged from -10.3 ‰ to 4.14 ‰ (Table 3). A similar trend was observed for the $\delta^{13}C_{DIC}$ with a value of -9.23 ‰ at 20 m of depth (RC_20) and 0.78 ‰ at 68 m of the depth (Table 3).

Microbial diversity

Even though multiple depths representative of the water column were sampled in the present study, only two sites yielded sufficient DNA for microbial diversity analysis. The low depth site, RC_20, located at a depth of 22 meters, was dominated by members belonging to the Cyanobacteria and Plancomycetota phyla, as well as Proteobacteria at lower abundances. Among Cyanobacteria, the order *Syneccocales* was the only one to be identified, more specifically the genus *Cyanobium* PCC-6307 (average abundance of 27 %). Among Planctomycetota, the most abundant classes were assigned to OM190 (average abundance of 18.5 %), with no further assigned taxonomy, and Planctomycetes, more specifically to the family *Rubinisphaeraceae* (average abundance of 10.4 %); however, it was not possible to resolve them down to the genus level. As to the low abundance Proteobacteria phyla, the most prevalent ASVs were classified into the order *Burkholderiales*, and the genus *Gallionella* (average abundance of 8 %).

The higher depth site, RC_68, located at a depth of 68 meters, was dominated by members belonging to the Proteobacteria and Verrucomicrobiota phylum, as well as Plactomycetota, Proteobacteria and Bacteroidota, but these latter at lower abundances. Within Proteobacteria, the order *Burkholderiales* was the most prevalent order, more specifically the genus *Paucibacter* (average abundance of 25.6 %). The Verrucomicrobiota phylum was dominated by members belonging to the *Omnitrophales* order, more specifically the uncultured genus Candidatus *Omnitrophus* (average abundance of 14 %). The less abundant Planctomycetota and Bacteroidota phyla were dominated by the order *Pirellulales* (average abundance of 9 %), with no identification down to the genus level, and the genus *Ideonella* (average abundance of 8 %), belonging to the *Burkholderiales* order, respectively .

Functional capabilities of Rio Cuarto microbial communities

The functional capabilities of the microbial communities of Rio Cuarto were inferred by searching for genes coding for functions related to energy metabolism and carbon fixation. Genes coding for carbon, hydrogen, iron, methane, nitrogen, oxygen, photosynthesis, and sulfur pathways, were found in all sampled sites, as well as in the mesocosm experiment. However,

their abundance varied among the different sites, suggesting the use of different major metabolic strategies by the microbial communities. For instance, the low depth site (20 m), was characterized by a higher abundance of genes related to Photosynthesis and Sulfur metabolism, such as *Psb*/*PsaA*, and *Sir,* respectively. *PsbA* codes for the photosystem II P680 reaction center, whereas *PsaA* codes for photosystem I. Additionally, *Sir* codes for a sulfite reductase (ferredoxin as catalytic center). On the other hand, at the higher depth site, we observed a transition of the metabolic properties of the microbial communities, with the increase in abundance of genes coding for functions related to hydrogen and methane, such as *frh* complex that codes for a CoenzymeF420 hydrogenase and *mcrA* and *fwd* complexes, coding for a methyl-coenzyme M reductase and formylmethanofuran dehydrogenase, respectively. Regarding carbon fixation pathways, we could identify genes coding for multiple carbon fixation pathways, such as the Calvin cycle, Wood-Ljungdahl cycle, Arnon-Buchanan cycle, and 3-Hydroxypropionate bi-cycle, at the different depths.

Mesocosm experiment

During the first two weeks, approximately, gas analysis of the three mesocosms revealed a decrease in CO_2 and an increase in CH_4 concentrations in groups A and B (Figure 9 a and b) while the control group (C) showed a momentaneous increase in CH_4 followed by a decrease and a steady state (C; Figure 9 c). Between the second and the third week, a decrease in CH_4 and an increase in CO_2 was observed in all three groups. $\delta^{13}C$ of CH_4 was consistent with the negative values found in the deepest layers of the lake, where biogenic methane was actively produced, as shown by Cabassi et al., (2014). $\delta^{13}C$ of CO_2 displayed small variations remaining around positive values.

Microbial diversity of the mesocosm 'A' was investigated using taxonomic assignment of metagenomic reads by Kaiju. At order level, of the 70 % classified, the community was dominated by members belonging to the *Burkholderiales* (45.3 %) and *Nitrosomonadales* (11.3 %). However, a small fraction (7 %) of the community was made up of low abundant members with less than 0.5% of all reads, while a larger number of reads was not assigned to known groups different from viruses (13 %).

The mesocosm experiment was enriched in genes coding for carbon fixation, iron, nitrogen, and sulfur. Regarding carbon fixation pathways, the 3-Hydroxypropionate bi-cycle was identified in higher abundances, given the presence of genes coding for a key enzyme of the cycle (*acuI*). Iron metabolism was enriched with genes coding for an iron chelator, *FeR*, coding for a ferric-chelate reductase. Additionally, it was possible to identify the presence of multiple key genes for Nitrogen metabolism related to denitrification through the presence of the *Nir* and *Nor* complexes (nitrite reductase and nitric oxide reductase, respectively), to assimilatory nitrate reduction due to the presence of the *Nar* and complexes (nitrate reductase and nitrite reductase, respectively), and to nitrogen fixation with the detection of the *Nif* (nitrogenase). The sulfur metabolism was dominated by genes related to sulfite oxidation (*sorA)* and thiosulfate oxidation by the SOX complex (*SOX)*.

Peptidase diversity

The occurrence of proteolytic enzymes in shallow and deep lake layers and at the end of the mesocosm experiment was investigated by blasting the assembled DNA contigs against the MEROPS database. In total, there are 141 MEROPS family annotations predicted to be secreted, M (Metallo) with 66, S (Serine) with 44, C (Cysteine) with 17, T (Threonine) with 6, A (Aspartic) with 5, N (Asparagine) with 1, U (Unknown) with 1, P (Mixed) with 1, and G (Glutamic) with 0. The comparison of the three samples showed the highest abundance of peptidase in MES_A, followed by RC_68, and RC_20 (Figure 10a), remaining consistent with the alpha diversity of the entire peptidases pool ('Observed' index) that showed a clear difference between the three samples. However, no particular differences were found between the three samples when looking at the most abundant peptidase ('Simpson' index).

Discussion

Meromictic lakes are incredible ecosystems that host huge carbon reservoirs which sustain complex microbial communities across the water column (Rouwet *et al.*, 2015; Tassi *et al.*, 2018). Lake Rio Cuarto, a low activity volcanic lake situated on the Caribbean side of the Cordillera Volcánica Central, is the deepest lake in Costa Rica (Alvarado *et al.*, 2011). The CO_2 injection from the bottom and the loading of organic matter from the upper layers favor the oxygen depletion along the lower hypolimnion, where anoxygenic microbial chemosynthesis is deeply involved in affecting water and gas chemistry (Cabassi *et al.*, 2014). Meromictic lakes have become largely studied since their potential in going through limnic eruptions, releasing large amounts of dissolved gasses in the atmosphere (Rouwet *et al.*, 2015). Although many investigations have tried to explain the causes of the overturning processes, dissolved gas reservoirs resulting from microbial activities have not been considered as a potential trigger for these events.

Here, we report gas and aqueous geochemistry and microbiology analysis across several depths of Lake Rio Cuarto, sampled in August 2019. This investigation aims to highlight secondary effects of the biosphere in driving the carbon cycling in complex ecosystems like meromictic lakes with low volcanic activity.

Physico-chemical variation along the Lake Rio Cuarto water column

The evaluation of water column physico-chemical conditions indicated a thermocline located below the layer of 30-40 m depth, which was different from those reported in previous studies (Alvarado *et al.*, 2011; Cabassi *et al.*, 2014), likely because the measurements were carried out in a different season of the year. A similar trend observed by Cabassi et al., (2014) was found for the pH, which tends to become weakly acidic after 30-40 m of depth, enough to support a constant dissolution of CO_2 in H^+ and HCO_3^-. At the same depth of the thermocline, a potential chemocline was also identified given the linear increase of concentration in Ca^{2+}, Mg^{2+}, Fe, Mn, and Co, and it was probably related to processes of water-rock interactions that mobilized soluble elements from the bottom through the upper layers (Balistrieri *et al.*, 1994; Cabassi *et al.*, 2014). In the past, the high content of Fe and Mn on the bottom of the lake was attributed both to *in situ* sedimentary minerogenic processes and to allochthonous input of insoluble Fe- and Mn-hydroxides, which become successively reduced in the hypolimnion (Göcke, 1996;

Hongve, 1997). The role of Fe- and Mn-hydroxide is crucial in controlling carbon recycling in meromictic lakes because their reduction to Fe(II) and Mn(II) is generally coupled with the degradation of organic matter, as well as methane oxidation, in which the two metals replace SO_4^{2-} and NO_3^- (*Environmental Microbe-Metal Interactions*, 2000; Crowe *et al.*, 2011; Rincón-Tomás *et al.*, 2016).

Chemical variation across the water column implies spatial changes in microbial community structure characterized by a predominance of phototrophs in the shallower layers, where light energy and CO_2 are used to produce biomass. A progressive succession with depth brings the community to be dominated by chemotrophic microorganisms that use inorganic and organic carbon compounds as source of carbon (Azam, 1998; Pjevac *et al.*, 2015; Alfreider *et al.*, 2017; Diao *et al.*, 2017). The carbon cycling in Lake Rio Cuarto is driven by different carbon sources, like the bottom input of CO_2, due to mantle degassing (Alvarado *et al.*, 2011), that facilitates the lake stratification process and provides substrates for hydrogenotrophic methane production. In addition to the geogenic CO_2, a broader range of organic carbon compounds sinks to the bottom, becoming potentially available for methylotrophic and acetoclastic methanogenesis, sustained by the biogenic $\delta^{13}C$ signature (<−60 ‰) in the water column until the upper layer (Whiticar, 1999). In addition, the isotopic discrimination factor between CH_4 and DIC ($\Delta\delta^{13}C_{CH4-DIC}$ < - 55 ‰) supported a biogenic methane formation rather than abiotic or thermogenic sources (Whiticar, 1999; Oswald *et al.*, 2016). The light increase in $\delta^{13}C$ values for the near sediment CH_4 may indicate partial consumption and C isotope fractionation by anaerobic methane oxidizers within the sediments (Oswald *et al.*, 2016). Given the absence of analysis regarding the CH_4 concentrations and isotopic composition in the sediments, we can not estimate whether the production of methane was higher in the water column rather than the sediments (Oswald *et al.*, 2016). The observed decrease of CH_4 during its rise from the lower layers towards the water surface leads to a decoupling between the methane production at the bottom and the emission across the water-atmosphere interface (Fuchs *et al.*, 2022). In addition to the whole carbon budget of the lake, a reliable source of carbon coming from autochthonous refractory microbial necromass can be recycled by microbes through mineralization and fermentation processes (Carstens *et al.*, 2012; Morrissey *et al.*, 2014; Chen *et al.*, 2022), leading to a net increase of CO_2 in the hypolimnion lower part (Cabassi *et al.*, 2014).

Microbial diversity and functional capabilities

The shallower site, RC_20, located at the epilimnion region of the lake, was dominated by members belonging to Cyanobacteria, Plancomycetota, and Proteobacteria phyla. The most abundant genus belonged to *Cyanobium gracile* PCC-6307, a picoprokaryotic cyanophyte commonly found in turbid freshwater shallow lakes (Bernát *et al.*, 2021). The high abundance of this photosynthetic group goes in line with our observation of a higher abundance of genes related to photosystem I and photosystem II at this location, suggesting that oxygenic photosynthesis is a major metabolic strategy to fix carbon dioxide from both the atmosphere and the subsurface degassing in the epilimnion region. The photosynthetic machinery has high demands for trace metals since they act as essential cofactors for the functioning of several proteins (Raven *et al.*, 1999). For instance, Fe is required for the electron chain complexes, and Mn for the functioning of photosystem II (Huertas *et al.*, 2014). Interestingly, both of these metals were present in very low concentrations (Fe 5.59 ppm and Mn 26.14 ppm; Figure 4 a and b) in this region of the lake, thus suggesting a depletion of them based on their demands.

Additionally, members belonging to Planctomycetota were highly abundant in the epilimnion regions of the lake, such as the OM 160 class and the family *Rubinisphaeraceae.* Planctomycetes have been observed in a diverse set of environments, ranging from hot springs (Memory Tekere, 2011), acidic environments (Ivanova & Dedysh, 2012), and glacial waters (Liu *et al.*, 2006), and are thought to be mainly aerobic, mesophilic and neutrophilic organisms (Lage & Bondoso, 2014). Additionally, previous studies have also identified Planctomycetes along the water column of the meromictic lake Sælenvannet, Norway (Storesund *et al.*, 2020). The OM 160 class, an uncultured group of Planctomycetes, and the family *Rubinisphaeraceae,* have been found in association with phototrophic organisms, such as Cyanobacteria and algae, through the utilization of a range of high-molecular weight sugars as carbon sources (Lage & Bondoso, 2014; Kaboré *et al.*, 2020). The high abundance of both Cyanobacteria and Planctomycetes suggests that carbon photoautotrophically fixed in the epilimnion regions is being recycled by heterotrophs belonging to the Planctomycetes phylum. At lower abundances, it was possible to identify the genus *Gallionella.* This genus comprises microorganisms capable of iron oxidation, as a source of energy, with a mixotrophic lifestyle, being able to switch from organic to inorganic carbon (Hallbeck & Pedersen, 1991; Kadnikov *et al.*, 2016). Both *Gallionella* and *Cyabobium* rely on iron to fulfill their metabolic needs, thus they could compete for soluble iron present in the environment. However, iron acquisition by Cyanobacteria relies on the transportation of Fe (III) into the cell, which is later reduced to Fe (II) (Jiang *et al.*, 2015), while *Gallionella* uses Fe (II) as a source of electrons (Kadnikov *et al.*, 2016). This suggests that both these microbial taxa are able to coexist by occupying two different niches based on the availability of different oxidation states of Fe, which further corroborates the low measurement of total Fe in the upper layers.

The high depth site, RC_68, was dominated by members belonging to the *Paucibacter* genus. This microbial group is known to be capable of degrading cyanobacterial hepatotoxins microcystins and nodularin produced in aquatic environments (Le *et al.*, 2022). The presence of these toxins in the deep layers of the lake could be due to their sinking in the water column over time. Members of this genus have been shown to grow heterotrophically under anaerobic conditions by coupling the oxidation of organic matter to nitrate reduction (Rapala *et al.*, 2005; Pheng *et al.*, 2017). Additionally, the genus Candidatus *Omnitrophus* has been previously described in the groundwater of the Homestake mine (Momper *et al.*, 2017). Genomic analysis revealed the presence of genes coding for hydrogen oxidation and methane oxidation, and for carbon fixation through the Wood-ljungdahl (WL) pathway. The presence of genes coding for such pathways in our analysis suggests that this group is able to grow chemolithoautotrophically in the deepest, anoxic region of the lake using hydrogen to reduce inorganic carbon to methane, in line with the high concentrations of methane measured in this region. It was also possible to identify members belonging to the Plactomycetota phylum, more specifically, the family *Pirellulaceae.* Cultured representatives of this family have been shown to be capable of growing heterotrophically through the oxidation of ammonia, and uncultured representatives by carbohydrate fermentation and sulfur reduction (Kaboré *et al.*, 2020). Unlike Plactomycetes growing on the surface, oxygenated layer, this group may be growing in the deepest region of the lake where oxygen is lacking by utilizing either organic matter produced by the chemolithoautotrophic groups, such as Candidatus *Omnitrophus,* or sinking particulate organic matter from the surface layers (necromass) (Figure 11), coupled to inorganic electron acceptors. On the same line, members belonging to the genus *Ideonella* were also identified in the deepest region of the lake. Representatives of this genus are known to be chemo-organotrhophs, that

utilize organic acids and amino acids as sole carbon sources (Tanasupawat *et al.*, 2016). However, under anaerobic conditions, *Ideonella* is able to oxidize organic matter through the reduction of chlorate as an electron acceptor (Malmqvist *et al.*, 1994). The high abundance of peptidases found at this site (Figure 10) also supports heterotrophy as a major metabolic strategy at the lower, anoxic regions of the lake in parallel with chemolithoautotrophy.

Three mesocosms of lake deep fluids were analyzed to understand the CH_4 production rate for three weeks. In the beginning, the gas analysis showed a slight increase in CH_4 against CO_2, indicating positive methanogenic activity, consistent with the negative values of the $\delta^{13}C_{CH4}$. However, the unexpected decrease of CH_4 against the increase of CO_2 observed at the end of the third week suggested a change in the chemical conditions of the fluids, which favor a pathway different from the methanogenesis, thus leading to the release of CO_2. At the end of the incubation, the community in group A was dominated by microbes of the order *Burkholderiales,* with *Ideonella* as the most abundant genus. *Ideonella*, as mentioned previously, is a group of chemo-organothrophs capable of utilizing different carbon sources in aerobic conditions, and of oxidizing organic matter through the reduction of chlorate under anaerobic conditions. Furthermore, the abundant peptidases found may support the hypothesis of an increase in heterotrophic pathways (Yamada *et al.*, 2012; Nguyen *et al.*, 2019), which caused the burst of CO_2 observed at the end of the experiment. Interestingly, we observed the presence of genes coding for pathways related to iron, methane, nitrogen, and sulfur metabolism. The oxidation of organic matter can be coupled to multiple electron acceptors (Sutton-Grier *et al.*, 2011). This may suggest that the microbial communities began exploring the oxidation of organic matter over time, by occupying different niches based on their ability to utilize different terminal acceptors (TEs). However, how the availability of multiple TEs and organic matter is constraining the growth of different microbial taxa has to be further explored.

Our results show that there is a shift in the identified microbial taxa, as well as in the major metabolic strategies employed by these microbial communities in the different layers of Rio Cuarto. At the surface layer, the majority of carbon is fixed by photosynthesis, by members belonging to the Cyanobacterial phyla, and further recycled by heterotrophs belonging to the Plactomycetora phylum. Interestingly, trace metals could also play important roles in niche stratification, for instance, to access available iron in the environment. At the lower, anoxic regions of the lake, the microbial communities grow by fixing carbon either chemolithautotrophycally or heterotrophically, based on the biomass produced by these groups, in addition to the necromass that sinks to the bottom overtime (Figure 11). This could imply an overproduction of carbon dioxide, which builds up in the system overtime, and is eventually released in major overturning events. Therefore, microorganisms inhabiting this region of the lake are key players contributing to CO_2 accumulation, in addition to the natural degassing.

Bibliography

Aeschbach-Hertig W, Hofer M, Kipfer R, Imboden DM, Wieler R (1999) Accumulation of mantle gases in a permanently stratified volcanic Lake (Lac Pavin, France). *Geochimica et Cosmochimica Acta* **63**, 3357–3372.

Alfreider A, Baumer A, Bogensperger T, Posch T, Salcher MM, Summerer M (2017) CO2 assimilation strategies in stratified lakes: Diversity and distribution patterns of chemolithoautotrophs. *Environmental Microbiology* **19**, 2754–2768.

Alvarado GE, Soto GJ, Salani FM, Ruiz P, Mendoza LH de (2011) The formation and evolution of Hule and Río Cuarto maars, Costa Rica. *Journal of Volcanology and Geothermal Research*, From maars to scoria cones: the enigma of monogenetic volcanic fields **201**, 342–356.

Azam F e (1998) Microbial Control of Oceanic Carbon Flux: The Plot Thickens.

Azam F, Malfatti F (2007) Microbial structuring of marine ecosystems. *Nature Reviews Microbiology* **5**, 782–791.

Balistrieri LS, Murray JW, Paul B (1994) The geochemical cycling of trace elements in a biogenic meromictic lake. *Geochimica et Cosmochimica Acta* **58**, 3993–4008.

Barry PH, Hilton DR, Fischer TP, De Moor JM, Mangasini F, Ramirez C (2013) Helium and carbon isotope systematics of cold "mazuku" CO2 vents and hydrothermal gases and fluids from Rungwe Volcanic Province, southern Tanzania. *Chemical Geology* **339**, 141–156.

Bernát G, Zavřel T, Kotabová E, Kovács L, Steinbach G, Vörös L, Prášil O, Somogyi B, Tóth VR (2021) Photomorphogenesis in the Picocyanobacterium Cyanobium gracile Includes Increased Phycobilisome Abundance Under Blue Light, Phycobilisome Decoupling Under Near Far-Red Light, and Wavelength-Specific Photoprotective Strategies. *Frontiers in Plant Science* **12**, 612302.

Boehrer B, Rohden C von, Schultze M (2017) Physical Features of Meromictic Lakes: Stratification and Circulation. In: *Ecology of Meromictic Lakes*, Ecological Studies (eds. Gulati RD, Zadereev ES, Degermendzhi AG). Springer International Publishing, Cham, pp. 15–34.

Boehrer B, Von Tuempling W, Mugisha A, Rogemont C, Umutoni A (2019) Reliable reference for the methane concentrations in Lake Kivu at the beginning of industrial exploitation. *Hydrology and Earth System Sciences* **23**, 4707–4716.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.

Brand A, Bruderer H, Oswald K, Guggenheim C, Schubert CJ, Wehrli B (2016) Oxygenic primary production below the oxycline and its importance for redox dynamics. *Aquatic Sciences* **78**, 727–741.

Buchfink B, Reuter K, Drost H-G (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature methods* **18**, 366–368.

Cabassi J, Tassi F, Mapelli F, Borin S, Calabrese S, Rouwet D, Chiodini G, Marasco R, Chouaia B, Avino R, Vaselli O, Pecoraino G, Capecchiacci F, Bicocchi G, Caliro S, Ramirez C, Mora-Amador R (2014) Geosphere-Biosphere Interactions in Bio-Activity Volcanic Lakes: Evidences from Hule and Rìo Cuarto (Costa Rica). *PLOS ONE* **9**, e102456.

Cabassi J, Tassi F, Vaselli O, Fiebig J, Nocentini M, Capecchiacci F, Rouwet D, Bicocchi G (2013) Biogeochemical processes involving dissolved CO2 and CH4 at Albano, Averno, and Monticchio meromictic volcanic lakes (Central–Southern Italy). *Bulletin of Volcanology* **75**, 683.

Caliro S, Chiodini G, Izzo G, Minopoli C, Signorini A, Avino R, Granieri D (2008) Geochemical and biochemical evidence of lake overturn and fish kill at Lake Averno, Italy. *Journal of Volcanology and Geothermal Research* **178**, 305–316.

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal* **11**, 2639–2643.

Carstens D, Köllner KE, Bürgmann H, Wehrli B, Schubert CJ (2012) Contribution of bacterial cells to lacustrine organic matter based on amino sugars and d-amino acids. *Geochimica et Cosmochimica Acta* **89**, 159–172.

Chen J, Zhou Y, Zhang Y (2022) New Insights into Microbial Degradation of Cyanobacterial Organic Matter Using a Fractionation Procedure. *International Journal of Environmental Research and Public Health* **19**, 6981.

Core R (2021) *Team. R: A Language and Environment for Statistical Computing, 2015*.

Cotel AJ (1999) A trigger mechanism for the Lake Nyos disaster. *Journal of volcanology and geothermal research* **88**, 343–347.

Crowe SA, Katsev S, Leslie K, Sturm A, Magen C, Nomosatryo S, Pack MA, Kessler JD, Reeburgh WS, Roberts JA, González L, Douglas Haffner G, Mucci A, Sundby B, Fowle DA (2011) The methane cycle in ferruginous Lake Matano. *Geobiology* **9**, 61–78.

Diao M, Sinnige R, Kalbitz K, Huisman J, Muyzer G (2017) Succession of Bacterial Communities in a Seasonally Stratified Lake with an Anoxic and Sulfidic Hypolimnion. *Frontiers in Microbiology* **8**.

Dokulil MT, Teubner K (2000) Cyanobacterial dominance in lakes. *Hydrobiologia* **438**, 1–12.

Environmental Microbe-Metal Interactions (2000) , 1st edn. John Wiley & Sons, Ltd.

Falkowski PG, Fenchel T, Delong EF (2008) The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science* **320**, 1034–1039.

Fuchs A, Casper P, Lewandowski J (2022) Dynamics of Greenhouse Gases (CH4 and CO2) in Meromictic Lake Burgsee, Germany. *Journal of Geophysical Research: Biogeosciences* **127**, e2021JG006661.

Giggenbach WF, Goguel RL (1989) Collection and analysis of geothermal and volcanic water and gas discharges. Report No. CD 2401. *Chemistry Division, DSIR, Petone, New Zealand*.

Göcke K (1996) Basic morphometric and limnological properties of Laguna Hule, a caldera lake in Costa Rica. *Revista de Biología Tropical* 537–548.

Gocke K, Bussing W, Cortés J (1990) The annual cycle of primary productivity in Laguna de Río Cuarto, a volcanic lake (maar) in Costa Rica. *Revista de Biología Tropical* 387–394.

Hallbeck L, Pedersen K (1991) Autotrophic and mixotrophic growth of Gallionella ferruginea. *Journal of General Microbiology* **137**, 2657–2661.

Hongve D (1997) Cycling of iron, manganese, and phosphate in a meromictic lake. *Limnology and Oceanography* **42**, 635–647.

Huertas M, López-Maury L, Giner-Lamia J, Sánchez-Riego A, Florencio F (2014) Metals in Cyanobacteria: Analysis of the Copper, Nickel, Cobalt and Arsenic Homeostasis Mechanisms. *Life* **4**, 865–886.

Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics* **11**, 1–11.

Ivanova AO, Dedysh SN (2012) Abundance, Diversity, and Depth Distribution of Planctomycetes in Acidic Northern Wetlands. *Frontiers in Microbiology* **3**.

Jiang H-B, Lou W-J, Ke W-T, Song W-Y, Price NM, Qiu B-S (2015) New insights into iron acquisition by cyanobacteria: an essential role for ExbB-ExbD complex in inorganic iron uptake. *The ISME Journal* **9**, 297–309.

Jones N (2010) Battle to degas deadly lakes continues. *Nature* **466**, 1033–1033.

Juutinen S, Rantakari M, Kortelainen P, Huttunen JT, Larmola T, Alm J, Silvola J, Martikainen PJ (2009) Methane dynamics in different boreal lake types. *Biogeosciences* **6**, 209–223.

Kaboré OD, Godreuil S, Drancourt M (2020) Planctomycetes as Host-Associated Bacteria: A Perspective That Holds Promise for Their Future Isolations, by Mimicking Their Native Environmental Niches in Clinical Microbiology Laboratories. *Frontiers in Cellular and Infection Microbiology* **10**, 519301.

Kadnikov VV, Ivasenko DA, Beletskii AV, Mardanov AV, Danilova EV, Pimenov NV, Karnachuk OV, Ravin NV (2016) A novel uncultured bacterium of the family Gallionellaceae: Description and genome reconstruction based on metagenomic analysis of

microbial community in acid mine drainage. *Microbiology* **85**, 449–461.
Kits KD, Campbell DJ, Rosana AR, Stein LY (2015a) Diverse electron sources support denitrification under hypoxia in the obligate methanotroph Methylomicrobium album strain BG8. *Frontiers in microbiology* **6**, 1072.
Kits KD, Klotz MG, Stein LY (2015b) Methane oxidation coupled to nitrate reduction under hypoxia by the G ammaproteobacterium M ethylomonas denitrificans, sp. nov. type strain FJG1. *Environmental microbiology* **17**, 3219–3232.
Kopf SH, Henny C, Newman DK (2013) Ligand-enhanced abiotic iron oxidation and the effects of chemical versus biological iron cycling in anoxic environments. *Environmental science & technology* **47**, 2602–2611.
Lage OM, Bondoso J (2014) Planctomycetes and macroalgae, a striking association. *Frontiers in Microbiology* **5**.
Le VV, Ko S-R, Kang M, Park C-Y, Lee S-A, Oh H-M, Ahn C-Y (2022) The cyanobactericidal bacterium Paucibacter aquatile DH15 caused the decline of Microcystis and aquatic microbial community succession: A mesocosm study. *Environmental Pollution* **311**, 119849.
Liu Y, Yao T, Jiao N, Kang S, Zeng Y, Huang S (2006) Microbial community structure in moraine lakes and glacial meltwaters, Mount Everest: Microbial diversity in lakes of Mount Everest. *FEMS Microbiology Letters* **265**, 98–105.
Lyu Z, Shao N, Akinyemi T, Whitman WB (2018) Methanogenesis. *Current Biology* **28**, R727–R732.
Malmqvist Å, Welander T, Moore E, Ternström A, Molin G, Stenström I-M (1994) Ideonella dechloratans gen.nov., sp.nov., a New Bacterium Capable of Growing Anaerobically with Chlorate as an Electron Acceptor. *Systematic and Applied Microbiology* **17**, 58–64.
McMurdie PJ, Holmes S (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**, e61217.
Memory Tekere (2011) Metagenomic analysis of bacterial diversity of Siloam hot water spring, Limpopo, South Africa. *AFRICAN JOURNAL OF BIOTECHNOLOGY* **10**.
Menzel P, Ng KL, Krogh A (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature communications* **7**, 1–9.
Michiels CC, Darchambeau F, Roland FA, Morana C, Llirós M, García-Armisen T, Thamdrup B, Borges AV, Canfield DE, Servais P (2017) Iron-dependent nitrogen cycling in a ferruginous lake and the nutrient status of Proterozoic oceans. *Nature Geoscience* **10**, 217–221.
Milucka J, Kirf M, Lu L, Krupke A, Lam P, Littmann S, Kuypers MM, Schubert CJ (2015) Methane oxidation coupled to oxygenic photosynthesis in anoxic waters. *The ISME journal* **9**, 1991–2002.
Momper L, Jungbluth SP, Lee MD, Amend JP (2017) Energy and carbon metabolisms in a deep terrestrial subsurface fluid microbial community. *The ISME Journal* **11**, 2319–2333.
Moor JM de, Aiuppa A, Pacheco J, Avard G, Kern C, Liuzzo M, Martínez M, Giudice G, Fischer TP (2016) Short-period volcanic gas precursors to phreatic eruptions: Insights from Poás Volcano, Costa Rica. *Earth and Planetary Science Letters* **442**, 218–227.
Morrissey EM, Berrier DJ, Neubauer SC, Franklin RB (2014) Using microbial communities and extracellular enzymes to link soil organic matter characteristics to greenhouse gas production in a tidal freshwater wetland. *Biogeochemistry* **117**, 473–490.
Mortimer CH (1953) The resonant response of stratified lakes to wind. *Schweizerische Zeitschrift für Hydrologie* **15**, 94–151.
Mott RW, Woods AW (2010) A model of overturn of CO2 laden lakes triggered by bottom mixing. *Journal of volcanology and geothermal research* **192**, 151–158.
Nayar A (2009) Earth science: A lakeful of trouble. *Nature* **460**, 321–323.
Nguyen TTH, Myrold DD, Mueller RS (2019) Distributions of Extracellular Peptidases Across Prokaryotic Genomes Reflect Phylogeny and Habitat. *Frontiers in Microbiology* **10**.
Noguchi T, Niino H (2010) Multi-layered diffusive convection. Part 1. Spontaneous layer

formation. *Journal of fluid mechanics* **651**, 443–464.
Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: a new versatile metagenomic assembler. *Genome research* **27**, 824–834.
Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2018) *vegan: Community Ecology Package*.
Oswald K, Jegge C, Tischer J, Berg J, Brand A, Miracle MR, Soria X, Vicente E, Lehmann MF, Zopfi J, Schubert CJ (2016) Methanotrophy under Versatile Conditions in the Water Column of the Ferruginous Meromictic Lake La Cruz (Spain). *Frontiers in Microbiology* **7**.
Oswald K, Milucka J, Brand A, Littmann S, Wehrli B, Kuypers MM, Schubert CJ (2015) Light-dependent aerobic methane oxidation reduces methane emissions from seasonally stratified lakes. *PloS one* **10**, e0132574.
Pasche N, Dinkel C, Müller B, Schmid M, Wüest A, Wehrli B (2009) Physical and biogeochemical limits to internal nutrient loading of meromictic Lake Kivu. *Limnology and Oceanography* **54**, 1863–1873.
Pasche N, Schmid M, Vazquez F, Schubert CJ, Wüest A, Kessler JD, Pack MA, Reeburgh WS, Bürgmann H (2011) Methane sources and sinks in Lake Kivu. *Journal of Geophysical Research: Biogeosciences* **116**.
Pheng S, Lee JJ, Eom MK, Lee KH, Kim S-G (2017) Paucibacter oligotrophus sp. nov., isolated from fresh water, and emended description of the genus Paucibacter. *International Journal of Systematic and Evolutionary Microbiology* **67**, 2231–2235.
Pjevac P, Korlević M, Berg JS, Bura-Nakić E, Ciglenečki I, Amann R, Orlić S (2015) Community Shift from Phototrophic to Chemotrophic Sulfide Oxidation following Anoxic Holomixis in a Stratified Seawater Lake. *Applied and Environmental Microbiology* **81**, 298–308.
Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**, D590–D596.
Raanan H, Pike DH, Moore EK, Falkowski PG, Nanda V (2018) Modular origins of biological electron transfer chains. *Proceedings of the National Academy of Sciences* **115**, 1280–1285.
Rapala J, Berg KA, Lyra C, Niemi RM, Manz W, Suomalainen S, Paulin L, Lahti K (2005) Paucibacter toxinivorans gen. nov., sp. nov., a bacterium that degrades cyclic cyanobacterial hepatotoxins microcystins and nodularin. *International Journal of Systematic and Evolutionary Microbiology* **55**, 1563–1568.
Raven JA, Evans MCW, Korb RE (1999) The role of trace metals in photosynthetic electron transport in O2-evolving organisms.
Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD (2018) The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic acids research* **46**, D624–D632.
Rice A (2000) Rollover in volcanic crater lakes: a possible cause for Lake Nyos type disasters. *Journal of Volcanology and Geothermal Research* **97**, 233–239.
Rincón-Tomás B, Khonsari B, Mühlen D, Wickbold C, Schäfer N, Hause-Reitner D, Hoppert M, Reitner J (2016) Manganese carbonates as possible biogenic relics in Archean settings. *International Journal of Astrobiology* **15**, 219–229.
Rouwet D, Christenson B, Tassi F, Vandemeulebrouck J (eds.) (2015) *Volcanic Lakes*. Advances in Volcanology. Springer Berlin Heidelberg, Berlin, Heidelberg.
Rouwet D, Tassi F, Mora-Amador R, Sandri L, Chiarini V (2014) Past, present and future of volcanic lake monitoring. *Journal of Volcanology and Geothermal Research* **272**, 78–97.
Sabroux JC, Dubois E, Doyotte C (1987) The limnic eruption: a new geological hazard. In: *International Scientific Congress on Lake Nyos Disaster*. p. 20.
Schmid M, Busbridge M, Wü est A (2010) Double‐diffusive convection in Lake Kivu. *Limnology and Oceanography* **55**, 225–238.
Sheik CS, Reese BK, Twing KI, Sylvan JB, Grim SL, Schrenk MO, Sogin ML, Colwell FS

(2018) Identification and Removal of Contaminant Sequences From Ribosomal Gene Databases: Lessons From the Census of Deep Life. *Frontiers in Microbiology* **9**.

Shy SS, Breidenthal RE (1990) Laboratory experiments on the cloud-top entrainment instability. *Journal of Fluid Mechanics* **214**, 1–15.

Storesund JE, Lanzèn A, Nordmann E-L, Armo HR, Lage OM, Øvreås L (2020) Planctomycetes as a Vital Constituent of the Microbial Communities Inhabiting Different Layers of the Meromictic Lake Sælenvannet (Norway). *Microorganisms* **8**, 1150.

Sutton-Grier ArianaE, Keller JK, Koch R, Gilmour C, Megonigal JP (2011) Electron donors and acceptors influence anaerobic soil organic matter mineralization in tidal marshes. *Soil Biology and Biochemistry* **43**, 1576–1583.

Tanasupawat S, Takehana T, Yoshida S, Hiraga K, Oda K (2016) Ideonella sakaiensis sp. nov., isolated from a microbial consortium that degrades poly(ethylene terephthalate). *International Journal of Systematic and Evolutionary Microbiology* **66**, 2813–2818.

Tassi F, Fazi S, Rossetti S, Pratesi P, Ceccotti M, Cabassi J, Capecchiacci F, Venturi S, Vaselli O (2018) The biogeochemical vertical structure renders a meromictic volcanic lake a trap for geogenic CO2 (Lake Averno, Italy). *PLOS ONE* **13**, e0193914.

Tassi F, Vaselli O, Fernández E, Duarte E, Martinez M, Delgado Huertas A, Bergamaschi F (2009) Morphological and geochemical features of crater lakes in Costa Rica: an overview.

Tischer J, Zopfi J, Frey C, Magyar PM, Brand A, Oswald K, Jegge C, Frame CH, Miracle MR, Sòria-Perpinyà X, Vicente E, Lehmann MF (2022) Isotopic signatures of biotic and abiotic N2O production and consumption in the water column of meromictic, ferruginous Lake La Cruz (Spain). *Limnology and Oceanography* **67**, 1760–1775.

Toming K, Kotta J, Uuemaa E, Sobek S, Kutser T, Tranvik LJ (2020) Predicting lake dissolved organic carbon at a global scale. *Scientific Reports* **10**, 8471.

Torres IC, Inglett KS, Reddy KR (2011) Heterotrophic microbial activity in lake sediments: effects of organic electron donors. *Biogeochemistry* **104**, 165–181.

Vaselli O, Tassi F, Fischer TP, Tardani D, Fernández E, Mar Martínez M del, Moor MJ de, Bini G (2019) The last eighteen years (1998–2014) of fumarolic degassing at the Poás volcano (Costa Rica) and renewal activity. *Poás Volcano: The Pulsing Heart of Central America Volcanic Zone* 235–260.

Walter XA (2011) *Anaerobic iron cycling in a neoarchean ocean analogue*.

Walter XA, Picazo A, Miracle MR, Vicente E, Camacho A, Aragno M, Zopfi J (2014) Phototrophic Fe(II)-oxidation in the chemocline of a ferruginous meromictic lake. *Frontiers in Microbiology* **5**.

Whiticar MJ (1999) Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. *Chemical Geology* **161**, 291–314.

Wickham H (2011) ggplot2. *Wiley interdisciplinary reviews: computational statistics* **3**, 180–185.

Woods AW, Phillips JC (1999) Turbulent bubble plumes and CO2-driven lake eruptions. *Journal of volcanology and geothermal research* **92**, 259–270.

Yamada N, Fukuda H, Ogawa H, Saito H, Suzumura M (2012) Heterotrophic bacterial production and extracellular enzymatic activity in sinking particulate matter in the western North Pacific Ocean. *Frontiers in Microbiology* **3**.

Zeikus JG (1977) The biology of methanogenic bacteria. *Bacteriological reviews* **41**, 514–541.

Zhang Y, Kling GW (2006) Dynamics of lake eruptions and possible ocean eruptions. *Annu. Rev. Earth Planet. Sci.* **34**, 293–324.

Zhu C, Miller M, Marpaka S, Vaysberg P, Rühlemann MC, Wu G, Heinsen F-A, Tempel M, Zhao L, Lieb W (2018) Functional sequencing read annotation for high precision microbiome analysis. *Nucleic acids research* **46**, e23–e23.

Tables

Table 1. Date, Location (GPS coordinates), Depth and physico-chemical parameters measured in the different depths sampled.

ID	Month	Date	Location	Lat	Long	Depth (m)	T (°C)	pH	Salinity
RC_20									
RC_30									
RC_40									
RC_50									
RC_60									
RC_68									

Table 2. Concentration of the major ions and trace elements measured, alkalinity and inorganic carbon isotopic signatures. All values are reported in ppm for the major and trace elements. N.d. - not detected; bdl - below detection limits. Detection limits for all the investigated compounds are described in the respective SOP cited in the material and methods section.

ID	Na^+	K^+	Cl^-	NH_4^+	Mg^{2+}	Ca^{2+}	NO^{3-}	SO_4^{2-}	HCO_3^-	$\delta^{13}C_{DIC}$	Ni	Fe	Co	Mn	Mo
RC_20															
RC_30															
RC_40															
RC_50															
RC_60															
RC_68															

Table 3. Carbon 13 isotopic signatures and concentrations on methane and CO_2 gas sampled.

ID	$\delta^{13}C_{CH4}$ (□)	CH_4 (ppm)	CO_2 (ppm)	$\delta^{13}C_{CO2}$ (□)
RC_20				
RC_30				
RC_40				
RC_50				
RC_60				
RC_68				

Table 4. Chemical composition (μmol x L^{-1}) of dissolved gasses (CO_2 and CH_4) and $\delta^{13}CO_2$ (expressed as ‰ V-PDB), $\delta^{13}CH_4$ (expressed as ‰ V-PDB) of gas samples collected.

Group	Day	CH_4 (ppm)	$\delta^{13}C_{CH4}$ (□)	CO_2 (ppm)	$\delta^{13}C_{CO2}$ (□)
A	0				
A	3				
A	7				
A	13				
A	21				
B	0				
B	1				
B	3				
B	6				
B	9				
B	15				
B	21				
C	0				
C	1				
C	3				
C	6				
C	7				
C	9				
C	13				
C	15				
C	21				

Figures

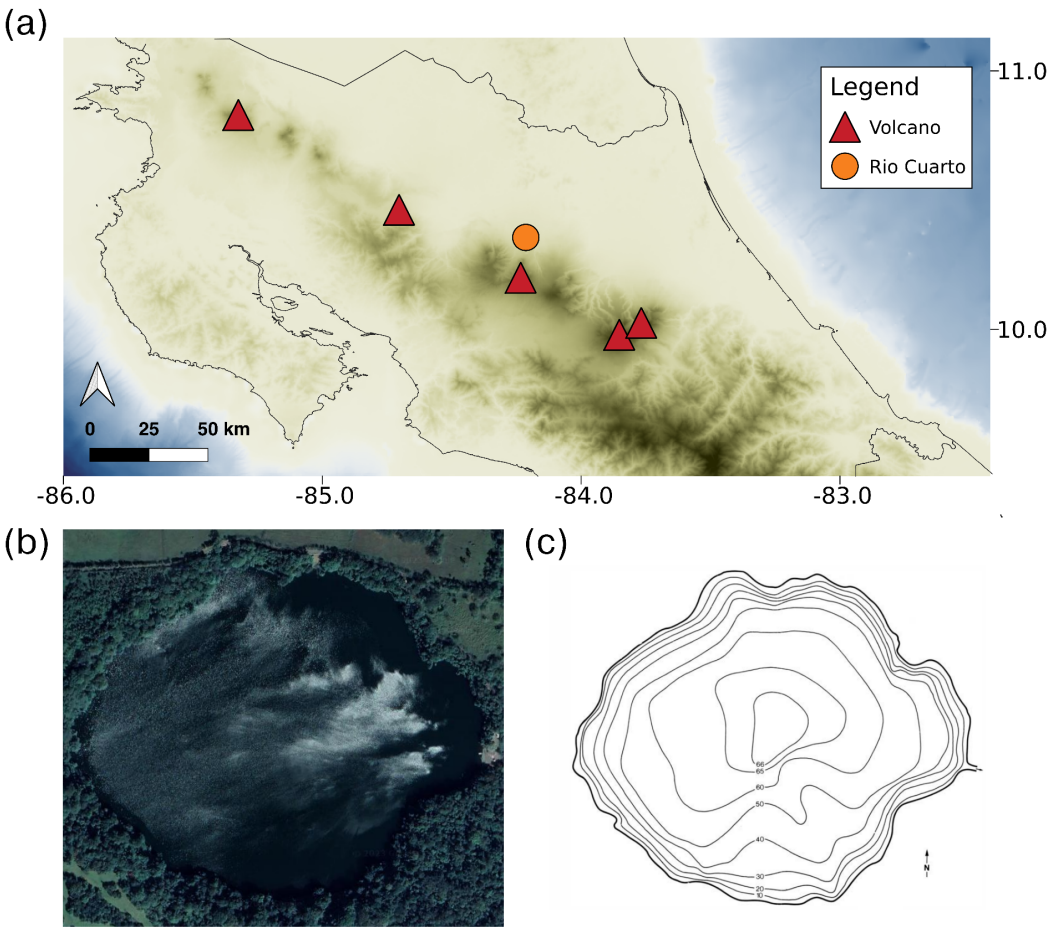


Figure 1. (a) Map of North-Central Costa Rica. Lake Rio Cuarto is indicated as an orange circle while volcanoes are indicated as red triangles (from North to South: Rincón de la Vieja, Arenal, Poás, Turrialba, and Irazú). (b) Lake Rio Cuarto, base maps: Google Earth, 2023. (c) Bathymetric map of Lake Río Cuarto, modified after Cabassi et al., (2014) and Gocke et al. (1990).

Figure 2. Distribution of temperature (a), pH (b), and salinity (c) along the Lake Rio Cuarto column water.

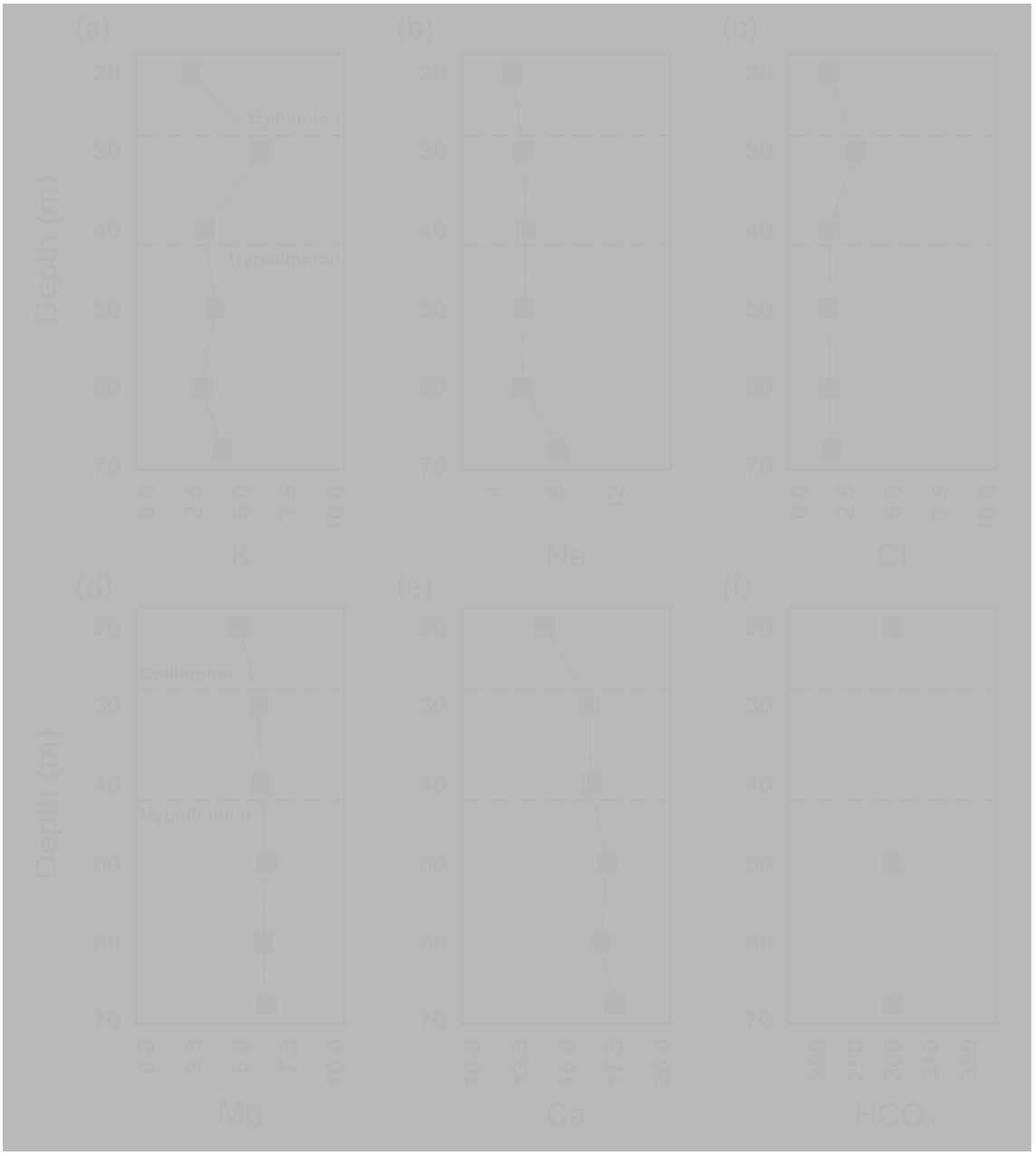


Figure 3. Distribution of major elements, K^+ , Na^+, Cl^-, Mg^{2+}, Ca^{2+}, and HCO^-_3 respectively from (a) to (f), along the Lake Rio Cuarto column water.

Figure 4. Distribution of different trace metals, (a) Fe tot, (b) Mn, (c) Ni, and (d) Co, along the Lake Rio Cuarto column water.

Figure 5. Vertical distribution of $\delta^{13}C_{CO2}$ and $\delta^{13}C_{CH4}$ of Lake Río Cuarto. The respective concentration of CO_2 and CH_4 is indicated by the size while the color indicates data from this study (green for CH_4 and black for CO_2) and from Cabassi et al., (2014) (gray).

Figure 6. Relative abundance of 16S rRNA libraries of two different depths (20 and 68 m) of Lake Rio Cuarto.

Figure 7. Relative abundance of metagenomic reads for two different depths of the Lake Rio Cuarto and for the mesocosm experiment.

Figure 8. Normalized abundance of shotgun metagenome functional read assignment for different key genes in the pathways of carbon, sulfur, nitrogen, hydrogen, photosynthesis, iron, and oxygen

Figure 9. Mesocosm experiment. CH_4 (green) and CO_2 (black) concentrations are reported for the three groups of mesocosms, two of which (a and b) CO_2 and H_2 were added in 1:4 ratio, while the third group (c) was the control. $\delta^{13}C$ for CH_4 and CO_2 was measured for all the three groups (from d to e).

Figure 10. (a) Relative abundance and (b) alpha diversity based on the Observed and Simpson indices on the peptidase genes found on long DNA sequences.

Figure 11. Lake Rio Cuarto carbon cycle infographic. Previous studies have pointed to hydrothermal mantle degassing as the only causative factor in meromictic lakes overturning (b). Conversely, the results of this study demonstrate biotic processes, such as necromass mineralization and fermentation, promote CO_2 accumulation (c). Together, both abiotic and biotic contributions to CO_2 production are believed to induce the described phenomenon.

Chapter 5

Recreational hot springs as environmental reservoir of potential multidrug-resistant pathogens

Keywords: hot springs, environmental pathogens, antibiotic resistance genes, recreational waters

Abstract

Background: Hot springs and spas are emerging as potential health hazards due to the presence of conditions supporting a diverse flora of opportunistic pathogens and the lack of regulations in numerous countries. In particular, skin infections related to strains of the bacterial genus *Acinetobacter* and *Pseudomonas* represent a growing threat.
Objectives and Methods: We used amplicon sequencing, shotgun metagenomic and microbial isolates to investigate the presence, distribution and antimicrobial resistance profile of putative pathogens in 29 naturally occurring and recreational hot springs from Costa Rica.
Results: Our results show that sequences related to the genus *Acinetobacter* and *Pseudomonas* were preferentially present in fluid samples from recreational hot springs relative to springs with less human influence, with abundances up to 68 % and 2.6 %, respectively, in the 16S rRNA data and 50 % and 5 % in metagenomic data for recreational hot springs. Strains of *P. aeruginosa* isolated from 4 recreational hot springs show a more diverse profile of antibiotic resistance compared to laboratory reference strains, with multidrug resistant strains comprising up to 30 % of the isolates. The potential for multidrug resistant pathogens in the wider microbial community is confirmed by the high diversity of antibiotic resistance genes (ARG) identified in the metagenomic sequences.
Conclusions: Our results indicate that human activities play a key role in the occurrence of opportunistic pathogens, with recreational hot springs harboring significantly higher abundances of multidrug resistant opportunistic pathogens. Our data show that recreational hot springs represent an environmental reservoir of potential multidrug-resistant pathogens, with important implications for public health.

Introduction

The past decade of research on the interdependence of human, animal, and environmental health has highlighted the importance of investigating and interpreting the spread of pathogens at the ecosystem level, in a "One Health" framework (Destoumieux-Garzón et al. 2018). The connection between environmental status and human health appears even stronger in the wake of the recent SARS-CoV2 pandemic, which likely emerged from human interactions with the environment and unsafe food handling (Sadeleer and Godfroid 2020). In this context, antibiotic resistance has become a serious problem in clinical settings (French 2005), with the development of new multidrug resistance in several groups of pathogens representing a public health threat (Tanwar et al. 2014). According to the most recent Center for Disease Control (CDC) report, more than 2.8 million antibiotic-resistant infections occur every year (CDC report 2019), with trends projected to increase for the next decade. Some clinical resistance genes have also been found to originate from environmental strains (Wright 2010). The myriad interactions between pathogens and environmental bacterial communities favors the acquisition of new resistance (Amalfitano et al. 2015). Ecosystems that come into contact with contamination vectors (for example sewage runoffs) may become reservoirs for the spreading and the evolution of new emerging drug resistant pathogens (Harris et al. 2012; Wang et al. 2021) that subsequently make their way back to the general population (Kraemer et al. 2019). Ingestion or contact with water contaminated by bacteria, viruses or protozoa can lead to several infections (Leclerc et al. 2002). In this way, new outbreaks of emerging opportunistic pathogens may occur, like those reported in Central Europe where the emerging pathogenic genus *Arcobacter* has been shown to be one of the main causes of enteric disease (Pérez-Cataluña et al. 2017; Prouzet-Mauléon et al. 2006). Together with emerging strains, the main healthcare-associated pathogens are represented by *Acinetobacter baumannii* and *Pseudomonas aeruginosa* (Motbainor et al. 2020), which cause a broad spectrum of infections from skin and wounds (Fleming et al. 2017; Kim et al. 2015), and in rare cases necrotizing fasciitis (Charnot-Katsikas et al. 2009; Reisman et al. 2012). *Pseudomonas aeruginosa* alone is responsible for 10-20% of nosocomial infections in intensive care units (Nicastri et al. 2003), and a high prevalence of these infections are caused by drug resistant phenotypes (Gill et al. 2016). The capacity of several groups of pathogens to persist in different kinds of environments and spread through them remains a worldwide concern (Exner et al. 2005). Often, the discovery of environmental pathogens is associated with water bodies like alpine rivers and lakes (Di Cesare et al. 2017a; Eckert et al. 2018), urban rivers (Cui et al. 2019) and coastal areas where fecal bacteria can diffuse out from areas of major concentrations of human land-based activities (Manini et al. 2022; Walters et al. 2011; Wang et al. 2017). Intense adverse weather events, likely to increase in frequency and intensity due to the current climate crisis (Ebi et al. 2021), are also correlated with increased pathogens and microbial resistance genes in the environment (Di Cesare et al. 2017b; Manini et al. 2022). Intensive agriculture also contributes to water quality impairment due to animal feeding operation or land applied manure (Rosen et al. 2000). Berg et al. (2005) showed that the soil can host a huge number of well known and emerging pathogens which cover a broad spectrum of clinical syndromes, underscoring the need to understand more about the colonization and transmission strategies of the opportunistic human pathogens through the rhizosphere. Recent studies focused their attention on potential reservoirs like hospital sewage where pathogens can acquire antibiotic resistance (Zhang et al. 2013). In the same way, the release of sewage sludge as soil fertilizer containing antibiotics, antibiotic-resistant bacteria, and

antibiotic resistance genes increase the environmental pool of antibiotic resistance genes and consequently favor the horizontal gene transfer of resistance among environmental microorganisms, constituting a significant reservoir of potential antibiotic-resistant pathogens (Bondarczuk et al. 2016; Calero-Cáceres et al. 2014). Besides soils, recreational water can be an important source of disease caused by pathogenic organisms, (Purnell et al. 2020). Water quality monitoring and management are the main approaches to reduce the potential infection risk, but the high variety of potential pathogenic organisms complicates the quantification and detection procedures. The most common indicator of recreational water contamination are fecal bacteria like *E. coli* or Enterococci, which can be tracked with conventional or alternative methods (Manini et al. 2022; Rodrigues and Cunha 2017). However, the efficacy of water treatment to reduce the infection rate by pathogens is not always clear; according to the US Center for Disease Control, 41% of outbreaks related to recreational water that had been treated to decrease pathogens were associated with hot tubs/spas and can be caused by pathogens from the genera *Cryptosporidium*, *Legionella* and *Pseudomonas*. On the other hand, outbreaks related to untreated recreational waters like those of lakes, rivers or oceans are mostly caused by enteric pathogens when fecally contaminated water is ingested. Sources of contamination can be stormwater runoff, sewage treatment plant discharge, and animal or boating waste (Graciaa et al. 2018; Hlavsa et al. 2018). In this context, water monitoring strategies are of primary importance to protect the public from infection risks due to pathogens harbored in recreational environments. The World Health Organization (WHO) has established specific guidelines for the monitoring of water quality (World Health Organization 2003), which have been adopted with specific legislations by numerous countries. Water types monitored worldwide include drinking water and treated (often chlorinated) or untreated recreational waters. According to the WHO, the levels of bacteria like *E. coli*, *P. aeruginosa* and *Legionella spp.* in public pools should be less than 1 colony forming unit (CFU) per 100 mL of water (World Health Organization 2003). Hot springs arise from geothermally-heated groundwater from Earth's crust (Des Marais and Walter 2019) and have been important to human society since Greek and Roman times. There is evidence for the use of hotsprings for recreational purposes in the earliest civilizations, where hot springs were considered sacred sites with powerful healing properties (Lamoreaux 2005; van Tubergen and van der Linden 2002). Currently, the recreational use of hot springs supports a growing sector of tourism worldwide with an estimated industry worth in excess of 250 billion US dollars (Erfurt-Cooper and Cooper 2009). Given the hundreds of millions of people that use naturally-occurring hot springs and *spas* annually, it does not come as a surprise that hot springs might be connected with infections. Several cases have been reported in recent years, frequently linked to bacteria belonging to genus *Bacillus* (Pandey et al. 2015), *Legionella* (Ji et al. 2014) and *Pseudomonas* (Mukherjee et al. 2012) and free-living protozoa (Marshall et al. 1997). The presence of *P. aeruginosa* as a cause of folliculitis linked to recreational water was described for the first time in 1975 (McCausland and Cox 1975). Since then, several cases of infection have been caused by *P. aeruginosa,* related to recreational water usage (Tate et al. 2003). Infections caused by the free-living ameba, *Naegleria fowleri*, have recently made news due to deadly cases of meningoencephalitis linked to hot spring use (Abrahams-Sandí et al. 2015, 2015; Vugia et al. 2019). Despite growing reports of infection cases linked with the recreational use of hot springs, information relative to the diversity and distribution of potential pathogens in naturally-occurring hot springs are limited (Ghilamicael et al. 2018; Jardine et al. 2017; Mavridou et al. 2018).

Here, we investigated the distribution of potential pathogens and antimicrobial resistance genes in several Costa Rican hot springs, comparing hot springs used for recreational purposes with hot springs in more remote areas. Our results show the presence of diverse potential pathogens linked with the recreational use of the springs. Metagenomic analysis and pure cultures isolated from the springs suggest that hot springs represent a major environmental reservoir of potential multidrug-resistant pathogens and emphasize that additional management measures need to be introduced to increase public safety.

Methods

Sampling approach

We sampled 29 individual hot springs spanning Costa Rica during two different campaigns (2017 and 2019). Some of the hot springs were sampled multiple times (see Table 1). Two types of pools were sampled: naturally occuring hot springs (n=24) and hot springs whch had been developed for recreational use (n=5) located within spas and resorts. Several of the natural hot springs had a detectable amount of anthropization, presumably due to occasional bathing of the local populations or due to use for washing, cleaning and other household activities. However, the number of people accessing the natural systems was presumably limited compared to recreational springs. In 2017, we sampled hot spring fluids and sediments (where present) from the spring inlet, defined here as the point in which the fluids with the more extreme values of temperature and/or pH were seeping into the pool (Figure 1). On average 1.5 L of fluids was sampled using a 0.22 µm Sterivex filter (Millipore Sigma) in replicate. Where present, 30 g of sediments were sampled into a sterile falcon tube. Both samples (fluids and sediments) were immediately frozen using liquid nitrogen in a cryogenic dry shipper (ThermoFisher Scientific, Arctic Express 20) for transport back to the laboratory. In 2019, fluid samples were collected in sterile 50 mL falcon tubes and kept at +4°C for isolation purposes. One replicate aliquot was also stored at -20°C after the addition of 15% glycerol (final concentrations). Additional samples were collected for the determination of the total prokaryotic counts, water major ions and dissolved organic carbon as previously reported (Barry et al. 2019; Fullerton et al. 2021).

Total prokaryotic counts and colony forming units

Samples for cell counts consisted of 1 ml fluids preserved with 3% final concentration of paraformaldehyde solution in phosphate-buffered-saline (PBS). Cell count samples were kept at room temperature during their return to the University of Tennessee and were weighed once back in the laboratory. Fluids for colony forming unit (CFU) analysis were kept at 4°C until processing. Briefly, 7-10 mL per sample were filtered on a 47 mm diameter 0.22 µm filter using a sterile filtering apparatus. The filter was placed on top of a defined agar plate, either using Leeds Acinetobacter Medium (Jawad et al. 1994) or Pseudomonas Selective Agar CN media (Goto and Enomoto 1970). For each condition, two replicate samples were incubated overnight at 37°C in accordance with standardized protocols (APHA, 2017). CFUs were manually counted on a colony counter and normalized to the fluid volume filtered.

Genomic DNA extraction and sequencing

Genomic DNA was extracted from Sterivex filters and from sediment following a modified protocol from (Giovannelli et al. 2016). Genomic DNA was visualized on 1.2 % w/v agarose gel and quantified using a Nanodrop spectrophotometer. Extracted DNA was subjected to both amplicon sequencing and shotgun metagenome sequencing. Amplicon sequencing was carried out after amplifying the bacteria-specific V4-V5 region of the 16S rRNA gene using primers 518F (AATTGGANTCAACGCCGG) and 926R (CCGYCAATTYMTTTRAGTTT) (Parada et al. 2016). Amplicon and shotgun metagenome sequencing was performed as part of the Census of Deep Life initiative with the Deep Carbon Observatory and performed at the Marine Biological Laboratory sequencing facility (https://www.mbl.edu/) on an Illumina MiSeq platform for samples from 2017. In both cases, the Illumina Nextera Flex kit for MiSeq+NextSeq, which requires a very small amount of starting material (1 ng) was used. Obtained shotgun metagenomes varied from 25 to 150 million base pairs. Amplicon sequencing data is available from the NCBI SRA archive with accession numbers PRJNA579365, while the shotgun metagenomes are deposited in PRJNA627197.

Strain isolation and testing

Small aliquots of hot spring fluids used for recreational purposes (Table 1) were inoculated in Nutrient Broth (BD, Difco) and incubated at 37 °C for several days. Afterward, dilutions to extinction were performed, following a further incubation at 37 °C. Small volumes of enrichment showing positive growth were transferred on a Nutrient -agar plate and incubated at 37 °C for 12 - 18 h. From them, singular bacterial colonies were systematically picked and streaked on Nutrient - agar plates in order to obtain isolated colonies. To identify *Pseudomonas* spp. strains, DNA was extracted and the 16S rRNA gene was amplified by PCR with universal primers 16S-8F (AGA GTT TGA TCC TGG CTC AG) and 16S-1517R (ACG GCT ACC TTG TTA CGA CTT). PCR products were sequenced with Bio-Fab Research, and identities were confirmed by blasting DNA sequences against the NCBI database.

Antimicrobial Susceptibility Testing

Antibiotic susceptibility testing was performed on all the *Pseudomonas* spp. identified strains, according to the European Committee and Antimicrobial Susceptibility Testing (EUCAST, 2018) and Clinical and Laboratory Standards Institute (CLSI, 2018) providing tables that list breakpoint guidelines to interpret antimicrobial resistance. *Pseudomonas* PAO1 was included in the analysis as a reference strain. Pure cultures were obtained on Mueller-Hinton agar plates after 12 h at 37°C and individual colonies were used to prepare the suspension for inoculation in liquid Mueller-Hinton media, two replicates for each sample. After an overnight incubation, an aliquot of the liquid culture was measured on the UV/VIS Spectrophotometer at the absorbance of 625 nm, until reaching 0.5 McFarland turbidity standard. Then, the suspension was inoculated on Mueller-Hinton agar plates and test disks were applied over the surface. The antibiotics were organized in two groups, specific and generic, the first group refers to

recommended antibiotics for antimicrobial susceptibility testing by EUCAST and CLSI instead the second group refers to generic antibiotics of common use. Meropenem (M, 10 µg), Ciprofloxacin (Cp, 5 µg), Gentamicin (G, 10 µg), Cefepime (Cf, 30 µg) are part of the specific group. Erythromycin (E, 30 µg), Cefsulodin (C, 30 µg), Ampicillin (A, 10 µg), Kanamycin (K, 30 µg) are part of the generic group. After overnight incubation at 37°C, the inhibition zone diameter was measured according to the EUCAST and CLSI BreakPoint Table (https://www.eucast.org/, https://clsi.org/).

Bioinformatic and statistical approaches

Amplicon sequencing reads were processed using *mothur* (Schloss et al. 2009), following the Miseq standard operating procedure for the identification of Amplicon Sequencing Variants (ASVs). Taxonomy was assigned using the RDP naive Bayesian classifier against the Silva v132 release (Quast et al. 2013). All statistical analyses, data processing and plotting were carried out in the R statistical software (R. development Core Team 2010), using the phyloseq (McMurdie and Holmes 2013) and ggplot2 (Wickham 2009) packages. Briefly, the obtained count table, taxonomy assignment and phylogenetic tree were combined together with the environmental variables into a phyloseq object. Low prevalence ASVs, mitochondria and chloroplast related sequences were removed. In both fluids and sediments, common laboratory contaminants from DNA processing, feces, and skin (Sheik et al. 2018) were largely absent (< 0.04% in the entire dataset and less than 0.01% in any individual library), and no ASV was shared by all samples. After QC and filtering a total of 2,346,695 reads comprising 40,690 ASVs were obtained. Genera from putative pathogenic species were used for downstream analysis. A list of the putative pathogenic genera searched in the libraries is available as Supplementary Table 1. Shotgun metagenomic reads were quality checked and trimmed using Trimmomatic (Bolger et al. 2014). Obtained post-qc reads were assembled using metaSPAdes (Nurk et al. 2017) with standard parameters. Obtained contigs were binned using MetaWRAP. And checked using CheckM. Post QC raw reads were used to assess the taxonomic composition with a sensitive profiling tool called Kaiju (Menzel et al. 2016). Assembled contigs were searched against the CARD database (Alcock et al. 2023) to identify antimicrobial resistance genes. From the blast, only the hits with almost 70% of coverage and 70% of similarity were retained. Complete 16S rRNA genes from the MAGs identified as *Acinetobacter* spp. and 16S rRNA partial sequences with individual abundance above 0.1% from the amplicon libraries were used for phylogenetic analyses. Representative sequences of *Acinetobacter* spp. type strains were downloaded from NCBI. Partial sequences of 16S rRNA from *Pseudomonas* spp. isolates and 16S rRNA partial sequences with individual abundance above 0.05% from the amplicon libraries were used for phylogenetic analyses while the representative sequences of *Pseudomonas* spp. type strains were downloaded from NCBI. Both the trees were processed separately. Alignment of the 16S rRNA sequences was carried out using MAFFT (Katoh et al. 2019) while the trimming step was performed with ClipKit (Steenwyk et al. 2020). Phylogenetic distances matrix were calculated using the GRT model and the maximum-likelihood trees were constructed with IQ-TREE (Minh et al. 2020). The branch support was estimated using the approximate likelihood ratio test (aLRT) (Anisimova and Gascuel 2006). A complete R script containing all the steps to reproduce our analysis is available at

https://github.com/giovannellilab/Selci_et_al_Hot_springs_pathogens and released as a permanent version using Zenodo under the DOI: https://doi.org/XXXXXXXXXX.

Results

Abundance of putative pathogenic sequences in Costa Rica hot springs

Putative pathogens were identified in nearly all sites that yielded amplifiable DNA (n=13 sites) with the exception of S1, S12, S17, and S18. Sequences associated with known putative pathogens constitute the 2.1% of the whole ASVs identified (Figure 2). Among them, the number of ASVs related to known pathogenic taxa accounted for more groups in fluid samples compared to the sediments (Kruskal test, p value < 0.001; Supplementary figure 1). In hot spring fluid samples (Figure 2), putative pathogens were represented, on average, by ASVs assigned to the genus *Acinetobacter* (52.8%) followed by *Flavobacterium* (9.5%) and *Pseudomonas* (5.6%). Less abundant sequences were related to the genera *Bacillus* (5.5%), *Halomonas* (3.9%), *Treponema* (3.2%), *Roseomonas* (3.1%), and *Comamonas* (3.0%). They were followed by genera like *Legionella* (2.3%), *Deinococcus* (1.8%), Staphylococcus *and* Enterococcus (1.6%), while *Leptospira*, *Stenotrophomonas*, and *Coxiella* had an abundance of 1.0%. Sampled fluids of recreational hot springs were characterized by higher concentrations of putative pathogens when compared to natural hot springs located in more remote locations (Kruskal test, p value < 0.05).

Acinetobacter was the main represented genus with a relative abundance of 94.3% in S3, 91.5% in S5, 88.2% in S6, 83.2% in S2, and 71.1% in S7. The genus *Pseudomonas* was second in abundance with 12% in S2 and S7, while in S3 was around 5%. The lowest *Pseudomonas* abundance was found in S5 and S6 fluids (0.3% and 1.7%, respectively) which were characterized by higher abundance of *Comamonas* (~ 8% for both). In natural hot springs not associated with spas and recreational centers, on the other hand, *Acinetobacter* was found with a high relative abundance only in S13 (86.7%), probably due to the presence of a farm next to the spring. Other natural hot springs like S10, S11, and S16 showed an average abundance of 51%, while for sites like S8, S9, S15, and S19 the abundance of *Acinetobacter* was 4.25% on average. *Pseudomonas* was present in natural hot springs with an abundance of 8.73% and 8.33% for S8 and S9, while for S11, S16, S10, and S19 the abundance had an average of 6.1%. The remaining springs (S13 and S15) were under 1%.

The presence of putative pathogens was also confirmed using shotgun metagenomic data. Reads classified as *Acinetobacter* were mainly present in recreational hot springs metagenomes, especially in S4 (50%), S3 (23.5%), S6 (2.2%), and S7 (0.65% of the total classified reads) while they were in low abundance or absent in natural hot springs, with abundances of 1.64% in S16. Reads assigned to *Comamonas* were found only in S6 (0.5%) while *Pseudomonas* was identified in all the investigated recreational hot springs with a relative abundance of 5% in S3, 1.42% in S6, 1.14% in S4, and 0.74% in S2, while it was not found in the natural springs (Supplementary figure 2). Colony forming units (CFU) count on *Pseudomonas* Selective Agar CN media (PCN agar) and Leeds *Acinetobacter* Medium (LCM) were used to confirm the viability of the putative pathogens identified through sequencing. Results for S5, the spring with the highest number of sequences assigned to *Acinetobacter* and *Pseudomonas*, was 2,680 CFU/100 mL (Figure 3A). No CFUs were obtained on LCM.

Phylogenetic analyses focused on the two most abundant genera identified in the 16S rRNA analysis, *Acinetobacter* and *Pseudomonas*. DNA sequences from the CR17 campaign were used to extract complete *Acinetobacter* spp. 16S rRNA genes (MAGs) and 16S rRNA gene partial sequences (amplicon libraries). The phylogenetic tree of *Acinetobacter* spp. (Figure 4) revealed that almost all the 16S rRNA ASVs fall close to *Acinetobacter* type strains known as human pathogens. In particular, *Acinetobacter* sp. CR1 clusters with *Acinetobacter nosocomialis* HQ180192, *Acinetobacter seifertii* FJ860878, and *Acinetobacter pittii* HQ180184, while *Acinetobacter* sp. CR10 is positioned close to *Acinetobacter baumannii* X81660. Moreover, all the complete 16S rRNA genes recovered from the MAGs and 104 ASVs fall within the *Acinetobacter junii* X81664 cluster. The abundance of amplicon sequence variants related to *Acinetobacter* spp. was also evaluated for hot spring type, with more ASVs in recreational hot springs relative to natural ones (Kruskal test, p value < 0.01). A *Pseudomonas* phylogenetic tree was constructed with the same approach as described above (Figure 5). Most of the 16S rRNA ASVs are related to *Pseudomonas* type strains known in literature as human pathogens, while all the isolated type strains fall in a cluster related to *Pseudomonas aeruginosa*. The abundance of ASVs related to *Pseudomonas* spp. was evaluated for hot spring type, with more ASVs in recreational hot springs compared to natural ones (Kruskal test, p value < 0.01).

Antibiotic resistance of the putative pathogens

Antibiotic resistance profiles were determined for 48 *P. aeruginosa* strains isolated from four different recreational springs. All 48 identified strains showed different antibiotic susceptibility to the two groups of antibiotics tested, specific and generic, respectively. In accordance to EUCAST, the distribution of the antibiotic resistance profiles obtained against Meropenem, Gentamicin, Ciprofloxacin, and Cefepime (specific group) showed different patterns both in the fluids coming from the recreational hot spring and in comparison with PAO1 reference strain (Figure 6). More than 12 % of the *P. aeruginosa* type strains identified in each study site showed resistance against Gentamicin, Ciprofloxacin, and Cefepime while only the strains isolated from S5 showed resistance against Meropenem. Multi-drug resistances (MDRs) were found in all the study sites with a relative abundance higher than 30 % in S5 and S3. PAO1 reference strain showed resistance only against the Meropenem and Cefepime and no MDR was detected. In accordance with CLSI, antibiotic resistance profiles obtained against Meropenem, Gentamicin, Ciprofloxacin, and Cefepime (specific group) showed a higher sensitivity of the *P. aeruginosa* type strains identified (Supplementary Figure 3). Only the strains coming from S5 showed resistance against two antibiotics while strains from S4 and S2 were resistant against Gentamicin. Strains isolated from S3 were completely sensitive to the antibiotics tested while PAO1 reference strain showed resistance only against the Meropenem.

In all the investigated sites, 100% of the *Pseudomonas* sp. isolates were resistant against Ampicillin and Erythromycin, while a lower percentage were resistant against Cefsulodin and Kanamycin (Supplementary Figure 4). Also for the generic group of antibiotics, MDR were found in each study site, and S5 showed the highest relative abundance in MDR strains (62.5 % of 24 total strains were MDR) while S3 showed the lowest (16.7 % of 6 total strains were MDR). PAO1 reference strain showed resistance only against the Cefsulodin and Kanamycin, and no MDR were detected.

The occurrence of Antimicrobial Resistance Genes (ARGs) in natural and recreational hot springs was investigated by blasting the assembled DNA contigs against the CARD database

(Figure 7). Natural hot springs fluids and sediments (S8, S13, S16, and S18) were characterized by a low number of ARGs that primarily belong to the group of *mux* and *sme* which codify for resistance against , while almost all recreational hot springs displayed a major diversity of ARGs (Figure 7). The fluids of S3, S4, and S2 showed the higher number of resistance genes which comprised the categories of *OXA*, *ADC*, and *ade*. In S5 and S6 fluids, in contrast, only a *nov* gene and a *ade* gene were found, respectively.

Discussion

Environmental reservoirs of antibiotic resistance represent a growing concern for global public health as potential sources for human infection (Barrett 2012; Mills and Lee 2019; Tello et al. 2012). Remote environments as well as recreational waters have been identified as current reservoirs of antibiotic resistance bacteria and genes (Eckert et al. 2018; Knapp et al. 2012; Overbey et al. 2015; Segawa et al. 2013; Zhang et al. 2022). The increased antimicrobial resistance of opportunistic pathogens together with increased human population densities in areas subject to habitat deterioration represent a global threat to public health (Myers and Patz 2009).

Hot springs harbor a plethora of microorganisms that include both naturally-occurring and introduced opportunistic pathogens. Previous studies reported the presence of several infective microorganisms including free living protozoa (Abrahams-Sandí et al. 2015; De Jonckheere 2005; Montalbano Di Filippo et al. 2017) and bacteria (Aburto-Medina et al. 2020; Mukherjee et al. 2012; Thorolfsdottir and Marteinsson 2013, Ghilamicael et al. 2018; Rupasinghe et al. 2022; Sheehan et al. 2005) in hot springs and spas. Opportunistic pathogens of the genus *Acinetobacter* and *Pseudomonas* have been reported worldwide as infective agents associated with the recreational use of hot springs (Barben et al. 2005; Falkinham Joseph O. et al. 2015; Lutz and Lee 2011; Martins et al. 1995; Moore et al. 2002; Niewolak and Opieka 2000). While their presence is well documented in recreational hot springs (Rahel et al. 2021) information related to their origin, distribution and antimicrobial resistance profiles are limited (Ghilamicael et al. 2018; Lutz and Lee 2011).

Here, we report sequencing and isolation data from 22 hot springs from Costa Rica, which include natural hot springs in remote areas as well as recreational hot springs. Our results indicate that the occurrence of opportunistic pathogens belonging to the *Acinetobacter* and *Pseudomonas* genera correlates with recreational springs associated with human activities whose physicochemical properties promote the persistence mesophilic microorganism. Species related to *Acinetobacter* were found only through in-silico analysis and phylogenetic investigations showed that most of them were closely related to human pathogens like *A. junii* and *A. baumannii*, two species well known in health care systems for their role in nosocomial infections (Hung et al. 2009; Visca et al. 2011). Species related to *Pseudomonas*, instead, were found both in in-silico and laboratory culture analysis. In particular, phylogenetic evaluations of DNA sequences from 16S rRNA libraries showed a similarity to species such as *P. fluorescens* and *P. fulva* related to human infections (Liu et al. 2014; Scales et al. 2014) while all laboratory isolates were closely related to *Pseudomonas aeruginosa* strains that are well known for the high rate of acute and chronic infections in hospital settings (Crone et al. 2020; Huber et al. 2016; Obritsch et al. 2005) as well as spas, hot springs and swimming pools (Lutz and Lee

2011; Moore et al. 2002). *P. aeruginosa* has an evolutionary advantage due to its ability to form biofilms (Hall-Stoodley and Stoodley 2005) and survive in different environmental niches (Berg et al. 2005, 1990) (i.e., recreational hotsprings), (Guida et al. 2016; Mena and Gerba 2009). The concentrations of *P. aeruginosa* colonies we found in recreational springs were orders of magnitude higher than the normal threshold accepted for natural spas (World Health Organization 2003), a result in line with findings from other environments exposed to anthropogenic pressure (Crone et al. 2020; Deredjian et al. 2014; Mena and Gerba 2009).

The ARGs detected in these recreational hot springs indicated a low distribution of genes related to drug efflux functions, compared to the most abundant occurrence of enzymatic degradation mechanisms by β-lactamases associated to *A. baumannii* and *P. aeruginosa* related genes (*e.g.*, OXA, ADC, and PDC) which confer resistance to a broad range of antibiotics like carbapenems and cephalosporins (Evans and Amyes 2014; Rao et al. 2020). These results were also consistent with the potential resistance found in multiple *P. aeruginosa* type strains (see below), suggesting a correlation between the relative proportion of resistant bacteria with the human activities present in recreational hot springs, an observation already reported for other environments (Amos et al. 2014; Di Cesare et al. 2015, 2017; Graham et al. 2011; Hatosy and Martiny 2015; Luo et al. 2010; Popowska et al. 2012; Pruden et al. 2012).

The antibiotic resistance profiles, for the isolated *P. aeruginosa* type strains, showed varying results when different standard references (EUCAST vs. CLSI) were used. In accordance with the EUCAST guidelines, most of the *P. aeruginosa* strains displayed resistance against almost all the tested antibiotics, with a particular resistance rate against general antibiotics like erythromycin, cefsulodin, kanamycin, and ampicillin, already known for their mild efficacy against *P. aeruginosa* infections (Alam et al. 2019; Bălăşoiu and Bălăşoiu 2014; Khan et al. 2015). Resistance was found also for more specific antibiotics like ciprofloxacin, cefepime, gentamicin, and meropenem, probably linked with *P. aeruginosa*'s ability to acquire mutations on specific targets, for example on the DNA gyrase and topoisomerase IV enzymes in fluoroquinolones resistance (Drlica and Zhao 1997; Jacoby 2005) or using different mechanisms that include antibiotic cleavage, efflux, or reduced drug uptake in *β*-Lactams resistance (Pfeifer et al. 2010; Poole 2004, 2011). Multidrug-resistant *P. aeruginosa* strains were detected in over half of the investigated sites, which may suggest the presence of a selective pressure able to escalate antibiotic resistance in the ecosystem (Li et al. 2017). This was previously observed by Sharma et al. (2022), who investigated the presence of resistomes in different Himalayan hot springs. In this context, metals as well as biocide contamination have been demonstrated to involve environmental multidrug resistance acquisition (Bengtsson-Palme et al. 2016; Catao et al. 2021; Thomas et al. 2020). This induces a hyperexpression of the drug efflux pumps in Gram-negative bacteria (Amsalu et al. 2020; Khan et al. 2018; Piddock 2006). When using the CLSI instead of EUCAST guidelines, the number of *P. aeruginosa* strains resistant to the tested antibiotics decreased (Supplementary Figure 1). The only resistance found was against gentamicin with the highest number of resistant isolates from the site S2, followed by S4 and S5 isolates. The latter was the only site where a low number *P. aeruginosa* type strains resistant to meropenem was found, suggesting a strong response to the antibiotic exposure since the CLSI has more stringent diameter breakpoints. The discrepancies between the two antimicrobial susceptibility testing systems (Supplementary Figure 3) are well known, and the lack of agreement in the antibiotic breakpoints interpretation have critical implications on surveillance

initiatives (Bork et al. 2017; Cusack et al. 2019, 2019; Hombach et al. 2013; Machuca et al. 2016; Rodríguez-Baño et al. 2012; Rodríguez-Martínez et al. 2011).

Overall our data suggest that hot springs represent an environmental reservoir of antimicrobial resistant opportunistic pathogens. The low abundance of putative pathogens and ARGs in natural occurring springs as well as the higher multidrug resistant pathogens observed only in the recreational hot springs suggests that the presence of pathogens is linked to human activities. Thus, anthropic pressure drives the differential occurrence of putative multidrug-resistant pathogens, making recreational hot springs a suitable reservoir for pathogen proliferation. Recreational springs have a more narrow range of pH and temperatures that makes them more suitable for human bathing compared to naturally occurring hot springs. These conditions however also fall in the physiological range of most known pathogens, supporting their higher abundances at recreational sites. In addition, it is possible that the combined effect of anthropic pressure and rainfall rate, which characterizes Central America during the rainy seasons, may affect the ARG dynamics, facilitating the drainage from contaminated soil to the hot spring fluids and contributing to the spreading of antibiotic resistance genes, as was previously reported for other pathogens (Di Cesare et al. 2017). Additionally, the presence of putative pathogens in recreational hot springs might be supported by insufficient management practices. The application of more stringent management protocols (Eurosurveillance editorial team Collective 2006), which include more frequent drainage and cleaning operations, better water quality management and regular microbiological testing, might mitigate the risk. This is however a temporary solution to the rising threat of antimicrobial resistance in the environment, which will require direct and decisive legislative interventions to limit antibiotic use together with new investment in antimicrobial research (Majumder et al. 2020).

Conclusion

In conclusion, our study presents a combined in-silico and laboratory culture survey to investigate the potential occurrence of putative pathogens and antibiotic resistance genes in several recreational and naturally occurring hot springs of Central America. Our results indicate that human activities play a key role in the occurrence of opportunistic pathogens, with recreational hot springs harboring significantly higher abundances of multi drug resistant opportunistic pathogens. This highlights the need for a better understanding of the hot springs' role as reservoirs of potential multi resistant pathogens in the environment. Given the rise in popularity of hot springs as tourist attractions globally, more effective management guidelines and prevention measures are necessary to ensure public safety and preserve the cultural and health legacy of this millenia-old leisure activity.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

Sequences are available through the NCBI SRA archive with accession numbers PRJNA579365 for the amplicon data, and PRJNA627197 for the shotgun metagenomes. A complete R script containing all the steps to reproduce our analysis is available at https://github.com/giovannellilab/Selci_et_al_Hot_springs_pathogens permanently stored with DOI: https://doi.org/10.5281/zenodo.XXXXXX. This work was partially supported by the Alfred P. Sloan Foundation through a Deep Carbon Observatory (G-2016-7206) grant to P.H.B., J.M.d.M, D.G. and K.G.L., with DNA sequencing provided from the Census of Deep Life.

References

Alcock, B. P., Huynh, W., Chalil, R., Smith, K. W., Raphenya, A. R., Wlodarski, M. A., ... & McArthur, A. G. (2022). CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Research.

Abrahams-Sandí E, Retana-Moreira L, Castro-Castillo A, Reyes-Batlle M, Lorenzo-Morales J. 2015. Fatal Meningoencephalitis in Child and Isolation of *Naegleria fowleri* from Hot Springs in Costa Rica. Emerg Infect Dis 21:382–384; doi:10.3201/eid2102.141576.

Aburto-Medina A, Shahsavari E, Cohen M, Mantri N, Ball AS. 2020. Analysis of the Microbiome (Bathing Biome) in Geothermal Waters from an Australian Balneotherapy Centre. Water 12:1705; doi:10.3390/w12061705.

Alam ST, Le TAN, Park J-S, Kwon HC, Kang K. 2019. Antimicrobial biophotonic treatment of ampicillin-resistant Pseudomonas aeruginosa with hypericin and ampicillin cotreatment followed by orange light. Pharmaceutics 11: 641.

Amalfitano S, Coci M, Corno G, Luna GM. 2015. A microbial perspective on biological invasions in aquatic ecosystems. Hydrobiologia 746:13–22; doi:10.1007/s10750-014-2002-6.

Amsalu A, Sapula SA, De Barros Lopes M, Hart BJ, Nguyen AH, Drigo B, et al. 2020. Efflux Pump-Driven Antibiotic and Biocide Cross-Resistance in Pseudomonas aeruginosa Isolated from Different Ecological Niches: A Case Study in the Development of Multidrug Resistance in Environmental Hotspots. Microorganisms 8:1647; doi:10.3390/microorganisms8111647.

Bălăşoiu M, Bălăşoiu AT. 2014. Pseudomonas Aeruginosa Resistance Phenotypes and Phenotypic Highlighting Methods. Curr Health Sci J 85–92; doi:10.12865/CHSJ.40.02.01.

Barben J, Hafen G, Schmid J. 2005. Pseudomonas aeruginosa in public swimming pools and bathroom water of patients with cystic fibrosis. J Cyst Fibros 4:227–231; doi:10.1016/j.jcf.2005.06.003.

Barrett JR. 2012. Preventing Antibiotic Resistance in the Wild: A New End Point for Environmental Risk Assessment. Environ Health Perspect 120:a321–a321; doi:10.1289/ehp.120-a321a.

Barry PH, Moor JM de, Giovannelli D, Schrenk M, Hummer DR, Lopez T, et al. 2019. Forearc

carbon sink reduces long-term volatile recycling into the mantle. Nature 568:487; doi:10.1038/s41586-019-1131-5.

Bengtsson-Palme J, Hammaren R, Pal C, Östman M, Björlenius B, Flach C-F, et al. 2016. Elucidating selection processes for antibiotic resistance in sewage treatment plants using metagenomics. Sci Total Environ 572: 697–712.

Berg G, Eberl L, Hartmann A. 2005. The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. Environ Microbiol 7:1673–1685; doi:10.1111/j.1462-2920.2005.00891.x.

Berg G, Seech AG, Lee H, Trevors JT. 1990. Identification and characterization of a soil bacterium with extracellular emulsifying activity. J Environ Sci Health Part A 25: 753–764.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120; doi:10.1093/bioinformatics/btu170.

Bondarczuk K, Markowicz A, Piotrowska-Seget Z. 2016. The urgent need for risk assessment on the antibiotic resistance spread via sewage sludge land application. Environ Int 87:49–55; doi:10.1016/j.envint.2015.11.011.

Bork JT, Heil EL, Leekha S, Fowler RC, Hanson ND, Majumdar A, et al. 2017. Impact of CLSI and EUCAST Cefepime breakpoint changes on the susceptibility reporting for Enterobacteriaceae. Diagn Microbiol Infect Dis 89: 328–333.

Calero-Cáceres W, Melgarejo A, Colomer-Lluch M, Stoll C, Lucena F, Jofre J, et al. 2014. Sludge As a Potential Important Source of Antibiotic Resistance Genes in Both the Bacterial and Bacteriophage Fractions. Environ Sci Technol 48:7602–7611; doi:10.1021/es501851s.

Catao EC, Gallois N, Fay F, Misson B, Briand J-F. 2021. Metal resistance genes enrichment in marine biofilm communities selected by biocide-containing surfaces in temperate and tropical coastal environments. Environ Pollut 268: 115835.

Charnot-Katsikas A, Dorafshar AH, Aycock JK, David MZ, Weber SG, Frank KM. 2009. Two Cases of Necrotizing Fasciitis Due to Acinetobacter baumannii. J Clin Microbiol 47:258–263; doi:10.1128/JCM.01250-08.

Crone S, Vives-Flórez M, Kvich L, Saunders AM, Malone M, Nicolaisen MH, et al. 2020. The environmental occurrence of Pseudomonas aeruginosa. APMIS 128:220–231; doi:10.1111/apm.13010.

Cui Q, Huang Y, Wang H, Fang T. 2019. Diversity and abundance of bacterial pathogens in urban rivers impacted by domestic sewage. Environ Pollut 249:24–35; doi:10.1016/j.envpol.2019.02.094.

Cusack TP, Ashley EA, Ling CL, Rattanavong S, Roberts T, Turner P, et al. 2019. Impact of CLSI and EUCAST breakpoint discrepancies on reporting of antimicrobial susceptibility and AMR surveillance. Clin Microbiol Infect 25:910–911; doi:10.1016/j.cmi.2019.03.007.

De Jonckheere JF. 2005. The isolation of Naegleria italica from Peru indicates that this potentially pathogenic species occurs worldwide. Parasitol Int 54:173–175; doi:10.1016/j.parint.2005.03.004.

Deredjian A, Colinon C, Hien E, Brothier E, Youenou B, Cournoyer B, et al. 2014. Low occurrence of Pseudomonas aeruginosa in agricultural soils with and without organic amendment. Front Cell Infect Microbiol 4.

Des Marais DJ, Walter MR. 2019. Terrestrial Hot Spring Systems: Introduction. Astrobiology 19:1419–1432; doi:10.1089/ast.2018.1976.

Destoumieux-Garzón D, Mavingui P, Boetsch G, Boissier J, Darriet F, Duboz P, et al. 2018. The One Health Concept: 10 Years Old and a Long Road Ahead. Front Vet Sci 5; doi:10.3389/fvets.2018.00014.

Di Cesare A, Eckert EM, Rogora M, Corno G. 2017a. Rainfall increases the abundance of antibiotic resistance genes within a riverine microbial community. Environ Pollut 226:473–478; doi:10.1016/j.envpol.2017.04.036.

Di Cesare A, Eckert EM, Rogora M, Corno G. 2017b. Rainfall increases the abundance of antibiotic resistance genes within a riverine microbial community. Environ Pollut 226:473–478; doi:10.1016/j.envpol.2017.04.036.

Drlica K, Zhao X. 1997. DNA gyrase, topoisomerase IV, and the 4-quinolones. Microbiol Mol Biol Rev 61: 377–392.

Ebi KL, Vanos J, Baldwin JW, Bell JE, Hondula DM, Errett NA, et al. 2021. Extreme Weather and Climate Change: Population Health and Health System Implications. Annu Rev Public Health 42:293–315; doi:10.1146/annurev-publhealth-012420-105026.

Eckert EM, Di Cesare A, Coci M, Corno G. 2018. Persistence of antibiotic resistance genes in large subalpine lakes: the role of anthropogenic pollution and ecological interactions. Hydrobiologia 824:93–108; doi:10.1007/s10750-017-3480-0.

Erfurt-Cooper P, Cooper M. 2009. *Health and Wellness Tourism: Spas and Hot Springs*. Channel View Publications.

Eurosurveillance editorial team Collective. 2006. Comprehensive guidance to reduce infection risk from spa pools and whirlpool baths. Euro Surveill 11:2925; doi:https://doi.org/10.2807/esw.11.11.02925-en.

Exner M, Kramer A, Lajoie L, Gebel J, Engelhart S, Hartemann P. 2005. Prevention and control of health care–associated waterborne infections in health care facilities. Am J Infect Control 33:S26–S40; doi:10.1016/j.ajic.2005.04.002.

Falkinham Joseph O., Hilborn Elizabeth D., Arduino Matthew J., Pruden Amy, Edwards Marc A. 2015. Epidemiology and Ecology of Opportunistic Premise Plumbing Pathogens: Legionella pneumophila, Mycobacterium avium, and Pseudomonas aeruginosa. Environ Health Perspect 123:749–758; doi:10.1289/ehp.1408692.

Fleming ID, Krezalek MA, Belogortseva N, Zaborin A, Defazio J, Chandrasekar L, et al. 2017. Modeling Acinetobacter baumannii wound infections: The critical role of iron. J Trauma Acute Care Surg 82:557–565; doi:10.1097/TA.0000000000001338.

French G. 2005. Clinical impact and relevance of antibiotic resistance. Adv Drug Deliv Rev 57:1514–1527; doi:10.1016/j.addr.2005.04.005.

Fullerton KM, Schrenk MO, Yücel M, Manini E, Basili M, Rogers TJ, et al. 2021. Effect of tectonic processes on biosphere–geosphere feedbacks across a convergent margin. Nat Geosci 14:301–306; doi:10.1038/s41561-021-00725-0.

Ghilamicael AM, Boga HI, Anami SE, Mehari T, Budambula NLM. 2018. Potential human pathogenic bacteria in five hot springs in Eritrea revealed by next generation sequencing. PLOS ONE 13:e0194554; doi:10.1371/journal.pone.0194554.

Gill JS, Arora S, Khanna SP, Kumar KH. 2016. Prevalence of Multidrug-resistant, Extensively Drug-resistant, and Pandrug-resistant Pseudomonas aeruginosa from a Tertiary Level Intensive Care Unit. J Glob Infect Dis 8:155; doi:10.4103/0974-777X.192962.

Giovannelli D, d'Errico G, Fiorentino F, Fattorini D, Regoli F, Angeletti L, et al. 2016. Diversity and Distribution of Prokaryotes within a Shallow-Water Pockmark Field. Front Microbiol 7; doi:10.3389/fmicb.2016.00941.

Goto S, Enomoto S. 1970. Nalidixic acid cetrimide agar. A new selective plating medium for the selective isolation of Pseudomonas aeruginosa. Jpn J Microbiol 14:65–72; doi:10.1111/j.1348-0421.1970.tb00492.x.

Graciaa DS, Cope JR, Roberts VA, Cikesh BL, Kahler AM, Vigar M, et al. 2018. Outbreaks Associated with Untreated Recreational Water — United States, 2000–2014. 67: 6.

Guida M, Di Onofrio V, Gallè F, Gesuele R, Valeriani F, Liguori R, et al. 2016. Pseudomonas aeruginosa in Swimming Pool Water: Evidences and Perspectives for a New Control Strategy. Int J Environ Res Public Health 13:919; doi:10.3390/ijerph13090919.

Harris SJ, Cormican M, Cummins E. 2012. Antimicrobial Residues and Antimicrobial-Resistant Bacteria: Impact on the Microbial Environment and Risk to Human Health—A Review. Hum Ecol Risk Assess Int J 18:767–809; doi:10.1080/10807039.2012.688702.

Hlavsa MC, Cikesh BL, Roberts VA, Kahler AM, Vigar M, Wade TJ, et al. 2018. Outbreaks Associated with Treated Recreational Water — United States, 2000–2014. 67: 5.

Hombach M, Wolfensberger A, Kuster SP, Böttger EC. 2013. Influence of clinical breakpoint changes from CLSI 2009 to EUCAST 2011 antimicrobial susceptibility testing guidelines on multidrug resistance rates of Gram-negative rods. J Clin Microbiol 51: 2385–2387.

Huber P, Basso P, Reboud E, Attrée I. 2016. Pseudomonas aeruginosa renews its virulence factors. Environ Microbiol Rep 8: 564–571.

Hung Y-T, Lee Y-T, Huang L-J, Chen T-L, Yu K-W, Fung C-P, et al. 2009. Clinical characteristics of patients with Acinetobacter junii infection. J Microbiol Immunol Infect Wei Mian Yu Gan Ran Za Zhi 42: 47–53.

Jacoby GA. 2005. Mechanisms of resistance to quinolones. Clin Infect Dis 41: S120–S126.

Jardine JL, Abia ALK, Mavumengwana V, Ubomba-Jaswa E. 2017. Phylogenetic Analysis and Antimicrobial Profiles of Cultured Emerging Opportunistic Pathogens (Phyla Actinobacteria and Proteobacteria) Identified in Hot Springs. Int J Environ Res Public Health 14:1070; doi:10.3390/ijerph14091070.

Jawad A, Hawkey PM, Heritage J, Snelling AM. 1994. Description of Leeds Acinetobacter Medium, a new selective and differential medium for isolation of clinically important Acinetobacter spp., and comparison with Herellea agar and Holton's agar. J Clin Microbiol 32: 2353–2358.

Ji W-T, Hsu B-M, Chang T-Y, Hsu T-K, Kao P-M, Huang K-H, et al. 2014. Surveillance and evaluation of the infection risk of free-living amoebae and Legionella in different

aquatic environments. Sci Total Environ 499:212–219; doi:10.1016/j.scitotenv.2014.07.116.

Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160–1166; doi:10.1093/bib/bbx108.

Khan HA, Ahmad A, Mehboob R. 2015. Nosocomial infections and their control strategies. Asian Pac J Trop Biomed 5:509–514; doi:10.1016/j.apjtb.2015.05.001.

Khan R, Roy N, Choi K, Lee S-W. 2018. Distribution of triclosan-resistant genes in major pathogenic microorganisms revealed by metagenome and genome-wide analysis. PloS One 13: e0192277.

Kim M, Christley S, Khodarev NN, Fleming I, Huang Y, Chang E, et al. 2015. Pseudomonas aeruginosa wound infection involves activation of its iron acquisition system in response to fascial contact. J Trauma Acute Care Surg 78:823–829; doi:10.1097/TA.0000000000000574.

Knapp C, Lima L, Olivares-Rieumont S, Bowen E, Werner D, Graham DW. 2012. Seasonal Variations in Antibiotic Resistance Gene Transport in the Almendares River, Havana, Cuba. Front Microbiol 3.

Kraemer SA, Ramachandran A, Perron GG. 2019. Antibiotic Pollution in the Environment: From Microbial Ecology to Public Policy. Microorganisms 7; doi:10.3390/microorganisms7060180.

Lamoreaux PE. 2005. History and classification of springs. Geol Soc Am Abstr Programs 37: 324.

Leclerc H, Schwartzbrod L, Dei-Cas E. 2002. Microbial Agents Associated with Waterborne Diseases. Crit Rev Microbiol 28:371–409; doi:10.1080/1040-840291046768.

Li L-G, Xia Y, Zhang T. 2017. Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection. ISME J 11: 651–662.

Liu Y, Liu K, Yu X, Li B, Cao B. 2014. Identification and control of a Pseudomonas spp (P. fulva and P. putida) bloodstream infection outbreak in a teaching hospital in Beijing, China. Int J Infect Dis 23: 105–108.

Lutz JK, Lee J. 2011. Prevalence and Antimicrobial-Resistance of Pseudomonas aeruginosa in Swimming Pools and Hot Tubs. Int J Environ Res Public Health 8:554–564; doi:10.3390/ijerph8020554.

Machuca J, Briales A, Díaz-de-Alba P, Martínez-Martínez L, Rodríguez-Martínez J-M, Pascual Á. 2016. Comparison of clinical categories for Escherichia coli harboring specific qnr and chromosomal-mediated fluoroquinolone resistance determinants according to CLSI and EUCAST. Enfermedades Infecc Microbiol Clínica 34: 188–190.

Majumder MAA, Rahman S, Cohall D, Bharatha A, Singh K, Haque M, et al. 2020. Antimicrobial stewardship: Fighting antimicrobial resistance and protecting global public health. Infect Drug Resist 13: 4713.

Manini E, Baldrighi E, Ricci F, Grilli F, Giovannelli D, Intoccia M, et al. 2022. Assessment of Spatio-Temporal Variability of Faecal Pollution along Coastal Waters during and after Rainfall Events. Water 14:502; doi:10.3390/w14030502.

Marshall MM, Naumovitz D, Ortega Y, Sterling CR. 1997. Waterborne protozoan pathogens. Clin Microbiol Rev 10: 67–85.

Martins MT, Sato MIZ, Alves MN, Stoppe NC, Prado VM, Sanchez PS. 1995. Assessment of microbiological quality for swimming pools in South America. Water Res 29: 2417–2420.

Mavridou A, Pappa O, Papatzitze O, Dioli C, Kefala AM, Drossos P, et al. 2018. Exotic Tourist Destinations and Transmission of Infections by Swimming Pools and Hot Springs—A Literature Review. Int J Environ Res Public Health 15:2730; doi:10.3390/ijerph15122730.

McCausland WJ, Cox PJ. 1975. Pseudomonas infection traced to motel whirlpool. J Environ Health 37: 455–459.

McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217; doi:10.1371/journal.pone.0061217.

Mena KD, Gerba CP. 2009. Risk Assessment of Pseudomonas aeruginosa in Water. In: *Reviews of Environmental Contamination and Toxicology Vol 201* (D.M. Whitacre, ed). Vol. 201 of *Reviews of Environmental Contamination and Toxicology*. Springer US:Boston, MA. 71–115.

Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat Commun 7:11257; doi:10.1038/ncomms11257.

Mills MC, Lee J. 2019. The threat of carbapenem-resistant bacteria in the environment: Evidence of widespread contamination of reservoirs at a global scale. Environ Pollut 255:113143; doi:10.1016/j.envpol.2019.113143.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol 37:1530–1534; doi:10.1093/molbev/msaa015.

Montalbano Di Filippo M, Novelletto A, Di Cave D, Berrilli F. 2017. Identification and phylogenetic position of Naegleria spp. from geothermal springs in Italy. Exp Parasitol 183:143–149; doi:10.1016/j.exppara.2017.08.008.

Moore JE, Heaney N, Millar BC, Crowe M, Elborn JS. 2002. Incidence of Pseudomonas aeruginosa in recreational and hydrotherapy pools. Commun Dis Public Health 5: 23–26.

Motbainor H, Bereded F, Mulu W. 2020. Multi-drug resistance of blood stream, urinary tract and surgical site nosocomial infections of Acinetobacter baumannii and Pseudomonas aeruginosa among patients hospitalized at Felegehiwot referral hospital, Northwest Ethiopia: a cross-sectional study. BMC Infect Dis 20; doi:10.1186/s12879-020-4811-8.

Mukherjee S, ArunimaSaha AKR, Chowdhury AR, Mitra AK. 2012. Identification and characterization of a green pigment producing bacteria isolated from Bakreshwar Hot Spring, West Bengal, India. Int J Environ Sci Res 2: 126–129.

Myers SS, Patz JA. 2009. Emerging Threats to Human Health from Global Environmental Change. Annu Rev Environ Resour 34:223–252; doi:10.1146/annurev.environ.033108.102650.

Nicastri E, Petrosillo N, Martini L, Larosa M, Gesu GP, Ippolito G, et al. 2003. Prevalence of nosocomial infections in 15 Italian hospitals: first point prevalance study for the INF-NOS project. Infection 31 Suppl 2: 10–15.

Niewolak S, Opieka A. 2000. Potentially pathogenic microorganisms in water and bottom sediments in the Czarna Hancza River. Pol J Environ Stud 9: 183–194.

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834; doi:10.1101/gr.213959.116.

Obritsch MD, Fish DN, MacLaren R, Jung R. 2005. Nosocomial infections due to multidrug‐resistant Pseudomonas aeruginosa: epidemiology and treatment options. Pharmacother J Hum Pharmacol Drug Ther 25: 1353–1364.

Overbey KN, Hatcher SM, Stewart JR. 2015. Water quality and antibiotic resistance at beaches of the Galápagos Islands. Front Environ Sci 3.

Pandey A, Dhakar K, Sharma A, Priti P, Sati P, Kumar B. 2015. Thermophilic bacteria that tolerate a wide temperature and pH range colonize the Soldhar (95 °C) and Ringigad (80 °C) hot springs of Uttarakhand, India. Ann Microbiol 65:809–816; doi:10.1007/s13213-014-0921-0.

Parada AE, Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol 18:1403–1414; doi:10.1111/1462-2920.13023.

Pérez-Cataluña A, Tapiol J, Benavent C, Sarvisé C, Gómez F, Martínez B, et al. 2017. Antimicrobial susceptibility, virulence potential and sequence types associated with Arcobacter strains recovered from human faeces. J Med Microbiol 66:1736–1743; doi:10.1099/jmm.0.000638.

Pfeifer Y, Cullik A, Witte W. 2010. Resistance to cephalosporins and carbapenems in Gram-negative bacterial pathogens. Int J Med Microbiol 300: 371–379.

Piddock LJ. 2006. Multidrug-resistance efflux pumps? not just for resistance. Nat Rev Microbiol 4: 629–636.

Poole K. 2011. Pseudomonas Aeruginosa: Resistance to the Max. Front Microbiol 2.

Poole K. 2004. Resistance to β-lactam antibiotics. Cell Mol Life Sci CMLS 61: 2200–2223.

Prouzet-Mauléon V, Labadi L, Bouges N, Ménard A, Mégraud F. 2006. *Arcobacter butzleri* : Underestimated Enteropathogen. Emerg Infect Dis 12:307–309; doi:10.3201/eid1202.050570.

Purnell S, Halliday A, Newman F, Sinclair C, Ebdon J. 2020. Pathogen infection risk to recreational water users, associated with surface waters impacted by de facto and indirect potable reuse activities. Sci Total Environ 722:137799; doi:10.1016/j.scitotenv.2020.137799.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596; doi:10.1093/nar/gks1219.

R. development Core Team. 2010. R: A Language and Environment for Statistical Computing.

Rahel C, Adriyani R, Nurfitria HA. 2021. Health Risk in Hot Springs: A Literature Review. J

Health Sci Prev 5:88–99; doi:10.29080/jhsp.v5i2.524.

Reisman JS, Weinberg A, Ponte C, Kradin R. 2012. Monomicrobial Pseudomonas necrotizing fasciitis: A case of infection by two strains and a review of 37 cases in the literature. Scand J Infect Dis 44:216–221; doi:10.3109/00365548.2011.626441.

Rodrigues C, Cunha MÂ. 2017. Assessment of the microbiological quality of recreational waters: indicators and methods. Euro-Mediterr J Environ Integr 2:25; doi:10.1007/s41207-017-0035-8.

Rodríguez-Baño J, Picón E, Navarro MD, López-Cerero L, Pascual A, Group E-R. 2012. Impact of changes in CLSI and EUCAST breakpoints for susceptibility in bloodstream infections due to extended-spectrum β-lactamase-producing Escherichia coli. Clin Microbiol Infect 18: 894–900.

Rodríguez-Martínez JM, Briales A, Velasco C, Díaz de Alba P, Martínez-Martínez L, Pascual A. 2011. Discrepancies in fluoroquinolone clinical categories between the European Committee on Antimicrobial Susceptibility Testing (EUCAST) and CLSI for Escherichia coli harbouring qnr genes and mutations in gyrA and parC. J Antimicrob Chemother 66: 1405–1407.

Rupasinghe R, Amarasena S, Wickramarathna S, Biggs PJ, Chandrajith R, Wickramasinghe S. 2022. Microbial diversity and ecology of geothermal springs in the high-grade metamorphic terrain of Sri Lanka. Environ Adv 7:100166; doi:10.1016/j.envadv.2022.100166.

Sadeleer ND, Godfroid J. 2020. The Story behind COVID-19: Animal Diseases at the Crossroads of Wildlife, Livestock and Human Health. Eur J Risk Regul 11:210–227; doi:10.1017/err.2020.45.

Scales BS, Dickson RP, LiPuma JJ, Huffnagle GB. 2014. Microbiology, genomics, and clinical significance of the Pseudomonas fluorescens species complex, an unappreciated colonizer of humans. Clin Microbiol Rev 27: 927–948.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541; doi:10.1128/AEM.01541-09.

Segawa T, Takeuchi N, Rivera A, Yamada A, Yoshimura Y, Barcaza G, et al. 2013. Distribution of antibiotic resistance genes in glacier environments. Environ Microbiol Rep 5: 127–134.

Sharma N, Kumari R, Thakur M, Rai AK, Singh SP. 2022. Molecular dissemination of emerging antibiotic, biocide, and metal co-resistomes in the Himalayan hot springs. J Environ Manage 307:114569; doi:10.1016/j.jenvman.2022.114569.

Sheehan KB, Henson JM, Ferris MJ. 2005. *Legionella* Species Diversity in an Acidic Biofilm Community in Yellowstone National Park. Appl Environ Microbiol 71:507–511; doi:10.1128/AEM.71.1.507-511.2005.

Steenwyk JL, Iii TJB, Li Y, Shen X-X, Rokas A. 2020. ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. PLOS Biol 18:e3001007; doi:10.1371/journal.pbio.3001007.

Tanwar J, Das S, Fatima Z, Hameed S. 2014. Multidrug Resistance: An Emerging Crisis.

Interdiscip Perspect Infect Dis 2014:1–7; doi:10.1155/2014/541340.

Tate D, Mawer S, Newton A. 2003. Outbreak of *Pseudomonas aeruginosa* folliculitis associated with a swimming pool inflatable. Epidemiol Infect 130:187–192; doi:10.1017/S0950268802008245.

Tech T. Monitoring for Microbial Pathogens and Indicators.

Tello A, Austin B, Telfer TC. 2012. Selective Pressure of Antibiotic Pollution on Bacteria of Importance to Public Health. Environ Health Perspect 120:1100–1106; doi:10.1289/ehp.1104650.

Thomas JC, Oladeinde A, Kieran TJ, Finger JW, Bayona-Vásquez NJ, Cartee JC, et al. 2020. *Co-occurrence of antibiotic, biocide, and heavy metal resistance genes in bacteria from metal and radionuclide contaminated soils at the Savannah River Site. Microb Biotechnol 13: 1179–1200*.

Thorolfsdottir B, Marteinsson V. 2013. Microbiological Analysis in Three Diverse Natural Geothermal Bathing Pools in Iceland. Int J Environ Res Public Health 10:1085–1099; doi:10.3390/ijerph10031085.

van Tubergen A, van der Linden S. 2002. A brief history of spa therapy. Ann Rheum Dis 61:273–275; doi:10.1136/ard.61.3.273.

Visca P, Seifert H, Towner KJ. 2011. Acinetobacter infection – an emerging threat to human health. IUBMB Life 63:1048–1054; doi:10.1002/iub.534.

Vugia DJ, Richardson J, Tarro T, Vareechon C, Pannaraj PS, Traub E, et al. 2019. *Notes from the Field:* Fatal *Naegleria fowleri* Meningoencephalitis After Swimming in Hot Spring Water — California, 2018. MMWR Morb Mortal Wkly Rep 68:793–794; doi:10.15585/mmwr.mm6836a3.

Walters SP, Thebo AL, Boehm AB. 2011. Impact of urbanization and agriculture on the occurrence of bacterial pathogens and stx genes in coastal waterbodies of central California. Water Res 45:1752–1762; doi:10.1016/j.watres.2010.11.032.

Wang Y, Chen Y, Zheng X, Gui C, Wei Y. 2017. Spatio-temporal distribution of fecal indicators in three rivers of the Haihe River Basin, China. Environ Sci Pollut Res 24:9036–9047; doi:10.1007/s11356-015-5907-3.

Wang Z, Gao J, Zhao Y, Dai H, Jia J, Zhang D. 2021. Plastisphere enrich antibiotic resistance genes and potential pathogenic bacteria in sewage with pharmaceuticals. Sci Total Environ 768:144663; doi:10.1016/j.scitotenv.2020.144663.

Wickham H. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag:New York.

World Health Organization, ed. 2003. *Guidelines for safe recreational water environments*. World Health Organization:Geneva.

Wright GD. 2010. Antibiotic resistance in the environment: a link to the clinic? Curr Opin Microbiol 13:589–594; doi:10.1016/j.mib.2010.08.005.

Zhang C, Qiu S, Wang Y, Qi L, Hao R, Liu X, et al. 2013. Higher Isolation of NDM-1 Producing Acinetobacter baumannii from the Sewage of the Hospitals in Beijing. R. Schuch, ed PLoS ONE 8:e64857; doi:10.1371/journal.pone.0064857.

Zhang T, Ji Z, Li J, Yu L. 2022. Metagenomic insights into the antibiotic resistome in freshwater

and seawater from an Antarctic ice-free area. Environ Pollut 309: 119738.

Tables

Table 1. Location and main environmental features of the hot springs sampled.

Hot spring	Sampling year	Hot spring type	Altitude (m)	Temperature (°C)	pH	Salinity (g/L)
S1	2017/2019					
S2	2017/2019					
S3	2017/2019					
S4	2017/2019					
S5	2017/2019					
S6	2017					
S7	2017					
S8	2017					
S9	2017					
S10	2017					
S11	2017					
S12	2017					
S13	2017					
S14	2017					
S15	2017					
S16	2017					
S17	2017					
S18	2017					
S19	2017					
S20	2017					
S21	2017					
S22	2017					

Table 2. Results of the antibiogram on the isolated *P. aeruginosa*

Hot spring	Sample	Meropenem	Ciprofloxacin	Gentamicin	Cefepime	Meropenem	Ciprofloxacin	Gentamicin	Cefepime
S5	36								
S5	38								
S5	40								
S5	41								
S5	42								
S5	43								
S5	44								
S5	45								
S5	46								
S5	47								
S5	48								
S5	49								
S5	50								
S5	51								
S5	52								
S5	53								
S5	54								
S5	55								
S5	56								
S5	57								
S5	58								
S5	59								
S5	60								
S5	61								
S4	62								
S4	63								

S4	64
S4	65
S4	66
S4	67
S4	68
S4	69
S4	70
S4	71
S4	72
S4	73
S2	84
S2	85
S2	86
S2	87
S2	88
S2	89
S3	90
S3	91
S3	92
S3	93
S3	94
S3	95

R=Resistant, S=Sensitive, I=Intermediate. EUCAST Zone diameter breakpoint: Meropenem S ≥ 24, R < 18. Ciprofloxacin S ≥26, R <26. Gentamicin S ≥15, R <15. Cefepime S ≥ 21, R < 21. * The Meropenem intermediate value has been arbitrarily defined because it is absent in the EUCAST breakpoint table. CLSI Zone diameter breakpoint: Meropenem S ≥19, I 16-18, R ≤15. Ciprofloxacin S ≥21, I 16-20, R ≤15. Gentamicin S ≥15, I 13-14, R ≤12. Cefepime S ≥18, I 15-17, R ≤14.

Figures

Figure 1. Temperature and pH of the sampled hot springs. Recreational hot springs placed within spa and resort are colored in blue while natural hot springs are in red. Spa/resort hot springs not linked with recreational waters usage are indicated in green.

Figure 2. Relative abundance of putative pathogens in the 16S rRNA libraries of hot spring fluid samples. Recreational and natural hot springs are indicated by the red and the blue bars, respectively.

Figure 3. A) Colony forming units of *Pseudomonas aeruginosa* (as identified by 16S rRNA sequencing) from the hot spring waters on a plate of PCN agar selective media. B) Example of antimicrobial susceptibility test on *Pseudomonas* spp. Strains isolated from a recreational hot spring showing multidrug resistance.

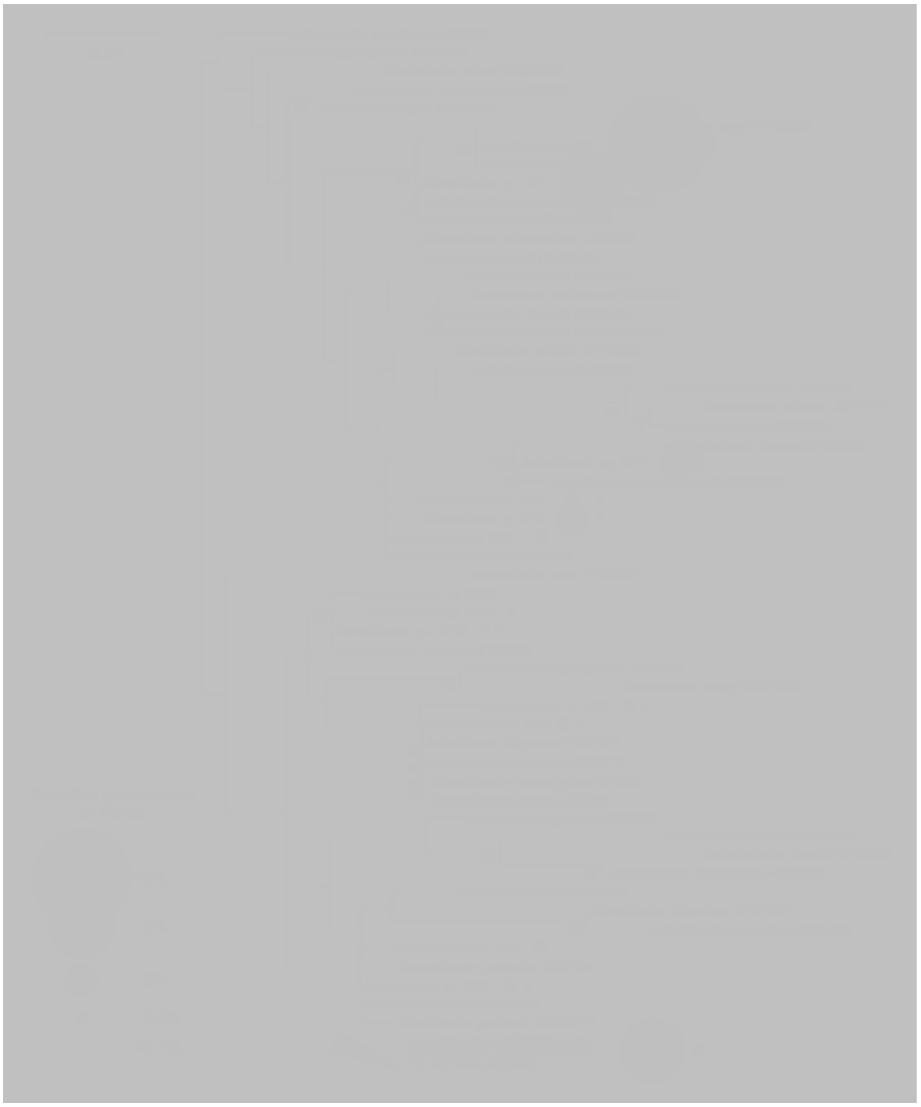


Figure 4. Neighbor-joining phylogenetic tree showing the relative position of the ASVs and the MAGs' 16S rRNA genes related to the genus *Acinetobacter* relative to known *Acinetobacter* species. The reference sequences of known human pathogens are displayed in bold and sequences from this study are followed by CR#. The average relative 16S rRNA gene amplicon abundance in the fluid 16S rRNA libraries is reported for the natural (blue) and recreational (red) hot springs. The *A. junii* cluster also contains the sequences of *A. gyllenbergii* AJ293694, *A. proteolyticus* KT997475, *Acinetobacter* sp. X81659, *A. modestus* KT997474, *A. viviani* KT997477, *A. courvalinii* KT997472. Tree based on 1,000 replicated bootstrap.

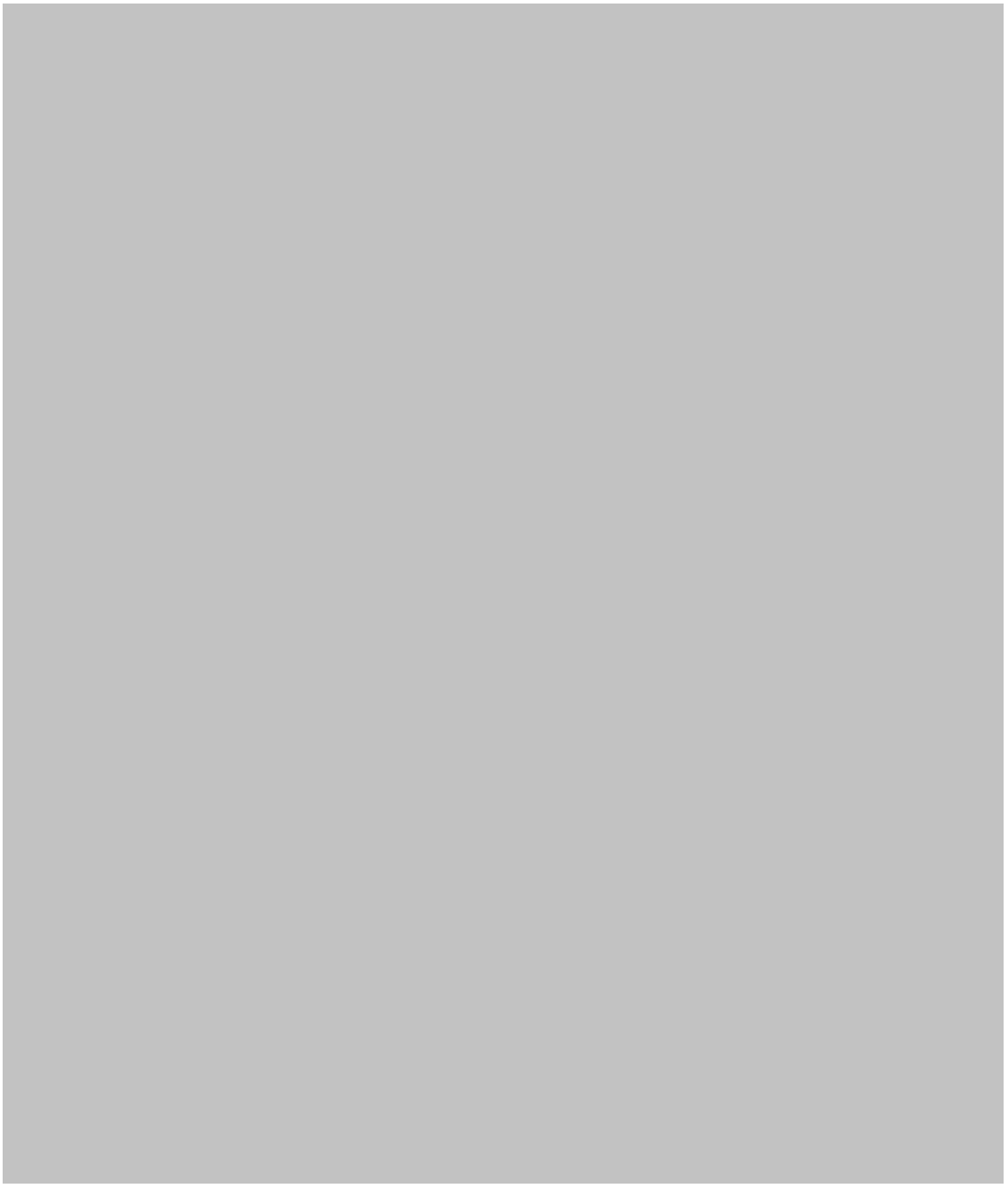


Figure 5. Neighbor-joining phylogenetic tree showing the relative position of the ASVs and the 16S rRNA genes of isolated type strains related to the genus *Pseudomonas* relative to known *Pseudomonas* species. The reference sequences of known human pathogens are displayed in bold. Average relative abundances in the fluid 16S rRNA libraries are reported for the natural (blue) and recreational (red) hot springs. Tree based on 1,000 replicated bootstraps.

Figure 6. Antibiogram for *P.aeruginosa* isolates according to EUCAST. Antibiotic resistance profile was developed by testing *P. aeruginosa* against Meropenem (M), Gentamicin (G), Ciprofloxacin (Cp), and Cefepime (Cf). Strains of *P. aeruginosa* with more than two antibiotic resistances were indicated as MDR (Multi-Drugs Resistance).

Figure 7. Absolute and relative abundances of Antimicrobial Resistance Genes (ARGs) found in Costa Rica 2017 sampled sites. Recreational and natural hot springs are indicated by the red and the blue bars, respectively.

Supplementary Information

Supplementary Tables 1. List of the putative pathogenic genera used to screen the 16S rRNA libraries.

Genera	
Acinetobacter	Klebsiella
Abiotrophia	Kocuria
Achromobacter	Lawsonia
Actinobacillus	Legionella
Arcanobacterium	Leptospira
Arcobacter	Listeria
Bacillus	Merkel_cell
Bartonella	Micrococcus
Bordetella	Morganella
Borrelia	Mycobacterium
Bordetella	Mycoplasma
Brucella	Neisseria
Burkholderia	Nocardia
Campylobacter	Pasteurella
Capnocytophaga	Photobacterium
Chlamydia	Plesiomonas
Clostridium	Propionibacterium
Comamonas	Proteus
Corynebacterium	Providencia
Coxiella	Pseudomonas
Cronobacter	Rhodococcus
Deinococcus	Rickettsiae
Dermatophilus	Roseomonas
Ehrlichia	Rothia
Enterococcus	Salmonella
Erysipelothrix	Serratia
Escherichia	Shewanella
Escherichia/Shigella	Shigella
Flavobacterium	Sphaerophorus
Francisella	Staphylococcus
Gardnerella	Stenotrophomonas
Granulicatella	Streptococcus
Haemophilus	Treponema
Hafnia	Vibrio
Halomonas	Yersinia
Helicobacter	

Supplementary Tables 2. Antibiotic Resistance Genes' (ARGs) absolute abundance found in Costa Rica hot springs assembled contigs.

ARG	S2	S3	S4	S5	S6	S8	S13	S16	S18
aba									
abe									
ADC									
ade									
amp									
amv									
OXA									
cpx									
mex									
opm									
PDC									
pmp									
sme									
sul									
mux									
nov									
ceo									
rph									
amr									

Supplementary figure 1. Relative abundance of putative pathogens in the 16S rRNA libraries of hot spring fluids (F) and sediments (S) samples. Recreational and natural hot springs are indicated by the red and the blue bars, respectively.

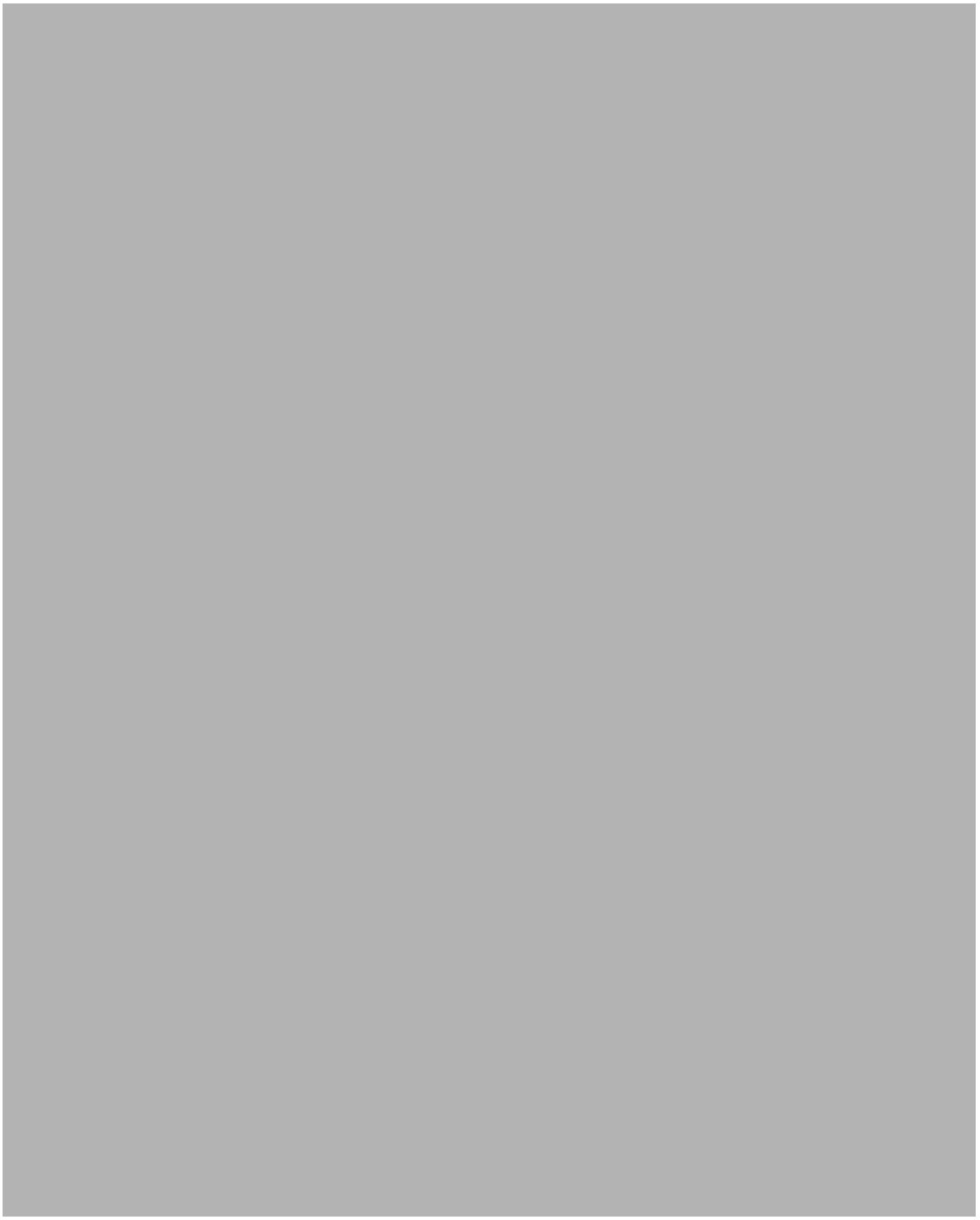
Supplementary figure 2. Relative abundance of putative pathogens in the shotgun metagenomic reads of hot spring fluids.

Supplementary figure 3. Antibiogram for *P.aeruginosa* isolates according to CLSI. Antibiotic resistance profile was developed by testing *P.aeruginosa* against Meropenem (M), Gentamicin (G), Ciprofloxacin (Cp), Cefepime (Cf).

Supplementary figure 4. Antibiogram for *P.aeruginosa* isolates. Antibiotic resistance profile was developed by testing *P.aeruginosa* against generic antibiotic Erythromycin (E), Kanamycin (K), Ampicillin (A), Cefsulodin (C).

Supplementary figure 5. EUCAST (dark grey) and CLSI (light grey) breakpoint discrepancies on antibiotic susceptibility profiles based on the 48 *P. aeruginosa* type strains isolated from four different recreational hot springs and *P. aeruginosa* PAO1 as control.

Cover Image



Chapter 6

Conclusions

In the course of this doctorate project I tried to give my contribution to the study of the interactions between biosphere and geosphere, with a first focus on prokaryotic overprinting of volatiles that are cycled at the convergent margins, and a second focus that link the occurrence of multi-drug resistant pathogens as an effect of anthropic activities around the hot springs. The idea is to provide new informations on the microbial composition and the fundamental functions on numerous deeply sourced seeps and secondary geothermal emissions that are spread along the Central America convergent margin, by the application of a mix of comparative genomic and geochemistry techniques combined with molecular assays. I started investigating the distribution of methane-cycling microorganisms associated with the isotopic signature of different methane seeps spread across different geological settings (Chapter 3). Since methane plays a significant role in regulating Earth's climate and it is also important for the understanding of the balance of reducing equivalents in Earth's interior, investigating the sources and sinks of methane at plate boundaries is crucial to constrain carbon cycling and Earth redox balance. Our results show that geological settings is one of the main factors in controlling the last steps of the carbon path through the Earth's crust before the release within the atmosphere, and that geological and biologically mediated secondary processes can in part overprint the deeper signal and influence the fate of methane at convergent margins. However, further investigations are necessary to assess the quantitative contribution of the different processes to the convergent margins methane flux. As a second step of my doctorate, I moved the focus on the role of microbial communities in controlling the carbon cycle in a meromictic lake (Chapter 4). These particular environments host incredible amounts of dissolved gases (e.g. CO_2 and CH_4) which can be released in the atmosphere after destabilization events. However, what leads the over accumulation of dissolved gases to critical conditions is still not clear. In this context, our results show a cryptic microbial loop where biological necromass, sinked in the deepest layers, works as a trigger for heterotrophic pathways, bringing to an overproduction of CO_2 over time. Finally, in the last chapter we propose a study (Chapter 5) about the role of the hot springs as potential reservoirs for antibiotic resistant pathogens as an effect of anthropic pressure. Results from sequence-based analysis showed a large difference between recreational and natural hot springs in terms of presence of putative pathogens and antibiotic resistance genes. Culture-dependent analyses were helpful in characterizing several pathogenic strains, showing a high diversity at the strain level and multiple resistances at antibiotics. We propose that anthropic pressure plays as the main driver for a differential occurrence of putative multidrug-resistant pathogens in recreational hot springs.

Chapter 7

Future perspectives

Convergent margins are regions where volatile elements like carbon, water, hydrogen, and sulfur are exchanged between the Earth's surface and interior. At these locations, carbon is recycled from deep reservoirs in two main forms: oxidized carbon in the form of carbon dioxide, and reduced carbon in the form of methane. While CO_2 is quantitatively more important and its volcanic fluxes have been better constrained, CH_4 represents the most reduced possible form of carbon which greatly contributes to greenhouse effects and Earth's climate stability. The quality and speciation of carbon subducted and recycled to the surface influence the redox potential of the Earth and have modified the atmospheric composition and climate over geological timescales (Hayes and Waldbauer, 2006; Galvez and Pubellier, 2019). Biological processes such as oxygenic photosynthesis, the burial and catalysis of organic carbon, actively contribute to the production and maintenance of the current Earth's breathable atmosphere (Bergin et al., 2015b). Microbes populating the subsurface can oxidize and reduce inorganic compounds such as methane and sulfur, which can have a significant impact on the global carbon and sulfur cycles. Sulfur can transition from different oxidation states into different compounds ranging from 6+ to 2- with strong redox ability in geological and biological reservoirs e.g., S^0, S^{2-} in H_2S, S^{4+} in SO_2 and S^{6+} in SO_4^{2-}. These properties made sulfur one of the fundamental players in understanding the co-evolution of Earth systems. On the other hand, methane can be oxidized by methanotrophic communities in presence of oxygen or anaerobically using different electron acceptors like sulfate, iron, manganese, nitrite and nitrate (Guerrero-Cruz et al., 2021). Currently, large spatial-scale correlation analyses between geological and biological processes as well as about constraints clarifying what feedbacks are present between deep Earth and terrestrial surface in controlling the methane degassing across convergent margins have not been performed yet. Therefore, our future perspectives include the investigation and correlation of the subducting zones microbial diversity and the large variability in isotopic signatures observed in geothermal systems, in order to give a clear picture of the origin of methane and the secondary geological and biological processes altering its quantity and speciation at convergent margins.

References

Bergin, E. A., Blake, G. A., Ciesla, F., Hirschmann, M. M., and Li, J. (2015). Tracing the ingredients for a habitable earth from interstellar space through planet formation. Proceedings of the National Academy of Sciences 112, 8965–8970. doi: 10.1073/pnas.1500954112.

Galvez, M. E., and Pubellier, M. (2019). "How do subduction zones regulate the carbon cycle?," in Deep carbon: Past to present (Cambridge University Press), 276–312.

Guerrero-Cruz, S., Vaksmaa, A., Horn, M. A., Niemann, H., Pijuan, M., and Ho, A. (2021). Methanotrophs: Discoveries, Environmental Relevance, and a Perspective on Current and Future Applications. Frontiers in Microbiology 12. Available at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.678057 [Accessed January 31, 2023].

Hayes, J. M., and Waldbauer, J. R. (2006). The carbon cycle and associated redox processes through time. Philosophical Transactions of the Royal Society B: Biological Sciences 361, 931–950. doi: 10.1098/rstb.2006.1840.

Publications

Peer reviewed

Rogers TJ, Buongiorno J, Jessen GL, Schrenk MO, Fordyce JA, de Moor JM, Ramírez CJ, Barry PH, Yücel M,Selci M, Cordone A, Giovannelli D, Lloyd KG. 2022. Chemolithoautotroph distributions across the subsurface of a convergent margin. ISME J, 1–11. doi: 10.1038/s41396-022-01331-7

Barry P, De Moor JM, A Chiodi, F Aguilera, MR Hudak, DV Bekaert, S J Turner, J Curtice, AM Seltzer, G L Jessen, E Osses, JM Blamey, M Amenabar, M Selci, M Cascone, A Bastianoni, M Nakagawa, RFilipovich, E Bustos, M Schrenk, J Buongiorno, CJ Ramírez, TJ Rogers, KG Lloyd and D Giovannelli. 2022. The helium and carbon isotope characteristics of the Andean Convergent Margin. Frontiers in Earth Sciences, 10. doi: 10.3389/feart.2022.897267

Cordone A, D'Errico G, Magliulo M, Bolinesi F, Selci M, Basili M, de Marco R, Saggiomo M, Rivaro P, Giovannelli D* and Mangoni O*. 2022. Bacterioplankton Diversity and Distribution in Relation to Phytoplankton Community Structure in the Ross Sea Surface Waters. Front. Microbiol. 13:722900. doi:10.3389/fmicb.2022.722900

Petrillo C, Castaldi S, Lanzilli M, Selci M, Cordone A, Giovannelli D, Isticato R. 2021. Genomic and physiological characterization of Bacilli isolated from salt-pans with Plant Growth Promoting features. Frontiers in Microbiology, 12:2517. doi: 10.3389/fmicb.2021.715678

Book chapter:

Cordone, A., Coppola, A., Severino, A., Correggia, M., Selci, M., Cascone, A., Vetriani C, Giovannelli D*. (2021). From sequences to enzymes: comparative genomics to study evolutionary conserved protein functions in marine microbes. Marine Genomics. Methods in Molecular Biology, vol 2498. Humana, New York, NY. https://doi.org/10.1007/978-1-0716-2313-8_5

Under revision:

Barosa B., Ferrillo A., Selci M., Giardina M., Bastianoni A., Correggia M., Di Iorio L., Bernardi G., Cascone M., Capuozzo R., Intoccia M., Price R. E., Vetriani C., Cordone A., and Giovannelli D. Mapping the microbial diversity associated with different geochemical regimes in the shallow-water hydrothermal vents of the Aeolian Archipelago, Italy. Frontiers in Microbiology.

Raegan P., Rogers TJ, Fullerton K. M., Selci M., Cascone M., Stokes M. H., Steen A. D., de Moor J. M., Chiodi A., Stefánsson A., Halldórsson S. A., Ramirez C. J., Jessen G. L., Barry P. H., Cordone A., Giovannelli D., and Lloyd K. G. Complex organic matter degradation by secondary consumers in chemolithoautotrophy-based subsurface geothermal ecosystems. PLOS ONE.

Selci, M., Cordone, A., Barosa, B., Bastianoni, A., Bastoni, D., Bolinesi, F., Capuozzo, R., Cascone, M., Corso, D., di Iorio, L. and Misic, C., 2023. Surface bacterioplankton community structure crossing the Antarctic Circumpolar Current Fronts. MDPI Microorganisms.

Preprint:

Selci, M., Cordone, A., Barosa, B., Bastianoni, A., Bastoni, D., Bolinesi, F., Capuozzo, R., Cascone, M., Corso, D., di Iorio, L. and Misic, C., 2023. Surface bacterioplankton community structure crossing the Antarctic Circumpolar Current Fronts. EarthArXiv https://doi.org/10.31223/X58Q04.