Organic-rich shales as archive and resource for microbial extremophiles

Dissertation

for the award of the degree

"Doctor of Philosophy" Ph.D. Division of Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

within the doctoral program Biology of the Georg-August University School of Science (GAUSS)

submitted by

Sania Arif

from Gujranwala, Pakistan

Göttingen, 2021

'All living beings, not just animals, but plants and microorganisms perceive. To survive, an organic being must perceive - it must seek or at least recognize, food and avoid environmental danger'.

Lynn Margulis (1938 – 2011)

Table of Contents

Chapter 4

Chapter 5

Appendix

Abstract

Organic-rich shales of sedimentary basins are the most abundant carbon-rich rocks throughout the world, formed as a consequence of different geological events. The repetitive sedimentation of the mudstones, clay particles, biological organic matter, microbial decomposition of the organic matter, and evaporation formed organic-rich shales of diverse origins in different geological eras. These organic-rich shales are an important source of gas, oil, and occasionally precious metals such as copper, gold, silver, uranium. Moreover, these shales may also provide environmental microbial communities with the sole source of carbon from their organic component.

Unlike the air-exposed black shales in subterrestrial environments, deep-sea subsurface black shales are colonized by anaerobic halotolerant bacteria and methanogens. The subterrestrial organic-rich shales enrich completely different microbial communities belonging mainly to aerobic/microaerophilic chemoorganotroph and chemolithotroph categories. The subterrestrial black shales provide essential organics characterized as kerogen and bitumen, accessible for microbial growth through fractures and microfissures. DNA analysis performed on the various black shales of different origins confirmed the microbial colonization by *Firmicutes, Actinobacteria*, and *Proteobacteria* communities. The inhabitant microorganisms are supposed to devour the hydrocarbons as a sole carbon source. Their contribution to the global carbon cycle, remediation of hydrocarbons, and facilitation of oil recovery from shales are being validated.

In this study, the inhabitant microbial diversity of the subterrestrial organic-rich environments was metagenomically investigated to identify if the colonized microbiota is specialized to degrade hydrocarbons and could cope with different environmental stress factors. It was further established by analysing microbial colonization of organic-rich shales with or without high heavy metal contents. In the first chapter, the organic-

rich shale samples with low metal content from Monte San Giorgio were analysed. The microbial structure inhabiting the organic component of the shale was distinguished from that of the rocky shale surface and limestone rocks. Furthermore, the enrichment effect of the media of choice on the native shale and limestone microbiome gave insights into the metabolic capabilities of the oligotrophic versus fastidious communities. The organics of the shale are not easily degradable and accessible carbon sources which facilitate the specialized oligotrophic microorganisms to colonize slowly. The functional profiling of the inhabitant microbiota clearly shows the ability to degrade complex aromatic hydrocarbon compounds.

The second chapter explored the native microbial communities in the metalliferous organic-rich shale of Marsberg Kilianstollen copper mine. The low temperature, toxic copper-enriched leachate, and low availability of organics adapted heavy metal resistant consortia that could also degrade the complex hydrocarbons. *Ktedonobacteria* biofilm formed near the acid mine drainage in Kilian copper mine was further investigated to reveal functional pathways from the metagenome-assembled genomes (MAGs).

Due to the common observation of *Ktedonobacteria* high abundance in the cold Marberg Kilianstollen, Germany and other extreme hot springs environment Sasso Pisano, Italy, the diversity of newly discovered *Ktedonobacteria* in these ecosystems was further elucidated. The inhabiting *Ktedonobacteria* strains were found to be phylogenetically distinct at Marsberg and Sasso Pisano. However, it is hypothesized that they show functional conservation for the carbon monoxide oxidation (Cox operons), heavy metal resistance, aromatic compounds metabolism, and sporulation mechanisms.

1. Introduction

1.1 Organic-rich Shales

Shale rocks are formed by the sedimentation of at least 50% fine-grained particle spherical diameter size less than 0.062-mm. The organic-rich shale contains organic content ranging from >0.5 wt organic carbon (Huyck, 1990). A general definition of organic-rich shales is that they are fine-grained laminated sedimentary rocks, mainly comprised of a mixture of quartz silt, clay minerals, and organic particles of biological origin. Nevertheless, one definition can't fit all organic-rich shales as the complex interplay of various geologic processes and variables determine the variable constituents and types (Schieber, 2003). Organic-rich shales are also called black or carboniferous shales and form a variety of sediments and sedimentary rocks. They may be related to high organic activities and biogeochemical cycling of carbon, oxygen, phosphorous, nitrogen, and sulphur; heavy metals such as molybdenum, iron, uranium, and vanadium may get included (Huyck, 1990).

Organic-rich shales could be categorized into different types based upon the magnitude of organic matter maturation. The organic-rich shales deposits from the Cambrian period are already matured for extraction of oil and natural gas and are described as oil shales. Oil shales are heated to recover considerable crude oil and combustible gas through the destructive distillation of the solvent-insoluble organic component (Dyni, 2006). Although the term "oil shale" incongruently hints that shale contains oil, it is a transformed form of ancient biomass/organic matter into pre-oil solid called kerogen that must be processed before usable crude oil could be synthesized (Foght, 2009). In contrast, bituminous shale is another form of oil shale that contains bitumen, which is solvent-soluble asphaltene-like colloidal hydrocarbons in patches or fissures. Bitumen is composed of organic solvent-soluble fractions of easily biodegradable oil-like

saturated and aromatic hydrocarbons that may facilitate the colonization of the bituminous shales by the subsurface microbial communities (Matlakowska & Sklodowska, 2011; Meslé et al., 2015). Note, the bitumen component of oil shale or bituminous shale has a distinctive biological origin than bitumen formed from the substantial biodegradation of the crude oil (McGenity, 2019).

1.2 Origin of Organic-rich Shales

Mostly black shales are marine sediments (Potter et al., 2012). In addition, lacustrine successions could also form prominent black shale deposits (Bohacs et al., 2000). The presence of the finely disseminated pyrite along the black organic matter in these shales led the geologists to conclude the requirement of the anoxic reducing bottom waters for the formation of ancient black shales (Schieber, 2003).

Sedimentary rocks are formed by deposition of mainly mud mineral particles (silicates, carbonates) at the bottom of the water bodies (lakes, oceans) that are compacted and solidified over a geologic time scale (Nijenhuis et al., 1999). During the formation of sedimentary rocks, organic debris of biological origin (mostly decayed microbial biomass of marine/terrestrial origin) may get included. Though most of the biomass gets degraded in the "microbial loop" during sedimentation in anaerobic processes, a substantial quantity may eventually get buried and compacted when a high surplus of organic carbon is produced (e.g., during algal blooms) (He et al., 2019). Fundamentally, organic-rich black shales are formed from marine sediment layers (sapropels) of >1 cm thickness and >2% C_{org} by weight, under variable oceanic conditions (Kidd et al., 1978). During the formation of the sediments, various microbial processes occur. If sulphate (from seawater) is available, sulphate reducing bacteria dissimilatory oxidize organic biomass with sulphate as an oxidative agent according to the following equation (Jørgensen et al., 2019; Rabus et al., 2006).

$$SO_4^{2-} + 2CH_2O \text{ (“Short” formula for organic biomass)} \longrightarrow H_2S + 2HCO_3^{-} \qquad (1)$$
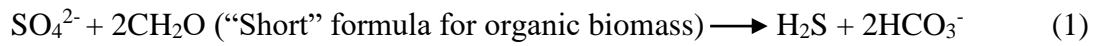
$$H_2S + Me2^{+} \longrightarrow MeS + 2 H^{+} \qquad (2)$$
Me-Metal

The very reactive H_2S precipitates ions of transition metals (like Fe^{2+}) and insoluble metal sulphides (FeS_2) are formed (Berg et al., 2020). Eventually, metal-rich ores, along with organic matter, are deposited in sediments according to equation (2) (Biswas et al., 2009).

Generally, rapid transgressions (sea-level rise and flooding of land) could also result in high input of organic material and fertilizing mineral compounds like P and N from land to sea. Several cycles of refilling, evaporation (and formation of the respective sediments) occurred until the basin is filled or the sea disappeared. It happened multiple times in earth history, for instance, during the Middle-Upper Devonian to Early Carboniferous times in the back-arc basin of the Rhenohercynian zone (360-300 Mio years ago) (von Raumer et al., 2017). Currently, the carbonaceous mud accumulation could also be observed in the Black Sea (Kaiser et al., 2017).

The closure and sub-duction of the Rheic Ocean by the continental intraplate geological events along with consequential sedimentation and volcanism during the Late Silurian - Early Devonian and Early Carboniferous included many biozones and faunal assemblages (Nance et al., 2010) and constituted present Rhenish Massif (Rheinisches Schiefergebirge) in the Rhenohercynian zone (Königshof et al., 2016; von Raumer et al., 2017). Several repetitive cycles of organic matter sedimentation and biomass microbial anoxic degradation led to the formation of several intercalated basic volcanic rocks with metalliferous Lower Carboniferous black shale series in the Rhenish Massif (Urban et al., 1995). These metalliferous Lower Carboniferous black shales have been mined extensively for their metal deposits, for instance, the copper-enriched black shales of Marsberg in the north-eastern Rhenish Massif region (Germany). The present

exposed terrestrial Lower Carboniferous black shale series in the Marsberg copper mines offers a unique opportunity to analyse the native microbial consortium in relation to their geochemical role in heavy metal tolerance and complex hydrocarbon catabolism.

1.3 Organic-rich Shales as a Substrate for Microbes

More than 80% of fossilized organic material in the sedimentary rocks is characterized as relatively inert kerogen. The rest is the lower molecular weight, solvent-extractible fraction of colloidal structure, bitumen (Behar et al., 2008; Hedges & Oades, 1997; Tissot & Welte, 1984).

Kerogen is a highly heterogeneous solid by-product of the residual detritus organic matter whose composition and structure vary considerably within a source over a small distance. It is found in coal and mineral oil deposits and can be classified as types I–IV. The dispersed organic carbon of kerogen is not a polymer but contains a complex agglomerate of randomly arranged macromolecules (may contain metal ions coordinated to charged groups) (Philp, 2003). With the increasing maturity of heterogeneous macromolecular structures of kerogen, the apolarity, molecular weight, fused ring size, carbon content, cross-linking, density, and glass-to-rubber transition temperature tend to increase (Xing & Pignatello, 2005).

The other relatively less abundant bitumen is composed of saturates, aromatics, resins, and asphaltenes (Atlas, 1981; Vandenbroucke & Largeau, 2007). The saturates are generally considered to be easily biodegradable, followed by monoaromatics (Jones et al., 2008; Leahy & Colwell, 1990), long-chain alkylmonoaromatics, and polycyclic aromatic hydrocarbons whereas recalcitrant and heavy compounds such as resins and asphaltenes are difficult to biodegrade (Gray et al., 2010; Walker et al., 1976).

In contrast, marine microbes in a polluted harbor were found to degrade complex aromatic hydrocarbons before the saturated compounds were attacked (Leahy & Colwell, 1990). A similar observation was made in the basins of San Juan and Powder River by the coal-bed methane microbial consortium (Formolo et al., 2008). It could be because the aromatic compounds are considered to be more reactive than resilient apolar sigma bonds of saturates and hydrocarbons biodegradability is also influenced by the type of microbes and abiotic factors (Gray et al., 2010). Strict anaerobes are reported to metabolize resins and asphaltenes as the sole electron donors and carbon sources (Magot et al., 2000). It affects the asphaltenes composition, structure, and spatial organization by the microbially produced oxidizable and fermentable metabolites, namely organic and fatty acids and methanogenic substrates (Liao et al., 2009; Wawrik et al., 2012).

Although kerogen is considered relatively inert and more difficult to degrade than bitumen due to its three-dimensional complex and compact structure, evidence of *in situ* aerobic microbial oxidation of kerogen at the exposed surface of Kentucky ''New Albany'' black shale was reported. The microbes could metabolize kerogen as a sole carbon source and assimilate carbon from it into cellular membrane lipids synthesis (S. Petsch et al., 2005). Meslé et al. (2013) observed microbial enrichment and biodegradation of the kerogen type II rich black shales from the Lower Jurassic era of the eastern Paris basin. Based on the biostimulation technique, the anaerobic microcosms facilitated the mineralization of the kerogen into methane by a complex interplay of syntrophic firmicutes and methanogens (Meslé et al., 2013). Biostimulation is a culture-based approach that enhances specific microbial populations and their metabolic yield above detection levels by providing specified nutrients. In another follow-up study, the anaerobic microbial consortium produced methane from the same black shale rocks of different maturities. The solvent-soluble fractions of easily

degradable maltenes and asphaltenes (bitumen) were depleted to enrich kerogen substrate in the shale. The microbial growth on kerogen-enriched shale was compared with that of unmodified organic-rich shale. Both bitumen and kerogen substrates were effectively converted to biogenic methane by the microbial consortium (Meslé et al., 2015).

Huda et al. (2020) isolated indigenous bacterial strains *Bacillus flexus* and *Enterobacter cloacae* from the Campanian bulk oil shale, Egypt. Theses microbial strains performed high biodegradation of different polycyclic aromatics such as iso-paraffins under aerobic conditions in the oil shale containing microcosms (El-Sheshtawy et al., 2020). Hence, it could be concluded that the organic-rich shales provide bitumen and kerogen, consisting of diverse organic compounds, among them hopanes, hopenes, and methylhopanoids that serve as microbial substrates to facilitate the microbial growth irrespective of aerobic and anaerobic conditions (McEvoy & Giger, 1986).

1.4 Microbial Degradation of Organic-rich Shale

Despite their low porosity and permeability, organic-rich shales can facilitate microbial growth as the sole carbon source (S. Petsch et al., 2005). These microbial communities eventually contribute to the black shale weathering process (Matlakowska et al., 2010; Matlakowska & Sklodowska, 2009). The diversity of black shale inhabiting microbial communities differs considerably in different geological settings; therefore, they have been broadly classified as anaerobic and aerobic environments to conclude shared microbial consortium and degradation pathways.

1.4.1 Anaerobic microbial degradation

In the anaerobic deep subsurface organic-rich shale reservoirs, the anaerobic breakdown of complex hydrocarbon to methane occurs irrespective of the terrestrial and marine origin of the organic-rich shales. For instance, methane of microbial origin has been

confirmed in various studies at the lacustrine sedimentary basins (Martini et al., 2003; Römer et al., 2020; Vuillemin et al., 2018), deep marine water black shales (Meyers et al., 2004), and exposed deep terrestrial black shales (S. T. Petsch et al., 2005) as a consequence of tectonic perturbations and hydraulic fracturing (Ulrich et al., 2018). The major microbial processes related to anaerobic black shale decomposition are characterized primarily as fermentation, Fe/Mn reduction, sulfate reduction, and methanogenesis (Li et al., 2014; Takai et al., 2003). The associated methanogens diversity is attributed to three families: Methanobacteriaceae, Methanomicrobiaceae, and Methanosarcinaceae (Penner et al., 2010; Zengler et al., 1999). The anaerobic bacterial diversity was found to be dominated by *Firmicutes* (Meslé et al., 2013)*, Proteobacteria* (Struchtemeyer et al., 2011), *Actinobacteria*, (Li et al., 2014), and *Bacteroidetes* (Mouser et al., 2016) depending on the type of organic-rich shale and other abiotic factors. The anaerobic degradation of the fossilized organic matter can be expanded to diverse substrates such as coal formation waters (Jones et al., 2010), petroleum reservoirs production waters (Grabowski et al., 2005), and the sedimentary organic carbon of shale-sandstone resources (Takai et al., 2003). Although various abundant microbial populations are specialized to metabolize coal, oil, and organics-rich shale, the underlying functional pathways are conserved for the anaerobic degradation of highly complex, recalcitrant organic polymers into methane (Meslé et al., 2015).

1.4.2 Aerobic microbial degradation

Usually, black shale reservoirs are deep buried subsurface anaerobic environments. However, these organic-rich strata may get exposed to lithosphere weathering conditions as a consequence of tectonic plate movements, changes in the sea level, mining, and other anthropogenic activities. Under the influence of exogenous abiotic

factors such as oxidation of organic matter or pyritic sulfur by the atmospheric oxygen and low pH, the high content of organic matter from the black shale is prone to degradation (Peucker-Ehrenbrink & Hannigan, 2000). In comparsion to anaerobic mineralization of the organic-rich shale to methane, the complex interplay of microbial communities involved in aerobic biodegradation is not extensively studied. Li et al. (2014) compared the innate microbial profile of the unexposed bottom zone to the surface oxidized weathered black shales of Chengkou County, Southwest China (Li et al., 2014). In all zones, 33 different taxa groups were detected in which *Proteobacteria*, *Actinobacteria*, and *Firmicutes* related bacteria *Bacillus, Thiobacillus*, *Sulfobacillus, Serratia, Streptococcus, Yonghaparkia, Ferrimicrobium,* and *Nocardioides* were abundant. The surface oxidized top zone of black shale contained more facultative and aerobic bacteria such as abundant strictly aerobic bacteria *Nocardioides* and *Segetibacter*. The middle region acidic zone was inhabited by many acid-tolerant or acidophilic microbes, such as *Anoxybacillus*, *Ferrimicrobium*, *Bacillus*, *Sulfobacillus*, and *Thiobacillus* involved in pyrite oxidation. Some of the microaerophiles/aerobes were identified in the anaerobic bottom zones due to oxygen diffusion through the micro-fissures. The black shale weathering processes were attributed to the organic matter degradation by heterotrophs and pyrite oxidation by chemolithotrophs, involved in S and Fe biogeochemical cycling (Li et al., 2014). In comparison to Li et al. 2014, Brazilian Permian Irati Formation shale by-products found 15 common bacterial phyla. *Proteobacteria*, *Firmicutes*, and *Actinobacteria* were the most abundant in both shales in China and Brazil (Goes et al., 2018). The subterrestrial copper-rich black shales from the Lubin copper mine were dominated by *Proteobacteria* and *Actinobacteria,* represented by several bacterial genera like *Limnobacter*, *Pseudomonas*, *Thiobacillus*, *Yonghaparkia*, *Sulfuricaulis,* and *Bradyrhizobium*. The chemoorganotrophic bacteria, archaea (*Nitrososphaera*), and fungi (*Aspergillus*) metabolized aliphatic and aromatic

hydrocarbons. On the other hand, the chemolithotrophic bacteria primarily oxidized reduced sulfur compounds and contributed to the dehydrogenation and oxidation of kerogen and sulfide minerals (Włodarczyk et al., 2018).

Some cultivable bacteria from Late Devonian New Albany Shale black shales were *gamma-Proteobacteria* (*Pseudomonas*, *Acinetobacter* sp.), *Firmicutes* (*Clostridium* sp.), and *beta-Proteobacteria* (*Dechloromonas*, *Comamonas* sp.) (S. Petsch et al., 2005). *Ralstonia picketii* was cultured from the Lower Autunian Autun black shale (France) (Berlendis et al., 2014). Five strains isolated from Kupferschiefer black shale of Lubin copper mine were identified as *gamma-Proteobacteria* (*Pseudomonas*, *Acinetobacter* spp.), *Firmicutes* (*Bacillus* sp.), and two *Actinobacteria* (*Microbacterium* sp.) (Matlakowska & Sklodowska, 2009).

Moreover, the yeast *Rhodotorula mucilaginosa* has also been isolated from Lubin copper mine subterrestrial black shale (Rajpert et al., 2013). There are not enough studies that analyze the wholistic bacterial biodiversity of the subterrestrial desiccated black shale and characterize their functional roles.

1.5 Microbial Pathways Involved

Maltakowska et al. (2009) examined the mechanism involved in the successful microbial colonization, biofilm development, and metabolism of fossilized organics in copper-bearing black shale by testing indigenous isolates (Matlakowska & Sklodowska, 2009). The isolates were able to grow in the mineral medium supplemented with copper-bearing black shale as a sole carbon and energy source. The bacterial community also accumulated phosphorus inside the cells from the organic matter and produced chelating agents, siderophores (Matlakowska & Sklodowska, 2011).

1.5.1 Heavy metal resistance

The isolates from copper-bearing black shale showed high levels of resistance to inorganic arsenic [As(III) and As(V)] up to 500 mmol l^{-1}. Multi-resistance against 2-7 mmol l^{-1} concentration of Co, Cu, Ni, and Zn was also observed. The biodegradation of the pyrrole ring of the metalloporphyrins present in kerogen was confirmed by the complete or partial degradation of the haem ring (Matlakowska & Sklodowska, 2009). From metalloorganic compounds and metalloporphyrins such as Cu(II) meso-tetra(4-carboxyphenyl)porphine and Co(III) protoporphyrin-IX chloride microbes were able to retrieve organic carbon as the sole energy source while metals (Cu, Co, Ni, V, As) were released in the aqueous phase of bacterial culture. Up to 74% of copper was mobilized from black shale substrate to the aqueous phase by bacterial activity that indicates the potential of precious metals biorecovery from tailings that contain up to 16% black shale (Matlakowska et al., 2010).

1.5.2 Aliphatic and Aromatic Compounds Metabolism

The dioxygenase activity was important to degrade phenanthrene, an aromatic hydrocarbon present in black shale ore (Sklodowska et al., 2005). All isolated strains possessed esterase (such as naphthol phosphohydrolase) and leucine/valine arylamidase activities that were also important to degrade the complex heavy chain hydrocarbons, organic acids, and esters present in black shale (Matlakowska & Sklodowska, 2009). The released organic carbon compounds, identified by GC-MS, as carboxylic acids, aromatic acids, alcohols, and long-chain aliphatic hydrocarbons indicated the potential to bioremediate the hydrocarbon contaminated sites (Matlakowska et al., 2013). In addition to assimilation of organic acids and esters, the biodegradation of polycyclic aromatic compounds was confirmed with the released intermediates and products such as phosphonic acid dioctadecyl ester and isoindole-1,3 in the aqueous phase of cultures (Matlakowska & Sklodowska, 2011). Proteomic analysis of the cellular proteins

isolated from the bacterial cultures identified the enzymes responsible for the initial aerobic decomposition of aliphatic small and long alkanes/alkenes (C_2-C_{10}, $>C_{10}$) (e.g., cytochrome P450, FAD-binding monooxygenase) and polycyclic aromatic hydrocarbons (3,4-dihydroxyphenylacetate 2,3-dioxygenase, catechol 2,3-dioxygenase, 2,3-dihydroxybiphenyl 1,2-dioxygenase, etc.). The hydroxylated aromatic ring in the intermediates (aromatic acids, bisphenols, etc.) could be cleaved by cellular enzymes such as aromatic acid decarboxylase, arylesterase (Stasiuk et al., 2017; Włodarczyk et al., 2016).

In the end, the byproducts, alcohols, and aldehydes could be utilized by the enzymes (aldehyde dehydrogenase, zinc-binding alcohol dehydrogenase, aldehyde oxidase) to produce carboxylic acids that could generate energy and assimilate into biomass through tricarboxylic acid (TCA) cycle (Włodarczyk et al., 2016). Detected lipases and hydrolases could be responsible for the enzymatic hydrolysis of esters and lipids (Stasiuk et al., 2017).

1.5.3 Iron sulphur oxidation

Pyrite is usually present in black shale as a major component and plays a crucial role in shaping the microbial community structure. The GC-MS detection of the iron-binding proteins (ferric siderophore receptor protein, ferrous iron transporter B, TonB-dependent siderophore receptor, bacterioferritin) in the cellular extract of bacterial culture supplemented with the black shale ore indicated that bacterial cells incorporate iron (Włodarczyk et al., 2016). The microbial consortia are involved in the iron and sulphur biogeochemical cycling during the black shale weathering process. Heterotrophic and chemolithoautotrophic microorganisms perform active biooxidative dissolution of pyrite through a variety of metabolic approaches, such as acidophilic S

and Fe oxidation and neutral Fe oxidation, as well as acidophilic Fe reduction (Li et al., 2014)

Acidithiobacillus oxidizes sulphur compounds and decreases pH via producing sulphuric acid during growth on reduced sulfur compounds. The sulfuric acid dissolves metals from metalliferous black shale according to equations (3) and (4) (Chen & Lin, 2001), resulting in the release of acidic, metalliferous fluids from rocks (acid rock drainage ARD) into the surrounding environment (Kwong et al., 2009).

S-oxidizing bacteria

$$2S + 3O_2 + 2H_2O \xrightarrow{\hspace{2cm}} 2SO_4^{2-} + 4H^+ \tag{3}$$

$$H_2SO_4 + Sediment{-}Me \xrightarrow{\hspace{1.5cm}} Sediment{-}2H + MSO_4 \tag{4}$$
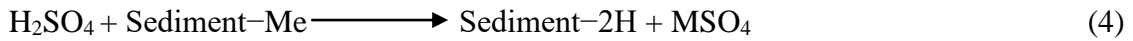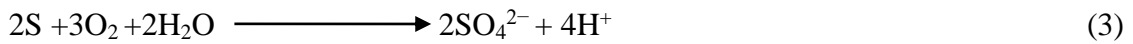
Me-Metal

At low pH, the bacteria perform acid-consuming oxidation of ferrous ion (Fe^{2+}) (Eq. (5)) that results in the pH increase (Marhual et al., 2008). Commonly, soluble metal sulfide minerals are oxidized by ferric ion (Fe^{3+}) to soluble metal sulfates and reduced ferrous ion (Fe^{2+}) (Eqs. 6 and 7) (Kinnunen et al., 2006). Thus, low pH promotes cyclic oxidation/reduction reaction and raises the solubilized metal concentrations with time (Marhual et al., 2008).

Fe-oxidizing bacteria

$$2Fe^{2+} + 0.5O_2 + 2H^+ \xrightarrow{\hspace{2cm}} 2Fe^{3+} + H_2O \tag{5}$$

$$FeS_2 + 8H_2O + 14Fe^{3+} \xrightarrow{\hspace{1.5cm}} 15Fe^{2+} + 2SO_4^{2-} + 16H^+ \tag{6}$$

$$CuFeS_2 + 4Fe^{3+} \xrightarrow{\hspace{1.5cm}} Cu^{2+} + 5Fe^{2+} + 2S^0 \tag{7}$$
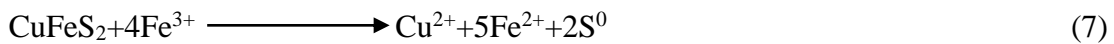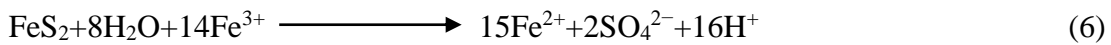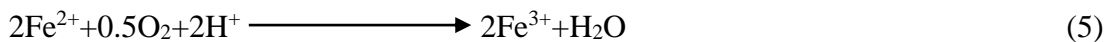
The produced elemental sulfur (S^0) forms a nonreactive layer covering metal sulfides. Moreoever, a high concentration of heavy metals and lower pH will eventually inactivate even acidophilic and heavy metal resistant bacteria and cause decreased metal solubilization (Amin et al., 2018; Chen & Lin, 2001).

In short, in order to successfully metabolize organic-rich shales, the microbes should be able to

1) break complex organics such as phenanthrene in black shale as sole carbon and energy source,

2) resist toxic concentrations of metalloids and metals (in case of metalliferous organic-rich shales),

3) assimilate organic acids proficiently, and

4) express diverse catabolic enzymes such as esterase, lipases, and dioxygenase

 (Matlakowska & Sklodowska, 2009).

1.6 Economic Importance and Environmental Impact

Organic-rich shales are of historical and commercial importance as they form roughly 70 % of the sedimentary rock carbon-rich rocks deposits worldwide (Klemme & Ulmishek, 1991). The 3.2-billion-year-old Fig Tree Series of South Africa is the oldest-known Carboniferous shale. All of the organic-rich shales have not been completely transformed into oil or coal, still, a significant proportion of fossilized organic matter remained immature and offers a huge and yet untapped biological reactive fossil carbon resource, which exceeds both the oil and coal resources (Rice & Claypool, 1981). Today, oil shales are being extensively exploited to generate oil and gas to meet current energy demand. The Green River formation, an oil-shale formation in Utah, Colorado, and Wyoming is a valuable source of synthetic crude oil. In Poland and eastern Germany, the Kupferschiefer, a bituminous shale is used for mining zinc, lead, and copper.

Microbial consortia associated with the organic-rich shale have economic and environmental importance in recovering potential hydrocarbon or shale gas energy sources. Microbially produced methane accounts for 20% of the total methane resource

and it may get trapped and accumulated over time in the low permeable rock of tight siltstones, sandstones, and shales (Rice & Claypool, 1981). Although it is a common perception that the natural gas produced from black shales has a thermogenic origin, Martini et al. (2003) proved that methane from the economic gas deposits located near the Michigan shale basin margins is produced by the microbial degradation of the thermogenic ethane and propane (Martini et al., 2003). Microbes serve as a double edge; they help to recover hydrocarbon sources and aid in methane generation but also degrade/spoil hydrocarbon sources. For instance, deep subsurface shale microbiology is important for maximizing production from hydrocarbon formations from hydraulic fracturing. Despite the addition of biocides during hydraulic fracturing, H_2S-producing, fermenting, and methanogenic prokaryotes still cause fouling of the organics in shale (Ulrich et al., 2018). This process inoculates and enriches halotolerant microorganisms in these reservoirs over time, resulting in a saline ecosystem.

Eventually, the biotransformation of organic-rich shale impacts global climate change, water quality, the biogeochemical cycles of various metals, and human health by redistributing elements and toxicants into the environment (Dalai et al., 2002; Georgiev et al., 2012). Being important reservoirs of pyrite deposits, black shale weathering is expected to produce sulphuric acid and release toxic trace metals and metalloids (Cd, Pb, and As) as a consequence of pyrite oxidation (Li et al., 2020; Tuttle et al., 2009). Subsequently, the discharge of acid rock drainage containing toxic heavy metals may contaminate and damage the local water resources, land environments, and biodiversity (Lavergren et al., 2009). Therefore, characterizing the prevalence and role of microbes in these ecosystems is essential for understanding their biogeochemical cycles and impacts. For instance, nine percent of rivers in England and Wales may pose serious threats to human health and the ecosystem due to acid mine discharge from abandoned mines (Johnston & Rolley, 2008).

1.7 Aims of the Study

The contribution of the microbial degradation of organics in shale under subsurface anaerobic conditions has been widely studied, however comprehensive investigation of the microbial colonization of the terrestrial and subterrestrial black shale and how they thrive energy from the organics is pending. Although several bacterial isolates have been studied for their potential to utilize black shale ore as a sole carbon source, metagenomic tools had not been extensively utilized to give a wholistic microbial diversity of oil shales. If we want to deeply characterize the microbial functions in overall black shale weathering processes, detailed information regarding the community structure and spatial variations of these microorganisms are essential. Therefore, we investigated the native lithobiontic microbial communities in the outcropping Monte San Giorgio organic-rich shale to address several major questions.

  (i)   the structure of the microbial populations in the rocks

  (ii)   the influence of abiotic factors on the microbial diversity

  (iii)   the effect of artificial media on the native microbiome enrichment

  (iv)   the processes of the aerobic oxidation of fossil hydrocarbons

Similarly, the influence of microorganisms on the downstream of acid mine drainage produced from Kupferschiefer black shale is even more limitedly investigated. The Marsberg copper mine offers a subterrestrial copper-bearing black shale environment where microbial communities around acid mine drainage and mine waters could be thoroughly investigated to

  (i)   define the native taxonomic microbial profile

  (ii)   characterization of unidentified *Chloroflexi Ktedonobacteria* members

(iii)   understand the underlying survival and colonization mechanisms

(iv)   metabolic processes involved in the fossil hydrocarbons degradation

The bacterial communities were investigated metagenomically through Ilumina sequencing. The metagenomes were assembled, annotated, and compared. The pathways related to hydrocarbons metabolism and heavy metal tolerance were hypothesized to be an important approach to the microbial survival and biodegradation of black shale organics. The identification of *Ktedonobacteria* at Marsberg cold copper mine and Sasso Pisano geothermal field led to more investigation for their phylogenetic characterization, heavy metal resistance, and hydrocarbon degradation.

1.8 References

Amin, N., Schneider, D., and Hoppert, M. (2018) Bioleaching potential of bacterial communities in historic mine waste areas. *Environmental Earth Sciences* **77**: 542.

Atlas, R.M. (1981) Microbial degradation of petroleum hydrocarbons: an environmental perspective. *Microbiological Reviews* **45**: 180-209.

Behar, F., Lorant, F., and Lewan, M. (2008) Role of NSO compounds during primary cracking of a Type II kerogen and a Type III lignite. *Organic Geochemistry* **39**: 1-22.

Berg, J.S., Duverger, A., Cordier, L., Laberty-Robert, C., Guyot, F., and Miot, J. (2020) Rapid pyritization in the presence of a sulfur/sulfate-reducing bacterial consortium. *Scientific Reports* **10**: 8264.

Berlendis, S., Beyssac, O., Derenne, S., Benzerara, K., Anquetil, C., Guillaumet, M. et al. (2014) Comparative mineralogy, organic geochemistry and microbial diversity of the Autun black shale and Graissessac coal (France). *International Journal of Coal Geology* **132**: 147-157.

Biswas, K.C., Woodards, N.A., Xu, H., and Barton, L.L. (2009) Reduction of molybdate by sulfate-reducing bacteria. *Biometals* **22**: 131-139.

Bohacs, K.M., Carroll, A.R., Neal, J.E., and Mankiewicz, P.J. (2000) Lake-basin type, source potential, and hydrocarbon character: an integrated sequence-stratigraphic-geochemical framework. *Lake basins through space and time: AAPG Studies in Geology* **46**: 3-34.

Chen, S.-Y., and Lin, J.-G. (2001) Bioleaching of heavy metals from sediment: significance of pH. *Chemosphere* **44**: 1093-1102.

Dalai, T.K., Singh, S.K., Trivedi, J., and Krishnaswami, S. (2002) Dissolved rhenium in the Yamuna River System and the Ganga in the Himalaya: Role of black shale weathering on the budgets of Re, Os, and U in rivers and CO2 in the atmosphere. *Geochimica et Cosmochimica Acta* **66**: 29-43.

Dyni, J.R. (2006) Geology and resources of some world oil-shale deposits. U.S. Geological Survey Scientific Investigations Report 2005–5294, 42 pp.

El-Sheshtawy, H.S., Khalil, N., and Farouk, S. (2020) Biodegradation Effect on the Campanian Oil Shale of Egypt. *Geomicrobiology Journal* **37**: 746-752.

Foght, J. (2009) Microbial communities in oil shales, biodegraded and heavy oil reservoirs, and bitumen deposits. In *Handbook of hydrocarbon and lipid microbiology.* Timmis, K.N. (ed). Springer, Berlin, pp. 2159-2172.

Formolo, M., Martini, A., and Petsch, S. (2008) Biodegradation of sedimentary organic matter associated with coalbed methane in the Powder River and San Juan Basins, USA. *International Journal of Coal Geology* **76**: 86-97.

Georgiev, S., Stein, H.J., Hannah, J.L., Weiss, H.M., Bingen, B., Xu, G. et al. (2012) Chemical signals for oxidative weathering predict Re–Os isochroneity in black shales, East Greenland. *Chemical Geology* **324**: 108-121.

Goes, K.C.G.P.d., Lovato, G.M., and Andrade, D.S. (2018) Composition of bacterial community in enrichment cultures of shale by-products from Irati Formation, Brazil. *Brazilian Journal of Microbiology* **49**: 742-748.

Grabowski, A., Nercessian, O., Fayolle, F., Blanchet, D., and Jeanthon, C. (2005) Microbial diversity in production waters of a low-temperature biodegraded oil reservoir. *FEMS Microbiology Ecology* **54**: 427-443.

Gray, N., Sherry, A., Hubert, C., Dolfing, J., and Head, I. (2010) Methanogenic degradation of petroleum hydrocarbons in subsurface environments: remediation, heavy oil formation, and energy recovery. *Advances in Applied Microbiology* **72**: 137-161.

He, W., Sun, Y., Guo, W., Shan, X., Su, S., Zheng, S. et al. (2019) Organic geochemical characteristics of the upper Cretaceous Qingshankou Formation oil shales in the Fuyu oilfield, Songliao Basin, China: Implications for oil-generation potential and depositional environment. *Energies* **12**: 4778.

Hedges, J.I., and Oades, J.M. (1997) Comparative organic geochemistries of soils and marine sediments. *Organic Geochemistry* **27**: 319-361.

Huyck, H.L. (1990) When is a metalliferous black shale not a black shale. *Metalliferous Black Shales and Related Ore Deposits, Proc 1989 US Working Group Meet, IGCP* **254**: 42-56.

Johnston, D., and Rolley, S. (2008) Abandoned mines and the water framework directive in the United Kingdom. *Mine water and the environment VSB–technical University of Ostrava, Czech Republic*: 529-532.

Jones, D., Head, I., Gray, N., Adams, J., Rowan, A., Aitken, C. et al. (2008) Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs. *Nature* **451**: 176-180.

Jones, E.J., Voytek, M.A., Corum, M.D., and Orem, W.H. (2010) Stimulation of methane generation from nonproductive coal by addition of nutrients or a microbial consortium. *Applied and Environmental Microbiology* **76**: 7013-7022.

Jørgensen, B.B., Findlay, A.J., and Pellerin, A. (2019) The Biogeochemical Sulfur Cycle of Marine Sediments. *Frontiers in Microbiology* **10**.

Kaiser, D., Konovalov, S., Schulz-Bull, D.E., and Waniek, J.J. (2017) Organic matter along longitudinal and vertical gradients in the Black Sea. *Deep Sea Research Part I: Oceanographic Research Papers* **129**: 22-31.

Kidd, R.B., MB, C., and WBF, R. (1978) Stratigraphy of eastern Mediterranean sapropel sequences recovered during DSDP Leg 42A and their paleoenvironmental significance. *Initial Reports of the Deep Sea Drilling Project* **42**: 421-443.

Kinnunen, P.-M., Heimala, S., Riekkola-Vanhanen, M.-L., and Puhakka, J. (2006) Chalcopyrite concentrate leaching with biologically produced ferric sulphate. *Bioresource Technology* **97**: 1727-1734.

Klemme, H., and Ulmishek, G.F. (1991) Effective petroleum source rocks of the world: stratigraphic distribution and controlling depositional factors. *AAPG Bulletin* **75**: 1809-1851.

Kwong, Y.J., Whitley, G., and Roach, P. (2009) Natural acid rock drainage associated with black shale in the Yukon Territory, Canada. *Applied Geochemistry* **24**: 221-231.

Königshof, P., Becker, R.T., and Hartenfels, S. (2016) The Rhenish Massif as a part of the European Variscides. *Münstersche Forschungen zur Geologie und Palaeontologie l* **108**: 1-13.

Lavergren, U., Åström, M.E., Bergbäck, B., and Holmström, H. (2009) Mobility of trace elements in black shale assessed by leaching tests and sequential chemical extraction. *Geochemistry: Exploration, Environment, Analysis* **9**: 71-79.

Leahy, J.G., and Colwell, R.R. (1990) Microbial degradation of hydrocarbons in the environment. *Microbiological Reviews* **54**: 305-315.

Li, J., Sun, W., Wang, S., Sun, Z., Lin, S., and Peng, X. (2014) Bacteria diversity, distribution and insight into their role in S and Fe biogeochemical cycling during black shale weathering. *Environmental Microbiology* **16**: 3533-3547.

Li, Q., Zhu, B., and Li, J. (2020) A comparative study on the micro-surface characteristics at black shale initial oxidation stage. *Scientific Reports* **10**: 1-9.

Liao, Y., Geng, A., and Huang, H. (2009) The influence of biodegradation on resins and asphaltenes in the Liaohe Basin. *Organic Geochemistry* **40**: 312-320.

Magot, M., Ollivier, B., and Patel, B.K. (2000) Microbiology of petroleum reservoirs. *Antonie van Leeuwenhoek* **77**: 103-116.

Marhual, N., Pradhan, N., Kar, R., Sukla, L., and Mishra, B. (2008) Differential bioleaching of copper by mesophilic and moderately thermophilic acidophilic consortium enriched from same copper mine water sample. *Bioresource Technology* **99**: 8331-8336.

Martini, A.M., Walter, L.M., Ku, T.C., Budai, J.M., McIntosh, J.C., and Schoell, M. (2003) Microbial production and modification of gases in sedimentary basins: A geochemical case study from a Devonian shale gas play, Michigan basin. *AAPG Bulletin* **87**: 1355-1375.

Matlakowska, R., and Sklodowska, A. (2009) The culturable bacteria isolated from organic-rich black shale potentially useful in biometallurgical procedures. *Journal of Applied Microbiology* **107**: 858-866.

Matlakowska, R., and Sklodowska, A. (2011) Biodegradation of Kupferschiefer black shale organic matter (Fore-Sudetic Monocline, Poland) by indigenous microorganisms. *Chemosphere* **83**: 1255-1261.

Matlakowska, R., Narkiewicz, W., and Sklodowska, A. (2010) Biotransformation of Organic-Rich Copper-Bearing Black Shale by Indigenous Microorganisms Isolated from Lubin Copper Mine (Poland). *Environmental Science & Technology* **44**: 2433-2440.

Matlakowska, R., Ruszkowski, D., and Sklodowska, A. (2013) Microbial transformations of fossil organic matter of Kupferschiefer black shale–elements mobilization from metalloorganic compounds and metalloporphyrins by a community of indigenous microorganisms. *Physicochemical Problems of Mineral Processing* **49**: 223-231.

McEvoy, J., and Giger, W. (1986) Origin of hydrocarbons in triassic serpiano oil shales: Hopanoids. *Organic Geochemistry* **10**: 943-949.

McGenity, T.J. (2019) *Microbial Communities Utilizing Hydrocarbons and Lipids: Members, Metagenomics and Ecophysiology*: Springer, Cham (Switzerland).

Meslé, M., Périot, C., Dromart, G., and Oger, P. (2013) Biostimulation to identify microbial communities involved in methane generation in shallow, kerogen-rich shales. *Journal of Applied Microbiology* **114**: 55-70.

Meslé, M., Dromart, G., Haeseler, F., and Oger, P.M. (2015) Classes of organic molecules targeted by a methanogenic microbial consortium grown on sedimentary rocks of various maturities. *Frontiers in Microbiology* **6**: 589.

Meyers, P.A., Forster, A., and Sturt, H. (2004) Microbial gases in Black Shale sequences on the Demerara Rise. *Proceedings of the Ocean Drilling Program Initial Reports* **207**:1-18.

Mouser, P.J., Borton, M., Darrah, T.H., Hartsock, A., and Wrighton, K.C. (2016) Hydraulic fracturing offers view of microbial life in the deep terrestrial subsurface. *FEMS Microbiology Ecology* **92**.

Nance, R.D., Gutiérrez-Alonso, G., Keppie, J.D., Linnemann, U., Murphy, J.B., Quesada, C. et al. (2010) Evolution of the Rheic ocean. *Gondwana Research* **17**: 194-222.

Nijenhuis, I.A., Bosch, H.J., Sinninghe Damsté, J.S., Brumsack, H.J., and De Lange, G.J. (1999) Organic matter and trace element rich sapropels and black shales: a geochemical comparison. *Earth and Planetary Science Letters* **169**: 277-290.

Penner, T.J., Foght, J.M., and Budwill, K. (2010) Microbial diversity of western Canadian subsurface coal beds and methanogenic coal enrichment cultures. *International Journal of Coal Geology* **82**: 81-93.

Petsch, S., Edwards, K., and Eglinton, T. (2005a) Microbial transformations of organic matter in black shales and implications for global biogeochemical cycles. *Palaeogeography, Palaeoclimatology, Palaeoecology* **219**: 157-170.

Petsch, S.T., Edwards, K.J., and Eglinton, T.I. (2005b) Microbial transformations of organic matter in black shales and implications for global biogeochemical cycles. *Palaeogeography, Palaeoclimatology, Palaeoecology* **219**: 157-170.

Peucker-Ehrenbrink, B., and Hannigan, R.E. (2000) Effects of black shale weathering on the mobility of rhenium and platinum group elements. *Geology* **28**: 475-478.

Philp, R.P. (2003) 7.09 - Formation and Geochemistry of Oil and Gas. In *Treatise on Geochemistry*. Holland, H.D., and Turekian, K.K. (eds). Pergamon, Oxford, pp. 223-256.

Potter, P.E., Maynard, J.B., and Pryor, W.A. (2012) *Sedimentology of shale: study guide and reference source*: Springer, New York.

Rabus, R., Hansen, T.A., and Widdel, F. (2006) Dissimilatory Sulfate- and Sulfur-Reducing Prokaryotes. In *The Prokaryotes: Volume 2: Ecophysiology and*

Biochemistry. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H., and Stackebrandt, E. (eds). Springer, New York, pp. 659-768.

Rajpert, L., Skłodowska, A., and Matlakowska, R. (2013) Biotransformation of copper from Kupferschiefer black shale (Fore-Sudetic Monocline, Poland) by yeast Rhodotorula mucilaginosa LM9. *Chemosphere* **91**: 1257-1265.

Rice, D.D., and Claypool, G.E. (1981) Generation, accumulation, and resource potential of biogenic gas. *AAPG Bulletin* **65**: 5-25.

Römer, M., Sahling, H., dos Santos Ferreira, C., and Bohrmann, G. (2020) Methane gas emissions of the Black Sea—mapping from the Crimean continental margin to the Kerch Peninsula slope. *Geo-Marine Letters* **40**: 467-480.

Schieber, J. (2003) Black shales. In *Encyclopedia of Sediments and Sedimentary Rocks*. Middleton, G.V., Church, M.J., Coniglio, M., Hardie, L.A., and Longstaffe, F.J. (eds). Springer, Dordrecht (Netherlands), pp. 83-85.

Sklodowska, A., Matlakowska, R., and Bal, K. (2005) Extracellular polymer produced in the presence of copper minerals. *Geomicrobiology Journal* **22**: 65-73.

Stasiuk, R., Włodarczyk, A., Karcz, P., Janas, M., Skłodowska, A., and Matlakowska, R. (2017) Bacterial weathering of fossil organic matter and organic carbon mobilization from subterrestrial Kupferschiefer black shale: long-term laboratory studies. *Environmental Microbiology Reports* **9**: 459-466.

Struchtemeyer, C.G., Davis, J.P., and Elshahed, M.S. (2011) Influence of the drilling mud formulation process on the bacterial communities in thermogenic natural gas wells of the Barnett Shale. *Applied and Environmental Microbiology* **77**: 4744-4753.

Takai, K., Mormile, M.R., McKinley, J.P., Brockman, F.J., Holben, W.E., Kovacik Jr, W.P., and Fredrickson, J.K. (2003) Shifts in archaeal communities associated with lithological and geochemical variations in subsurface Cretaceous rock. *Environmental Microbiology* **5**: 309-320.

Tissot, B.P., and Welte, D.H. (1984) *Petroleum formation and occurrence*: Springer, Berlin.

Tuttle, M.L., Breit, G.N., and Goldhaber, M.B. (2009) Weathering of the New Albany Shale, Kentucky: II. Redistribution of minor and trace elements. *Applied Geochemistry* **24**: 1565-1578.

Ulrich, N., Kirchner, V., Drucker, R., Wright, J.R., McLimans, C.J., Hazen, T.C. et al. (2018) Response of Aquatic Bacterial Communities to Hydraulic Fracturing in Northwestern Pennsylvania: A Five-Year Study. *Scientific Reports* **8**: 5683.

Urban, H., Stribrny, B., Zereini, F., and Ye, Y. (1995) Geochemistry and metallogenesis of Lower Carboniferous black shale-hosted ore deposits, NE Rhenish Massif, FR Germany. *Ore Geology Reviews* **9**: 427-443.

Vandenbroucke, M., and Largeau, C. (2007) Kerogen origin, evolution and structure. *Organic Geochemistry* **38**: 719-833.

von Raumer, J.F., Nesbor, H.-D., and Stampfli, G.M. (2017) The north-subducting Rheic Ocean during the Devonian: consequences for the Rhenohercynian ore sites. *International Journal of Earth Sciences* **106**: 2279-2296.

Vuillemin, A., Ariztegui, D., Horn, F., Kallmeyer, J., Orsi, W.D., and Team, t.P.S. (2018) Microbial community composition along a 50 000-year lacustrine sediment sequence. *FEMS Microbiology Ecology* **94**.

Walker, J., Petrakis, L., and Colwell, R. (1976) Comparison of the biodegradability of crude and fuel oils. *Canadian Journal of Microbiology* **22**: 598-602.

Wawrik, B., Mendivelso, M., Parisi, V.A., Suflita, J.M., Davidova, I.A., Marks, C.R. et al. (2012) Field and laboratory studies on the bioconversion of coal to methane in the San Juan Basin. *FEMS Microbiology Ecology* **81**: 26-42.

Włodarczyk, A., Szymańska, A., Skłodowska, A., and Matlakowska, R. (2016) Determination of factors responsible for the bioweathering of copper minerals from organic-rich copper-bearing Kupferschiefer black shale. *Chemosphere* **148**: 416-425.

Włodarczyk, A., Lirski, M., Fogtman, A., Koblowska, M., Bidziński, G., and Matlakowska, R. (2018) The oxidative metabolism of fossil hydrocarbons and sulfide minerals by the lithobiontic microbial community inhabiting deep subterrestrial Kupferschiefer Black Shale. *Frontiers in Microbiology* **9**:972.

Xing, B., and Pignatello, J.J. (2005) Sorption | Organic Chemicals. In Hillel, D. (ed)., *Encyclopedia of Soils in the Environment*. Elsevier, Oxford, pp. 537-548.

Zengler, K., Richnow, H.H., Rosselló-Mora, R., Michaelis, W., and Widdel, F. (1999) Methane formation from long-chain alkanes by anaerobic microorganisms. *Nature* **401**: 266-269.

2. Composition, Diversity and Functional Analysis of the Modern Microbiome of the Middle Triassic Cava Superiore Beds (Monte San Giorgio, Switzerland)

Sania Arif[1*], Joachim Reitner[2], Michael Hoppert[1],

[1] Institute of Microbiology & Genetics, Department of General Microbiology;

[2] Geosciences Centre, Department of Geobiology; Georg-August-Universität Göttingen

*Correspondence to sania.arif@stud.uni-goettingen.de

Author Contributions:

M.H. and J.R. designed the study and collected the samples. S.A. performed the experiments, analyzed data, interpreted the results, and wrote the manuscript. M.H. contributed reagents/material/analysis tools, interpreted the results, and corrected the manuscript. J.R. critically reviewed and corrected the manuscript.

SCIENTIFIC REPORTS

natureresearch

OPEN

Composition, Diversity and Functional Analysis of the Modern Microbiome of the Middle Triassic Cava Superiore Beds (Monte San Giorgio, Switzerland)

Sania Arif[1]*, Joachim Reitner[2] & Michael Hoppert[1]

Organic-rich laminated shales and limestones from the Monte San Giorgio (Lugano Prealps, Switzerland) are known as famous fossil lagerstätten for excellently preserved fossils from the Middle Triassic Period. The various bituminous shales from Monte San Giorgio are thermally immature and rich in diverse organic compounds, which provide unique substrates for active soil microbial communities. We selected the Cava superior beds of the Acqua del Ghiffo site for this study. To investigate its microbial structure and diversity, contig assembly, Operational Taxonomic Units (OTUs) clustering, and rarefaction analysis were performed for bacterial 16S rDNA preparations from bituminous and non-bituminous limestone strata with the MetaAmp pipeline. Principal coordinates analysis shows that the microbial communities from the bituminous strata differ significantly from limestone samples ($P < 0.05$ Unifrac weighted). Moreover, metagenomic tools could also be used effectively to analyze the microbial communities shift during enrichment in specific growth media. In the nutrient-rich media, one or few taxa, mainly *Proteobacteria* and *Firmicutes*, were enriched which led to the drastic diversity loss while oligotrophic media could enrich many taxa simultaneously and sustain the richness and diversity of the inoculum. Piphillin, METAGENassist and MicrobiomeAnalyst pipeline also predicted that the Monte San Giorgio bituminous shales and oligotrophic enriched microbiomes degrade complex polycyclic aromatic hydrocarbons.

Monte San Giorgio, located near Lake Lugano in Canton Ticino, Switzerland, is listed as a UNESCO World Heritage site because of its best preserved paleontological record from the Middle Triassic Period (approx. 245–230 million years ago)[1,2]. Many abundant and exceptionally detailed fossils of reptiles, fish, bivalves, ammonites, echinoderms and crustaceans, insects, and plants have been reported[3]. During the later Middle Triassic, the region around Monte San Giorgio formed restricted basins that were largely isolated from the open ocean with an off-shore reef[4–6]. Being close to land, this basin included both diverse marine and terrestrial life under nearly perfect stagnant, euxinic, or even anoxic conditions. Successive phases of marine transgression and regression created different environments that allowed for the deposition of fine-grained, mostly laminated mudstones, limestones, and dolomites with varying content of organic material (up to 40%)[7–9]. The Middle Triassic sequence consists of approximately 1,000 metres of reef limestones, dolomites, and bituminous shales which formed in marine conditions at the southern margin of the Triassic 'Tethys' Ocean. At least five distinct fossiliferous formations within the Middle Triassic Lower Meride limestone are the Grenzbitumenzone, Cava Inferior, Cava Superior, Cassina Beds and Kalkschieferzone, each yielding different vertebrate assemblages[5,10,11] and consisting of finely laminated limestones and marls with intercalated volcanic ash layers of andesitic and rhyolitic composition.

Many scientific reports document the fossils of Monte San Giorgio but no studies have been conducted to evaluate its extant geomicrobiology, though bituminous shales may provide substrates suitable for microbial growth. These shales provide rather an immature bitumen and partly kerogen, consisting of diverse organic

[1]Institute of Microbiology & Genetics, Department of General Microbiology, Georg-August-Universität Göttingen, Göttingen, Germany. [2]Geosciences Centre, Department of Geobiology, Georg-August-Universität Göttingen, Göttingen, Germany. *email: sania.arif@stud.uni-goettingen.de
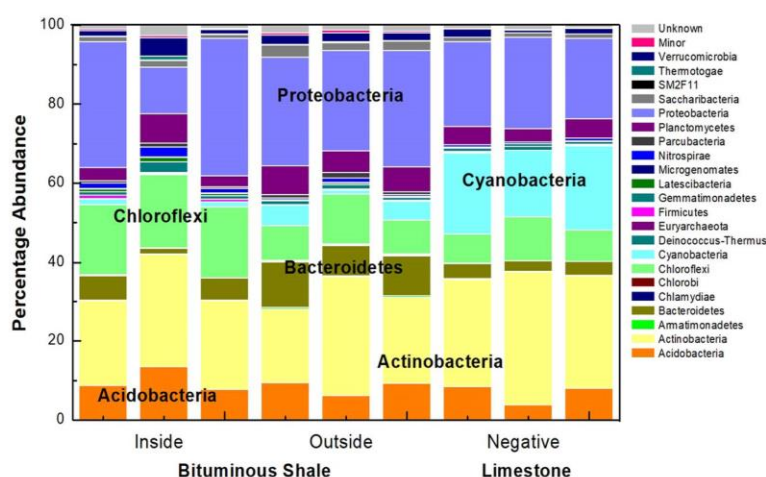
32

Figure 1. Relative abundance and taxonomic assignments of microbial flora at phylum level of inside and outside surfaces of the bituminous shale and limestone as the negative control in triplicates. Minor phyla having <0.02 relative abundance are represented as "Minor".

compounds, among them hopanes, hopenes, and methylhopanoids[12] and their metagenomic biodiversity analysis of the non-culturable microbiome is interesting from a microbiological, geochemical as well as a biotechnological point of view. The aim of this study was to explore the microbial structure and functional diversity of the Monte San Giorgio organic-rich shale and understand its contribution towards the carbon cycle through degradation of the associated complex hydrocarbons. The other objective was to investigate how the microbial growth and diversity can be influenced by different enrichment media and which hydrocarbon-degrading metabolic groups may be enriched from the environment. For this purpose, we examined the microbial communities of Monte San Giorgio organic-rich shale and limestone rocks samples via Illumina MiSeq sequencing, followed by microbial growth in enrichment media to elucidate the correlation of the microbial communities and enrichment effect of media. Paired-end sequencing was performed by Göttingen Genomics Laboratory after indexing of the V3-V4 PCR amplicons with the Nextera XT DNA library prep kit (Illumina, San Diego, Cal, USA). In addition, the geochemical and bioremediation metabolic potential of the microbiomes were predicted with the online available tools METAGENassist[13] and Piphillin[14].

Results

Metagenomics of Monte San Giorgio samples. A total of 4113 distinct OTUs were affiliated to 37 known phyla in the Monte San Giorgio samples. *Actinobacteria* ($25.8\% \pm 4.6$) and *Proteobacteria* ($25.08\% \pm 6.5$) were the most abundant phyla in all samples, followed by *Bacteroidetes* ($5.9\% \pm 3.25$), *Acidobacteria* ($8.2\% + 2.5$), *Planctomycetes* ($5.1\% \pm 1.6$), *Verrucomicrobia* ($3.6\% \pm 1.1$), and *Saccharibacteria* ($1.6\% \pm 0.65$) (Fig. 1). *Cyanobacteria* was the most prevalent phylum in the limestone samples while *Chloroflexi* and *Latescibacteria* were abundant in the organic-rich shale samples. The analysis of variance (ANOVA) validated that the outside rock surface sample of the shale was prevalent in *Armatimonadetes*, *Saccharibacteria*, *SM2F11*, *Chlorobi*, *Bacteroidetes*, and *TM6* ($p < 0.05$). *Proteobacteria* subclass *Gammaproteobacteria* was relatively abundant in the inside surface of a freshly cleaved organic-rich shale sample (54%), whereas in limestone samples, *Alphaproteobacteria* (84%) were prevalent. More in-depth analysis at the order level with the hierarchical clustering and heatmap visualization illustrated the unique geomicrobiological fingerprint of each biospecimen (Fig. S1). The outside surface of the bituminous shale contained at least seven strictly aerobic or facultative orders (*Desulfurellales*, *Flavobacteriales*, *Rhizobiales*, *Xanthomonadales Sphingobacteriales*, *Sphingomonadales*, and *Caulobacteriales*) which are known for aerobic hydrocarbon metabolism[15]. In contrast, the organic-rich shale contained nine strictly anaerobic orders (*Thermotogales*, *Thermoplasmatales*, *Desulfobacterales*, *Chromatiales*, *Desulfuromonadales*, *Hydrogenophilales*, *Campylobacterales*, *Rhodocyclales*, and *Clostridiales*) exclusively.

These anaerobic orders may be potentially involved in the hydrocarbon metabolism along with the only three abundant aerobic orders (*Bacillales*, *Lactobacillales*, and *Enterobacteriales*)[15]. The limestone rocks were abundant in the orders *Rubrobacterales*, *Deinococcales*, *Rhodobacterales*, *Streptomycetales*, which have the ability to survive harsh environmental conditions such as high ionizing radiation[16–18].

Effect of enrichment media on the relative abundance of microbes. After the first round of enrichment, *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, and *Chloroflexi* were found to be relatively abundant in almost every media. The selective effect of the culture media was not evident at this stage since the taxa of the natural rock environment were not displaced yet (Fig. 2a). After the final round of enrichment, the number of OTUs decreased from 5906 to 3296 which indicates the loss of diversity and microbial richness as the microbes most favorably adapted to the culture media were selected over others during the serial inoculum enrichment. *Firmicutes* was significantly enriched in the LB media, while *Proteobacteria* predominated in the succinate minimal and iron basal media. Interestingly, 9 K media, PYGV, and 1% crude oil minimal media were still rich in biodiversity and
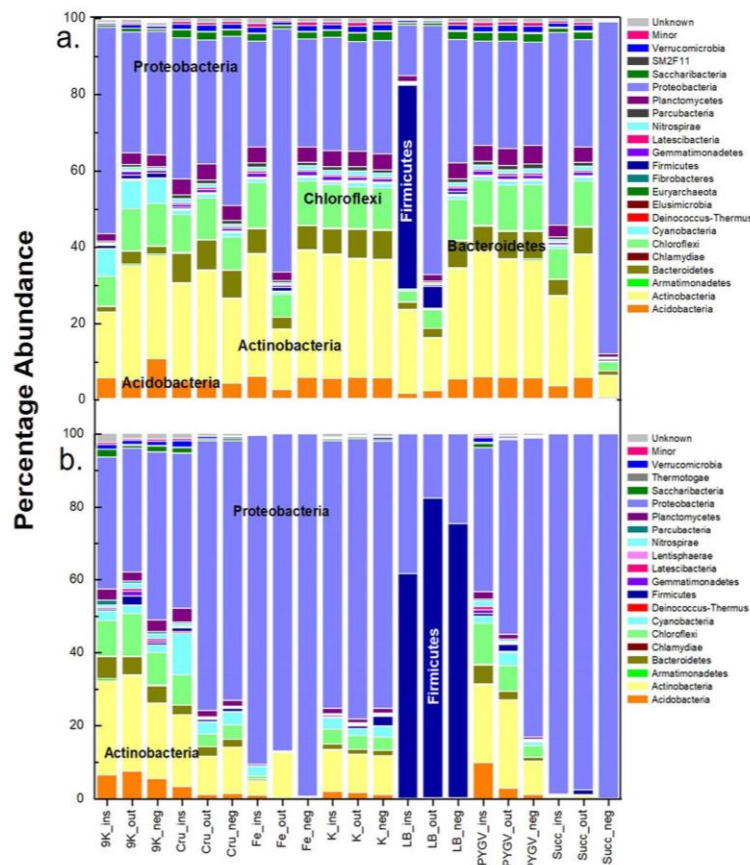
Figure 2. Relative abundance of microbial flora in the different enrichment media. (**a**) First and (**b**) third round of enrichments of the outside (out) and inside (ins) organic-rich shale and limestone negative control (neg) sample inoculums in 9 K (Silverman and Lundgren medium) Cru (Crude oil minimal media), K (Keller medium), LB (Luria-Bertani broth), PYGV (Peptone, Yeast extract, Glucose and vitamin mixture and mineral salts), Fe (Iron basal media), and Succ (Succinate minimal media).

accommodated more diverse phyla and orders than other media (Fig. 2b). In the heatmap, the orders enriched in the nutrient-rich complex and oligotrophic media were not only distinguishable but also more abundant in the oligotrophic media (Fig. S2). The effect of each medium in enriching specific organisms was also observable, for instance, the LB media has enriched *Pseudomonadales*, and *Bacillales*, while the K media has enriched the soil/cave related orders *Rhizobiales, Ktedonobacterales, Acidobacteriales*, and *Armatimonadales*.

Alpha diversity. Diversity indices provide important quantitative information about rarity and commonness of species which helps to understand community structure. The alpha diversity index Chao1 and abundance-based coverage estimators (ACE) indicated that richness of the samples was outside shale surface > inside shale surface > negative limestone control. The microbial species (OTUs) were more abundant on the outside surface of the shale sample than on inside shale surface and on limestone samples (Table S1). Following the serial inoculum enrichment, the subcultures drastically lost their alpha diversity especially when carbon/nitrogen sources were very rich, such as in LB complex media. The loss of OTUs also reduced both richness (Chao1 and ACE) and diversity (Shannon and Simpson) indexes, indicating that only a few adapted microbes were selected over others in the enrichment media (Fig. 3a). Nevertheless, the oligotrophic media provided more stringent conditions where the fast-growing bacteria did not utilize the carbon source to simply outgrow the rest; hence the species richness and diversity seem to be preserved (Fig. 3b).

Beta diversity. Principal Coordinates Analysis (PCoA) is a multidimensional scaling method to explore and to visualize similarities or dissimilarities of complex metagenomic data into 2–3 dimensions. The PCoA of Monte San Giorgio samples showed that the microbial communities of inside, outside shale surface, and limestone samples were distinct to each other. The sample triplicates resemble each other more closely, though a considerable variance is obvious (Fig. 4a). This hypothesis was tested with the Unifrac weighted algorithm[19]. The significance between the inside and outside surfaces of the bituminous shale was <*0.0010*, same as the outside and limestone samples indicating that both populations are similar to each other but the organic-rich (inside) and limestone
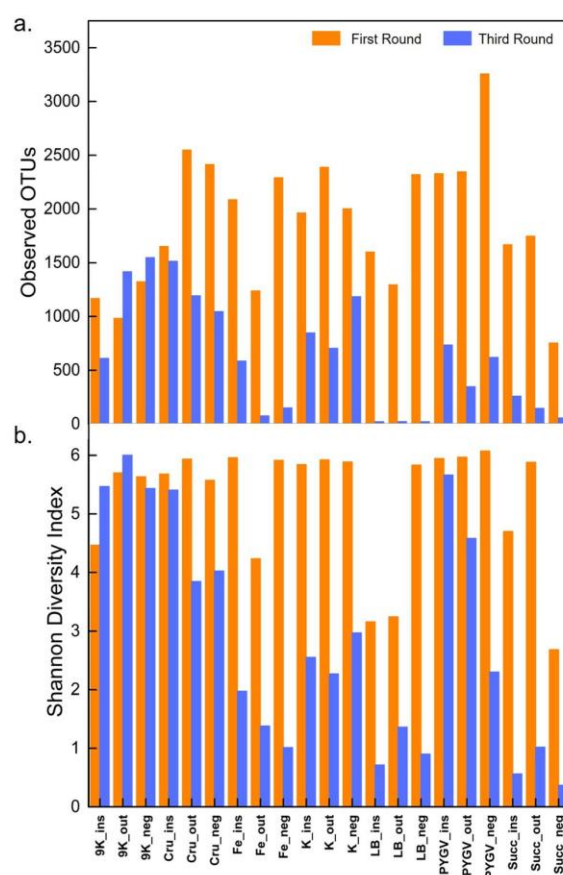
Figure 3. The number of observed OTUs (**a**) and Shannon diversity index (**b**) after the first and third round of enrichments. See Fig. 2 for abbreviations. Note that the nutrient-rich complex media have lost OTUs by more than 80%. The oligotrophic media are higher in alpha diversity than the nutrient-rich complex media.

samples differ from each other ($p = 0.0190$). The results were also validated with the analysis of molecular variance (AMOVA) nonparametric method[20]. AMOVA interprets whether genetic diversity within two populations is significantly different from their pooled populations or not (Table S2). It can be assumed that the different microorganisms from soil are being selected by these diverse environmental conditions as the soil-contaminated outside surface of shale's microbiome shared similarity with both organic-rich bituminous shale and nutrient-starved limestone microbiomes.

Effect of media and time in the enrichment process. The PCoA plots after the first round of incubation showed that 9 K media selected similar sets of microbes from the different samples, forming a separate cluster, distinct from all other media which promote growth of a set of further microbial communities, very similar to each other. Obviously, the effect of these other media in selecting different microbes was not substantial at this stage of the enrichment process (Fig. 4b). After the third round of serial subculturing, the effect of media in selecting the same set of organisms from different samples was profound, implying the same environmental conditions selected the same microbes over the passage of time. 9 K, K, crude oil, and, PYGV selected sets of microbes forming a cluster distinct from that of Fe basal media, succinate minimal, and LB media (Fig. 4c).

Functional profile analysis. *Taxonomic to phenotypic mapping.* The taxonomic profile of the Monte San Giorgio microbiome depicts distinct communities related to the bituminous and non-bituminous samples which may possibly indicate distinct metabolic potentials (Fig. S1). To predict the metabolic capabilities of these microorganisms, their taxonomic profiles were fed to METAGENassist, which matches the taxonomic data with the phenotype databases[13]. The clustering analysis (heatmap) indicates the Monte San Giorgio shale microbiome is capable of degrading complex organics such as aromatic hydrocarbons, naphthalene, chlorophenol and atrazine; sulfate and nitrate metabolism is prominent in contrast to the limestone samples (Fig. 5a). The oligotrophic media in contrast to the nutrient-rich media have clearly enriched the shale microbiome with much diverse hydrocarbon metabolizing potential. For instance, 9 K, K, and crude oil minimal media enrich the microbes that can degrade chlorophenol, aromatic hydrocarbons, and naphthalene respectively.
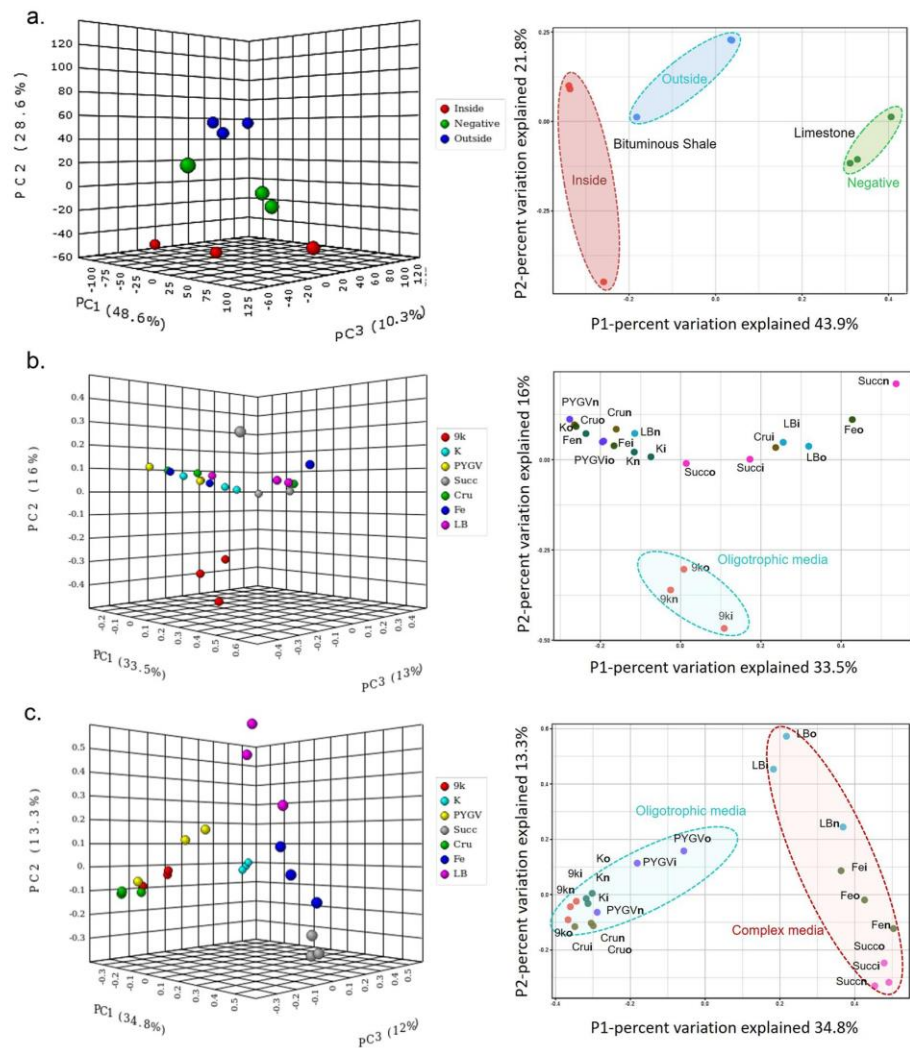
35

Figure 4. Principal Coordinates Analysis (PCoA) of the Monte San Giorgio samples and its enrichment cultures. PCoA of (**a**) inside, outside of bituminous shale and limestone rocks (negative), (**b**) enrichment cultures after the first round of incubation (**c**) third and final round of enrichments from inside, outside of bituminous shale and negative control: limestone rocks. See Fig. 2 for abbreviations of media; **i** (inside), **o** (outside) and **n** (negative control) are appended to an abbreviation.

Metagenomics data analysis. The abundance of functional genes was estimated from the 16s rRNA metagenomic data that allowed inference of possible function and mechanisms. Genome contents were predicted for each OTU and then the gene copy numbers were summarized by KO identifier (also called K number) retrieved from the KO database of functional orthologs (KEGG Kyoto Encyclopedia of Genes and Genomes Orthology)[21]. The PCoA based on these KO identifiers correlated with the PCoA predicted from the OTUs since some of the functional proficiencies (KO) are unique to the OTUs nearest matched genomes (Figs. 4 and 6). The functional profiles of the inside, outside, and negative limestone control samples differ from each other and the effect of the oligotrophic and complex media is also profound in selecting the microbes with the different functional profiles.

To further elucidate the inferred key functional pathways in these biospecimen, differential abundance analysis for 16S rRNA Seq data was done with the statistical R packages DESeq2[22]. The gene abundance data retrieved from the FASTA sequences of the submitted 16s rRNA OTUs was summarized with KO abundances, KEGG reference database[21]. The significantly abundant features for bituminous shale (inside) against the outer surface of bituminous shale and limestone rock samples over 1-fold abundance difference were considered. The enzymes such as 4-hydroxybenzoyl-CoA thioesterase, benzaldehyde dehydrogenase, benzoate/toluate dioxygenase and vanillate monooxygenase, involved in the benzoate, xylene, fluorobenzoate, dioxin, aminobenzoate, and
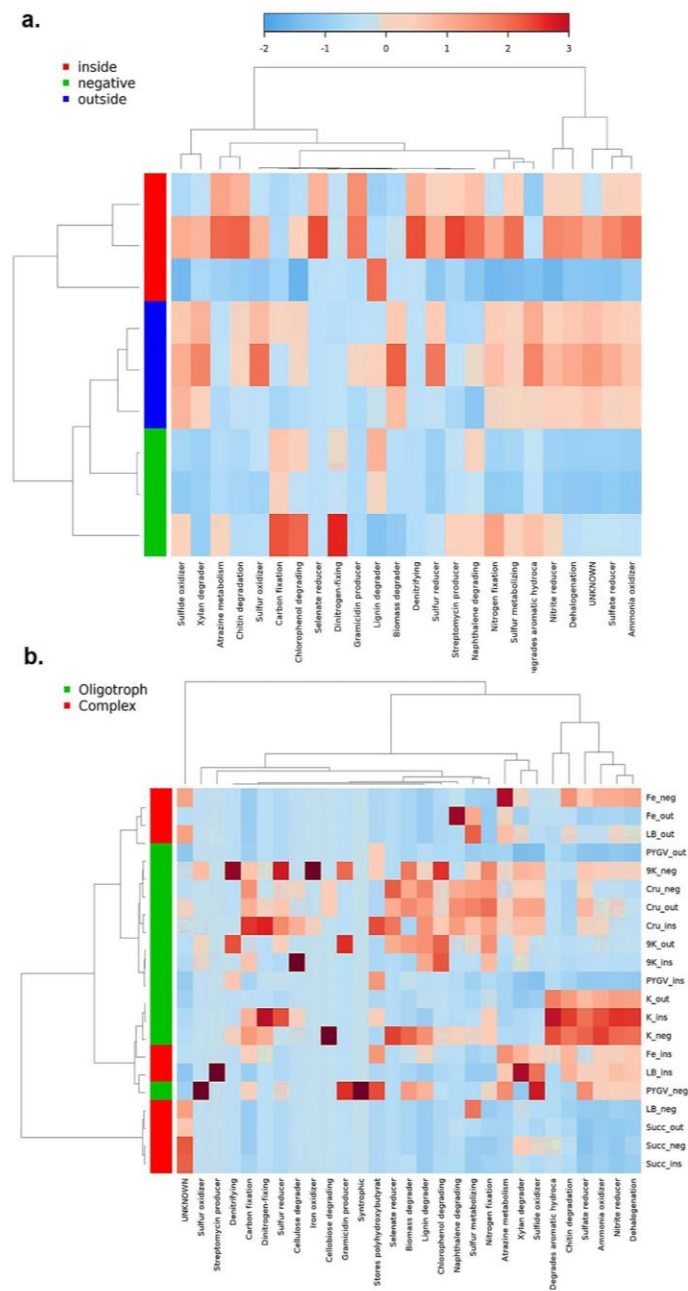
Figure 5. The metabolic profile of (**a**) environmental samples and (**b**) after the third round of enrichments from the Monte San Giorgio samples.

nitrotoluene degradation pathways were differentially abundant in the bituminous shale microbiome while the phenol hydroxylase, hydratase, 2 keto-4-pentenoate cobaltochelatase involved in the toluene, chlorocyclohexane, chlorobenzene, porphyrin, and chlorophyll metabolism were relatively abundant in the outside of the shale and limestone microbiome (Fig. 7). In order to know which enrichment media were more suitable for the selection and growth of microbes with bioremediation potential, the predicted key proteins and enzymes involved in the polycyclic aromatic hydrocarbon degradation pathways were analyzed in the oligotrophic and nutrient-rich complex media as described above. After the third round of enrichment, the oligotrophic media have selected microbes with key enzymes such as phenol hydroxylase, p-cumate 2,3-dioxygenase, toluene monooxygenase,

37

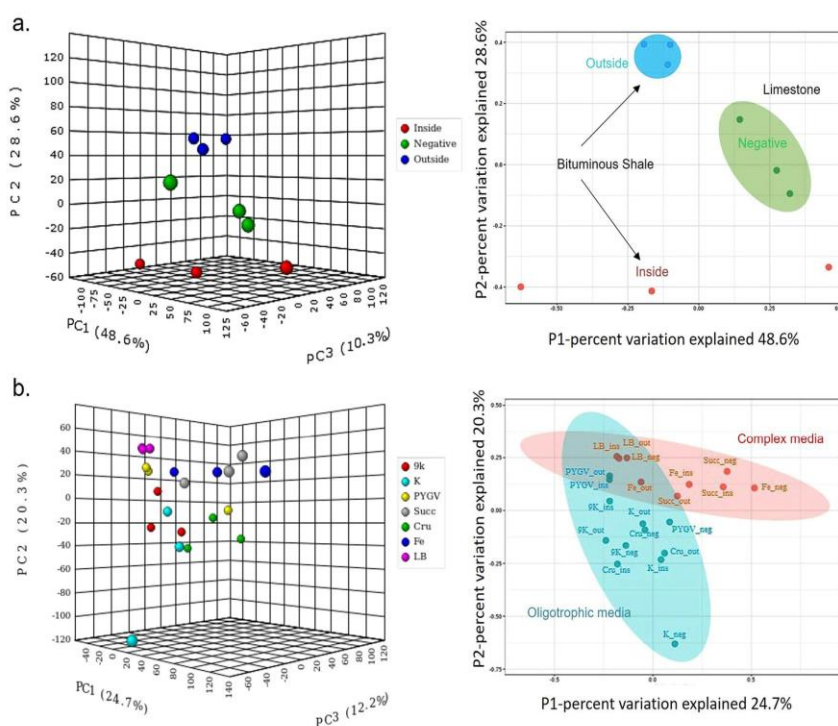**Figure 6.** PCoA of the (**a**) Monte San Giorgio environmental samples and (**b**) after the third round of enrichments based on gene abundance data (K numbers). See Fig. 2 for abbreviations.

p-hydroxybenzoate 3-monooxygenase, carbazole 1,9a-dioxygenase etc., potentially involved in the degradation of chlorocyclohexane, chlorobenzene, toluene, xylene, benzoate, naphthalene, and dioxin (Fig. 8).

Discussion

About 95 million years ago, the progressive northwards movement of the African plate compressed the Eurasian plate, raising the seabed about 1,100 meters over sea level resulting in the emergence of the pyramid-like mountain, Monte San Giorgio. Most impressive is the fossil- and organic-rich series of Triassic and Jurassic sedimentary rocks (limestones, dolomites, marls, bituminous shales, etc.). Understanding its extant geomicrobiology would also expand our knowledge of shale bioweathering and long-term carbon cycling. With the 16s rRNA gene amplification and sequencing advances, the metagenomic structure and function of microbial communities inhabiting its bituminous shale and non-bituminous (limestone) rocks were investigated. The geomicrobiology of the Monte San Giorgio shale is similar to that of oil shale by-products - despite their geologically different origins - from Brazil (Irati Formation)[23]. The identified phyla *Acidobacteria, Actinobacteria, Bacteroidetes, Firmicutes, Nitrospirae, Proteobacteria, Chloroflexi, Cyanobacteria, Deinococcus-Thermus, Spirochaetes Gemmatimonadetes, Planctomyces*, and *Thermotogae* were also present in the black shale samples of Chengkou County, China, which covered the bottom zone toward the surface regolith[24]. However, bacteria involved in Fe-S cycling (*Acidithiobacillus, Sulfobacillus, Thiobacillus, Ferrimicrobium,* and *Ferrithrix*) and in anaerobic conversion of hydrocarbons to methane and CO_2, catalyzed by syntrophic bacteria and methanogenic archaea, which are an important part of the dominant processes of black shale bioweathering, were absent in the Monte San Giorgio shale samples[15,24].

Moreover, halotolerant *Proteobacteria* and *Firmicutes* (*Idiomarina, Marinobacter, Marinobacterium, Halomonas, Vibrio* and *Halanaerobium*) occurring in microbial mats from oil shale hydraulic fracturing wastewater treatment[25–27] were also absent because the bioweathering of the Monte San Giorgio shale takes place under different topological and geochemical factors where desiccation stress at the exposed, terrestrial site and aerobic conditions should influence the microbial community. We suggest that the aerobic/facultative anaerobic orders (*Desulfurellales, Flavobacteriales, Rhizobiales, Xanthomonadales, Sphingobacteriales, Sphingomonadales, Caulobacterales, Bacillales, Lactobacillales,* and *Enterobacteriales*) and anaerobic orders (*Thermotogales, Thermoplasmatales, Desulfobacterales, Chromatiales, Desulfuromonadales, Hydrogenophilales, Campylobacterales, Rhodocyclales, Clostridiales*) found on the outside and inside surface of bituminous shale are responsible for the bioweathering and hydrocarbon degradation (Fig. S1)[15]. The majority of these orders belong to Proteobacteria. The sub-classes Alpha, Beta, Delta-, and Gammaproteobacteria have metabolisms capable of oxidizing a range of complex polymeric carbon compounds (including sugars, alcohols, organic acids, amino acids and carbohydrates) and complex hydrocarbons under aerobic and anaerobic conditions; underscoring the enhanced capacity

Figure 7. K numbers involved in the differentially abundant pathways (Ko numbers)[21] of the bituminous shales (inside), against its outside surface and limestone samples (Controls). Enzymes shared by different pathways are shown as overlapping color boxes.

for organic matter degradation[28–30]. The functional profiling predicted from the 16srRNA data also shed light on metabolic pathways with enzymes such as dioxygenases, thioesterases, hydrolases, aldoses, involved in the complex polycyclic hydrocarbon degradation that contributes to bioweathering of these bituminous shales. Of course, the validity of this assumption remains to be clarified and further studies are required to determine the underlying bioweathering mechanism.

In addition, we also studied how serial inoculum enrichment process enriches some microorganisms of the Monte San Giorgio shale communities over others. Choice of minimal media and extended incubation times facilitates the slow-growing, rarely isolated groups[31]. Complex media, such as LB, are preferred choice for the sake of convenience and their richness in carbon sources. Complex media allow the fast-growing copiotrophic microbes to consume nutrients and outgrow the slow growing chemoheterotrophic microbes, resulting in the substantial diversity loss. This general feature is also observed as a consequence of agricultural or other human activities, when increased input of nutrients in, e.g., water bodies, results in elevating the risk of eutrophication and reduced biodiversity[32,33]. Even just the nitrogen source addition can shift the bacterial diversity from oligotrophic towards a copiotrophic community[27,29,30,32,34,35], but if the nutrients are limited, the fastidious microbes cannot simply outgrow the slow growing microbes, for instance in the crude oil medium, where the carbon source is not easily accessible. Thus, the oligotrophic media provide stringent conditions that keep a check on the growth of fast-growing microbes but let the slow-growing microbes to grow and proliferate. Consequently, the diversity loss in the oligotrophic media is not as drastic as in the complex media. Additionally, oligotrophic media are more suitable for the selection and growth of slow-growing specialized microbes in bioremediation and polycyclic aromatic hydrocarbon degradation that also differentially expressed key enzymes such as phenol hydroxylase, p-cumate 2,3-dioxygenase, toluene monooxygenase, p-hydroxybenzoate 3-monooxygenase, carbazole 1,9a-dioxygenase involved in the chlorocyclohexane, chlorobenzene, toluene, xylene, benzoate, nitrotoluene, naphthalene, and dioxin degradation.
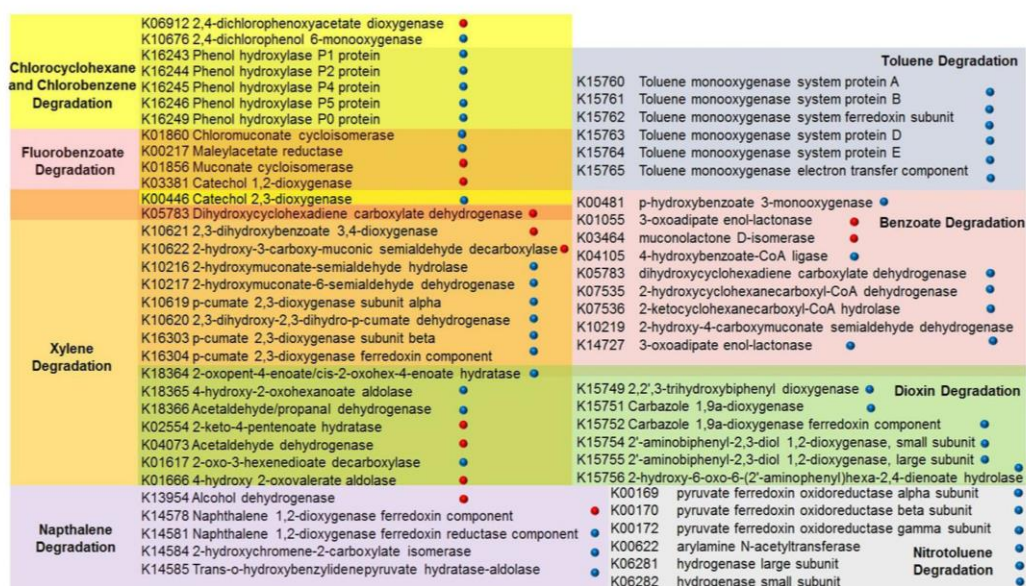
Figure 8. Differential abundance analysis of enriched microbes in the oligotrophic (●) and nutrient-rich media (●) after the third round of enrichments. Enzymes shared by different pathways are shown as overlapping color boxes.

Conclusion

Our study provides the first insight into modern microbial diversity and metabolic types of the bituminous shale and limestone rocks from Monte San Giorgio. It explores how different environmental conditions (organic-rich shale and nutrient-deprived limestone) aid in selection and habituation of the different microbiomes from the environment (soil contaminated outer surface of shales). Moreover, this study expands our understanding about serial subculturing of the microbial communities in general, which will improve our choice of media for enriching microbes with biotechnological and bioremediation potential. The computational tools anticipate further that the Monte San Giorgio shale microbiome has specialized metabolic pathways involved in the complex hydrocarbon degradation and the oligotrophic media are preferred for enriching such microorganisms with particular biochemical and bioremediation applications.

Material and Methods

Sampling site. Rock samples were collected in Jul-2017 from the site Acqua del Ghiffo (45°54′20.4″N 8°55′54.4″E), situated at Monte San Giorgio (Fig. S3). Samples were taken from an outcropping bedrock with a rock hammer and chisel and immediately wrapped in sterile aluminium foil until further use. The MSG-17-1 samples were organic-rich bituminous shale derived from the Cava superiore and Cava inferiore strata while MSG-17-2 were derived from non-bituminous Meride Limestone and used as the negative control for this study. Samples were not directly touched, in order to avoid contamination with organisms adhering to skin and clothes. If necessary, tools were treated with ethanol and cleaned with a sterile cloth prior to use. Care was taken that the sterilizing liquid was completely evaporated before sampling. The sampling sites, as well as the samples, did not have direct contact with a soil cover, in order to avoid direct contamination with soil particles and leaf litter. Rock masses of 700–1200 g were collected, in order to reduce desiccation of subsamples to be taken just prior to DNA extraction. Samples were kept wrapped at ambient temperature until further analysis. At least three samples of the MSG-17-1 and MSG-17-2 were used for the DNA extraction. Samples for DNA extraction (0.4 g each) were taken by scratching the outer surface of a rock piece (outside sample) with a sterile scalpel in the vicinity of a bunsen burner. The slabby rock was then carefully cleaved and the freshly exposed surface (exhibiting a characteristic diesel smell) was again sampled with a sterile scalpel (inside sample).

DNA extraction and Illumina sequencing of the Monte San Giorgio samples. The genomic DNA of the microbial community from inside and outside samples of the MSG-17-1 rock samples was extracted by the Powersoil DNA isolation kit (Qiagen, Venlo, The Netherlands) according to the manufacturer's instructions. For each sample, 0.4 g of scratched material was used to extract soil microbial genomic DNA while blanks were also assessed for each sample to eliminate possible reagents contamination with microbial DNA. Essentially, in DNA isolation with this kit, the total genomic DNA was captured on a silica membrane in a spin column after inhibitors removal and protein precipitation. The same procedure was performed for the limestone rocks which served as the negative control. Followed by washing and elution, the quality and quantity of the extracted genomic DNA were checked with gel electrophoresis using 0.8% agarose in Tris-acetate-EDTA (TAE) buffer, pH (8.3)[36] and photometrically in a Nanodrop ND-1000 spectrophotometer (PeqLab, Germany), respectively. No DNA contamination

was observed in the blanks after the DNA extraction and the subsequent PCR amplification. The variable regions V3-V4 of the 16s rRNA gene was amplified with the Illumina overhang adapters on their 5′ end of the 16S Amplicon Miseq PCR Forward 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3′ and Reverse 5′-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3′ primers[37] with the PCR profile with some modifications[38]. A GeneRead Size Selection Kit (Qiagen, Germany) was used to remove primers and dimers from the resulting amplicons. Indexing of these PCR products was performed with Nextera XT DNA library prep kit (Illumina, San Diego, Cal, USA) according to manufacturer's instructions. Paired-end sequencing was performed in collaboration with Göttingen Genomics Laboratory with an Illumina MiSeq sequencer (Illumina). Blanks were automatically discarded during the processing of sequences due to very low read counts.

Enrichment media on microbial communities.

The inoculum from the inside, outside, and negative MSG samples were enriched in different media (9 K[39,40], K[41], PYGV[42], succinate minimal media[43], minimal media supplemented with 1% crude oil[44], iron basal media, and LB media[45]) (Table S3). The culture media were re-inoculated after every week and after 3 cycles of inoculum subculture and incubation at room temperature with rotary shaking at 120 rounds per minute, a subset of microbes promoted by the specific growth medium is supposed to be fully enriched. The genomic DNA was extracted from the enrichments after 1st and 3rd round of incubation with the Powersoil DNA isolation kit (MoBio Laboratories) subjected to DNA extraction and sequencing as described above.

Metagenomic analysis via MetaAmp pipeline.

The Illumina amplicon sequencing data was fed to the online available MetaAmp automated pipeline for metagenomic analysis (http://ebg.ucalgary.ca/metaamp/)[46]. First, usearch -fastq_mergepairs assembled the demultiplexed and uncompressed fastq format sequence files as paired-end reads via USEARCH software package[47]. The read pairs shorter than 350 bp length, misaligned and mismatched in the overlap region were removed. Next, the forward and reverse primers were identified and trimmed in the Mothur software package via trim.seqs command[48]. The reads without the forward and reverse primers or with mismatches in the primer region were discarded. In the quality filtering step, the low-quality reads were removed to minimize the influence of sequencing errors using usearch -fastq_filter command with -fastq_trunclen -fastq_maxee -fastaout options in USEARCH. The high-quality reads were labeled with the unique sample ids and the reads from different files are pooled together. Next, the UPARSE software[49] was used to dereplicate reads, to discard singletons and chimaeras, and to cluster the pooled high-quality reads into operational taxonomic units (OTUs). The OTU clustering threshold is 0.97 that is 97% identity. The OTUs were assigned taxonomic status with the classify.seqs command in Mothur that utilize the reference SILVA training dataset (http://www.mothur.org/wiki/Taxonomy_outline). In the end, Mothur generated rank-abundance data, rarefaction curves, alpha-diversity indexes and beta-diversity, followed by rarefying samples into subsamples. In addition, principal coordinate analysis (PCoA) was also calculated using the Bray-Curtis index to compute dissimilarities among different samples. Hypothesis testing was done with the unifrac.weighted[19] and Analysis of Molecular Variance (AMOVA)[20] tests.

Functional profiling.

The taxonomic profile data obtained from the MetaAmp was further processed by the METAGENassist online Server to give an overview of the functional profiles of different samples[13]. The comparative metagenomic analysis was performed with the clustering algorithm ward in relation to interpret the metabolic profiles of the predicted microbiomes. More in-depth details of the enzymes and metabolic pathways involved were obtained from another recently developed online tool Piphillin (http://secondgenome.com/Piphillin)[14] which directly searches each representative FASTA sequence of the submitted OTUs against USEARCH version 8.0.1623[47]. Gene copy numbers for each genome were summarized with KO abundances, KEGG reference database[21]. The comprehensive statistical and meta-analysis including differential abundance analysis of gene abundance data was completed with the online tool MicrobiomeAnalyst[50] and its R packages DESeq2[22]. The significant features for bituminous shale (inside) against the outer surface of bituminous shale and limestone rock samples over 1-fold difference were considered to summarize the results.

References

1. Demelia, P. P. & Flaviani, A. Aspects of the palynostratigraphy of the triassic sardinian sequences (preliminary report). *Rev. Paleobot. Palynol* **37**, 329–343 (1982).
2. Rieber, H. P. Monte San Giorgio und Besano, mittlere Trias, Schweiz und Italien [Monte San Giorgio and Besano, Middle Triassic, Switzerland and Italy] in *Europäische* Fossillagerstätten (ed. Meischner, D.) 83–90 (Springer, 2000).
3. Seilacher, A. *et al.* Sedimentological, ecological and temporal patterns of fossil Lagerstätten. *Philos. Transact. Roy. Soc. London. B, Biol. Sci.* **311**, 5–24 (1985).
4. Rieber, H. Ergebnisse paläontologisch-stratigraphischer Untersuchungen in der Grenzbitumenzone (Mittlere Trias) des Monte San Giorgio (Kanton Tessin, Schweiz) [Results of paleontological-stratigraphic investigations in the Grenzbitumenzone (Middle Triassic) of Monte San Giorgio]. *Eclogae Geol. Helv.* **66**, 667–685 (1973).
5. Furrer, H. The Kalkschieferzone (Upper Meride Limestone; Ladinian) near Meride (Canton Ticino, southern Switzerland) and the evolution of intraplatform basin. *Eclogae Geol. Helv.* **88**, 827–852 (1995).
6. Stockar, R., Adatte, T., Baumgartner, P. O. & Föllmi, K. B. Palaeoenvironmental significance of organic facies and stable isotope signatures: the Ladinian San Giorgio Dolomite and Meride Limestone of Monte San Giorgio (Switzerland, WHL UNESCO). *Sedimentology* **60**, 239–269 (2013).
7. ter Voorde, M. & Bertotti, G. Thermal effects of normal faulting during rifted basin formation, 1. A finite difference model. *Tectonophysics* **240**, 133–144 (1994).

8. Bertotti, G. & ter Voorde, M. Thermal effects of normal faulting during rifted basin formation, 2. The Lugano-Val Grande normal fault and the role of pre-existing thermal anomalies. *Tectonophysics* **240**, 145–157 (1994).
9. Bertotti, G., Picotti, V., Bernoulli, D. & Castellarin, A. From rifting to drifting: tectonic evolution of the South-Alpine upper crust from the Triassic to the Early Cretaceous. *Sediment. Geol.* **86**, 53–76 (1993).
10. Stocker, M. R. & Butler, R. J. Phytosauria. *Geol. Soc. Spec. Publ* **379**, 91–117 (2013).
11. Sander, P. M. & Greenwood, P. H. The pachypleurosaurids (Reptilia: Nothosauria) from the Middle Triassic of Monte San Giorgio (Switzerland) with the description of a new species. *Philos. Transact. Roy. Soc. London. B, Biol. Sci* **325**, 561–666 (1989).
12. McEvoy, J. & Giger, W. Origin of hydrocarbons in triassic serpiano oil shales: hopanoids. *Org. Geochem.* **10**, 943–949 (1986).
13. Arndt, D. *et al.* METAGENassist: a comprehensive web server for comparative metagenomics. *Nucleic Acids Res* **40**, 88–95 (2012).
14. Iwai, S. *et al.* Piphillin: Improved prediction of metagenomic content by direct inference from human microbiomes. *PLOS ONE* **11**, e0166104 (2016).
15. An, D. *et al.* Metagenomics of hydrocarbon resource environments indicates aerobic taxa and genes to be unexpectedly common. *Environ. Sci. Technol.* **47**, 10708–10717 (2013).
16. Albuquerque, L. & da Costa, M. S. Actinobacteria in *The Prokaryotes* (eds. Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E. & Thompson, F.) 861–866 (Springer, 2014).
17. Garrity, G. M. & Holt, J. G. Deinococci class. nov. in *Bergey's Manual of Systematics of Archaea and Bacteria, vol. 1, 2nd ed.* (eds. Garrity, G. M., Noone, D. R., Castenholz, R. W) 395 (Springer, 2001).
18. Kopejtka, K. *et al.* Genomic analysis of the evolution of phototrophy among haloalkaliphilic Rhodobacterales. *Genome Biol. Evol* **9**, 1950–1962 (2017).
19. Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J. & Knight, R. UniFrac: an effective distance metric for microbial community comparison. *ISME J.* **5**, 169–172 (2011).
20. Mengoni, A. & Bazzicalupo, M. The statistical treatment of data and the analysis of MOlecular VAriance (AMOVA) in molecular microbial ecology. *Ann. Microbiol.* **52**, 95–101 (2002).
21. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2015).
22. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq. 2. *Genome Biol.* **15**, 550 (2014).
23. Goes, K. C. G. P. D., Lovato, G. M. & Andrade, D. S. Composition of bacterial community in enrichment cultures of shale by-products from Irati Formation, Brazil. *Braz. J. Microbiol.* **49**, 742–748 (2018).
24. Li, J. *et al.* Bacteria diversity, distribution and insight into their role in S and Fe biogeochemical cycling during black shale weathering. *Environ. Microbiol.* **16**, 3533–3547 (2014).
25. Murali Mohan, A., Hartsock, A., Hammack, R. W., Vidic, R. D. & Gregory, K. B. Microbial communities in flowback water impoundments from hydraulic fracturing for recovery of shale gas. *FEMS Microbiol. Ecol.* **86**, 567–580 (2013).
26. Mouser, P. J., Borton, M., Darrah, T. H., Hartsock, A. & Wrighton, K. C. Hydraulic fracturing offers view of microbial life in the deep terrestrial subsurface. *FEMS Microbiol. Ecol.* **92**, fiw166 (2016).
27. Akyon, B., Stachler, E., Wei, N. & Bibby, K. Microbial mats as a biological treatment approach for saline wastewaters: the case of produced water from hydraulic fracturing. *Environ. Sci. Technol.* **49**, 6172–6180 (2015).
28. Robinson, G. *et al.* Profiling bacterial communities associated with sediment-based aquaculture bioremediation systems under contrasting redox regimes. *Sci. Rep* **6**, 38850–38850 (2016).
29. Couradeau, E. *et al.* Prokaryotic and eukaryotic community structure in field and cultured microbialites from the alkaline Lake Alchichica (Mexico). *PLOS ONE* **6**, e28767 (2011).
30. Marín, I. Proteobacteria in *Encyclopedia of Astrobiology* (eds. Amils, R. *et al.*) 2036–2037 (Springer, 2014).
31. Davis, K. E. R., Joseph, S. J. & Janssen, P. H. Effects of growth medium, inoculum size, and incubation time on culturability and isolation of soil bacteria. *Appl. Environ. Microbiol.* **71**, 826–834 (2005).
32. Suding, K. N. *et al.* Functional- and abundance-based mechanisms explain diversity loss due to N fertilization. *Proc. Natl. Acad. Sci. USA* **102**, 4387–4392 (2005).
33. Ponce-Soto, G. Y. *et al.* Enrichment experiment changes microbial interactions in an ultra-oligotrophic environment. *Front. Microbiol* **6**, 246 (2015).
34. Cleland, E. E. & Harpole, W. S. Nitrogen enrichment and plant communities. *Ann. N. Y. Acad. Sci* **1195**, 46–61 (2010).
35. Fierer, N. *et al.* Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J.* **6**, 1007–1017 (2012).
36. Sambrook, J. & Russell, D. W. Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor Laboratory Press, 2001).
37. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1 (2012).
38. Amin, N., Schneider, D. & Hoppert, M. Bioleaching potential of bacterial communities in historic mine waste areas. *Environ. Earth Sci* **77**, 542 (2018).
39. Silverman, M. P. & Lundgren, D. G. Studies on the chemoautotrophic iron bacterium *Ferrobacillus ferrooxidans*. I. An improved medium and a harvesting procedure for securing high cell yields. *J. Bacteriol.* **77**, 642–647 (1959).
40. Yates, J. R. & Holmes, D. S. Two families of repeated DNA sequences in *Thiobacillus ferrooxidans*. *J. Bacteriol.* **169**, 1861–1870 (1987).
41. Rosson, R. A., Tebo, B. M. & Nealson, K. H. Use of poisons in determination of microbial manganese binding rates in seawater. *Appl. Environ. Microbiol.* **47**, 740–745 (1984).
42. Ghiorse, W. C. & Hirsch, P. Isolation and properties of ferromanganese-depositing budding bacteria from Baltic Sea ferromanganese concretions. *Appl. Environ. Microbiol.* **43**, 1464–1472 (1982).
43. Kepkay, P. E. & Nealson, K. H. Growth of a manganese oxidizing *Pseudomonas sp.* in continuous culture. *Arch. Microbiol.* **148**, 63–67 (1987).
44. Liu, Z., Jacobson, A. M. & Luthy, R. G. Biodegradation of naphthalene in aqueous nonionic surfactant systems. *Appl. Environ. Microbiol.* **61**, 145–151 (1995).
45. Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293–300 (1951).
46. Dong, X. *et al.* Fast and simple analysis of MiSeq amplicon sequencing data with MetaAmp. *Front. Microbiol* **8**, 1461 (2017).
47. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
48. Schloss, P. D. *et al.* Introducing mothur: open-Source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).
49. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Meth* **10**, 996–998 (2013).
50. Dhariwal, A. *et al.* MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucl. Acids Res.* **45**, W180–W188 (2017).

42

Author contributions

M.H. and J.R. designed the study and collected the samples. S.A. performed the experiments, analyzed data, interpreted the results, and wrote the manuscript. M.H. contributed reagents/material/analysis tools, interpreted the results and corrected the manuscript. J.R. critically reviewed and corrected the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-55955-5.

Correspondence and requests for materials should be addressed to S.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

43

Supplementary Information

Composition, Diversity and Functional Analysis of the modern Microbiome of the middle Triassic Cava Superiore Beds (Monte San Giorgio, Switzerland)

Sania Arif[1*], Joachim Reitner[2], Michael Hoppert[1],

[1] Institute of Microbiology & Genetics, Department of General Microbiology; [2]Geosciences Centre, Department of Geobiology; Georg-August-Universität Göttingen

*Correspondence to sania.arif@stud.uni-goettingen.de

Figure S1, Figure S2, Figure S3, Table S1, Table S2, Table S3

Supplementary Information
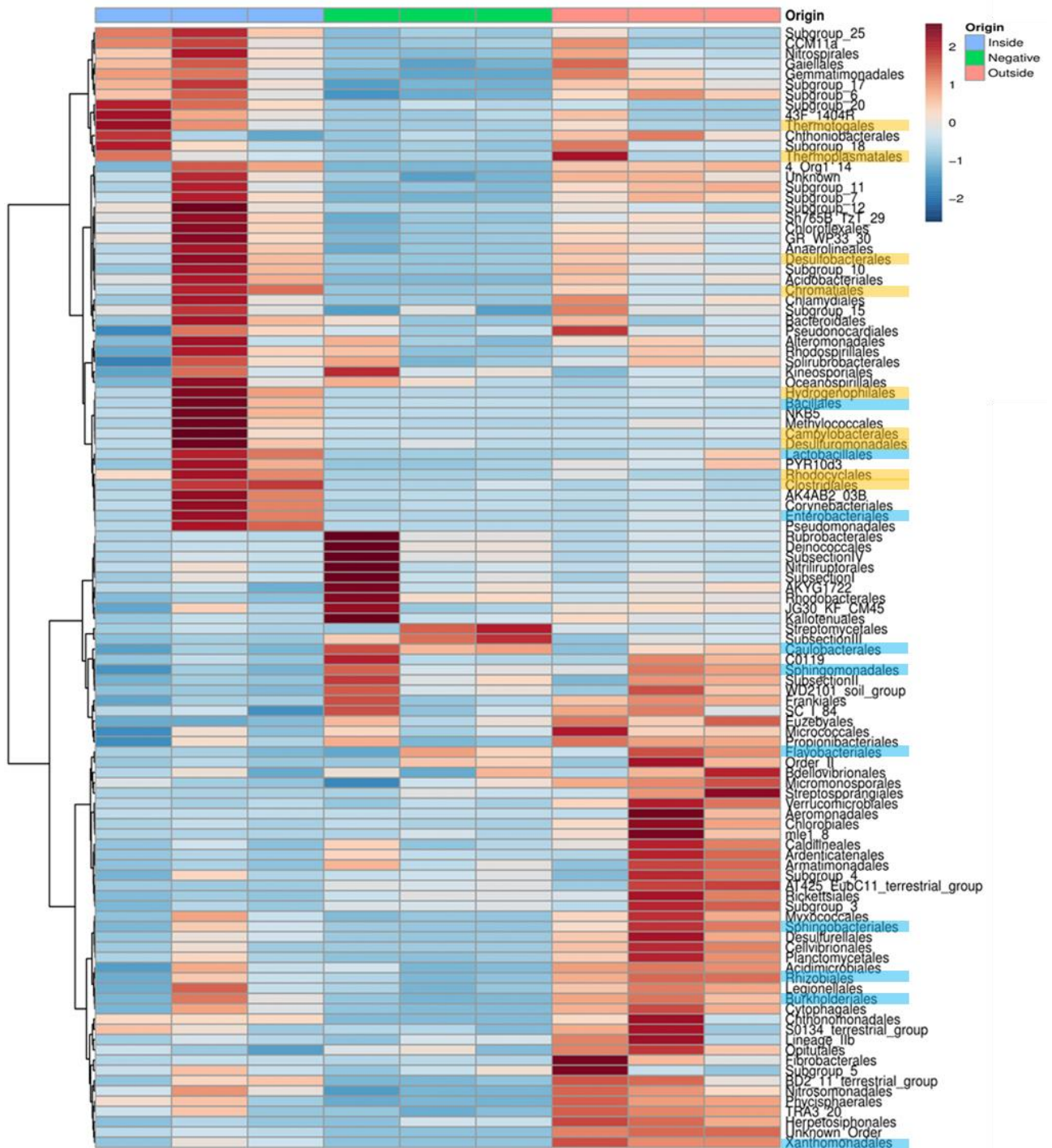
Figure S1. Heatmap of the Monte San Giorgio geomicrobiology at the order level. The known already reported anaerobic and the aerobic/facultative orders involved in the hydrocarbon metabolism are highlighted in yellow and blue respectively.
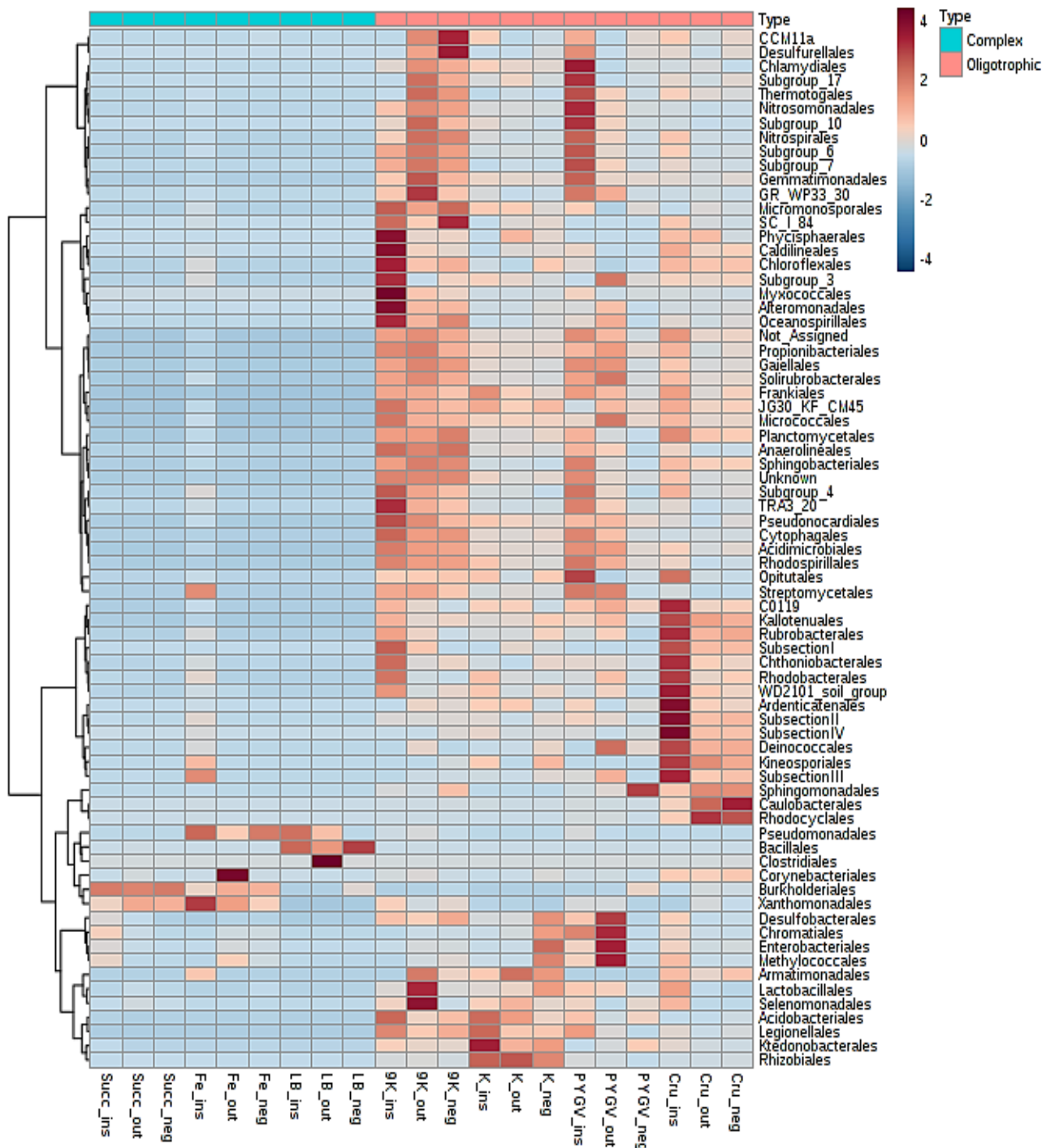
Figure S2. Heatmap of the orders, enriched in the nutrient-rich complex and oligotrophic media from Monte San Giorgio rock samples, after the third and final round of enrichment.
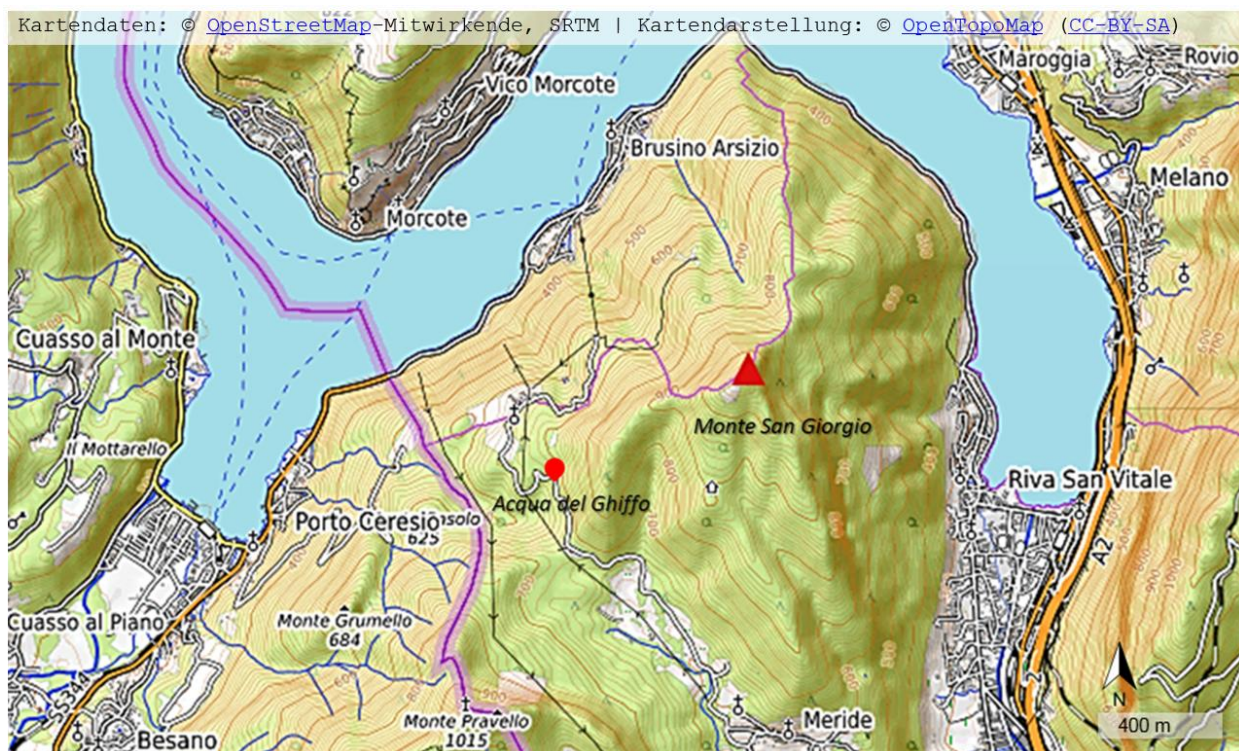
Figure S3. Map of sampling site (Acqua del Ghiffo) in the Monte San Giorgio mountain range. The map was reproduced from OpenTopoMap (https://opentopomap.org), under creative common license CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0/legalcode), subsequently marked with the sampling site and area.

Table S1. Alpha diversity indexes of the Monte San Giorgio environmental samples

	Observed OTUs	Chao1	ACE	Shannon	Simpson
Outside	2712+141.5	3105+55	3175+49	6.5+0.005	0.004995 +0.000087
Inside	2012+65	2522+18	2575+16	6.1+0.005	0.0071915+ 0.0001615
Negative	1991+18	2440+65	2483+62	6.04+0.017	0.007062+0.000346

Table S2. Analysis of molecular variance and unifrac analysis of the Monte San Giorgio environmental samples

Analysis of Molecular Variance (AMOVA)

C-P-T	Among	Within	Total
SS	0.719341	0.853931	1.57327
df	2	6	8
MS	0.359671	0.142322	

Fs: 2.52716
p-value: 0.005*

C-P	Among	Within	Total
SS	0.378582	0.537901	0.916483
df	1	4	5
MS	0.378582	0.134475	

Fs: 2.81526
p-value: 0.099

C-T	Among	Within	Total
SS	0.445195	0.636744	1.08194
df	1	4	5
MS	0.445195	0.159186	

Fs: 2.7967
p-value: 0.101

P-T	Among	Within	Total
SS	0.255234	0.533217	0.788451
df	1	4	5
MS	0.255234	0.133304	

Fs: 1.91468
p-value: 0.086

Unifrac.weighted

Tree#	Groups	WScore	WSig
1	C-P	1.000000	<0.0010
1	C-T	1.000000	0.0190
1	P-T	0.889840	<0.0010

C Negative control **T** Outside **P** Inside

Table S3. Details of Enrichment Media

Medium	Components per litre distil H$_2$O	Features	Growth	References
9K	(NH$_4$)$_2$SO$_4$, KCl, K$_2$HPO$_4$, MgSO$_4$·7H$_2$O Ca(NO$_3$) $_2$, FeSO$_4$·7H$_2$O H$_2$SO$_4$ (1N)	Acidic medium	Iron-oxidizing bacteria	Silverman, M.P. & Lundgren, D.G. 1959 [35] Yates, J.R. & Holmes, D.S. 1987 [36]
K	MnSO$_4$·H$_2$O, FeSO$_4$·7H$_2$O, peptone, yeast extract, and 10 mM HEPES buffer, pH 7.5	Manganese containing complex media	Heterotrophic Mn oxidizers	Rosson, R.A., Tebo, B.M. & Nealson, K.H. 1984 [37]
PYGV	Peptone, Yeast Extract, Mineral Salt and Vitamin Solution	Low-nutrient medium	Slow growing and sheathed bacteria	Ghiorse, W.C. & Hirsch, P. 1982 [38]
Succinate Minimal media	Succinate, FeSO$_4$, MnCl$_2$, NHCl, KCl, K$_2$HPO$_4$, Hepes Buffer, Vitamin Solution	Basal minimal medium	Mn binding and metabolism	Kepkay, P.E. & Nealson, K.H. 1987 [39]
Fe basal media	Na$_2$HPO$_4$·12H$_2$O, KH$_2$PO$_4$, NH$_4$Cl, MgSO$_4$·7H$_2$O, CaCl$_2$·2H$_2$O, Fe(III)-NH$_4$-Citrate, SL-10 Trace Elements Solution	Sodium acetate as carbon source	Fe(III)-NH$_4$-Citrtate as electron acceptor	Customized for this study
Crude oil with Minimal Media	MgSO$_4$, CaCl$_2$, KH$_2$PO$_4$, K$_2$HPO$_4$, NH$_4$NO$_3$, and FeCl$_3$, 1% crude oil and pH adjusted to 7-7.2	Complex organics containing media	Crude oil degrading Bacteria	Liu, Z., Jacobson, A.M. & Luthy, R.G. 1995 [40]
Luria-Bertani LB medium	Tryptone, yeast extract and NaCl	Nutritionally-rich medium	Fast growing bacteria, fastidious protein degraders	Bertani, G. 1951 [41]

3. Composition and Niche-Specific Characteristics of Microbial Consortia Colonizing a Copper Mine in the Rhenish Massif

Sania Arif[1*], Heiko Nacke[2], Elias Schliekmann[1], Andreas Reimer[3], Gernot Arp[3], and Michael Hoppert [1]

[1]Department of General Microbiology, Institute of Microbiology and Genetics,

[2]Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics,

[3]Geoscience Centre, Department of Geobiology, Georg-August-Universität Göttingen, Germany

Author Contributions:

Conceptualization, M.H., H.N., and S.A; Methodology, S.A., M.H., E.S., A.R., G.A., H.N.; Sample collection; M.H., S.A., E.S.; Formal analysis; S.A., E.S., A.R., G.A., H.N.; Writing—original draft, S.A., G.A., H.N..; Writing—review and editing, M.H., S.A., H.N., A.R., and G.A..; Funding acquisition, M.H, A.R., and S.A.

Composition and Niche-Specific Characteristics of Microbial Consortia Colonizing Marsberg Copper Mine in the Rhenish Massif

Sania Arif[1], Heiko Nacke[2], Elias Schliekmann[1], Andreas Reimer[3], Gernot Arp[3], and Michael Hoppert[1]

[1]Department of General Microbiology, Institute of Microbiology and Genetics, George August Universität, Göttingen, 37077, Germany
[2]Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, George August Universität, Göttingen, 37077, Germany
[3] Geoscience Centre, Department of Geobiology, Georg-August-Universität Göttingen, 37077, Germany

Correspondence to: Sania Arif (sarif@gwdg.de)

Abstract. The Kilianstollen Marsberg (Rhenish Massif, Germany) has been extensively mined for copper ores, dating from Early Medieval Period till 1945. The exposed organic-rich alum shale rocks influenced by the diverse mine drainages at an ambient temperature of 10 °C could naturally enrich biogeochemically distinct heavy metal resistant microbiota. This metagenomic study evaluates the microbially colonized subterranean rocks of the abandoned copper mine Kilianstollen to characterize the colonization patterns and biogeochemical pathways of individual microbial groups. Under the selective pressure of the heavy metal contaminated environment at illuminated sites, *Chloroflexi* (*Ktedonobacteria*) and *Cyanobacteria* (*Oxyphotobacteria*) build up whitish-greenish biofilms. In contrast, *Proteobacteria*, *Firmicutes* and *Actinobacteria* dominate rocks around the uncontaminated spring water streams. The metagenomic analysis revealed that the heavy metal resistant microbiome was evidently involved in redox cycling of transition metals (Cu, Zn, Co, Ni, Mn, Fe, Cd, Hg). No deposition of metals or minerals, though, was observed by transmission electron microscopy in *Ktedonobacteria* biofilms which may be indicative for the presence of different detoxification pathways. The underlying heavy metal resistance mechanisms, as revealed by analysis of metagenome-assembled genomes, were mainly attributed to transition metal efflux pumps, redox enzymes, volatilization of Hg^0, methylated intermediates of As(III) and reactive oxygen species detoxification pathways.

Key words

Copper mine, Ktedonobacteria, Rhenish Massif, Metagenomics, Heavy metal detoxification, Metagenome-Assembled Genome, Functional Profiling

1

1. Introduction

The historic copper mining area Marsberg is situated on the north-eastern edge of the Rhenish Schiefergebirge (Rhenish
30   Massif) which is composed of Variscan folded rocks of Devonian and Carboniferous age (Urban et al., 1995). The Marsberg
Upper Devonian sequence mainly consists of metamorphic clay shales, sandstones, siltstones and carbonate rocks, whilst the
Lower Carboniferous rocks contain a copper rich black shale series (Siegmund et al., 2002). Investigations of the Marsberg
copper ore deposits revealed insights in their geology, ore formation and recent re-mineralizations (Stribrny, 1987). The
copper-rich sediments formed about 380 million years ago in the Devonian on the southern edge of the Laurussia continent
35   (America and Europe). The Marsberg copper deposit originated from tectonic movements which caused disintegration of the
Lower Carboniferous Alum Shales and lydites and exposed the Upper Devonian rocks, resulting in fissures, faults and
breccia rich in metals (e.g., 7-16% Cu), sulphides (0.5-3.8%), carbonate carbon (0.35- 2.4 3%), and organic carbon (0.3-
2.5%) (Urban et al., 1995). In rock samples, mean contents of copper (81-1277 ppm) are found to be higher than those of
other metals (Pb 36-417, Zn 78-660, Co 34-63, Ni 79-450, V 49-160, and Cr 45-193 ppm) (Urban et al., 1995). The Upper
40   Devonian to Lower Carboniferous rocks are completely exposed in Marsberg Kilianstollen copper mine, offering conditions
for formation of diverse secondary minerals and mine drainages. The most important sulphide ore minerals present in fault
and fault-related breccia zones are described as chalcopyrite, neodigenite, chalcocite, bornite, and covellite (Stribrny, 1987).
These minerals in the sediments all shape the prevailing biogeochemical conditions in the Marsberg mine waters. The
biological and atmospheric oxidation of sulphides from pyrite (FeS_2) or chalcopyrite ($CuFeS_2$) could mobilize transition
45   metals (Amin et al., 2018). The reduced transition metals (Cu, Fe, Mn) from fault-bound breccias and black alum shales are
oxidized, resulting in copper-rich (acidic) sulphidic mine waters with high concentrations of copper and iron, but also other
dissolved ions (Fig. 1, green color). In addition, the fresh groundwater (flowing in NE direction) is enriched with calcium
carbonate from the Upper Permian Zechstein limestone, while sulphate is derived from gypsum and anhydrite of the same
formation group (Fig. 1, blue color). The calcium and sulphate levels in these mine waters can be higher than in fresh water,
50   up to 2/3 of the values for sea water. These naturally flowing water streams of the Kilianstollen mine ranging from fresh
water to heavy metal enriched acidic leachate offers unique subsurface cold heavy metal enriched habitat to study the
colonized microbial communities under influence of various mine waters.

2
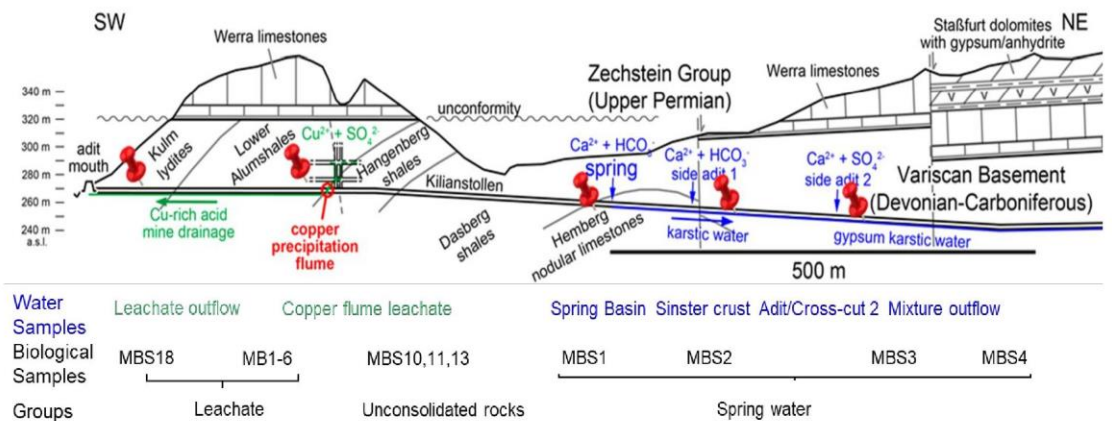
Figure 1. Simplified geological section of the Kilianstollen (Marsberg) and formation of different ground waters. Folded
55   Devonian-Carboniferous lydites, shales, and nodular limestones of the Variscan Basement are unconformably overlain by
Upper Permian carbonates and evaporites of the Zechstein Group. The reduced transition metals (Cu, Fe, Mn) from fault-
bound breccias and black shales are oxidized, resulting in copper-rich (acidic) sulphidic mine waters (Emmerich, 1987),
draining towards the adit entrance in SW direction. Some of the mine waters (flowing in NE direction) are enriched in
calcium, carbonate, and sulphate ions, depending on the rocks (limestone and gypsum) in contact. The geological section is
60   based on (Oskar and William, 1936; Farrenschon et al., 2008; Stribrny, 1987; Stribrny and Urban, 2000), and an unpublished
mining map. The red pins mark the location of the sampling sites along the mine drainage system for both water and
biological samples collections. The biological samples were further grouped into three groups based on their origin and
nearby waters.

65   An ambient temperature of 10 °C, a relative humidity of 98% and an appropriate abundance of organics (2-10%) in the alum
shale, these organic-rich copper shales also provide microscale spaces for microbial colonization and aromatic compounds
catabolism, perhaps because of their high content of kerogen, providing partially complex organic biodegradable compounds
such as long-chain and polycyclic aromatic hydrocarbons, esters, organic acids, thiophenes and metalloporphyrins (Dziewit
et al., 2015). The availability of the soluble sulphate (SO_4^{2-}) and transition metals ions from the nearby sulphuric waters
70   sources (Silver and Walderhaug, 1992) are important in shaping an epilithic but also heavy metal and/or acid-tolerant
bacterial community. Nevertheless, the emissions of the operational Marsberg mine railway diesel engine also provide
another resource of organic compounds and the regular visits of tourists, artificial ventilation and illumination are some man-
made impacts on the native microbial consortium. In this study, prokaryotic communities associated with the rocks around
the Marsberg Kilianstollen mine waters were metagenomically evaluated to observe whether the mine waters enriched in
75   transition metals may be toxic to microbial inhabitants or, conversely, support unique forms of metal respiration and enrich
resistant microbial consortia under oligotrophic conditions. To elucidate further key processes involved in their resistance

3

against high transition metal concentration and metabolism of the aromatic compounds, the metagenomically assembled genomes (MAGs) from a biofilm (MB1) nearby the copper-rich (acidic) sulphidic mine waters were assembled and analysed for the genetic targets related to toxic Hg(I) and As compounds reduction, Cu(I) oxidation, heavy metal ions extrusion,

80 dehalogenation, and hydrocarbon compounds catabolism. Understanding the selective pressure exerted by heavy metals on microbes and corresponding microbial resistance mechanisms could unveil their biogeochemical consequences and applications.

2. Materials and Methods

85 ### 2.1 Sampling site

The rock samples colonized with soft biomass/biofilm around mine drainages of the Kilianstollen, Marsberg, Germany (51.453502°N, 8.861703°E) were collected under sterile precautions: Hammer and chisel were disinfected with 70% ethanol prior to use, gloves were worn during sampling in order to reduce risks of contamination (Fig. 1, S1). The samples were taken from two locations in the alum shale region where mining activity was particularly high. For statistical analysis, the

90 samples (MB1-6 and MBS18), (MBS10,11,13) and (MBS1-4) were divided into leachate, unconsolidated rocks and spring water groups based on the nearby drainage water bodies and some of the representative samples are shown in Fig. 2. A detailed description of the biological features and origin of the collected samples has been published previously (Arif et al., 2021b). The biofilm and water samples were taken along the natural flowing water streams in the Kilianstollen mine, for instance, the biofilms (MBS1-4 grouped as spring water biofilms) were collected around the spring freshwater stream which

95 was gradually enriched in calcium, carbonate, and sulphate ions with the flow streams from the crosscut that branches off from the rear course of the mine (Fig. 1). The biological samples under the influence of the copper-rich acidic leachate stream were taken from the immediate vicinity of a copper precipitation flume 'Zement-Kupferplatte' (MB1-6 biofilms) and directly from the entrance of mine outflow stream (MBS18), were grouped as leachate samples (Fig. 1, 2). The copper precipitation flume is an iron plate which is continuously flooded by the copper-rich acidic leachate stream and the copper is

100 being deposited electrochemically. The other samples from a wooden plank (MBS10) next to mine unconsolidated rocks (MBS11, 13) (Fig. 2) were grouped as unconsolidated samples. Samples collected for microscopic analysis were refrigerated until preparation. For extraction of metagenomic DNA, freshly collected samples were stored at -20 °C till further use. The samples (0.4 g each in triplicates) for DNA extraction were obtained by scratching the biofilms of the rock piece and wooden plank with a sterile scalpel.

4
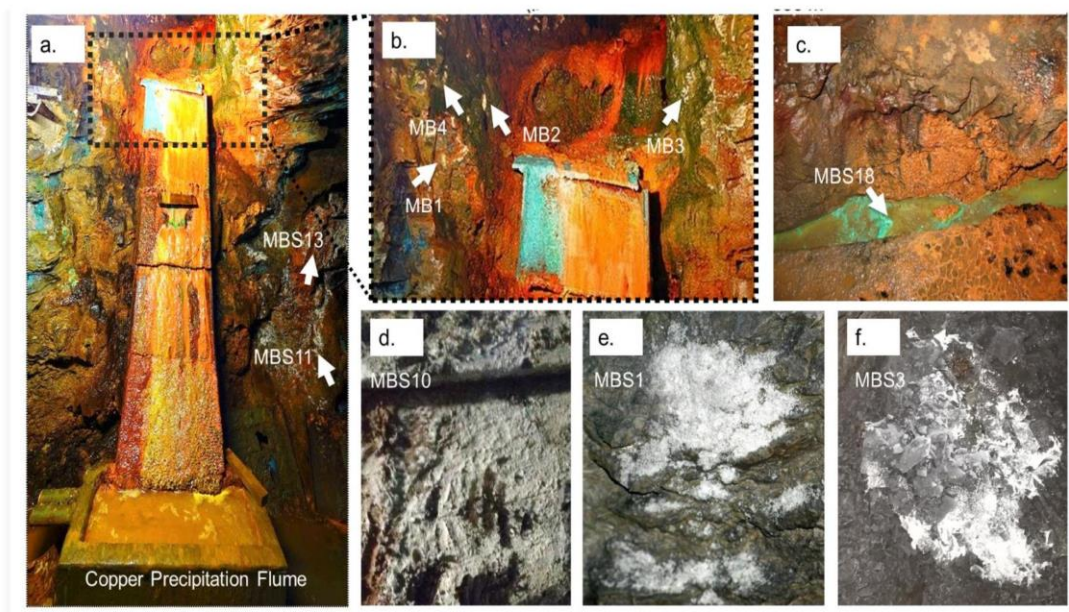
105

Figure 2. Kilianstollen biological samples. Samples were taken from sites in the vicinity of a copper precipitation flume (a and b) that is drained actively with the leachate water, from a wooden plank (d) located just right to the copper flume, directly from the outflow stream (c) near the opening and from rocky mine walls (e and f) exposed to the underlying groundwater and pit water stream at the rear end crosscut of the Kilianstollen.

110 **2.2 Hydro-chemical analysis**

Mine drainage water around sampling sites was analysed on site via a WTW Tetracon 925 conductivity probe equipped with a WTW Multi 3430 device, a WTW Sentix 940 electrode for pH and temperature, and a Schott PT61 redox electrode (Xylem, Rye Brook, NY, USA); calibration was performed with pH buffers 12.000, 10.010, and 7.010 (HI6012, HI6010 and HI6007, Hanna Instruments, RI, USA).

115 To determine the anions, cations, total organic carbon (TOC), total alkalinity (TA) and dissolved inorganic carbon (DIC), water samples were collected without headspace in Schott-Duran glass bottles (Schott, Mainz, Germany) and polyethylene (PE) bottles. Samples considered for cation analysis were filtered in separate 50 mL aliquots through 0.7 μm pore membrane filters and acidified with 100 μl HNO_3 (Suprapur; Merck, Darmstadt, Germany). For determination of total sulphide (ΣH_2S), aliquots were fixed with Zn-acetate. Total alkalinity (TA) was determined by acid-base titration within 2 hrs after sampling

120 using a hand-held titration device and 1.6 N H_2SO_4 cartridges as titrant (Hach Lange GmbH, Düsseldorf, Germany). All other samples were processed within 24 hrs after sampling.

Main cations (Li^+, Na^+, K^+, and Mg^{2+}, Ca^{2+}, Sr^{2+}) and anions (F^-, Cl^-, Br^-, NO_3^- and SO_4^{2-}) were analysed by ion chromatography with non-suppressed and suppressed conductivity detection (Metrohm 820 IC, Metrohm 883 Basic IC;

5

Metrohm, Herisau), respectively. The concentrations of NH_4^+, NO_3^-, PO_4^{3-}, ΣH_2S and dissolved silica were determined

125   photometrically, using a SI Analytics Uviline 9400 spectrophotometer. Dissolved inorganic carbon was determined with a TOC-L CPH analyser (Shimadzu, Kyoto, Japan). The PHREEQC software (version 3.5.0, 2019; (Parkhurst and Appelo, 2013)), with the phreeqc.dat and wateqf4.dat databases was used for calculation of ion activities, pCO_2 (partial pressure of CO_2) of samples and mineral saturation states. Saturation is given as $SI = \log (IAP/KSo)$, where IAP denotes ion activity product and KSo solubility product of the mineral phase.

130   **2.3 DNA extraction, 16S rRNA gene amplification and amplicon sequencing**

The microbial genomic DNA (gDNA) from 0.4 g scratched samples were extracted by using the DNeasy PowerSoil kit (Qiagen, Venlo, the Netherlands) as per manufacturer's instructions. In brief, the total gDNA released from cell lysis was treated for inhibitors removal and protein precipitation, then captured on and subsequently eluted from a silica membrane of a spin column. Blanks were also processed in addition to each sample to estimate DNA contamination. Following elution,

135   the extracted gDNA was visually observed with 0.8% agarose gel electrophoresis using TAE buffer, pH (8.3) (Sambrook and Russell, 2001) and photometrically quantified using a Nanodrop ND-1000 spectrophotometer (PeqLab, Erlangen, Germany). No gDNA contamination was measured in the blanks following the DNA extraction.

For Illumina MiSeq sequencing, V3-V4 hypervariable regions of 16S rRNA genes were amplified via polymerase chain reaction (PCR) and tagged to 5′ overhang adapter sequences (underlined) with the aid of MiSeq 16S amplicon PCR forward

140   primer 341F 5′-<u>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG</u>CCTACGGGNGGCWGCAG-3′ and reverse primer 805R 5′-<u>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG</u>GA CTACHVGGGTATCTAATCC-3′) (Klindworth et al., 2012). PCR reaction mixture (Amin et al., 2018) was modified to obtain a final volume of 50 µl in double-distilled nuclease-free water by mixing 5 × Phusion GC Buffer (10 µl), 10 µM forward and reverse primer (1.0 µl each), 10 mM dNTPs (1.0 µl), 5% DMSO (v/v, 2.5 µl), 50 mM $MgCl_2$ (0.15 µl), 0.5 µl of 2 U/µl Phusion HF DNA polymerase (Thermo-

145   Fisher Scientific, Waltham, MA, USA) and 25 ng template DNA (2.0 µl). The PCR profile comprised preheating at 94 °C for 3 min followed by 25 cycles of heating at 94 °C for 45 s, annealing at 60 °C for 45 s and extension at 72 °C for 30 s. The reaction ended with a final elongation step at 72 °C for 5 min. After PCR amplification, the amplicons were visually assessed with gel electrophoresis using 1.3% (w/v) agarose in 1× TAE buffer (Thermo-Fisher Scientific), pH 8.3 (Sambrook and Russell, 2001), and photometrically quantified in a Nanodrop ND-1000 spectrophotometer (PeqLab). The subsequent

150   purification was performed with the GeneRead Size Selection Kit (Qiagen) to remove primers and PCR reagents. After indexing of these PCR amplicons using the Nextera XT DNA library prep kit (Illumina, San Diego, CA, USA), paired-end sequencing was performed with an Illumina MiSeq sequencer in collaboration with the Göttingen Genomics Laboratory.

2.3 Amplicon sequencing data processing

The Illumina amplicon sequencing data was processed online by employing the automated pipeline for metagenomic

155   analysis MetaAmp (http://ebg.ucalgary.ca/metaamp/) (Dong et al., 2017). Using USEARCH software package, the

6

demultiplexed Fastq format sequence files were assembled as paired-end reads (Edgar, 2010). The misaligned and mismatched reads and paired end reads shorter than 350 bp length were discarded. Next, the primers were trimmed based on the Mothur software package (Schloss et al., 2009), and the reads without primers or with mismatched primer regions were removed. To minimize sequencing errors, the low-quality reads were scrapped using USEARCH. Dereplication, removal of

160 singletons and chimeras, and clustering of pooled high-quality reads into operational taxonomic units (OTUs) on the basis of 97% identity was done by UPARSE software (Edgar, 2013). The taxonomic status of the OTUs was assigned via Mothur by using the SILVA v138 as a reference (Glöckner, 2019). The taxonomic profile obtained from MetaAmp was further fed to the Microbiome online R-based server to plot respective graphs (Arndt et al., 2012). Based on the Bray-Curtis index, in principal coordinates analysis (PCoA) the data representing similarities of complex microbial communities were plotted into

165 2D and 3D graphs. Amplicon sequencing data has been published to the Sequence Read Archive (SRA) as SRR12876542-SRR12876555 under the Project accession number PRJNA670497 as announced previously (Arif et al., 2021b).

16S rRNA gene sequences related to already published *Ktedonobacteria* and *Actinobacteria* strains were included for phylogenetic analysis. In a first step, sequences belonging to *Ktedonobacteria* or *Actinobacteria*-related OTUs were aligned with the already published sequences through MUSCLE, implemented in MEGA-X software, by using default settings

170 (Stecher et al., 2020). Next, following the Kimura 2-parameter model, phylogenetic analyses and molecular evolutionary distances were calculated. The phylogenetic trees were constructed using the maximum likelihood algorithm and 1,000 bootstrap samplings to test tree topology.

2.4 Microscopy

The topographical features of the sampled biofilms (MB1-6) from the leachate group were observed using a Motic SMZ-171

175 stereo microscope (Motic GmbH, Germany) equipped with a Canon A650 camera. For transmission electron microscopy (TEM), the biofilm MB2 specimens were washed with phosphate buffered saline (50 mM, pH 7.0–7.5), followed by fixation in 2% (v/v) glutaraldehyde solution and incubation at 0 °C for 90 min. Subsequently, samples were dehydrated in a series of 15%, 30%, 50%, 70%, 95%, and 100% (v/v) aqueous ethanol solutions each for at least 30 min. After embedding samples with 66.6% LR White resin (London Resin CO Ltd., UK) in ethanol at 25 °C for 2 h and overnight incubation in 100% resin

180 at 4 °C, the samples were polymerized for 12 h at 55 °C. The 80-100 nm ultrathin sections were cut with diamond knives (DDK, Wilmington, DE, USA) in Reichert Ultracut E ultramicrotome (Leica Biosystems, Wetzlar, Germany). The sections stabilized by formvar-coated 300 mesh copper grids (Plano GmbH, Wetzlar, Germany) were stained with Uranyl Acetate Replacement Stain (Electron Microscopy Sciences, Hatfield, PA, USA) for 20 min. Images were captured with a Gatan Orius 4 K camera attached to a Jeol 1011 electron microscope (Jeol GmbH, Munich, Germany) and processed with the 314

185 Gatan Digital Micrograph software (Gatan Inc., Pleasanton, USA) and Adobe CS2 Photoshop (Adobe Systems Inc., San José, Cal., USA).

7

Biogeosciences

Discussions

EGU

2.5 Metabolic profiling based on Metagenome-Assembled Genomes (MAGs)

Extracted DNA from the biofilm sample MB1 abundant in Chloroflexi, Cyanobacteria and Actinobacteria as a representative of leachate group was submitted to the Göttingen Genomics Laboratory for shotgun metagenomic sequencing. The gDNA
190 extraction followed by quality and quantity assessment was performed as described above. Illumina paired-end sequencing libraries were prepared using the Nextera DNA sample preparation kit and subsequently sequenced on a MiSeq system with the reagent kit v3 with 600 cycles (Illumina). For pre-processing of sequencing data, quality control, per-read quality pruning, read filtering, adapter trimming, and base correction fastp v.0.19.4 (Chen et al., 2018) was used. The assembly of short read metagenomic data into metagenomic scaffolds was carried out by the metagenome assembler metaSPAdes
195 v.3.14.0 (Nurk et al., 2017). Subsequently, bins were determined using MaxBin v.2.2.7 (Wu et al., 2015). CheckM v.1.1.2 was used to evaluate the MAGs quality by providing robust estimates of genome completeness and contamination (Parks et al., 2015). Each high-quality MAG was then annotated using PROKKA v1.14.5 (Seemann, 2014). Genome wide orthologous clusters across multiple species were determined with a web server: OrthoVenn v2 (Xu et al., 2019), which assigned the protein sequence data to a high-level summary of functional categories such as biological process, molecular function, and
200 cellular component with GOSlim annotation and UniProt search. Finally, the PROKKA output was analysed by using the PathoLogic (Karpe et al., 2011) component of the Pathway Tools software v.23.5 (Karp et al., 2015) and the MetaCyc database v.23.5 (Caspi et al., 2019). The Metagenome-Assembled Genomes (MAGs) were classified taxonomically using GTDB-Tk v.1.0.2 and the Genome Taxonomy Database (GTDB) (release 89) (Chaumeil et al., 2020; Parks et al., 2019). The raw sequencing and assembly data have been already published in the SRA (SRR12886061) and Genbank
205 (JADEYI000000000 and JADMIG000000000-JADMN000000000) under the Project accession number PRJNA670497 (Arif et al., 2021a). The pathways maps of these MAGs showing their metabolic potentials and annotations can also be assessed at the Göttingen Research Online Database https://doi.org/10.25625/W9PWCX.

3. Results

3.1 Physiochemical parameters of Kilianstollen mine waters

210 The physiochemical parameters pH, electric potential and major ions concentrations were determined for the spring water samples (spring basin; source of fresh ground water), and after the intermixing of a stream rich in calcium/hydrogen carbonate (side adit 1), where a solid sinter crust was formed in the stream bed. Two other records of parameters were taken in a stream rich in calcium/sulphate ions (side adit 2, "gypsum karstic water"), before and after mixing with the karstic water of the mainstream. Further, two samples were taken from a copper flume leachate and a copper sulphate leachate, outflow
215 towards the adit mouth (Fig. 1). The concentration of dissolved ions indicated higher contents of Ca, Na, Cl, SO4, Fe, Cu, Zn, and Mn in the leachate samples (Fig. 3). The highest concentration of the transition metals Ni, Co, Fe, Cu, Zn, and Mn (0.62, 0.46, 35, 85, 2, and 4 mgL-1), was observed in the acidic copper flume leachate (pH 4.8) along with SO_4^-, Ca^+, Na^+, Cl^- and NO_3^- (1952, 403, 415, 499, and 188 mgL-1) (Fig. 3a). Due to the metal precipitation, the concentrations of these

8

transition metals towards the adit leachate outflow stream dropped to 0.2, 0.1, 2.9, 0.005, 0.8 and 1.3 mgL-1, respectively,

220   and pH raised to 7.26. The heavy metal content in spring water was considerably lower, i.e., in the range of 0.23 to 5.9 µgL-1. The efflux of the transition metal ions from the adit crosscut 2 drainage raised the heavy metal concentration of the spring water in the range of 3.8 to 262 µgL-1, particularly for Zn, Mn, Cu, and Ni (262, 76, 44 and 18 µgL⁻). To understand the copper toxicity and homeostasis with respect to microbial consortia, biofilms growing at the rocky mine walls were investigated nearby these water bodies.

225

Figure 3. Heavy metal content in Marsberg drainage mine waters. The concentration of major ions related to alkali and halogen groups (a) and heavy metals (b) have been compared. The insert shows the pH of each mine water, in the same order as depicted for major ions. The copper flume heavy metal measurements were separately plotted as the copper and iron ion concentrations were exceptionally high. The adit crosscut 2 mine drainage mixed with the sinter crust stream to form the

230   mixture outflow stream.

9

3.2 Distribution of bacterial taxa in Marsberg Kilian copper mine samples

The distribution of the predominant taxa at the collection sites varied drastically with the quality and type of nearby water sources (Fig. 4a). The most obvious sign of microbial growth, visible to the naked eye, were sub-aerial, whitish biofilms (MBS1-4) growing on the rocks, nearby the spring water stream. They were dominated by *Proteobacteria* (38%) and

235 *Actinobacteria* (21%). The relative abundance of *Proteobacteria* declined, whereas the relative abundance of *Actinobacteria* increased gradually as the sampling site moves from spring water (MBS1) to karstic water containing heavy metal discharge influx (MBS3). In the leachate samples group, *Chloroflexi* (30%), *Cyanobacteria* (23%) and *Actinobacteria* (19%) were abundant in the greenish-whitish biofilms (MB1-6) collected either in close vicinity of the copper flume leachate or directly from the heavy metal leachates streams (outflow water stream sample MBS18). Since these sites were more intensively

240 illuminated with light bulbs than other sampling sites, this could have facilitated the growth of *Cyanobacteria*. The biofilms collected next to the copper flume from wooden plank (MBS10) and moist unconsolidated rocks (MBS11,13) were enriched mainly in *Actinobacteria* (41%) and *Acidobacteria* (20%). When the three sample groups were compared in terms of abundant taxa, *Cyanobacteria* and *Proteobacteria* were significantly abundant in the leachate and spring water stream group samples, respectively ($p < 0.05$ ANOVA), while *Actinobacteria* and *Acidobacteria* seemed to be ubiquitous.

245 At the class taxonomic rank, 5657 OTUs were identified across all samples comprising *Actinobacteria* (25%), *Ktedonobacteria* (13%), *Oxyphotobacteria* (12%), *Acidobacteria* (9%) *Gammaproteobacteria* (8%), *Deltaproteobacteria* (5%), *Bacteroides* (5%) and *Deltaproteobacteria* (3%). MBS1-4 whitish biofilms growing nearby a spring water stream ere dominated by *Deltaproteobacteria* and *Gammaproteobacteria* (48%), *Bacilli* (29%), and *Actinobacteria* (51% and 25%). *Ktedonobacteria*, being the most abundant class (26%) in the leachate samples group (MB1-6 and MBS18), constituted 85%

250 of *Chloroflexi*. *Oxyphotobacteria* (a class within the phylum *Cyanobacteria*) and *Actinobacteria* also contributed 23% and 16% to the leachate group. Since the corresponding biofilms colonize rock surfaces in direct contact to the acidic as well as sulphidic and transition metal (Fe, Cu, Zn, and Mn) rich mine drainage water (Fig. 1 and 2), it is assumed that the low pH and the high heavy metal concentration of mine water contributed to the *Ktedonobacteria*, *Oxyphotobacteria*, and *Actinobacteria* natural enrichment. Statistical analysis revealed that the classes *Ktedonobacteria* and *Oxyphotobacteria* were

255 significantly abundant in the leachate group, while the *Bacilli* in the spring water stream group and *Actinobacteria*, *Acidobacteria Fimbriimonadia*, and *Gemmatimonadetes* classes in the unconsolidated rocks group were significantly abundant as compared to other groups classes ($p < 0.05$ ANOVA).

Biogeosciences Open Access

EGU

Discussions
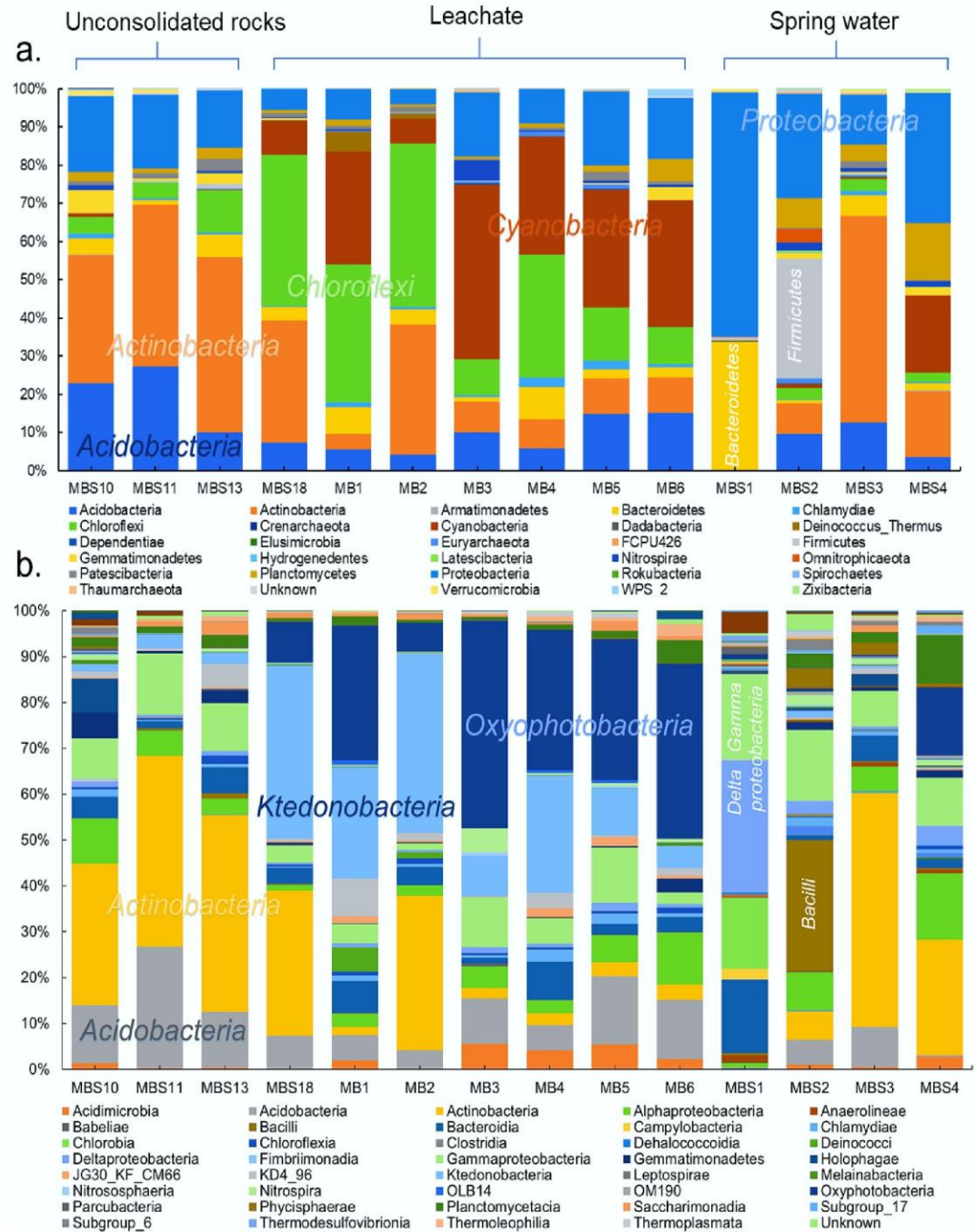
Figure 4. Relative abundances of Marsberg Kilianstollen bacterial taxa. At phylum (a) and class (b) level, the taxonomy and

260 relative abundance of the OTUs depicts the bacterial community composition and colonization at Kilianstollen sampling sites. The samples are shown in three groups based on their origin (unconsolidated rocks, leachate and spring water). The classes showing less than 2 % relative abundance are not mentioned. Besides bacterial taxa, the selected primers also led to the detection of *Crenarchaeota*, *Euryarchaeota*, and *Thaumarchaeota* in low abundance.

265 **3.3 Alpha and beta diversity**

The alpha diversity index Chao1 indicated that unique OTUs (richness) were abundant in the MBS1-4 spring water stream samples, followed by the MBS10,11,13 unconsolidated rocks and MB1-6, MBS18 leachate group samples (p<0.0003 ANOVA) (Fig. 5). The Shannon diversity index showed the same pattern when the sampling groups were statistically compared (p =0.025 ANOVA, Table S1). Alpha diversity is high in the MBS1-4 samples, possibly due to moisture and

270 neutral conditions from the adjacent spring water stream. With respect to the leachate samples group, low diversity indexes were observed as only a few adapted microbes could colonize, indicating an enrichment effect due to extreme environmental conditions (Fig. 5a).

The principal coordinates analysis (PCoA) showed that the microbial communities of spring water, leachate, and unconsolidated rocks samples were distinct to each other (Fig. 5c, d). According to the Unifrac weighted algorithm

275 (Lozupone et al., 2011) and analysis of molecular variance (AMOVA) nonparametric method (Table S2) (Mengoni and Bazzicalupo, 2002), the spring water and leachate samples were phylogenetically distinct to each other (p =0.034), conversely, the unconsolidated rocks microbiome was similar to both spring water and leachate groups (p >0.097). Conclusively, the unconsolidated rocks group is the intermediate between the other groups in terms of diversity and the diverse environmental conditions led to enrichment of different microorganisms. The hypotheses that the low pH and the

280 high heavy metal concentration of mine water contributed to the *Ktedonobacteria* and *Actinobacteria* natural selection and enrichment was further confirmed with the cannocial correspondence analysis (Fig. S2). The representative of leachate groups (MB1, MBS18) cluster closely with the low pH, high heavy metal concentration and *Ktedonobacteria* abundance as compared to the spring water samples which showed that after mixing with the adit leachate, the major abundant taxa *Proteobacteria* (MBS1-2) was replaced with *Actinobacteria* (MBS3).

12
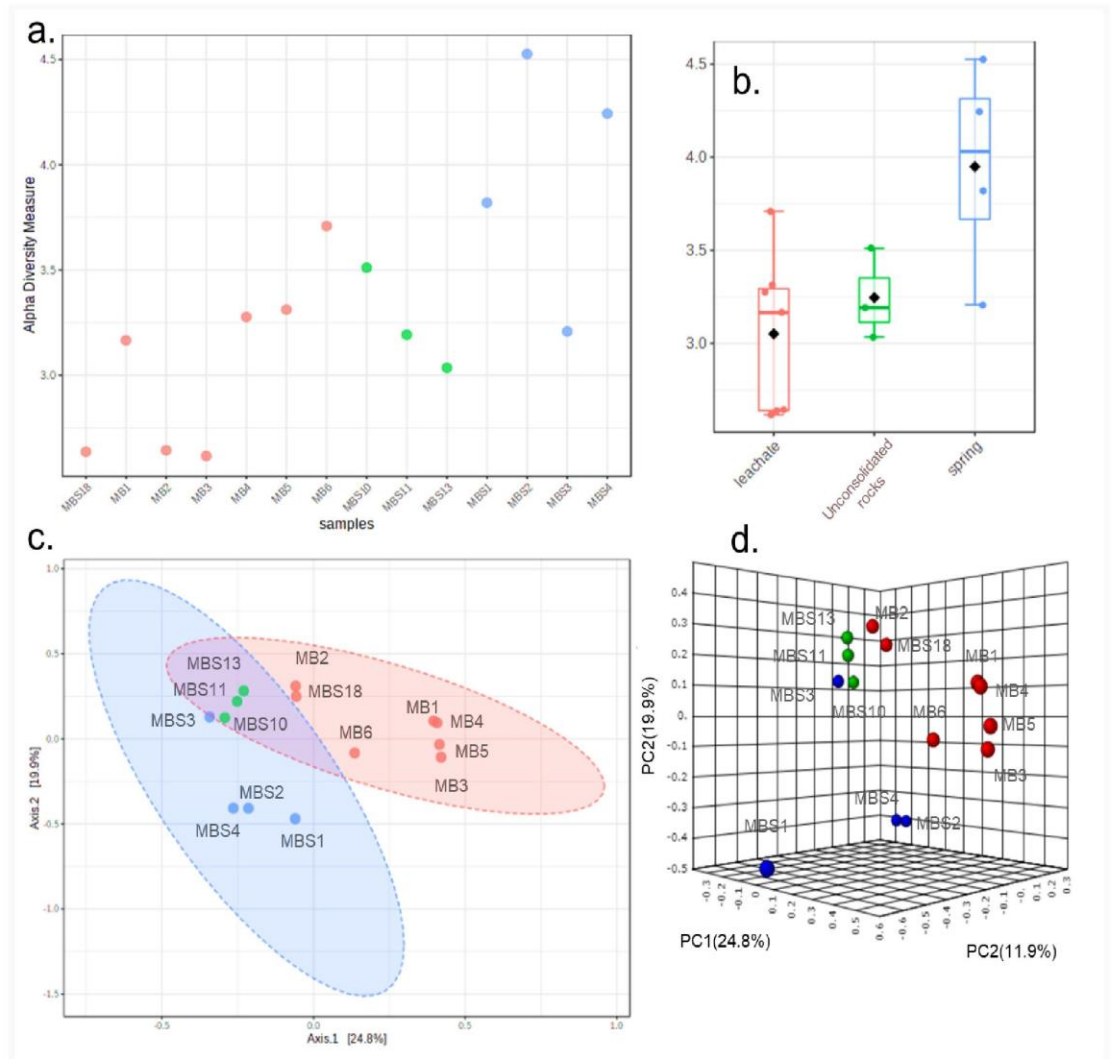
285

Figure 5. Alpha and beta diversity. Shannon diversity index (a) and boxplots (b) depict that the spring water samples have the highest alpha diversity. The PCoA 2d (c) and 3d (d) graphs indicated the unconsolidated rocks (⬤) microbial communities share similarity with both the spring water (⬤) and mine leachate (⬤) samples microbiome.

13

Biogeosciences

Discussions

EGU

Open Access

3.4 Appearance of sampled biofilms and phylogenetic analysis

290

Eukaryotic microalgae and aerial mycelia could be observed in light microscopy images when the collected leachate samples group was visualized under the light microscope (Fig. 6). Interestingly, a nematode related to *Poikilolaimus oxycercus* was also found which colonizes the deep subsurface sites (Borgonie et al., 2019) and one unicellular alga (*Coccomyxa subellipsoidea*) was highly abundant (unpublished previous data). TEM micrographs also showed eukaryotic (algal) along

295

with prokaryotic cells in the biofilm, indicating cohabitation. Mineral deposition around the microbial cell walls was not observed, suggesting the inhabiting microbiota has employed some other pathways to cope with the heavy metal toxicity under low pH instead of metal precipitation (Fig. 6). TEM micrographs of the MB3 biofilm from the leachate group also revealed the presence of sporulating hyphae which could be identified as either mycelia-like branched *Ktedonobacteria*, or *Actinobacteria*. Since *Ktedonobacteria* being the most abundant class had much higher relative abundance as compared to

300

the *Actinobacteria* and more importantly the distinguishing sporulation pattern of *Ktedonobacteria* (one spore per cell) ruled out the possibility of *Actinobacteria*.

Marsberg Kilianstollen offers a large reservoir of uncultured novel strains, comprising numerous *Ktedonobacteria* representatives (Fig. S3, S4). The genus level-based analysis led to the identification of at least 10 distinct uncultured genera affiliated with the *Ktedonobacteria* class and 80% of them could be classified to the *Ktedonobacteraceae* family. Within

305

Actinobacteria, 20 known genera were identified, whereby the *Pseudonocardiaceae* family represented 88% actual abundance of these genera members. The most abundant genus, designated C0119, has been frequently assigned to *Ktedonobacteria* based on reference database comparisons (Dube et al., 2019; Paul Chowdhury et al., 2019). Nevertheless, C0119 OTU sequences have been shown to not cluster with *Ktedonobacteria* based upon phylogenetic analysis of *Chloroflexi* sequences (Glöckner, 2019; Jones et al., 2017) (Fig. S5). The unclassified strains might need to be classified into

310

a new class of the cold climate adapted extremophiles belonging to the phylum *Chloroflexi*. In contrast, most of the *Actinobacteria* OTUs were classified as representatives of known genera such as the rare *Actinobacteria* (*Mycobacterium, Nocardioides, Micrococcus, Crossiella, Amycolatopsis, Nocardia* and *Pseudonocardia*) which could be potentially interesting with respect to novel bioactive compounds, biodegradation and biodeterioration pathways [28-30].

14
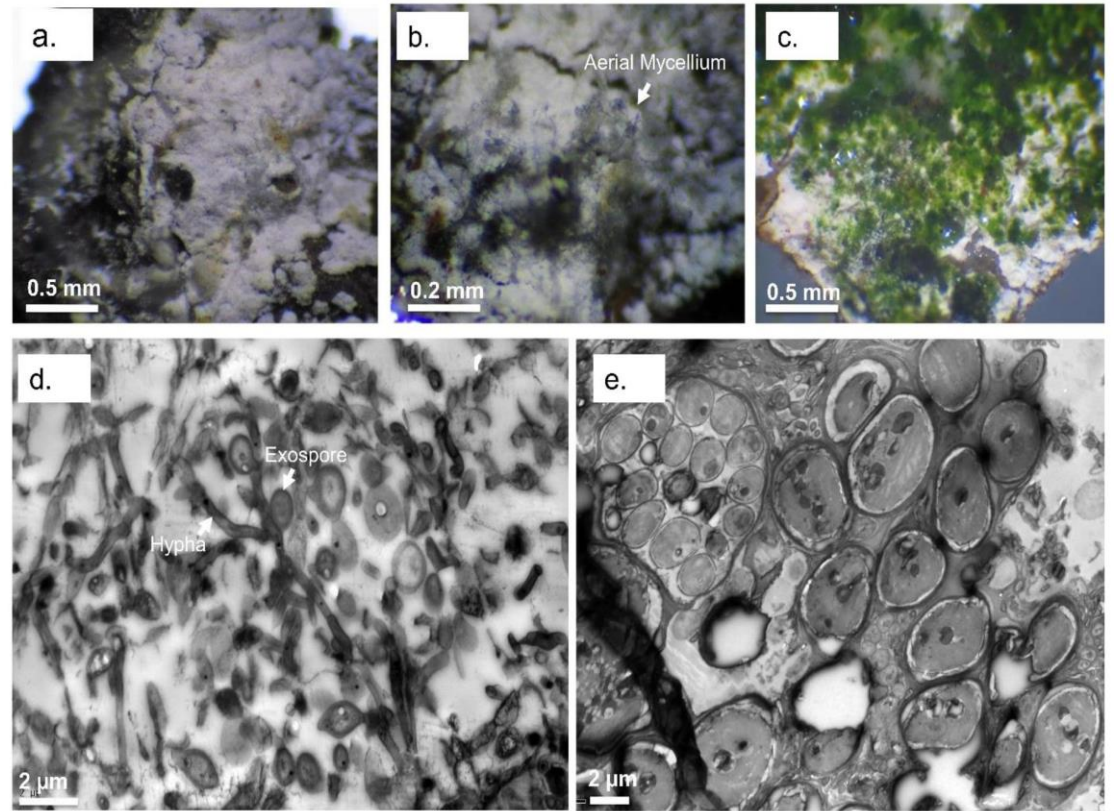
315 Figure 6. Light and transmission electron microscopy of Marsberg Kilianstollen leachate group biofilms. The images (a-c) captured under the light microscope depict the aerial mycelium and green algal biomass along with the white bacterial biofilm. TEM micrographs (d) and (e) show the mycelial growth and spores formed at short branches along a hypha resembling *Ktedonobacteria* sporulation pattern (hypha and exospore are marked in d) and the large eukaryotic algal cells (e), respectively.

320

3.5 Key metabolic pathways to survive under extreme conditions

3.5.1 Metagenome-assembled genomes (MAGs)

Biofilm sample MB1 as a representative of leachate group was selected for shotgun metagenomic sequencing based on the abundance of overall representative taxa of interest such as *Chloroflexi* and *Actinobacteria*. Eight relatively complete MAGs

325 (completeness, ≥89%) were obtained, and subsequently proteins and metabolic pathways predicted. The genome sizes of the

15

Biogeosciences
Discussions

EGU

eight MAGs (designated Mberg 002, 006, 008, 009, 010, 011, 015, and 019), showing a contamination rate ≤10%, ranged from 2.6 to 4.9 Mb. Furthermore, the number of identified genes ranged from 2516 to 4772. Based on phylogenetic analysis the MAGs were assigned to *Actinobacteria, Binatia, Deinococci, Chloroflexota, Dehalococcoidia, Chloroflexia and Ktedonobacteria* (Fig. S6).

330   **3.5.2 Orthologous gene clusters**

The preliminary comparison and analysis of GOSlim terms for core orthologous gene clusters for each MAG revealed unique survival pathways involved in extremophily (Fig. 7). The identified protein genes were grouped in 5540 distinct, 5395 orthologous and unique 145 single-copy gene clusters. The shared orthologous gene ontology GO terms correspond to common cellular functions such as respiration and cell wall synthesis. These orthologous clusters contribute to various
335   survival pathways such as aromatic and sulphur compounds metabolic processes, detoxification pathways for arsenic compounds and heavy metals ions transport, dehalogenation, and sporulation. The detected broad-spectrum heavy metal binding domains and associated proteins identified in the PROKKA annotations are also important to survive high concentrations of heavy metals in the environment. The metal binding proteins are mostly enzymes which require transition metals Zn, Cu, Co, Ni as cofactor to perform different biological processes. The cupredoxins superfamily proteins and
340   domains were frequently found in shared clusters and these contain type 1 Cu binding sites which are involved in oxidation reactions conferring resistance against Cu by various proteins (azurin, multicopper oxidases [MCO], laccases, and nitrosocyanin) (Arguello et al., 2013; Vita et al., 2016; Donaire et al., 2002; Redinbo et al., 1994; Zaballa et al., 2012). Mberg 010, assigned to *Binatia*, showed orthologous protein clusters related to aromatic compound degradation pathways. Mberg019, affiliated with *Ktedonobacteria*, encodes unique plasma membrane proteins, identified in unshared clusters. All
345   MAGs seem to mediate transition metal homeostasis as metal binding proteins and transporters were frequently identified. The process of spore formation seems to be prevalent in all MAGs as GO terms for sporulation were observed in shared and unshared orthologous clusters. To study the MAGs with respect to heavy metal homeostasis, efflux systems, as well as detoxification and aromatic degradation pathways in more detail, pathway maps (https://doi.org/10.25625/W9PWCX) were generated based on PROKKA outputs and subsequently inspected.

350

16

Mberg 006
GO:0030435; sporulation; P45693
GO:0008270; zinc ion binding; A4VLQ5
GO:0046872; metal ion binding; P67732
GO:0042594; response to starvation; P43811
GO:0006825; copper ion transport; Q5ZWR1 Copper-exporting P-type ATPase copA
GO:0055085; transmembrane transport; A0QQ71 Phosphate-import protein PhnD
GO:0009408; response to heat; Q8GWL1 Heat-stress-associated 32 HSA32
GO:0006811; ion transport; Q9WZS2 Mberg006
GO:0018414; nickel incorporation into metallo-sulfur cluster; P31897

Mberg008
GO:0019439; aromatic compound catabolic process; O52866
GO:0034765; regulation of ion transmembrane transport;
GO:0046872; metal ion binding; Q51739 Q8RXR2 O27296 Q9RY41 P9WMW2
GO:0006885; regulation of pH; Q1IY26

Mberg 002-006
GO:0006825; copper ion transport, Copper-transporting ATPase 1 ATP7A P49015

Mberg002
GO:0046872; metal ion binding; Uncharacterized metallohydrolase YodQ O34984, N-acyl homoserine lactonase ahlD, Q7X3T2 Probable quorum-quenching lactonase YtnP O34760

GO:0030435; sporulation; Stage II sporulation protein D spoIID P13251 Putative cytochrome bd menaquinol oxidase subunit I ythA C0SP90

Mberg 009
GO:0008271; secondary active sulfate transmembrane transporter activity; High affinity sulfate transporter 2 ST2 P53392
GO:0046685; response to arsenic-containing substance; Putative arsenical pump-driving ATPase MTH_1511 O27555
GO:0042207; styrene catabolic process; Phenylacetaldehyde dehydrogenase styD O06837
GO:0008270; zinc ion binding; Q06004 B8E160
GO:0005506; iron ion binding; P29604
GO:0046872; metal ion binding; O27549 P9WKS2 Q89VT8 Q9WWU5 Q7Z7P9
GO:0006805; xenobiotic metabolic process; Q9HBJ2

Venn diagram values: 450, 220, 111, 203, 156, 130, 71, 484, 74, 96, 129, 118, 159, 439, 224

Labels: Mberg002, Mberg006, Mberg008, Mberg009

Mberg002-009
GO:0030001; metal ion transport; Q9KFG3 Manganese-binding lipoprotein MntA

Mberg 006-008 GO:0009651; response to salt stress, Q9ZW31
All Mberg GO:0046914; transition metal ion binding; P0A673

Mberg011
GO:0022857; transmembrane transporter activity; Q58026
GO:0005506; iron ion binding; P29604 Ferredoxin
GO:0030435; sporulation; C0SP90

Mberg010
GO:0019439; aromatic compound catabolic process; Q51975 Q51974 O33477 O85673 O52379 Q53122 A5W4G5
GO:0046232; carbazole catabolic process; Q8G8B6 Q8GI14
GO:0019383;(+)-camphor catabolic process; H3JQW0
GO:0015673; silver ion transport; P38054 Cation efflux system protein CusA
GO:0046686; response to cadmium ion; P94177 Cation efflux system protein CzcA
GO:1901359; tungstate binding; O57890 Molybdate/tungstate-binding protein WtpA
GO:0018896; dibenzothiophene catabolic process; P54998
GO:1901170; naphthalene catabolic process; Q9X9Q6
GO:0010128; benzoate catabolic process via CoA ligation; Q51601
GO:0043640; benzoate catabolic process via hydroxylation P19076 P07770
GO:0042203; toluene catabolic process P49155
GO:0006790; sulfur compound metabolic process D3RPB9
GO:0019380; 3-phenylpropionate catabolic process; P0ABR6
GO:0046687; response to chromate; P17550
GO:0046872; metal ion binding; E7FHP1 P00200 Q89VT8 O34760 Q7X3T2
GO:0004497; monooxygenase activity; G2IN04

Mberg015
GO:0055085; transmembrane transport; P9WG00 Q6CZ31 P77156 P37624 Q9KEE9 O51924 Q57RB0 P39642 Q66FU7 O32155 O07017 P46904 Q7CH99 P21409 P9WG00 O32155 P40980 O51924 P40980 O32155 O58760 P55603 P06109
GO:0046872; metal ion binding; Q2YTD2 O34984 Q9RY41 P9WMW2 Q5SME3
GO:0030435; sporulation P13251 P45693
GO:0055072; iron ion homeostasis; P37009
GO:0016784; 3-mercaptopyruvate sulfurtransferase activity; Q9I452
GO:0008270; zinc ion binding; Q06004 A4VLQ5
GO:0006885; regulation of pH; Q1IY26

Venn diagram values: 107, 432, 29, 960, 238, 188, 37, 457, 358, 97, 28, 70, 144, 90, 432

Labels: Mberg010, Mberg011, Mberg015, Mberg019

Mberg019
GO:0030435; sporulation Q9K5N0 B6H9U8
GO:0018414; nickel incorporation into metallo-sulfur cluster; P31897
GO:0008270; zinc ion binding; Q06004
GO:0046872; metal ion binding; O27549 O27296 P67732 Q4L749
GO:0019439; aromatic compound catabolic process; O52866 GO:0006825; copper ion transport; Q5ZWR1 Copper-exporting P-type ATPase CopA
GO:0034765; regulation of ion transmembrane transport; Q8RXR2
GO:0009432; SOS response; Q5L0C2
GO:0005886; plasma membrane; O31603 A6T022 A6T022 P12752 A3DC75 O34578 Q55705
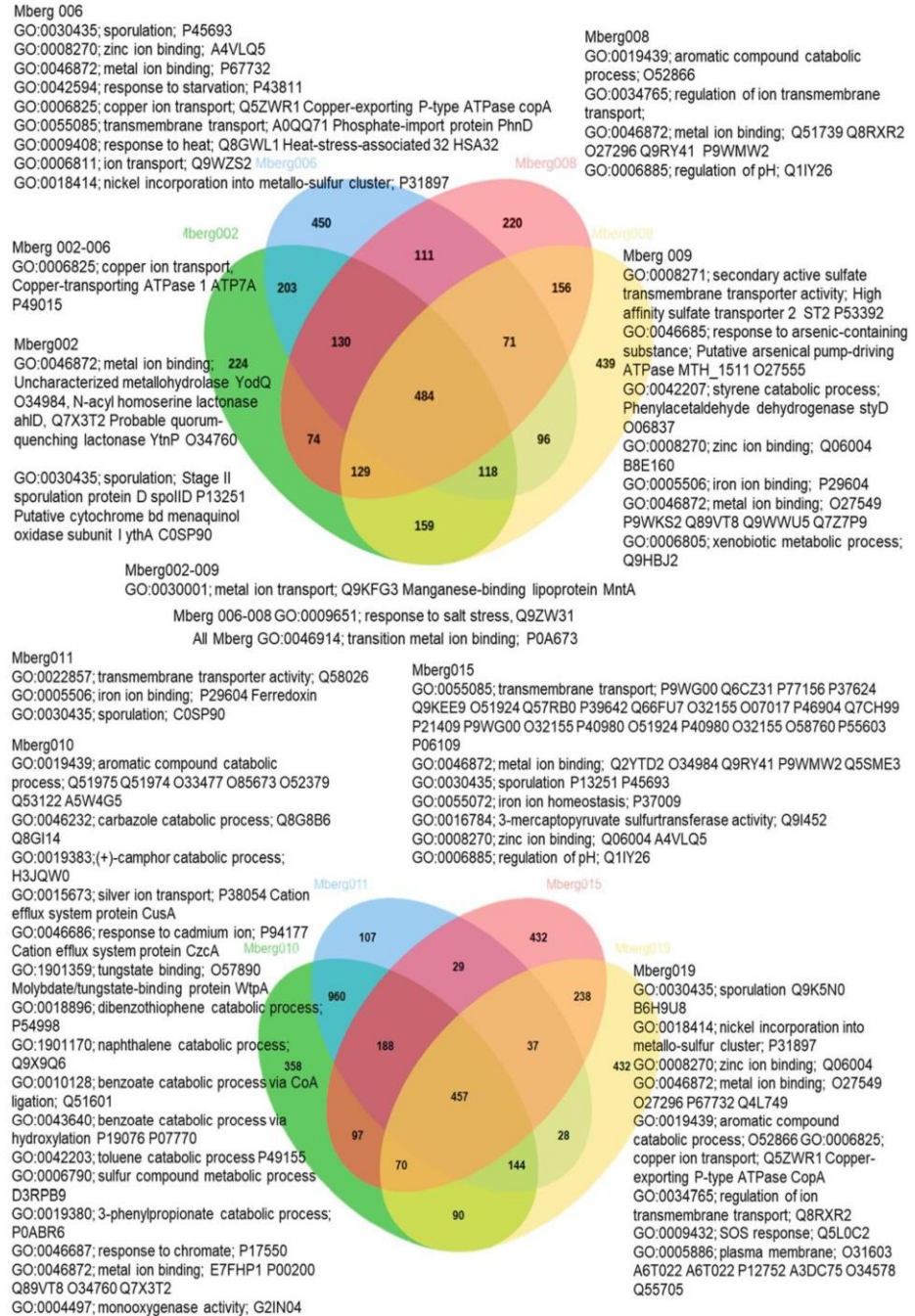
17

Figure 7. Venn diagram displaying the distribution of shared orthologous clusters among the eight assembled genomes. The orthologous clusters of similar proteins have been allocated the Go terms based on same function or process. Only the Go terms related to metals, aromatic compound, sporulation, and pH regulations are mentioned to identify the microbial survival

355   under extreme stress conditions. The closely matched proteins are identified as Swiss-prot accession numbers and names of some important hits are mentioned.

3.5.3 Heavy metal homeostasis and efflux systems

Transporters such as copper-exporting P-type and oxidation enzymes involved in transition metals homeostasis were identified in all selected MAGs (Fig. 8). The genes involved in copper homeostasis include copper-sensing transcriptional

360   repressors CsoR and RicR, copper-exporting P-type ATPases CptA, ActP and CopA, and oxidation enzymes, multicopper oxidases (MCOs) in all MAGs. Basically, upon Cu(I) ion binding, the dissociation of CueR, CsoR and RicR transcriptional repressors (Fu et al., 2014; Smaldone and Helmann, 2007) activates the copper regulon genes related to transporters CopA-like P-type ATPases, metallothioneins or copper binding chaperons CopC, CopZ and periplasmic multicopper oxidase MmcO, CueO (Osman et al., 2010; Shi et al., 2014). Under anaerobic conditions the multicopper oxidase enzymes become

365   inactive, and therefore another three-component channel/pore Cus complex controls Cu(I) efflux through CusA, an inner membrane energy-driving channel which is attached to the outer membrane pore CusC through the periplasmic CusB protein (Outten et al.).

The *cutA* locus, presumably involved in copper tolerance and homeostasis, was also characterized to affect tolerance levels to zinc, nickel, cobalt and cadmium ions (Fong et al., 1995). The metalloregulatory transcriptional response to di- and

370   multivalent heavy metal ions Cd(II), Pb(II), Bi(III), Zn(II) as well as Cu(II) is maintained by SmtB/ArsR family repressors CmtR and CadC (Busenlehner et al., 2003). MntH is a divalent metal cation transporter which displays broad substrate specificity and can regulate the intracellular accumulation of several divalent cations, including Mn(II), Cd(II), Co(II), Zn(II) and, to a lesser extent, Ni(II) and Cu(II) (Makui et al., 2000). CzcD, a heavy metal cation efflux transporter, mediates heavy metal resistance with respect to Cd(II)/Co(II)/Zn(II) in the absence of the high resistance CzcCBA system (Nies, 2003;

375   Anton et al., 1999; Papp-Wallace and Maguire, 2006).

Manganese homeostasis is maintained by a manganese efflux pump, MntP (Waters et al., 2011). A putative bacterial multicopper oxidase MoxA has been reported to compact the cellular Mn(II) toxicity through surface oxidation to the insoluble Mn(III) and Mn(IV) oxides (Ridge et al., 2007; Zhang et al., 2015). HoxN, a high-affinity nickel transporter facilitates the nickel translocation process as nickel permease (Wolfram et al., 1995). The influx of Mg(II), Ni(II) and Co(II)

380   is coordinated via an ubiquitous divalent metal ion transporter CorA (Kersey et al., 2012) and efflux of these divalent metal ions is directed through CorC, a magnesium/cobalt efflux protein (Gibson et al., 1991). The homeostasis of the transition metals cobalt, nickel and iron and Ni(II) and Co(II) detoxification is also regulated by a nickel-cobalt exporter, designated RcnA, through efflux (Koch et al., 2007).

18

385 Cellular zinc uptake is regulated through the energy intensive import system ZnuABC, where ZnuA binds Zn(II) in the periplasmic space and docks Zn(II) into the membrane permease ZnuB and the ZnuC of the pump finally catalyses ATP-dependent Zn(II) import into the cytosolic environment. In contrast, ZupT is involved in less energy intensive non-specific Zn(II) uptake along with the transition metals Fe(II), Co(II), Cd(II), and Mn(II) along the concentration gradient. Under high cytosolic divalent Zn(II), Cd(II), and Pb(II) concentrations, the MerR homologue, ZntR induces a Zn(II)-, Cd(II)-, and Pb(II)-transporting P-type ATPase ZntA (Rensing and Mitra, 2007).

390 The oxyanions molybdate and tungstate are taken up through the membrane by the high-affinity ModABC molybdate system along with sulphate ions (Markovich, 2001). Sulphate and thiosulphate are taken up by sulphate permeases, carriers belonging to the SulT family, encoded by the cysPTWA operon, and SulP family members (inorganic anion uptake carriers) (Kertesz, 2001).

19

3.5.4 Detoxification and aromatic degradation pathways

To successfully colonize the acid mine leachate downstream sites, microbes should harbour acid resistance, detoxification, and metabolizing systems as identified in all MAGs (Fig. 9). Arginine-dependent acid resistance, encoded by the MAGs, relies on arginine decarboxylase SpeA, which decarboxylates arginine (Arg) to produce agmatine (Tsai and Miller, 2013; Richard and Foster, 2004). Mercuric reductase MerB reduces divalent Hg(II) to volatile mercury Hg (Silver and Hobman, 2007) and arsenate reductase ArsC reduces the arsenate As(V) to arsenite As(III) which is delivered to ATP-dependent anion pump ArsAB by the metallochaperone ArsD (Martin et al., 2001). The enzymes related to aromatic compounds; chlorinated phenols, benzoate, atrazine, cinnamate, biphenyl, phenylacetate carbazol, catechol and 4-sulphocatechol phenylethylamine, naphthalene and 5-nitroanthranilate degradation were mainly identified as hydroxylases, dioxygenases, dehydrogenases, epoxidase etc. To detoxify reactive oxygen species (ROS), toxins, and antibiotics compounds with thiols, mycobacteria and some other *Actinomycetales* utilize mycothiol-mediated detoxification through Mca enzyme (Newton et al., 2008; Newton et al., 2000). Other detected important enzymes related to ROS detoxification were superoxide dismutase SOD and peroxidases (Broxton and Culotta, 2016).

21
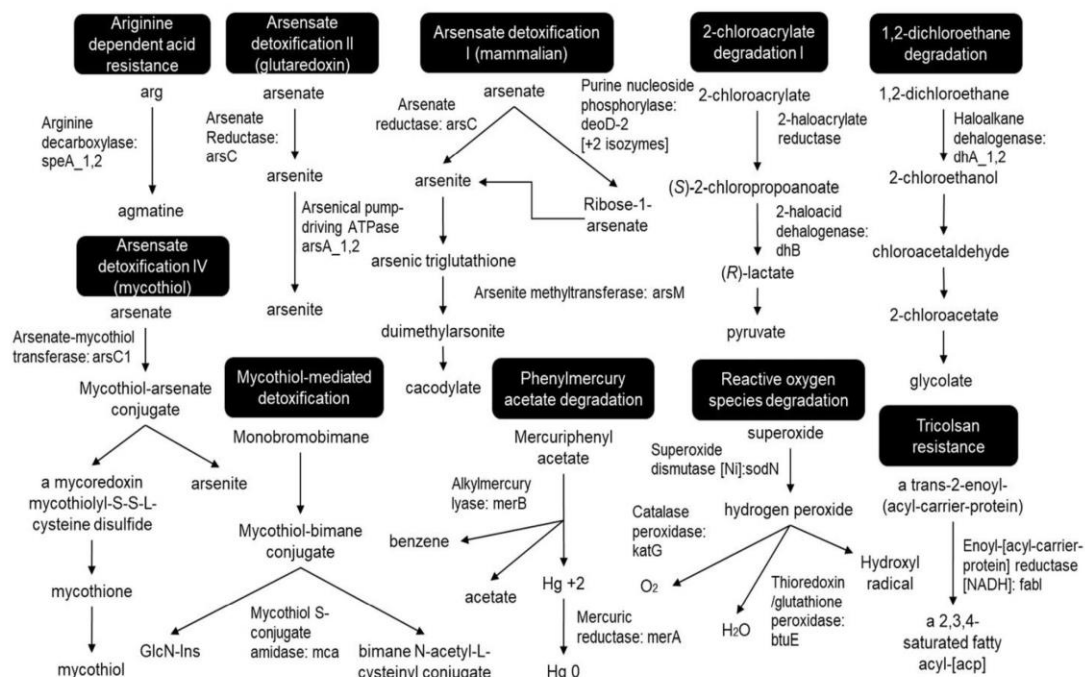
420

Figure 9. Detoxification pathways. The pathways have been redrawn from the pathway maps of all MAGs, provided in the supplementary pdf files.

4. Discussion

425   Generally, the habitats, which exceed most natural heavy metal concentrations, are dominated by prokaryotes with chemolithoautotrophic lifestyles, which produce sulphuric acid and are responsible for the mobilization of heavy metals from rocks through the oxidation of metal sulphide (Schippers and Sand, 1999). In accordance with this, a significant heavy metal mobilization of Cu, Zn, Mn, Pb, and Ni from copper slag debris has been already measured with enrichment cultures of iron/sulphur-oxidizing *Acidithiobacillus ferrivorans* and *Leptospirillum* obtained from the Marsberg Kilianstollen

430   iron/manganese sampled crusts (Amin et al., 2018). However, the present study aims to give a complete microbiome overview of Kilianstollen habitats influenced by mine drainage, comprising leachate downstream colonization sites, where

22

71

we could expect microbial metal reduction, deposition, and detoxification. To colonize the Marsberg mine subsurface habitat, the microbes must have developed resistance against heavy metals such as molybdenum, manganese, cobalt, zinc, and copper. Above optimal concentrations, copper confers its toxicity to microbes through its redox activity which catalyses

435   a Fenton-like reaction, resulting in the generation of reactive oxygen species that may cause protein damage and lipid peroxidation (Macomber and Imlay, 2009) and destabilizing the iron-sulphur clusters via Fe(II) displacement in key enzymes prosthetic groups (Azzouzi et al., 2013; Chillappagari et al., 2010; Dupont et al., 2011). Thus, copper may inhibit the growth of common fast-growing *Proteobacteria* or *Firmicutes* such as *Escherichia coli, Pseudomonas aeruginosa, Vibrio spp., Bacillus cereus*, and *Bacillus subtilis*, even at micromolar concentrations (Gordon et al., 1994). This metagenomic study of

440   the Marsberg copper mine also observed the colonization of heavy metal tolerant microbes and inhibition of *Proteobacteria* and *Firmicutes* growth under the influence of copper toxicity. The colonization sites near the copper-rich leachate outflow stream such as a copper flume (MB1-6 biofilms) enrich copper resistant groups belonging to abundant phyla *Chloroflexi* (*Ktedonobacteria*), *Actinobacteria* and *Cyanobacteria* (*Oxyophotobacteria*) as compared to the freshwater stream sites; mainly colonized by fastidious non-resistant *Proteobacteria*. A metatranscriptomic study of an abandoned Pb-Zn mine (Coto

445   Txomin, Spain) also determined that the heavy metal concentrations (up to 3220 and 97 g kg^{-1} of Pb, Cd, and Zn, respectively) and low pH (4-6) drastically influenced the soil microbial diversity, suppressed the relative abundance of *Actinobacteria, Acidobacteria*, and *Alphaproteobacteria*, and enhanced slow-growing metal and acid-tolerant taxa affiliated with *Chloroflexi* (*Ktedonobacteria*) (Epelde et al., 2015).

The same enrichment trend for *Chloroflexi* (*Ktedonobacteria* and KD4-96) was observed when a microcosm setup

450   supplemented with acid mine drainage contaminated soil and cysteine hydrochloride was incubated for 6 months at 30 °C, which decreased the abundance of the major taxa *Acidobacteria, Acidimicrobiia, Actinobacteria*, and *Thermoleophilia* (Gupta et al., 2018). A high abundance of *Ktedonobacteria* at the downstream arsenic deposits of the acid sulphate hot spring Tengchong, China, suggests that these aerobic heterotrophic mesophiles and thermophiles may have been involved in arsenic reduction/toleration along with iron and sulphur oxidation cycles (Jiang et al., 2016). The Marsberg cold-adapted

455   *Ktedonobacteria* also colonized the downstream acid mine leachate, indicating the ability to reduce heavy metal ions or tolerate the sediments along with sulphur compounds oxidation. The co-occurrence of *Cyanobacteria* (*Oxyphotobacteria*) and *Chloroflexi* (*Ktedonobacteria*) could facilitate the growth of heterotrophs by providing carbon nutrients and in return the heterotrophs may remediate the heavy metal contaminated sites, resulting in better microbial survival and colonization of the microbial consortium. The selective enrichment of *Ktedonobacteria* at Marsberg copper mine indicates the ability of these

460   psychrophiles to inhabit cold environments and could be linked to heavy metal tolerance along with iron and sulphur oxidation cycles. The extreme conditions survival is attributed to the high metabolic plasticity of *Ktedonobacteria,* a diverse class ranging from thermophilic to mesophilic isolates, and the type stain *Ktedonobacter racemifer* has an unusually large genome of 13 Mbp, containing 9539 genes, 601 of which are transposases (Chang et al., 2011). The identified KO07665 based on KEGG analysis belonging to the *Ktedonobacteriaceae* codes for the copper resistance phosphate regulon response

465   regulator CusR (Thomas Iv et al., 2020). The pathway map and annotation of the Marsberg *Ktedonobacteria* MAG also

23

indicates a detection mechanism with respect to Cu(I) by copper-sensing transcriptional repressor (RicR), oxidation to Cu(II) via multicopper oxidase (MCO) and finally export outside the cells through copper-exporting P-type ATPases (CopA, CtpA) and non-specific heavy metal cadmium, cobalt and zinc/H(+)-K(+) antiporter (CzcD).

High concentrations of heavy metal ions in the environment promotes selection for heavy metal-resistant microbes, with
470   either chromosomal or plasmid-level genes, maintaining heavy metal ion homeostasis inside cells (Rademacher and Masepohl, 2012). The organisms may reduce the sensitivity by employing the permeability barriers, by enhancement of active transport of metal ions from the cytoplasm (efflux) through a specific membrane transport system, by enzymatic detoxification, by reduction of metal ions by redox reactions, or by complexation of heavy metals resulting in extracellular and intracellular sequestration, etc. (Silver and Phung, 2005; Hobman et al., 2007). For all organisms in heavy metal
475   environments, one or more of these adaptations are prerequisites for survival. This study investigated the pathways crucial for microbial survival nearby acid mine drainage, and all MAGs explicated specialized adaptations to cope with heavy metal toxicity. Especially copper homeostasis is maintained by copper-exporting P-type ATPases efflux pumps ActP, CptA and CopA (Arguello et al., 2013; Kim et al., 2008; Festa and Thiele, 2011), Cu(I)-specific metalloregulatory proteins (CsoR, RicR (Fu et al., 2014) and CopR (Villafane et al., 2011)), resistance to acidity mediated copper toxicity by ActP (Reeve et
480   al., 2002), oxidation of Cu(I) to less toxic Cu(II) by the cupredoxins proteins superfamily such as periplasmic multicopper oxidases (MCOs) (Rowland and Niederweis, 2013) and expression of metallochaperones and metallothioneins with binding constants for Cu in the picomolar–femtomolar range (Rae et al., 1999; González-Guerrero and Argüello, 2008). The sensory cytosolic CopZ-like chaperones stimulate the transcription of several copper-stress related genes via trafficking the Cu(I) to the DNA-bound CsoR and CueR and the inner membrane-localized channels, especially copper-exporting P-type ATPases
485   such as ActP, CptA and CopA for efflux to the extracytosolic periplasmic environment (Arguello et al., 2013; Kim et al., 2008; Festa and Thiele, 2011; Novoa-Aponte et al., 2019). Mounting evidence also confirms the coexistence of the anaerobic Cus system with the aerobic CopA regulon in 44% of γ-proteobacteria group members (Hernández-Montes et al., 2012). The periplasmic CopC proteins maintain the bacterial copper homeostasis via binding both Cu(I) and Cu(II) along with CopD inner membrane protein (Cha and Cooksey, 1993). PCuAc-like chaperone PccA required for the biogenesis of the copper
490   centre assembly in the cbb3-type cytochrome c oxidases (CcO) have been demonstrated in various bacteria (Thompson et al., 2012; Andrei et al., 2020). CopA has also been shown to compete for Cu(I) with sub-femtomolar affinity against copper binding ligands and the efflux rate of this enzymes has been measured up to 27-130 $nmolmg^{-1}mn^{-1}$ (Wijekoon et al., 2017). Deletion of efflux channels render the cells susceptible to copper toxicity via intracellular accumulation of heavy metals (Sitthisak et al., 2007). The intracellular copper and zinc (10^{-6} and 10^{-4} M, correspondingly) is buffered at a bound state as
495   their transcriptional regulators CueR, Zur, and ZntR bind the cognate metals with very high affinities (10^{-21} and 10^{-15} M). This indicates that the free copper and zinc cytoplasmic concentrations are to be closely regulated to avoid toxicity as compared to manganese and iron (low regulators affinities 10^{-5} M and 10^{-6} M, respectively) (Porcheron et al., 2013). Additionally, manganese homeostasis is maintained by a manganese efflux pump (MntP), as Mn is essential for enzymatic

24

catalysis in the carbon metabolism, extracellular capsule polysaccharide synthesis and protection against oxidative stress
500   (Waters et al., 2011).

Since most of the heavy metal efflux pumps are ATPases, a constant supply of energy (as ATP) is met through converting various aromatic compounds to TCA cycle intermediates (Fuchs et al., 2011; Ghosal et al., 2016), in addition to several inorganic nutrients catabolism. To detoxify ROS, toxins, and antibiotic compounds, mycobacteria and some other Actinomycetales utilize mycothiol-mediated detoxification (Newton et al., 2008; Newton et al., 2000). Other important
505   enzymes related to ROS detoxification were superoxide dismutase SOD and peroxidases which degrade the superoxide anion radicals (Broxton and Culotta, 2016). The *arsRDABC* operon codes for an ATP-dependent anion pump that confers resistance to arsenite, arsenate, antimonite, and tellurite (Rosen, 2002). Arsenite methyltransferase ArsM catalyses the formation of several volatile methylated intermediates from As(III), which eventually results in loss of arsenic from the cells through passive diffusion (Huang et al., 2018; Zeng et al., 2018). The mercuric reductase MerB reduces divalent Hg(II) to
510   less bioavailable metallic mercury Hg^0 vapour which is also volatilized under aerobic conditions and leaves the cells through passive diffusion (Silver and Hobman, 2007). These detoxification pathways also explain the lack of mineral or metals deposition as visible in TEM images as either oxidation of metallic ions to less soluble ionic form or their reduction to volatile products makes them leave the bacteria through extrusion or diffusion under low pH. Hence, to inhabit the Marsberg acid mine leachate sites, the cold-adapted microbiome has the plasticity to express a wide range of heavy metal-specific
515   enzymes which could either oxidize or reduce different metallic ions to neutralize their toxic effects, regulate their intracellular concentrations and integrate the heavy metals as cellular components, cofactors for enzymatic functions, protection against oxidative stress. Culturing of specific strains especially *Ktedonobacteria* may further clarify mechanisms of heavy metal resistance (exporter systems, metal chelators, bioorganic compounds enhancing metal precipitation). This might also give implications to mine waste treatment, bioremediation and biomining.

5. Conclusion
520

The freshly collected samples from Marsberg copper mine (NE Rhenish Massif, Germany) were taken as a hitherto unexplored inventory of extremophilic organisms with largely unknown properties with respect to long-term survival, heavy metal tolerance, and degradation of complex organic compounds. The typical colonization patterns, mainly composed of *Firmicutes,* and *Proteobacteria* changed considerably towards uncultured *Chloroflexi*, including *Ktedonobacteria*
525   representatives, when the sampling sites around the spring water stream shifted to the acid mine leachate outflow. The acid mine drainage with influx of heavy metals altered the composition, drastically reduced the richness and evenness of microbial communities, and exerted selective pressure towards resistance to metal contamination. Consequently, the microbiome has evolved various survival pathways related to aromatic and sulphur compounds metabolism, toxic arsenic compounds reduction, copper ions oxidation and heavy metal ions reduction and extrusion, dehalogenation and sporulation.

25

Biogeosciences
Discussions

Open Access

EGU

535 **7. Contributions**

Conceptualization, M.H., H.N., and S.A; Methodology, S.A., M.H., E.S., A.R., G.A., H.N.; Sample collection; M.H., S.A., E.S.; Formal analysis; S.A., E.S., A.R., G.A., H.N.; Writing—original draft, S.A., G.A., H.N..; Writing—review and editing, M.H., S.A., H.N., A.R., and G.A..; Funding acquisition, M.H, A.R. and S.A.

8. Competing interests

540 The authors declare no competing interests

References

Amin, N., Schneider, D., and Hoppert, M.: Bioleaching potential of bacterial communities in historic mine waste areas, Environmental Earth Sciences, 77, 542, 10.1007/s12665-018-7714-x, 2018.

Andrei, A., Öztürk, Y., Khalfaoui-Hassani, B., Rauch, J., Marckmann, D., Trasnea, P.-I., Daldal, F., and Koch, H.-G.: Cu 545 Homeostasis in Bacteria: The Ins and Outs, Membranes, 10, 242, 2020.

Anton, A., Große, C., Reißmann, J., Pribyl, T., and Nies, D. H.: CzcD is a heavy metal ion transporter involved in regulation of heavy metal resistance in Ralstonia sp. strain CH34, Journal of bacteriology, 181, 6876-6881, 1999.

Arguello, J. M., Raimunda, D., and Padilla-Benavides, T.: Mechanisms of copper homeostasis in bacteria, Frontiers in cellular and infection microbiology, 3, 73, 2013.

550 Arif, S., Nacke, H., and Hoppert, M.: Metagenome-Assembled Genome Sequences of a Biofilm Derived from Marsberg Copper Mine, Microbiology Resource Announcements, 10, e01253-01220, 10.1128/MRA.01253-20, 2021a.

Arif, S., Schliekmann, E., and Hoppert, M.: 16S rRNA Amplicon Sequencing of Microbial Biofilms from Marsberg Copper Mine, Germany, Microbiology Resource Announcements, 10, e01315-01320, 10.1128/MRA.01315-20, 2021b.

Arndt, D., Xia, J., Liu, Y., Zhou, Y., Guo, A. C., Cruz, J. A., Sinelnikov, I., Budwill, K., Nesbø, C. L., and Wishart, D. S.: 555 METAGENassist: a comprehensive web server for comparative metagenomics, Nucleic Acids Research, 40, 88-95, 10.1093/nar/gks497, 2012.

Azzouzi, A., Steunou, A. S., Durand, A., Khalfaoui-Hassani, B., Bourbon, M. l., Astier, C., Bollivar, D. W., and Ouchane, S.: Coproporphyrin III excretion identifies the anaerobic coproporphyrinogen III oxidase HemN as a copper target in the Cu+-ATPase mutant copA− of R ubrivivax gelatinosus, Molecular microbiology, 88, 339-351, 2013.

560 Borgonie, G., Magnabosco, C., García-Moyano, A., Linage-Alvarez, B., Ojo, A. O., Freese, L. B., Van Jaarsveld, C., Van Rooyen, C., Kuloyo, O., Cason, E. D., Vermeulen, J., Pienaar, C., Van Heerden, E., Sherwood Lollar, B., Onstott, T. C., and Mundle, S. O. C.: New ecosystems in the deep subsurface follow the flow of water driven by geological activity, Scientific Reports, 9, 3310, 10.1038/s41598-019-39699-w, 2019.

Broxton, C. N. and Culotta, V. C.: SOD Enzymes and Microbial Pathogens: Surviving the Oxidative Storm of Infection, 565 PLoS pathogens, 12, e1005295-e1005295, 10.1371/journal.ppat.1005295, 2016.

Busenlehner, L. S., Pennella, M. A., and Giedroc, D. P.: The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance, FEMS microbiology reviews, 27, 131-143, 2003.

26

Biogeosciences

Discussions

EGU

Open Access

Caspi, R., Billington, R., Keseler, I. M., Kothari, A., Krummenacker, M., Midford, P. E., Ong, W. K., Paley, S., Subhraveti, P., and Karp, P. D.: The MetaCyc database of metabolic pathways and enzymes - a 2019 update, Nucleic Acids Research, 570 48, D445-D453, 10.1093/nar/gkz862, 2019.

Cha, J.-S. and Cooksey, D. A.: Copper hypersensitivity and uptake in Pseudomonas syringae containing cloned components of the copper resistance operon, Applied and environmental microbiology, 59, 1671-1674, 1993.

Chang, Y.-J., Land, M., Hauser, L., Chertkov, O., Del Rio, T. G., Nolan, M., Copeland, A., Tice, H., Cheng, J.-F., and Lucas, S.: Non-contiguous finished genome sequence and contextual data of the filamentous soil bacterium Ktedonobacter 575 racemifer type strain (SOSP1-21 T), Standards in genomic sciences, 5, 97-111, 2011.

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., and Parks, D. H.: GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database, 2020.

Chen, S., Zhou, Y., Chen, Y., and Gu, J.: fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics, 34, i884-i890, 10.1093/bioinformatics/bty560, 2018.

580 Chillappagari, S., Seubert, A., Trip, H., Kuipers, O. P., Marahiel, M. A., and Miethke, M.: Copper stress affects iron homeostasis by destabilizing iron-sulfur cluster formation in Bacillus subtilis, Journal of bacteriology, 192, 2512-2524, 2010.

Donaire, A., Jiménez, B., Fernández, C. O., Pierattelli, R., Niizeki, T., Moratal, J.-M., Hall, J. F., Kohzuma, T., Hasnain, S. S., and Vila, A. J.: Metal− Ligand Interplay in Blue Copper Proteins Studied by 1H NMR Spectroscopy: Cu (II)− 585 Pseudoazurin and Cu (II)− Rusticyanin, Journal of the American Chemical Society, 124, 13698-13708, 2002.

Dong, X., Kleiner, M., Sharp, C. E., Thorson, E., Li, C., Liu, D., and Strous, M.: Fast and simple analysis of MiSeq amplicon sequencing data with MetaAmp, Frontiers in Microbiology, 8, 1461, 10.1101/131631, 2017.

Dube, J. P., Valverde, A., Steyn, J. M., Cowan, D. A., and Van der Waals, J. E.: Differences in bacterial diversity, composition and function due to long-term agriculture in soils in the eastern free State of South Africa, Diversity, 11, 590 61, 2019.

Dupont, C. L., Grass, G., and Rensing, C.: Copper toxicity and the origin of bacterial resistance—new insights and applications, Metallomics, 3, 1109-1118, 2011.

Dziewit, L., Pyzik, A., Szuplewska, M., Matlakowska, R., Mielnicki, S., Wibberg, D., Schlüter, A., Pühler, A., and Bartosik, D.: Diversity and role of plasmids in adaptation of bacteria inhabiting the Lubin copper mine in Poland, an 595 environment rich in heavy metals, Frontiers in Microbiology, 6, 152, 2015.

Edgar, R. C.: Search and clustering orders of magnitude faster than BLAST, Bioinformatics, 26, 2460-2461, 10.1093/bioinformatics/btq461, 2010.

Edgar, R. C.: UPARSE: highly accurate OTU sequences from microbial amplicon reads, Nature Methods, 10, 996, 10.1038/nmeth.2604

600 https://www.nature.com/articles/nmeth.2604#supplementary-information, 2013.

Emmerich, H. P. H., A: Sekundärmineralbildung aus Grubenwässern im Kupferbergwerk Niedermarsberg [Secondary mineral formation from mine water in the Niedermarsberg copper mine], 1987.

Epelde, L., Lanzen, A., Blanco, F., Urich, T., and Garbisu, C.: Adaptation of soil microbial community structure and function to chronic metal contamination at an abandoned Pb-Zn mine, FEMS microbiology ecology, 91, 1-11, 2015.

605 Farrenschon, J., Oesterreich, B., Blumenstein, S., Holzinger, M., and Wrede, V.: Geologische Karte von Nordrhein-Westfalen 1:25 000, Erläuterungen zu Blatt 4519 Marsberg [Geological Map of North Rhine-Westphalia, Explanatory notes to sheet 4519 Marsberg], 2008.

Festa, R. A. and Thiele, D. J.: Copper: an essential metal in biology, Current Biology, 21, R877-R883, 2011.

Fong, S. T., Camakaris, J., and Lee, B. T. O.: Molecular genetics of a chromosomal locus involved in copper tolerance in 610 Escherichia coli K-12, Molecular microbiology, 15, 1127-1137, 1995.

Fu, Y., Chang, F.-M. J., and Giedroc, D. P.: Copper transport and trafficking at the host-bacterial pathogen interface, Accounts of chemical research, 47, 3605-3613, 10.1021/ar500300n, 2014.

Fuchs, G., Boll, M., and Heider, J.: Microbial degradation of aromatic compounds — from one strategy to four, Nature Reviews Microbiology, 9, 803-816, 10.1038/nrmicro2652, 2011.

615 Ghosal, D., Ghosh, S., Dutta, T. K., and Ahn, Y.: Current State of Knowledge in Microbial Degradation of Polycyclic Aromatic Hydrocarbons (PAHs): A Review, Frontiers in Microbiology, 7, 1369, 2016.

27

Gibson, M. M., Bagga, D. A., Miller, C. G., and Maguire, M. E.: Magnesium transport in Salmonella typhimurium: the influence of new mutations conferring Co2+ resistance on the CorA Mg2+ transport system, Molecular microbiology, 5, 2753-2762, 1991.

620 Glöckner, F. O.: The SILVA Database Project: An ELIXIR core data resource for high-quality ribosomal RNA sequences, Biodiversity Information Science and Standards, 3, e36125, 2019.

González-Guerrero, M. and Argüello, J. M.: Mechanism of Cu+-transporting ATPases: soluble Cu+ chaperones directly transfer Cu+ to transmembrane transport sites, Proceedings of the National Academy of Sciences, 105, 5992-5997, 2008.

625 Gordon, A. S., Howell, L. D., and Harwood, V.: Responses of diverse heterotrophic bacteria to elevated copper concentrations, Canadian journal of microbiology, 40, 408-411, 1994.

Gupta, A., Dutta, A., Sarkar, J., Panigrahi, M. K., and Sar, P.: Low-abundance members of the Firmicutes facilitate bioremediation of soil impacted by highly acidic mine drainage from the Malanjkhand copper project, India, Frontiers in Microbiology, 9, 2882, 2018.

630 Hernández-Montes, G., Argüello, J. M., and Valderrama, B.: Evolution and diversity of periplasmic proteins involved in copper homeostasis in gamma proteobacteria, BMC microbiology, 12, 249, 2012.

Hobman, J. L., Yamamoto, K., and Oshima, T.: Transcriptomic responses of bacterial cells to sublethal metal ion stress, in: Molecular microbiology of heavy metals, Springer, 73-115, 2007.

Huang, K., Xu, Y., Packianathan, C., Gao, F., Chen, C., Zhang, J., Shen, Q., Rosen, B. P., and Zhao, F.-J.: Arsenic methylation 635 by a novel ArsM As(III) S-adenosylmethionine methyltransferase that requires only two conserved cysteine residues, Molecular microbiology, 107, 265-276, 10.1111/mmi.13882, 2018.

Jiang, Z., Li, P., Jiang, D., Dai, X., Zhang, R., Wang, Y., and Wang, Y.: Microbial community structure and arsenic biogeochemistry in an acid vapor-formed spring in Tengchong Geothermal Area, China, PloS one, 11, e0146331, 2016.

Jones, D. S., Lapakko, K. A., Wenz, Z. J., Olson, M. C., Roepke, E. W., Sadowsky, M. J., Novak, P. J., and Bailey, J. V.: Novel 640 microbial assemblages dominate weathered sulfide-bearing rock from copper-nickel deposits in the Duluth complex, Minnesota, USA, Applied and environmental microbiology, 83, 2017.

Karp, P. D., Latendresse, M., Paley, S. M., Krummenacker, M., Ong, Q. D., Billington, R., Kothari, A., Weaver, D., Lee, T., Subhraveti, P., Spaulding, A., Fulcher, C., Keseler, I. M., and Caspi, R.: Pathway Tools version 19.0 update: software for pathway/genome informatics and systems biology, Briefings in Bioinformatics, 17, 877-890, 10.1093/bib/bbv079, 645 2015.

Karpe, P. D., Latendresse, M., and Caspi, R.: The pathway tools pathway prediction algorithm, Standards in genomic sciences, 5, 424-429, 2011.

Kersey, C. M., Agyemang, P. A., and Dumenyo, C. K.: CorA, the magnesium/nickel/cobalt transporter, affects virulence and extracellular enzyme production in the soft rot pathogen Pectobacterium carotovorum, Molecular plant 650 pathology, 13, 58-71, 2012.

Kertesz, M. A.: Bacterial transporters for sulfate and organosulfur compounds, Research in Microbiology, 152, 279-290, 2001.

Kim, B.-E., Nevitt, T., and Thiele, D. J.: Mechanisms for copper acquisition, distribution and regulation, Nature chemical biology, 4, 176-185, 2008.

655 Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F. O.: Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies, Nucleic Acids Research, 41, e1, 10.1093/nar/gks808, 2012.

Koch, D., Nies, D. H., and Grass, G.: The RcnRA (YohLM) system of Escherichia coli: a connection between nickel, cobalt and iron homeostasis, Biometals, 20, 759-771, 2007.

660 Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., and Knight, R.: UniFrac: an effective distance metric for microbial community comparison, The ISME Journal, 5, 169-172, 10.1038/ismej.2010.133, 2011.

Macomber, L. and Imlay, J. A.: The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity, Proceedings of the National Academy of Sciences, 106, 8344-8349, 2009.

28

Biogeosciences
Discussions

665   Makui, H., Roig, E., Cole, S. T., Helmann, J. D., Gros, P., and Cellier, M. F. M.: Identification of the Escherichia coli K-12 Nramp orthologue (MntH) as a selective divalent metal ion transporter, Molecular Microbiology, 35, 1065-1078, 10.1046/j.1365-2958.2000.01774.x, 2000.

Markovich, D.: Physiological roles and regulation of mammalian sulfate transporters, Physiological reviews, 81, 1499-1533, 2001.

670   Martin, P., DeMel, S., Shi, J., Gladysheva, T., Gatti, D. L., Rosen, B. P., and Edwards, B. F. P.: Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme, Structure, 9, 1071-1081, 2001.

Mengoni, A. and Bazzicalupo, M.: The statistical treatment of data and the Analysis of MOlecular VAriance (AMOVA) in molecular microbial ecology, Annals of Microbiology, 52, 2002.

Newton, G. L., Av-Gay, Y., and Fahey, R. C.: A Novel Mycothiol-Dependent Detoxification Pathway in Mycobacteria
675   Involving Mycothiol S-Conjugate Amidase, Biochemistry, 39, 10739-10746, 10.1021/bi000356n, 2000.

Newton, G. L., Buchmeier, N., and Fahey, R. C.: Biosynthesis and Functions of Mycothiol, the Unique Protective Thiol of Actinobacteria, Microbiology and Molecular Biology Reviews, 72, 471, 10.1128/MMBR.00008-08, 2008.

Nies, D. H.: Efflux-mediated heavy metal resistance in prokaryotes, FEMS microbiology reviews, 27, 313-339, 2003.

680   Novoa-Aponte, L., Ramírez, D., and Argüello, J. M.: The interplay of the metallosensor CueR with two distinct CopZ chaperones defines copper homeostasis in Pseudomonas aeruginosa, Journal of Biological Chemistry, 294, 4934-4945, 2019.

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A.: metaSPAdes: a new versatile metagenomic assembler, Genome research, 27, 824-834, 10.1101/gr.213959.116, 2017.

685   Osman, D., Waldron, K. J., Denton, H., Taylor, C. M., Grant, A. J., Mastroeni, P., Robinson, N. J., and Cavet, J. S.: Copper homeostasis in Salmonella is atypical and copper-CueP is a major periplasmic metal complex, Journal of Biological Chemistry, 285, 25259-25268, 2010.

Outten, F. W., Huffman, D. L., Hale, J. A., and O'Halloran, T. V.: - The independent cue and cusSystems confer copper tolerance during aerobic and anaerobic growth inEscherichia coli, - 30670,

690   Papp-Wallace, K. M. and Maguire, M. E.: Manganese transport and the role of manganese in virulence, Annu. Rev. Microbiol., 60, 187-209, 2006.

Parkhurst, D. L. and Appelo, C.: Description of input and examples for PHREEQC version 3: a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations, US Geological Survey2328-7055, 2013.

695   Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W.: CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes, Genome research, 25, 1043-1055, 2015.

Parks, D. H., Chuvochina, M., Chaumeil, P.-A., Rinke, C., Mussig, A. J., and Hugenholtz, P.: Selection of representative genomes for 24,706 bacterial and archaeal species clusters provide a complete genome-based taxonomy, BioRxiv, 771964, 2019.

700   Paul Chowdhury, S., Babin, D., Sandmann, M., Jacquiod, S., Sommermann, L., Sørensen, S. J., Fliessbach, A., Mäder, P., Geistlinger, J., and Smalla, K.: Effect of long-term organic and mineral fertilization strategies on rhizosphere microbiota assemblage and performance of lettuce, Environmental microbiology, 21, 2426-2439, 2019.

Porcheron, G., Garénaux, A., Proulx, J., Sabri, M., and Dozois, C. M.: Iron, copper, zinc, and manganese transport and regulation in pathogenic Enterobacteria: correlations between strains, site of infection and the relative importance of
705   the different metal transport systems for virulence, Frontiers in cellular and infection microbiology, 3, 90, 2013.

Rademacher, C. and Masepohl, B.: Copper-responsive gene regulation in bacteria, Microbiology, 158, 2451, https://doi.org/10.1099/mic.0.058487-0, 2012.

Rae, T. D., Schmidt, P. J., Pufahl, R. A., Culotta, V. C., and O'Halloran, T. V.: Undetectable intracellular free copper: the requirement of a copper chaperone for superoxide dismutase, Science, 284, 805-808, 1999.

710   Redinbo, M. R., Yeates, T. O., and Merchant, S.: Plastocyanin: structural and functional analysis, Journal of bioenergetics and biomembranes, 26, 49-66, 1994.

29

Biogeosciences
Discussions

Reeve, W. G., Tiwari, R. P., Kale, N. B., Dilworth, M. J., and Glenn, A. R.: ActP controls copper homeostasis in Rhizobium leguminosarum bv. viciae and Sinorhizobium meliloti preventing low pH-induced copper toxicity, Molecular microbiology, 43, 981-991, 2002.

715 Rensing, C. and Mitra, B.: Zinc, cadmium, and lead resistance and homeostasis, in: Molecular microbiology of heavy metals, Springer, 321-341, 2007.

Richard, H. and Foster, J. W.: Escherichia coli glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential, Journal of bacteriology, 186, 6032-6041, 10.1128/JB.186.18.6032-6041.2004, 2004.

720 Ridge, J. P., Lin, M., Larsen, E. I., Fegan, M., McEwan, A. G., and Sly, L. I.: A multicopper oxidase is essential for manganese oxidation and laccase-like activity in Pedomicrobium sp. ACM 3067, Environmental Microbiology, 9, 944-953, 2007.

Rosen, B. P.: Biochemistry of arsenic detoxification, FEBS Letters, 529, 86-92, https://doi.org/10.1016/S0014-5793(02)03186-1, 2002.

725 Rowland, J. L. and Niederweis, M.: A multicopper oxidase is required for copper resistance in Mycobacterium tuberculosis, Journal of bacteriology, 195, 3724-3733, 2013.

Sambrook, J. and Russell, D. W.: Molecular cloning: A laboratory manual, the third edition, Cold spring harbor laboratory press, cold spring harbor, New York2001.

Schippers, A. and Sand, W.: Bacterial leaching of metal sulfides proceeds by two indirect mechanisms via thiosulfate 730 or via polysulfides and sulfur, Applied and environmental microbiology, 65, 319-321, 1999.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., and Weber, C. F.: Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities, Applied and Environmental Microbiology, 75, 7537, 10.1128/AEM.01541-09, 2009.

735 Seemann, T.: Prokka: rapid prokaryotic genome annotation, Bioinformatics, 30, 2068-2069, 2014.

Shi, X., Festa, R. A., Ioerger, T. R., Butler-Wu, S., Sacchettini, J. C., Darwin, K. H., and Samanovic, M. I.: The copper-responsive RicR regulon contributes to Mycobacterium tuberculosis virulence, MBio, 5, 2014.

Siegmund, H., Trappe, J., and Oschmann, W.: Sequence stratigraphic and genetic aspects of the Tournaisian "LiegenderAlaunschiefer" and adjacent beds, International Journal of Earth Sciences, 91, 934-949, 10.1007/s00531-740 001-0252-9, 2002.

Silver, S. and Hobman, J. L.: Mercury Microbiology: Resistance Systems, Environmental Aspects, Methylation,and Human Health, in: Molecular Microbiology of Heavy Metals, edited by: Nies, D. H., and Silver, S., Springer Berlin Heidelberg, Berlin, Heidelberg, 357-370, 10.1007/7171_2006_085, 2007.

Silver, S. and Phung, L. T.: A bacterial view of the periodic table: genes and proteins for toxic inorganic ions, Journal of 745 Industrial Microbiology and Biotechnology, 32, 587-605, 2005.

Silver, S. and Walderhaug, M.: Gene regulation of plasmid-and chromosome-determined inorganic ion transport in bacteria, Microbiology and Molecular Biology Reviews, 56, 195-228, 1992.

Sitthisak, S., Knutsson, L., Webb, J. W., and Jayaswal, R. K.: Molecular characterization of the copper transport system in Staphylococcus aureus, Microbiology, 153, 4274-4283, 2007.

750 Smaldone, G. T. and Helmann, J. D.: CsoR regulates the copper efflux operon copZA in Bacillus subtilis, Microbiology (Reading, England), 153, 4123, 2007.

Stecher, G., Tamura, K., and Kumar, S.: Molecular Evolutionary Genetics Analysis (MEGA) for macOS, Molecular Biology and Evolution, 37, 1237-1239, 10.1093/molbev/msz312, 2020.

Stribrny, B.: Die Kupfererzlagerstätte Marsberg im Rheinischen Schiefergebirge-Rückblick und Stand der Forschung, 755 Erzmetall, 40, 423-427, 1987.

Stribrny, B. and Urban, H.: Zur Geologie und Lagerstättenbildung des Kupfererzvorkommens von Marsberg im Rheinischen Schiefergebirge [O n the geology and formation of the copper ore deposit of Marsberg in the Rhenish region Slate Mountains], 2000.

30

760 Thomas Iv, J. C., Oladeinde, A., Kieran, T. J., Finger Jr, J. W., Bayona-Vásquez, N. J., Cartee, J. C., Beasley, J. C., Seaman, J. C., McArthur, J. V., and Rhodes Jr, O. E.: Co-occurrence of antibiotic, biocide, and heavy metal resistance genes in bacteria from metal and radionuclide contaminated soils at the Savannah River Site, Microbial biotechnology, 2020.

Thompson, A. K., Gray, J., Liu, A., and Hosler, J. P.: The roles of Rhodobacter sphaeroides copper chaperones PCuAC and Sco (PrrC) in the assembly of the copper centers of the aa3-type and the cbb3-type cytochrome c oxidases, Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1817, 955-964, 2012.

765 Tsai, M.-F. and Miller, C.: Substrate selectivity in arginine-dependent acid resistance in enteric bacteria, Proceedings of the National Academy of Sciences, 110, 5893, 10.1073/pnas.1301442110, 2013.

Urban, H., Stribrny, B., Zereini, F., and Ye, Y.: Geochemistry and metallogenesis of Lower Carboniferous black shale-hosted ore deposits, NE Rhenish Massif, FR Germany, Ore Geology Reviews, 9, 427-443, https://doi.org/10.1016/0169-1368(94)00023-H, 1995.

770 Villafane, A., Voskoboynik, Y., Ruhl, I., Sannino, D., Maezato, Y., Blum, P., and Bini, E.: CopR of Sulfolobus solfataricus represents a novel class of archaeal-specific copper-responsive activators of transcription, Microbiology, 157, 2808, https://doi.org/10.1099/mic.0.051862-0, 2011.

Vita, N., Landolfi, G., Baslé, A., Platsaki, S., Lee, J., Waldron, K. J., and Dennison, C.: Bacterial cytosolic proteins with a high capacity for Cu(I) that protect against copper toxicity, Scientific Reports, 6, 39065, 10.1038/srep39065, 2016.

775 Waters, L. S., Sandoval, M., and Storz, G.: The Escherichia coli MntR miniregulon includes genes encoding a small protein and an efflux pump required for manganese homeostasis, Journal of bacteriology, 193, 5887-5897, 2011.

Wijekoon, C. J., Udagedara, S. R., Knorr, R. L., Dimova, R., Wedd, A. G., and Xiao, Z.: Copper ATPase CopA from Escherichia coli: quantitative correlation between ATPase activity and vectorial copper transport, Journal of the American Chemical Society, 139, 4266-4269, 2017.

780 Wolfram, L., Friedrich, B., and Eitinger, T.: The Alcaligenes eutrophus protein HoxN mediates nickel transport in Escherichia coli, Journal of Bacteriology, 177, 1840-1843, 1995.

Wu, Y.-W., Simmons, B. A., and Singer, S. W.: MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets, Bioinformatics, 32, 605-607, 10.1093/bioinformatics/btv638, 2015.

Xu, L., Dong, Z., Fang, L., Luo, Y., Wei, Z., Guo, H., Zhang, G., Gu, Y. Q., Coleman-Derr, D., and Xia, Q.: OrthoVenn2: a web 785 server for whole-genome comparison and annotation of orthologous clusters across multiple species, Nucleic acids research, 47, W52-W58, 2019.

Zaballa, M.-E., Abriata, L. A., Donaire, A., and Vila, A. J.: Flexibility of the metal-binding region in apo-cupredoxins, Proceedings of the National Academy of Sciences, 109, 9254-9259, 2012.

Zeng, X.-C., Yang, Y., Shi, W., Peng, Z., Chen, X., Zhu, X., and Wang, Y.: Microbially Mediated Methylation of Arsenic in the 790 Arsenic-Rich Soils and Sediments of Jianghan Plain, Frontiers in Microbiology, 9, 1389, 2018.

Zhang, Z., Zhang, Z., Chen, H., Liu, J., Liu, C., Ni, H., Zhao, C., Ali, M., Liu, F., and Li, L.: Surface Mn (II) oxidation actuated by a multicopper oxidase in a soil bacterium leads to the formation of manganese oxide minerals, Scientific reports, 5, 10895, 2015.

Oskar, G. and William, P.: Erläuterungen zur geologischen Karte von Preußen und benachbarten deutschen Ländern, 795 Blatt 4519 Marberg [Geological map of Prussia and neighboring German countries, 4519 Marsberg], 1936.

31

Supplementary Information

Composition and Niche-Specific Characteristics of Microbial Consortia colonizing a Copper Mine in the Rhenish Massif

Supplementary Information

Sania Arif[1], Heiko Nacke[2], Elias Schliekmann[1], Andreas Reimer[3], Gernot Arp[3], and Michael Hoppert [1]

1Department of General Microbiology, Institute of Microbiology and Genetics,

2Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics,

3Geoscience Centre, Department of Geobiology, Georg-August-Universität Göttingen, Germany

*Correspondence to sarif@gwdg.de

Figure S1, Figure S2, Figure S3, Figure S4, Figure S5, Figure S6, Figure S7, Table S1, Table S2

Figure S1. Map of sampling site (Kilianstollen, Marsberg) in the Rhenish Massif mountain range. Reproduced from OpenTopoMap (https://opentopomap.org), under creative common license CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0/legalcode), the map was subsequently modified with the sampling sites (red pushpins) and simplified underground Kilian copper mine adit.
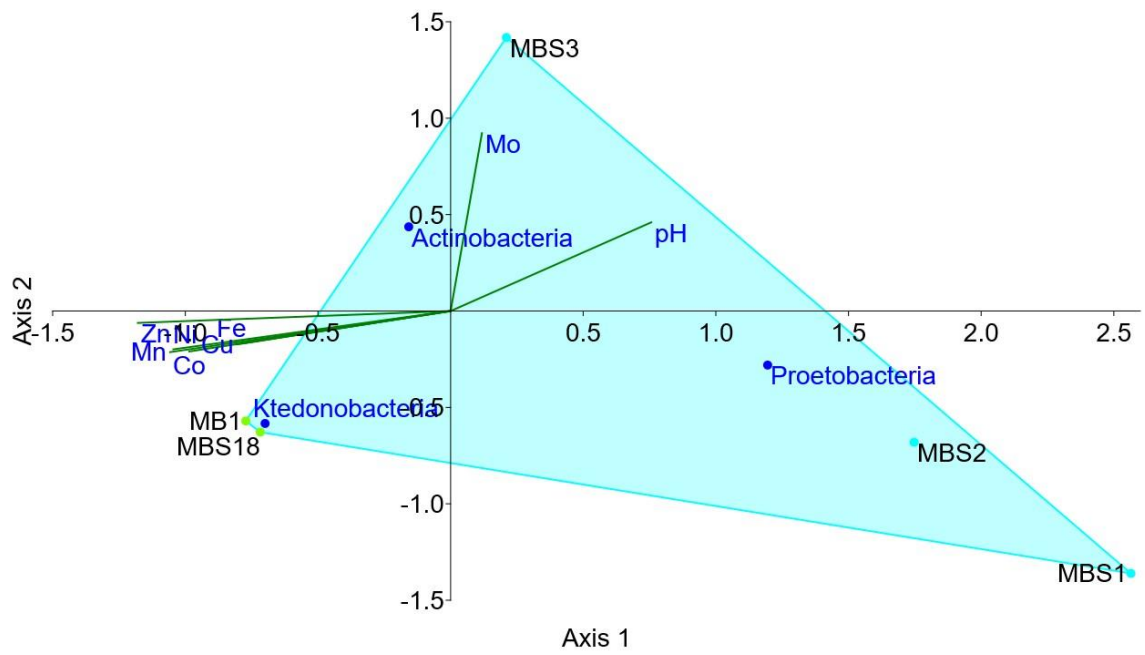
Figure S2. Canonical correspondence analysis of the biological samples collected at the leachate and spring water site. The heavy metals concentrations were taken as the abiotic factors against the abundance of *Proteobacteria*, *Actinobacteria*, and *Ktedonobacteria* for the representative of leachate (MB1, MBS18) and spring water (MBS1, MBS2, MBS3) groups.
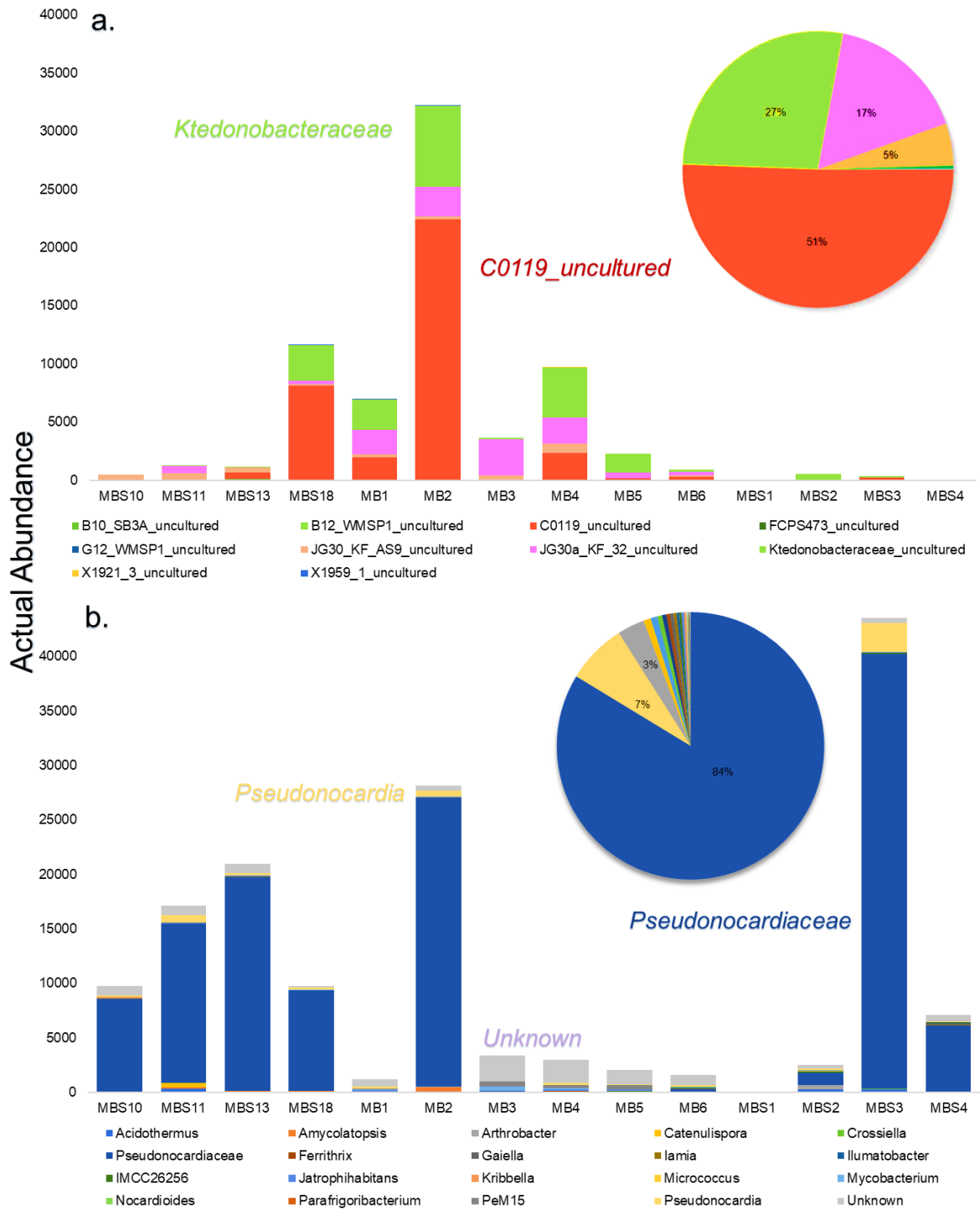
Figure S3. Actual abundance of *Ktedonobacteria* and *Actinobacteria* OTUs at genus level. The actual abundance of OTUs belonging to *Ktedonobacteria* (a) and *Actinobacteria* (b) class were observed at genus level, without any normalization, to analyse the low abundant OTUs. The pie chart insert shows the overall abundance of the major genus OTUs.
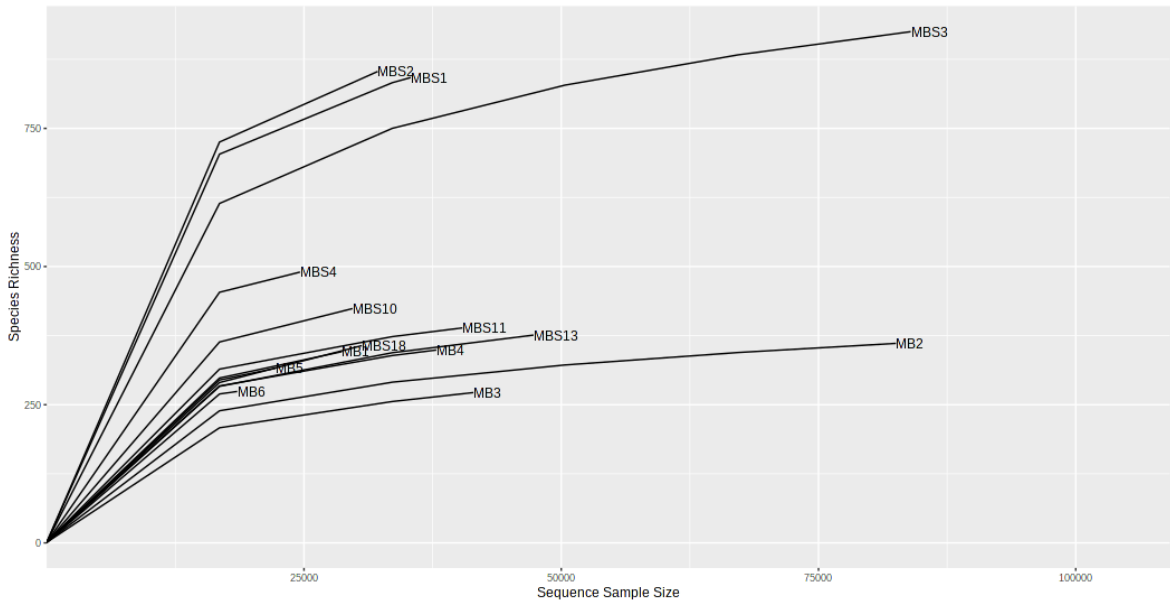
Figure S4. Rarefaction Curve Analysis. The samples are plotted in terms of OTUs observed (species richness) and sampled reads (sequence sample size).
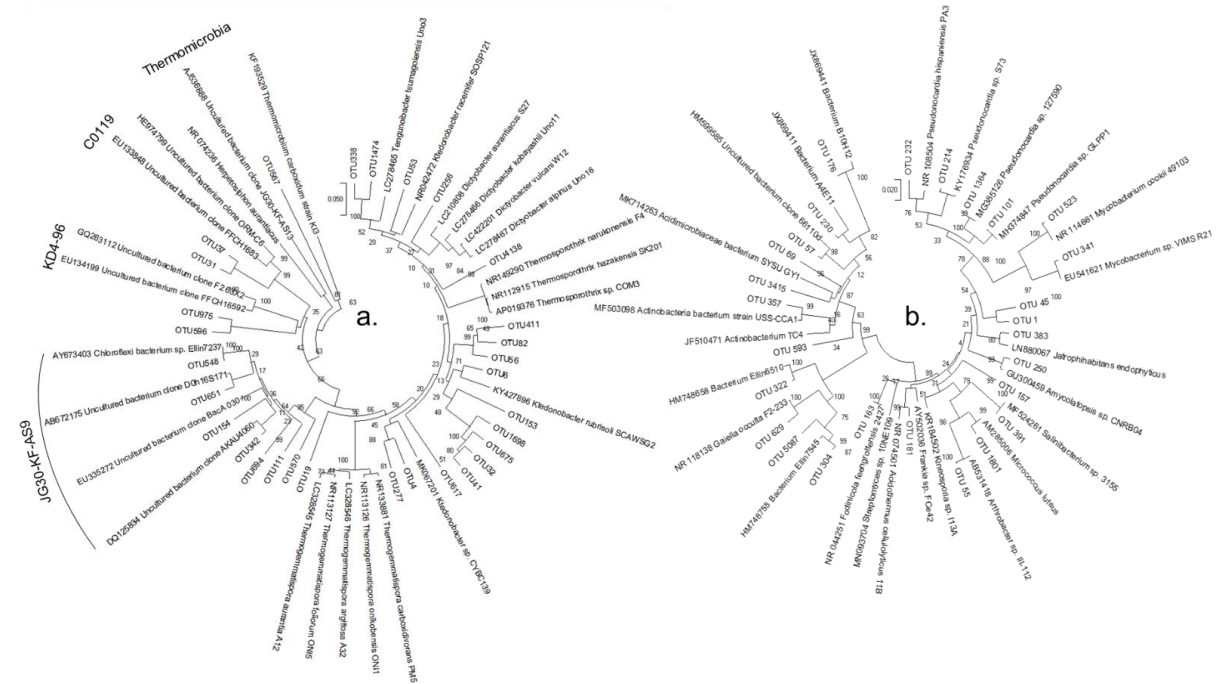
Figure S5. Phylogenetic tree of the small subunit SSU ribosomal RNA genes of OTUs. *Ktedonobacteria* **(a) and** *Actinobacteria* **(b).** Maximum-likelihood phylogenetic tree constructed based on V3-V4 amplicon sequences OTUs of *Ktedonobacteria* 16S rRNA gene (a) with the validly published *Ktedonobacteria* species. For the *Actinobacteria* class OTUs (b), the highest blast hits sequences were used. Bootstrap values are given at the branching points.
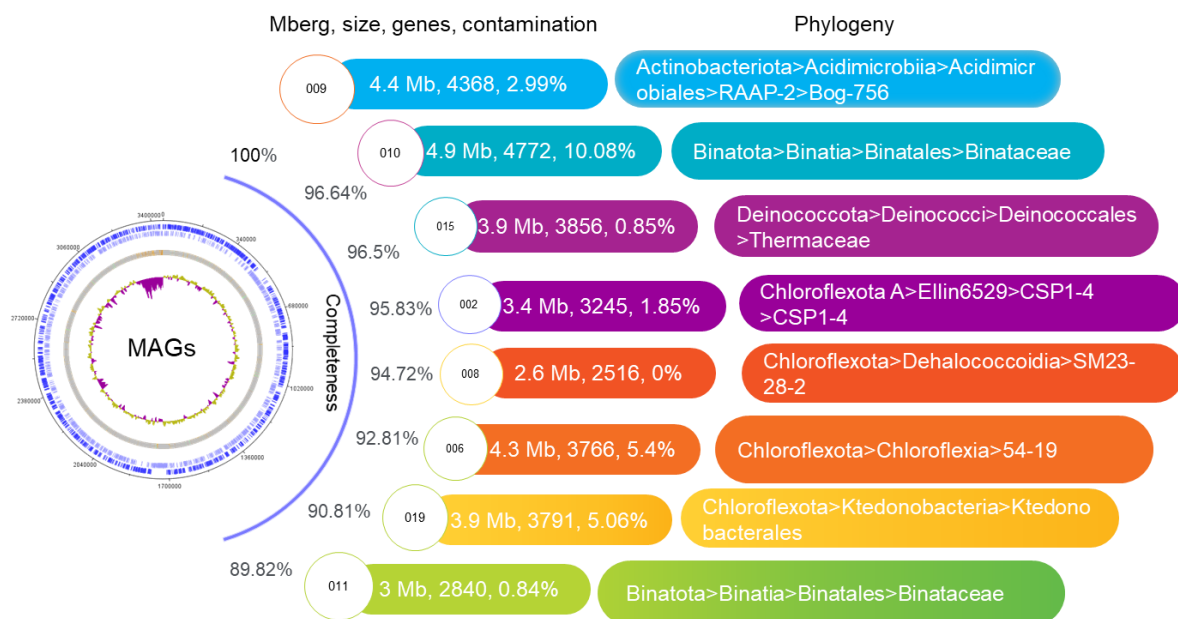
Figure S6. Overview of metagenome-assembled genomes and their key features

Table S1. Alpha diversity indexes of the Marsberg environmental samples

	p-value: [ANOVA]	F-value
Chao1	0.00030265	18.499
ACE	0.0002466	19.409
Shannon p-value	0.025179	5.242

Table S2. Analysis of molecular variance and unifrac analysis of the Marsberg samples

Unifrac.weighted

Tree#	Groups	UWScore	UWSig
1	L-M	0.953519	0.0970
1	L-W	0.980299	0.0340
1	M-W	0.981865	0.1450

AMOVA (Analysis of Molecular Variance)

L-M-W	Among	Within	Total
SS	1.37542	2.93434	4.30976
df	2	11	13
MS	0.687712	0.266758	

Fs: 2.57803
p-value: <0.001*

L-M	Among	Within	Total
SS	0.331903	1.82319	2.15509
df	1	8	9
MS	0.331903	0.227898	

Fs: 1.45637
p-value: 0.077

L-W	Among	Within	Total
SS	0.979114	2.52913	3.50824
df	1	9	10
MS	0.979114	0.281014	

Fs: 3.48422
p-value: 0.002*

M-W	Among	Within	Total
SS	0.738092	1.51637	2.25446
df	1	5	6
MS	0.738092	0.303274	

Fs: 2.43375
p-value: 0.024

L Leachate M unconsolidated rocks W Spring water

3.1 16S rRNA Amplicon Sequencing of Microbial Biofilms from Marsberg Copper Mine, Germany

16S rRNA Amplicon Sequencing of Microbial Biofilms from Marsberg Copper Mine, Germany

Sania Arif,[a] Elias Schliekmann,[a] Michael Hoppert[a]

[a]Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August-Universität, Göttingen, Germany

ABSTRACT The 16S rRNA amplicons from biofilms inhabiting rocks near various water bodies of Marsberg Copper Mine (Rhenish Massif, Germany) reveal the diversity of their microbial communities. The abundance of *Chloroflexi* and *Cyanobacteria* taxa in the biofilms near leachate streams indicated the selective enrichment of *Ktedonobacteria* and *Oxyphotobacteria* members.

Marsberg Kilianstollen Copper Mine (Hochsauerland District, North Rhine-Westphalia, Germany) is known for its copper-rich Lower Carboniferous alum shales and lydites as well as copper remineralizations; previous investigations revealed the geochemistry of heavy metals, copper in particular, in (acidic) sulfidic mine waters (1). In this study, selective enrichment of putatively heavy metal-tolerant microbial communities under copper stress is revealed by 16S rRNA amplicon sequencing of biofilms, sampled on rock surfaces near a spring water stream, on mine unconsolidated rock samples, and in acidic leachate stream sediment.

In February 2018, at Marsberg Copper Mine (51.453502°N, 8.861703°E), biofilms growing near or in diverse mine drainages, ranging from fresh stream water to copper-rich acidic leachate, were aseptically collected and stored at $-20°C$. The biofilm samples MB1 through MB6 were collected from colonized rocks of the adit wall, within 25 to 35 cm of an acidic leachate stream, MBS18 was taken directly from the acidic leachate outflow stream sediment, samples MBS1 through MBS4 were collected from colonized rocks near the fresh spring water stream, and the biofilms from unconsolidated mine rocks and a wooden plank used for timbering were labeled as MBS11, MBS13, and MBS10 (Table 1). The microbial genomic DNA extraction was carried out by following the standard procedures of the PowerSoil DNA isolation kit (Qiagen, Venlo, the Netherlands). The 16S rRNA V3 to V4 hypervariable regions were amplified with the MiSeq 16S amplicon PCR forward primer 5'-GTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3' and reverse primer 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATC-3') (2). The amplicons were purified from primers and PCR reagents via the GeneRead size selection kit (Qiagen), according to the manufacturer's instructions. The purified amplicons were sequenced using the MiSeq platform via the 2×300-bp paired-end method and a MiSeq v3 reagent kit (Illumina). FASTQ reads were submitted to the automated pipeline for metagenomic analysis using MetaAmp (http://ebg.ucalgary.ca/metaamp/) with optimized parameters. The minimum length of overlap was set to 24 bp, and the maximum number of mismatches in the overlap region and primer sequences was adjusted to 4 (3). The pipeline employed UPARSE to perform operational taxonomic unit (OTU) clustering. Each OTU was taxonomically assigned at 97% sequence identity against the SILVA v138 database (https://mothur.org/wiki/silva_reference_files/) by Mothur (4, 5).

The prevalent phylum taxa were *Actinobacteria* (27%), *Chloroflexi* (17%), *Proteobacteria* (16%), *Cyanobacteria* (12%), *Acidobacteria* (10%), and *Bacteroidetes* (6%) in all samples (MBS1 through MBS4, MB1 through MB6, MBS10, MBS11, MBS13, MBS18).

87

TABLE 1 Sampling sites and sample information[a]

Sampling site	Sample name	SRA accession no.	Sample type
Spring water stream	1MBS	SRR12876555	Colonized rock sample
	2MBS	SRR12876554	
	3MBS	SRR12876549	Whitish terrestrial biofilm
	4MBS	SRR12876550	
Zement-Kupferplatte copper precipitation flume drained with acidic orange-colored leachate	MB1	SRR12876544	Colonized rock sample
	MB2	SRR12876543	
	MB3	SRR12876542	Greenish-white terrestrial biofilm
	MB4	SRR12876553	
	MB5	SRR12876552	
	MB6	SRR12876551	
Supporting wooden plank	10MBS	SRR12876548	Soft surface biomass
Mine unconsolidated rock	11MBS	SRR12876547	Whitish terrestrial biofilm
	13MBS	SRR12876546	
Outflow greenish-blue-colored leachate stream	18MBS	SRR12876545	Liquid sample

[a] BioProject accession number PRJNA670497.

According to the UniFrac weighted algorithm (6), the microbial consortium of the spring water and leachate sample groups were distinct from each other ($P = 0.034$), while the unconsolidated rock group microbiome shared similarity with both the spring water and leachate samples ($P > 0.097$). *Ktedonobacteria*, *Cyanobacteria*, and *Actinobacteria* contributed 26%, 23%, and 16%, respectively, to the biofilm community (MB1 through MB6 and MBS18) growing near the copper acidic leachate stream. The genus-level analysis of all samples identified 10 distinct uncultured *Ktedonobacteria* genera; 80% of these could be classified within the *Ktedonobacteraceae* family. Marsberg Kilianstollen offers a large reservoir of uncultured novel heavy metal-resistant microbial strains, especially belonging to the class *Ktedonobacteria* of *Chloroflexi*.

Data availability. The 16S rRNA gene amplicon raw read data sets have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA670497, and the individual samples under the accession numbers SRR12876542, SRR12876543, SRR12876544, SRR12876545, SRR12876546, SRR12876547, SRR12876548, SRR12876549, SRR12876550, SRR12876551, SRR12876552, SRR12876553, SRR12876554, and SRR12876555 are publicly published.

REFERENCES

1. Emmerich HP, Heydemann A. 1987. Sekundärmineralbildung aus Grubenwässern im Kupferbergwerk Niedermarsberg [Secondary mineral formation from mine water in the Niedermarsberg copper mine]. Der Aufschluss 38:149–156. (In German.)
2. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41:e1. https://doi.org/10.1093/nar/gks808.
3. Dong X, Kleiner M, Sharp CE, Thorson E, Li C, Liu D, Strous M. 2017. Fast and simple analysis of MiSeq amplicon sequencing data with MetaAmp. Front Microbiol 8:1461. https://doi.org/10.3389/fmicb.2017.01461.
4. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541. https://doi.org/10.1128/AEM.01541-09.
5. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10:996–998. https://doi.org/10.1038/nmeth.2604.
6. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. 2011. UniFrac: an effective distance metric for microbial community comparison. ISME J 5:169–172. https://doi.org/10.1038/ismej.2010.133.

88

3.2 Metagenome-Assembled Genome Sequences of a Biofilm Derived from Marsberg Copper Mine, Germany

Metagenome-Assembled Genome Sequences of a Biofilm Derived from Marsberg Copper Mine

Sania Arif,ᵃ Heiko Nacke,ᵇ Michael Hoppertᵃ

ᵃDepartment of General Microbiology, Institute of Microbiology and Genetics, Georg-August-Universität, Göttingen, Germany
ᵇDepartment of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August University, Göttingen, Germany

ABSTRACT We sequenced the metagenome of a biofilm collected near a leachate stream of the Marsberg copper mine (Germany) and reconstructed eight metagenome-assembled genomes. These genomes yield copper resistance through Cu(I) oxidation via multiple copper oxidases and extrusion through copper-exporting P-type ATPases.

The historic Marsberg copper mine (51°27′12.6″N, 8°51′42.1″E) offers ambient natural conditions for the enrichment of heavy metal-resistant consortia under the influence of copper-rich (acidic) sulfidic mine waters at low temperature (10°C) (1). In February 2018, a biofilm near copper-rich leachate was aseptically collected from rock samples by using a sterile scalpel. Microbial DNA was extracted using the DNeasy PowerSoil kit (Qiagen, Venlo, Netherlands) according to the manufacturer's protocol. The purified DNA from the biofilm sample (designated MB1) was used to generate Illumina paired-end sequencing libraries with the Nextera DNA sample preparation kit (Illumina, San Diego, CA, USA); the libraries were sequenced by employing the MiSeq reagent kit v.3 and a MiSeq instrument as stated by the manufacturer (Illumina). Default parameters were used for all software unless otherwise specified. Quality trimming of reads was performed by employing fastp v.0.20.1 (qualified quality phred score, 20; minimal read length, 50 bp) (2) and yielded 18,235,972 paired-end reads. Base correction in overlapping regions (the correction option was selected concerning fastp-based quality trimming using default parameters; this option allows identification of overlapping regions of each pair of reads, and mismatched base pairs in these regions can be corrected if one base shows high quality and the other very low quality) and removal of the automatically detected adapter sequences were performed. Low-quality bases at the 5′ and 3′ ends of reads were trimmed once the mean quality score within a sliding window of 4 dropped below 20. Sequences were *de novo* assembled into 53,638 contigs of ≥1,000 bp via metaSPAdes v.3.14.0 (3). Binning was performed using MaxBin v.2.2.7 (minimum contig length, 1,000 bp; minimum probability for binning, 0.50) (4). Application of CheckM v.1.1.2 revealed eight relatively complete metagenome-assembled genomes (MAGs) (completeness, ≥89%; contamination rate, ≤10%) (5). Each MAG was annotated using Prokka v.1.14.5 (6). Subsequently, Prokka output was analyzed by using the Pathway Tools software v.23.5 (7) with the MetaCyc database v.23.5 (8). MAGs were classified taxonomically using GTDB-Tk v.1.0.2 and the Genome Taxonomy Database (GTDB) (release 89) (9, 10).

The MAGs were classified as members of *Actinobacteria* (Mberg 009), "*Candidatus* Binatota" (Mberg 010 and 011), *Chlorobacteria* (Mberg 002, 006, 008, and 019), and *Deinococcus-Thermus* (Mberg 015). Functional analysis revealed the presence of genes for copper-sensing transcriptional repressors CsoR and RicR, copper-exporting P-type ATPases such as ActP, CptA, and CopA, and oxidation enzymes, multicopper oxidases (MCOs), involved in copper homeostasis in all MAGs. The detoxification pathways for reactive oxygen species, toxins, and antibiotic compounds involve superoxide dismutase

(SOD) and peroxidases, which degrade superoxide anion radicals, and mycothiol-mediated detoxification through the enzyme Mca with thiols (11–13). All MAGs also include genes encoding aromatic compound degradation enzymes to generate ATP, which could potentially be used by copper-ATPase transporters.

Data availability. Raw sequencing data are available at the NCBI Sequence Read Archive (SRA) under accession number SRR12886061. The metagenome assembly is available at GenBank under accession number JADEYI000000000. The MAGs are available at GenBank under accession numbers JADMIG000000000, JADMIH000000000, JADMII000000000, JADMIJ000000000, JADMIK000000000, JADMIL000000000, JADMIM000000000, and JADMIN000000000. Prokka-based annotations of the eight MAG contigs are publicly available at the Göttingen Research Online Database (https://doi.org/10.25625/ODCARY).

REFERENCES

1. Emmerich HP, Heydemann A. 1987. Sekundärmineralbildung aus Grubenwässern im Kupferbergwerk Niedermarsberg [Secondary mineral formation from mine water in the Niedermarsberg copper mine]. Der Aufschluss 38:149–156. (In German.)
2. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
3. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834. https://doi.org/10.1101/gr.213959.116.
4. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638.
5. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
6. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
7. Karp PD, Latendresse M, Paley SM, Krummenacker M, Ong QD, Billington R, Kothari A, Weaver D, Lee T, Subhraveti P, Spaulding A, Fulcher C, Keseler IM, Caspi R. 2016. Pathway Tools version 19.0 update: software for pathway/genome informatics and systems biology. Brief Bioinform 17:877–890. https://doi.org/10.1093/bib/bbv079.
8. Caspi R, Billington R, Keseler IM, Kothari A, Krummenacker M, Midford PE, Ong WK, Paley S, Subhraveti P, Karp PD. 2020. The MetaCyc database of metabolic pathways and enzymes: a 2019 update. Nucleic Acids Res 48:D445–D453. https://doi.org/10.1093/nar/gkz862.
9. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848.
10. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. 2019. Selection of representative genomes for 24,706 bacterial and archaeal species clusters provide a complete genome-based taxonomy. bioRxiv 771964. https://doi.org/10.1101/771964.
11. Newton GL, Buchmeier N, Fahey RC. 2008. Biosynthesis and functions of mycothiol, the unique protective thiol of *Actinobacteria*. Microbiol Mol Biol Rev 72:471–494. https://doi.org/10.1128/MMBR.00008-08.
12. Newton GL, Av-Gay Y, Fahey RC. 2000. A novel mycothiol-dependent detoxification pathway in mycobacteria involving mycothiol *S*-conjugate amidase. Biochemistry 39:10739–10746. https://doi.org/10.1021/bi000356n.
13. Broxton CN, Culotta VC. 2016. SOD enzymes and microbial pathogens: surviving the oxidative storm of infection. PLoS Pathog 12:e1005295. https://doi.org/10.1371/journal.ppat.1005295.

90

4. Sasso Pisano Geothermal Field Environment Harbours Diverse *Ktedonobacteria* Representatives and Illustrates Habitat-Specific Adaptations

Sania Arif [1]*, Corinna Willenberg [1], Annika Dreyer [1], Heiko Nacke [2] and Michael Hoppert [1]

[1]Institute of Microbiology and Genetics, Department of General Microbiology, University of Göttingen, Grisebachstraße 8, 37077 Göttingen, Germany

[2]Institute of Microbiology and Genetics, Department of Genomic and Applied Microbiology, University of Göttingen, Grisebachstraße 8, 37077 Göttingen, Germany

*Correspondence: Email: sarif@gwdg.de

Author Contributions:

Article

Sasso Pisano Geothermal Field Environment Harbours Diverse *Ktedonobacteria* Representatives and Illustrates Habitat-Specific Adaptations

Sania Arif [1,*], **Corinna Willenberg** [1], **Annika Dreyer** [1], **Heiko Nacke** [2] **and Michael Hoppert** [1]

[1] Department of General Microbiology, Institute of Microbiology and Genetics, University of Göttingen, Grisebachstraße 8, 37077 Göttingen, Germany; corinna.willenberg@stud.uni-goettingen.de (C.W.); annika.dreyer@stud.uni-goettingen.de (A.D.); mhopper@gwdg.de (M.H.)

[2] Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, University of Göttingen, Grisebachstraße 8, 37077 Göttingen, Germany; hnacke@gwdg.de

* Correspondence: sarif@gwdg.de; Tel.: +49-551-39-9657; Fax: +49-551-39-3808

Abstract: The hydrothermal steam environment of Sasso Pisano (Italy) was selected to investigate the associated microbial community and its metabolic potential. In this context, 16S and 18S rRNA gene partial sequences of thermophilic prokaryotes and eukaryotes inhabiting hot springs and fumaroles as well as mesophilic microbes colonising soil and water were analysed by high-throughput amplicon sequencing. The eukaryotic and prokaryotic communities from hot environments clearly differ from reference microbial communities of colder soil sites, though *Ktedonobacteria* showed high abundances in various hot spring samples and a few soil samples. This indicates that the hydrothermal steam environments of Sasso Pisano represent not only a vast reservoir of thermophilic but also mesophilic members of this *Chloroflexi* class. Metabolic functional profiling revealed that the hot spring microbiome exhibits a higher capability to utilise methane and aromatic compounds and is more diverse in its sulphur and nitrogen metabolism than the mesophilic soil microbial consortium. In addition, heavy metal resistance-conferring genes were significantly more abundant in the hot spring microbiome. The eukaryotic diversity at a fumarole indicated high abundances of primary producers (unicellular red algae: *Cyanidiales*), consumers (*Arthropoda: Collembola* sp.), and endoparasite *Apicomplexa* (*Gregarina* sp.), which helps to hypothesise a simplified food web at this hot and extremely nutrient-deprived acidic environment.

Keywords: Sasso Pisano; hot vents; fumarole; 16S rRNA gene; 18S rRNA gene; microbial diversity

1. Introduction

The superheated steams and fumaroles near Sasso Pisano village (Castelnuovo Val di Cecina, Italy) are the central part of the Larderello geothermal field, located in the inner Northern Apennines (Southern Tuscany). The field exhibits extremely high subsurface temperatures (300–350 °C) at a pressure range of 4–7 MPa [1–3]. A vapour-dominated reservoir within a metamorphic basement at a depth of 3000 m is connected to a productive 1000 m deep carbonate-evaporite reservoir, which is almost sealed by overlying units. Many artificial steam wells were established for geothermal energy generation. The fumarolic fields of Sasso Pisano between Lagoni del Sasso and Monterotondo Marittimo represent a unique landscape of natural steam vents [3]. The overpressurised hot steam and springs release and/or redeposit CH_4, CO_2, sulphur compounds (H_2S, SO_4^{2-}), boric acid (H_3BO_3), ammonia (NH_4^+), and heavy metals at the surface [4,5], as water vapours flow (50 t/h) along fault zones/fractures through the sedimentary layers [3,6]. The condensation of sulphur dioxide-rich steam at the surface around fumaroles contributes to an acidic environment (pH 0.5–5) [7], which also affects the surrounding area [8,9].

The Sasso Pisano geothermal field comprises very extreme environmental constraints including heated surface soil (50 to 100 °C), extremely low pH values, low content of bioavailable nitrogen, carbon and phosphorus compounds, and high content of heavy metals as well as trace elements, especially around fumaroles [5,9,10]. These stringent conditions would limit the growth and productivity of microbes and select only a few of the most adaptive extremophiles to survive in the extreme geothermal fields. The bacterial communities of this extreme environment are barely explored, though the extremophilic microorganisms are of great interest due to their potential biotechnological importance [11]. To learn more about the bacterial and eukaryotic communities in the presence of hot steams, environmental DNA from the geothermal fields in Sasso Pisano was isolated for 16S and 18S rRNA gene sequencing and subsequent diversity analysis. The study of the Sasso Pisano fumaroles' rock and soil samples revealed high predominance of extremophilic novel uncultured *Ktedonobacteria* and eukaryotic *Cyanidiales*. Furthermore, the adaptive strategies of *Ktedonobacteria* and *Cyanidiales* in relation to environmental constraints were inspected. We also analysed if the diversity of *Ktedonobacteria* at Sasso Pisano is similar to that determined in other environments.

2. Materials and Methods

2.1. Sample Collection

The samples were taken at Sasso Pisano (Pisa, Tuscany Region, Italy) in April 2019 (43°09.632′ North, 10°51.538′ East) and stored at −20° until further use. For prokaryotic community analysis, the sampling sites at the hot springs and nearby soil were sampled (see sample designations in Figure 1). The designation of the samples was SP (Sasso Pisano). The hot springs varied greatly in intensity and discharged whitish to blackish tarnish, accompanied by steam bubbling. Sample SP1 was taken from a greenish/brownish biofilm growing at the top of the hot spring A vent, constantly exposed to steam. Sample SP2 was taken from a yellowish/beige, possibly elemental sulphur deposit formed by vapours at the hot spring A. SP11 (greyish muddy watery discharge) and SP12 (yellowish greyish sediments) were directly sampled from the hot spring B vent. The hot spring C was sampled at different distances from the openings SP24, SP26 and SP7. The hot spring D was sampled at the vent (SP310) and when the discharge was meeting hot spring C sediments and organic material (plant leaves: SP38) (Figure 1D). A description of the features and details with respect to all samples are provided in Table S1. Soil samples (Sp513–515, Sp3, Sp517, Sp719, Sp819–821, Sp122) with various textures were frequently collected around the different hot springs. To study a thermophilic eukaryotic consortium, growing under extremely acidic (pH 2) and hot conditions (44–55 °C), a greenish biofilm (SP5) was collected from the vertical walls of a small gorge filled with hot steam seeping out of the ground (Figure 1H). For comparison among the extreme sites, a water stream with a neutral pH and a lower temperature (15 °C) near the fumarole crossing site was sampled and labelled as SP4 (control).

Figure 1. Sampling sites for 16S rRNA and 18S rRNA gene-based analysis. (**A–D**) are hot springs sites, while (**E–G**) are representatives of the soil sampling sites. Each sample is labelled "SP" (Sasso Pisano), followed by numbers. The picture inserted in (**A**) represents the yellowish-orange SP2 sampling site, formed by the surface deposition of the sulphur vapours probably from a solfatara-type vent. Hot spring sites (**A,D**) were either of clear or whitish liquid discharge. Hot springs (**B,C**) were splashing muddy greyish/blackish water. The fumarole rocky wall was continuously wrapped in hot steam (**H**). The greenish biofilm forms a tight layer on the crumbly rock of the fumarole (SP5) and was compared with a cold and pH-neutral water stream (SP4) (**I**).

2.2. Environmental DNA Extraction, Amplicon Generation and Sequencing

The Power soil DNA isolation kit (Qiagen, Hilden, Germany) was used to extract the whole genomic DNA of the stored samples. The samples were subjected to three extractions (triplicates) each, according to the manufacturer's protocol. The extracted gDNA was then concentrated with the Eppendorf Concentrator Plus (Eppendorf, Hamburg, Germany) at 45 or 60 °C for 35 min to obtain the optimal gDNA concentration. Qualitative and quantitative determination of DNA was performed by gel electrophoresis and a Nanodrop instrument (NanoDrop 2000 c Spectrophotometer, ThermoFisher Scientific, Waltham, MA, USA). After quality assessment, partial 16S and 18S rRNA genes were amplified via polymerase chain reaction (PCR). For eukaryotes, the 18S amplicon PCR forward primer = 5′-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA GCA SCY GCG GTA ATT CC-3′ and the reverse primer = 5′-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GAC TTT CGT TCT TGA TYR A-3′ were used. For bacteria, the 16S amplicon forward primer 5′-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT ACG GGN GGC WGC AG-3′ and the reverse primer = 5′-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGA

CTA CHV GGG TAT CTA ATCC-3′ were applied [12,13]. These primers were additionally linked to the overhang adapter sequences to make the amplicon suitable for Illumina MiSeq sequencing. Both partial 16S and 18S rRNA genes were amplified via a modified PCR Master mixture: Double-distilled nuclease-free water (32 μL) was mixed with 10 μL Phusion GC Buffer-5×, 10 μM forward and reverse primers (1.0 μL each), 5% DMSO (2.5 μL), 50 mM MgCl$_2$ (0.15 μL), 10 mM dNTPs (1.0 μL), 25 ng template DNA (2.0 μL) and 0.5 μL of 2 U/μL Phusion HF DNA polymerase (ThermoFisher Scientific) to obtain a final volume of 50 μL reaction mixture. A thermal cycler (Biometra, Göttingen, Germany) was used to anneal the 18S and 16S rRNA gene-specific primers with the extracted gDNA at 56 and 60 °C, respectively. PCR was performed as described [13,14]. The amplification products were qualitatively and quantitatively controlled on a 0.8% agarose gel with 1× TAE buffer (ThermoFisher Scientific) and a NanoDrop spectrophotometer (ThermoFisher Scientific). PCR products were further processed by using the GeneRead Size Selection Kit (Qiagen) to wash off primers and PCR reagents from the resulting amplicons. Indexing of OCR products was executed with the Nextera XT DNA library prep kit (Illumina, San Diego, CA, USA); subsequently, paired-end sequencing at the Göttingen Genomics Laboratory using the 2 × 300 bp Paired-End mode with an Illumina MiSeq sequencer (Illumina) was performed. The obtained Fastq sequences are accessible at the NCBI database under the Sasso Pisano microbiome project number PRJNA725822.

2.3. Processing of 16S rRNA Gene-Based Amplicon Sequencing Data

The MetaAmp automated pipeline, an online resource for metagenomic analysis, was utilised to analyse 16S rRNA gene-based amplicon sequencing data (http://ebg.ucalgary.ca/metaamp/, accessed on 18 November 2020) [15]. Initially, the USEARCH software merged the uncompressed and demultiplexed sequence data and removed low-quality reads [16]. The subsequent step was the removal of non-matching and misaligned read pairs of small lengths (less than 350 bp). Moreover, the Mothur software package was utilised for the identification and trimming of forward and reverse primers [17]. Any reads that lacked forward and reverse primers or containing primer mismatches were discarded. In the next step, the reads were dereplicated using UPARSE software version 7.1 [18]. This software also discarded chimaeras and singletons. The resulting, high-quality reads were assembled into taxonomic units (OTUs) based on the 0.97 OTU clustering threshold (97% identity). The Mothur software package assigned taxonomic status to the OTUs with reference to the SILVA training dataset (http://www.mothur.org/wiki/Taxonomy_outline, accessed on 18 November 2020). Lastly, the aforementioned software generated relative abundance data, rarefaction curves, as well as alpha and beta diversity indexes. The samples were rarefied into subsamples. The lack of similarity within different samples was computed through the Bray–Curtis index and principal coordinate analysis (PCoA) was performed. The Permutational Multivariate Analysis of Variance (PERMANOVA) test allowed for hypothesis testing.

2.4. Analysis of 18S rRNA Gene-Based Amplicon Sequencing Data

With respect to analysis of 18S rRNA gene-based data, paired-end reads were merged using PEAR (Paired-End reAd mergeR) [19] and resulting sequences of small lengths (less than 400 bp) were removed. Furthermore, the BBMap tool was used to remove all sequences that did not match the primer sequences at the 5′ and 3′ ends [20]. Subsequently, chimeric sequences were removed with the VSEARCH tool [21]. Data cleaned in this way were converted from the FASTQ file format to the FASTA format and mapped with QIIME (Quantitative Insights into Microbial Ecology) against the reference genes from the SILVA 132 18S database [22] with a 97% match [23].

2.5. Transmission Electron Microscopy (TEM)

For TEM sample preparation, the samples SP4 and SP5 were placed in a fixing solution (0.2%, v/v, glutardialdehyde, 0.3%, w/v, formaldehyde). Subsequently, they were washed with a 50 mM K_2HPO_4 buffer solution and parts of the biofilm were embedded in agar (Agar-Agar, Kobe I, Carl Roth, Karlsruhe, Germany). Sliced small agar blocks of samples were dehydrated with ethanol in increasing concentration, followed by infiltration with synthetic resin (LR White Resin, London Resin Company, Berkshire, UK), and then, embedded in gelatine capsules, which were polymerised at 55 °C for 24 h.

To obtain ultra-thin sections, the polymerised samples were first pre-milled with a sample trimming device (TM 60, Reichert-Jung, Wetzlar, Germany), and sliced with the ultramicrotome (Ultracut E, Reichert-Jung) using a diamond knife (DuPont, Wilmington, DE, USA). The quality of the ultrathin sections was assessed based on their interference pattern and "fished" out of the water using coated grids (Plano, Wetzlar, Germany) as specimen support. For staining of ultrathin sections, uranyl acetate and replacement stain (UAR-EMS Uranyl Acetate Replacement Stain, Electron Microscopy Sciences, Hatfield, UK) were applied at different dilutions (undiluted, 1:3, 1:5, v/v). The grid with the section side facing downwards was placed on a drop of 20 μL of the respective staining agent and was incubated for ten minutes. The ultrathin sections were examined by transmission electron microscopy (JEM-1011 Electron Microscopes, JEOL, Akishima, Tokyo, Japan) at a voltage of 80 kV and electron micrographs were recorded.

2.6. Prokaryotic Community Functional Profiling

The MetaAmp-generated taxonomic profile data were utilised to generate a holistic overview of functional profiles of various samples. In this context, the Piphilin online server (http://secondgenome.com/Piphillin, accessed on 5 October 2020) [24] was used. Piphilin allows for the direct search of representative OTU sequences against a database composed of 16S rRNA gene sequences using USEARCH version 8.0.1623 [16]. In the next step, gene copy numbers of every inferred genome are summed to generate KO abundances (the KEGG reference database was used) [25]. The online available server MicrobiomeAnalyst [26] was utilised for the completion of statistical and meta-analysis of gene abundance data. The R package DESeq2 was used to compute a differential abundance analysis [27]. To summarise the findings, important features of hot sampling sites were considered in regard to soil samples that exhibited over a 1-fold difference.

2.7. Phylogenetic Analysis

The phylogenetic analysis was focused on 16S rRNA gene sequences, including those of different *Ktedonobacteria* type strains. MUSCLE, implemented in MEGA-X software version 7.0, was utilised to align these sequences and also *Ktedonobacteria* class-related OTU sequences [28]. Lastly, as part of the phylogenetic analyses, evolutionary distances were assessed. The Kimura 2-parameter model was used for this. The construction of phylogenetic trees was performed utilising the maximum likelihood method [29]. Bootstrap values were calculated based on 1000 replications.

3. Results

3.1. Prokaryotic Community Composition

A total of 788,808 high-quality 16S rRNA gene sequences were available for analysis of prokaryotic communities. Per sample, the sequence amount ranged from 14,219 to 92,423 and a total of 1797 OTUs were identified (Figure S1). The sampling sites at the hydrothermal field were divided into two groups: hot springs and soils. The relative abundances at hot springs varied greatly per sample, emphasising each hot spring microbiome is unique (Figure 2). For instance, the prokaryotic profile of the samples SP24 and SP26, emerging from the same blackish muddy hot spring C, was dominated mainly by *Aquificae* ($p = 0.001083$), and hot spring B SP11 and SP12 samples were dominated by *Deinococcus/Thermus* ($p = 1.8017 \times 10^{-15}$). *Chloroflexi* were abundant in the SP1 biofilm taken from hot spring A expelling clear water steam

fumes, and samples SP11 and SP12 from slightly muddy water hot spring B. The phylum-level profile also changed drastically when the hot spring streams (SP24, SP26 and SP310) were coming into contact with soil (SP27) or other organic matter (leaves, SP38) downstream, suggesting the integration of a mesophilic consortium (including *Firmicutes* or *Proteobacteria*). The overall composition of the hot spring samples comprised *Chloroflexi* (21%), *Proteobacteria* (19%), *Firmicutes* (17%), *Cyanobacteria* (8%), *Aquificae* (6%), *Acidobacteria* (5%), *Deinococcus Thermus* (4%), *Euryarchaeota* (3%), *Thermotogae* (6%) and *Actinobacteria* (2%). In contrast, in all soil samples, the most abundant taxa were *Cyanobacteria* (27%), *Proteobacteria* (19%), *Acidobacteria* (11%), *Actinobacteria* (10%), *Chloroflexi* (8%), *Euryarchaeota* (7%), *Planctomycetes* (6%) and *Thaumarchaeota* (4%) and the composition showed no drastic variations among most of the samples. An increased abundance of *Chloroflexi* was detected in soil samples that were collected nearby fumaroles or hot springs.
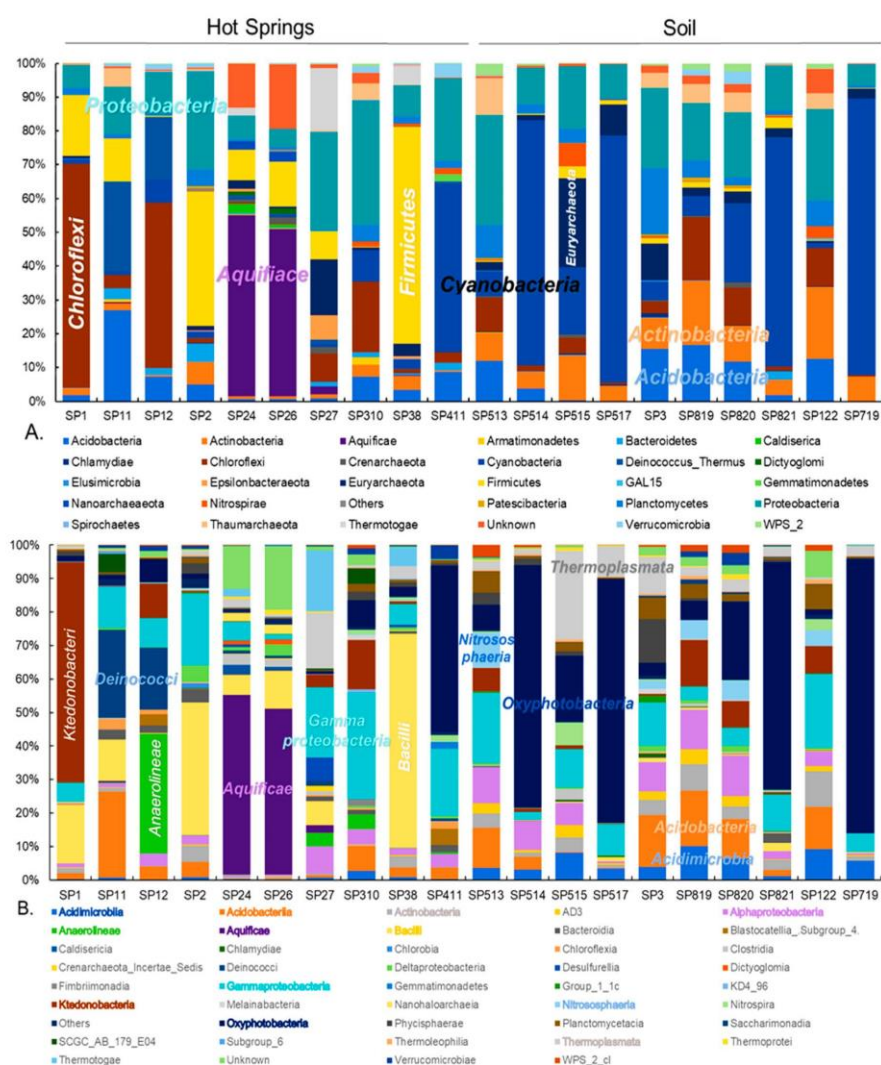

Figure 2. Relative abundance of prokaryotic taxa in Sasso Pisano hot spring and soil samples. Prokaryotic community composition at the phylum level (**A**) and class level (**B**) is shown for individual samples. Abundant taxa are highlighted within stacked bars or within taxa legends.

At the class level, the differences between the hot spring samples were more prominent. For instance, the SP1 biofilm from hot spring A showed a high abundance of the branched hyphae and spore forming *Ktedonobacteria* (67%), while the other hot spring sample SP12 was enriched with filamentous thermophilic *Anaerolineae* (36%). *Deinococci* (27%), *Aquificae* (54–50%), *Gammaproteobacteria* (21–32%), *Acidobacteria* (16%), and *Oxyphotobacteria* (54%) were the most abundant taxa in samples SP11, SP24, SP26, SP27 and SP310, respectively. *Bacilli* (*Firmicutes*) inhabited the surface of yellowish sulphur deposits (44%, SP2) and a downstream site (65%, SP38) where the whitish discharge mixed with leaves and blackish muddy discharge from hot spring C. In the group of soil samples, the abundant classes were, in different samples, *Oxyphotobacteria*, *Thermoplasmata*, *Alphaproteobacteria* and *Gammaproteobacteria*, *Acidobacteria* and *Actinobacteria*. The depicted pie charts show taxa in the hot spring and soil group in the order of ascending relative abundance (Figure 3). Overall, the hot spring group showed a higher relative abundance of thermophilic taxa.
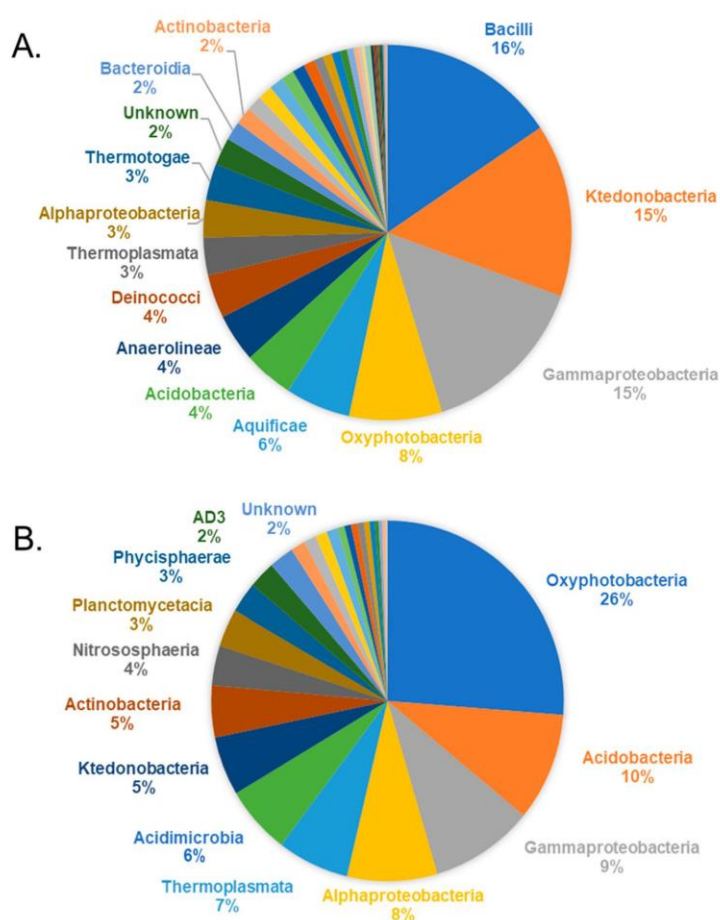

Figure 3. Overall bacterial community composition of the hot spring and soil microbiome. The actual abundances at the class level are plotted for the hot spring (**A**) and soil group (**B**). Taxa showing less than 2% relative abundance are not displayed.

3.2. Diversity of Ktedonobacteria

The 30 most abundant OTU sequences potentially affiliated with *Ktedonobacteria* were aligned with 16S rRNA gene sequences of isolated *Ktedonobacteria* strains and metagenomic sequences resulting from NCBI blast. In the phylogenetic tree, the OTUs are sandwiched between mesophilic (*Dictyobacter*) and thermophilic (*Thermogemmatispora*) groups of the *Ktedonobacteria*, indicating the presence of a large pool of uncultured mesophilic to thermophilic *Ktedonobacteria* strains. Most strains could be classified as members of *Ktedonobacteraceae*; however, OTUs related to JG30-KF-AS9 and B10-SB3A seem to be outliers and are more closely related to thermophilic *Thermogemmatisporaceae* (Figure 4). The detected OTUs indicate that the natural hydrothermal field of Sasso Pisano offers a rich reservoir of novel uncultured *Ktedonobacteria*, which should be targeted with respect to cultivation, genome sequencing, and further exploration. The highest abundance of *Ktedonobacteraceae* was detected in the SP1 biofilm, which indicates a natural enrichment of these bacteria at a hot spring vent (Figure S2). More diverse genus members belonging to *Ktedonobacteria* were identified in another hot spring sample (SP310; whitish discharge).

Figure 4. Maximum likelihood phylogenetic tree including *Ktedonobacteria* OTUs. The phylogenetic tree was constructed by aligning OTU sequences potentially affiliated with *Ktedonobacteria* with 16S rRNA gene sequences of known *Ktedonobacteria* species.

3.3. Alpha and Beta Diversity

For alpha and beta diversity analysis, sequence datasets were rarefied to the lowest detected read size per sample. The Shannon index indicated that some of the hot spring samples comprise a lower diversity than soil group samples (Figure 5). The soil group potentially exhibited a higher alpha diversity in some cases, as the environmental conditions are comparably less harsh and facilitate a broad range of microorganisms to grow. In contrast, due to the acidic pH, higher temperature, and constant washout at the hot spring spots, only acidic thermophilic microbes are supposed to grow. However, the *t*-test result does not show a significant difference between the two groups, indicating that the soil microbial communities may provide an active microbial influx to the hot springs and vice versa (Figure 5A).
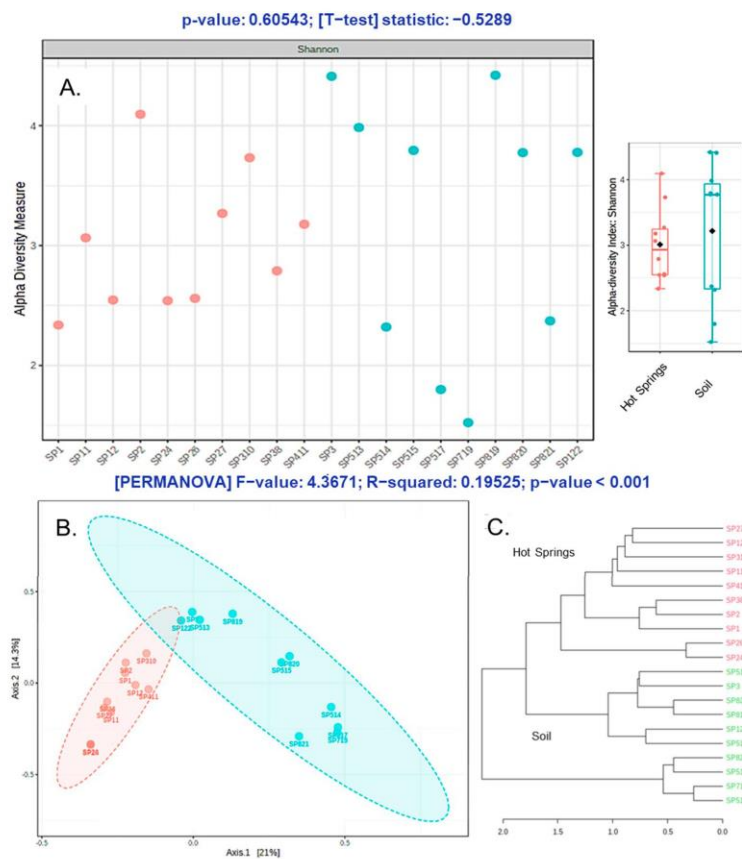
Figure 5. Comparative alpha and beta diversity analysis of the hot springs versus soil samples. The Shannon index (**A**) showed that both hot springs and soil samples were rich in terms of microbial alpha diversity, though the particular microbial composition for both habitats was distinct, based on the PCoA (**B**) and Bray-Curtis tree (**C**) analysis.

To visualize and explore the complex metagenomic data, the multidimensional scaling method—Principal Coordinates Analysis (PCoA)—was applied to observe the similarities at the OTU level. The PCoA of Sasso Pisano samples showed that the microbial communities of hot springs and soil samples were clustering separately, with a considerable variance among the two groups (Figure 5B). According to the Permutational Multivariate Analysis of Variance (PERMANOVA), the beta diversity among both groups also suggested the

inhabitant microbial communities are distinct, indicating the stringent effect of hot spring constraints in selection and enrichment of a unique microbial consortium as compared to the soil microflora. The tree diagram, calculated based on a Bray–Curtis index, also indicated that both groups are clearly distinct (Figure 5C).

3.4. Functional Profile

The abundance of functional genes summarised by KO identifiers (also called K numbers) was estimated from the OTU data. Based upon this estimation, the predicted function of genes was used to evaluate a metabolic profile of the microbial communities in different sites. The differential abundance analysis of KO numbers indicated that the hot spring microbiota was metabolically rich in terms of heavy metal resistance, methane, sulphur and nitrogen metabolism as well as aromatic compounds' degradation pathways as compared to the soil microbial consortium. Numerous transporters involved in the extrusion of Mn, Zn, Cu, W, Co, Ni and Mo were present in the hot spring microbial consortium (Table 1). The functional profile also showed that the different KO numbers associated with methane metabolism are differentially present in hot spring and soil samples (Figure 6). This indicates that the hot spring microbiome utilises different enzymes as compared to the soil microbiome to metabolise environmental methane. The soil microbiome shows preferences for aerobic methane metabolism such as methane oxidation to formaldehyde and formate dehydrogenation to CO_2, while the hot spring microbiome seems to favour anaerobic methane metabolism through reverse-methanogenesis by means of methyl-coenzyme M reductase (Mcr-AOM) [30,31] (Figure S3). Moreover, with respect to carbon fixation, mainly formaldehyde seems to be fixed via the ribulose-monophosphate pathway (RuMP) [32]. More KO numbers in the hot spring microbial communities were also observed to be differentially abundant, involved in sulphur and nitrogen metabolism (data not shown). In the aromatic compound degradation pathways, genes and enzymes involved in the aminobenzoate, benzoate, chlorocyclohexane, chlorobenzene, chloroalkane, chloroalkene, dioxin, fluorobenzoate, nitrotoluene, naphthalene, styrene, xylene and toluene degradation were also differentially abundant in the hot spring samples (Figure S4).

Table 1. Differentially abundant transporters required for heavy metal resistance in hot springs microbiome.

| Transport System	
Manganese	**Zinc**
K19973 mntA; manganese ATP-binding protein	
K19975 mntC; manganese substrate-binding protein	K11707 troA; manganese/zinc/iron substrate-binding protein
K19976 mntB; manganese permease protein	K11708 troC; manganese/zinc/iron permease protein
K11601 mntC; manganese substrate-binding protein	K11709 troD; manganese/zinc/iron permease protein
K11603 mntA; manganese ATP-binding protein	K11710 troB; manganese/zinc/iron ATP- binding protein
K11602 mntB; manganese permease protein	
Iron	**Copper**
K11604 sitA; manganese/iron substrate-binding protein	K19340 nosF; Cu-processing system ATP-binding protein
K11605 sitC; manganese/iron permease protein	K19341 nosY; Cu-processing system permease protein
K11606 sitD; manganese/iron permease protein	
K11607 sitB; manganese/iron ATP-binding protein	**Tungstate**
K02010 afuC; iron(III) ATP-binding protein	K05772 tupA; tungstate substrate-binding protein
K02011 afuB; iron(III) permease protein	K05773 tupB; tungstate permease protein
K02012 afuA; iron(III) substrate-binding protein	
Molybdate	**Nickel**
	K15584 nikA; nickel substrate-binding protein
K02017 modC; molybdate ATP-binding protein	K15585 nikB; nickel permease protein
K02018 modB; molybdate permease protein	K15586 nikC; nickel permease protein
K02020 modA; molybdate substrate-binding protein	K15587 nikD; nickel ATP-binding protein
Sulphate/Thiosulphate	**Cobalt**
K02045 cysA; sulfate/thiosulfate ATP-binding protein	K02006 cbiO; cobalt/nickel ATP-binding protein
K02046 cysU; sulfate/thiosulfate permease protein	K02007 cbiM; cobalt/nickel permease protein
K02047 cysW; sulfate/thiosulfate permease protein	K02008 cbiQ; cobalt/nickel permease protein
K02048 cysP; sulfate/thiosulfate substrate-binding protein	K02009 cbiN; cobalt/nickel transport protein
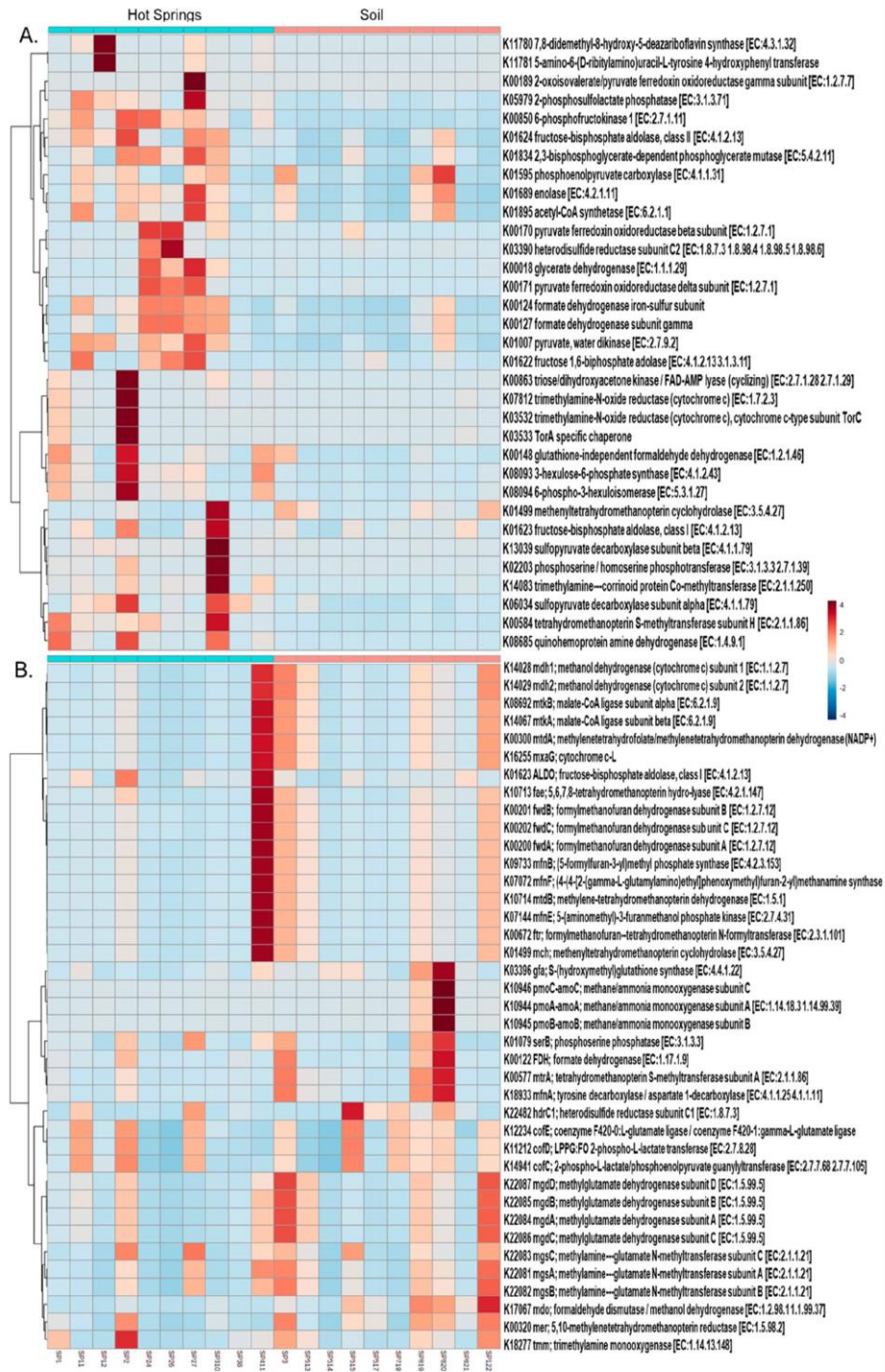
Figure 6. Differentially abundant genes involved in methane metabolism. The hot spring microbiome (**A**) was found to be abundant in the genes involved in the anaerobic methane degradation as compared to the aerobic methane degradation in the soil microbiome (**B**).

3.5. Eukaryotic Diversity at a Fumarole

According to rarefaction curves, 300–450 OTUs were identified in the neutral pH and mesophilic stream conditions; the extreme fumarole site comprised less than 50 OTUs (Figure S5). A total of 57 OTUs affiliated with known taxonomic groups were identified in stream sample (SP4) data, while the fumarole rock (SP5) data comprised a total of 21 groups, which shows a considerably higher alpha diversity in the water stream biome than in the fumarole biome. It can be concluded that the extreme conditions of the fumarole limit microbial alpha diversity, while microbial communities could easily populate the neutral pH water stream.

Organisms known to typically colonise freshwater habitats, including amoebae (11.2%) which can be divided into the groups *Dactylopodida* (1.1%), *Filamoeba* (7%) and *Ischnamoeba* sp. (2.3%), were detected. The *Incerta Sedis* represents a kind of placeholder for unspecified species in the genome database. The paraphyletic group of green algae is represented by *Chlorophyta* (0.75%) and *Charophyta* (4%). Among the multicellular organisms, *Chaetonotida* (10%), belonging to the phylum *Gastrotricha* ("hairybacks"), was detected. Furthermore, *Adineta vaga*, belonging to the rotifers, was found with 2.1% and the lower fungi were represented with a total of 5.7% relative abundance. The group of SAR (*Stramenopiles*, *Alveolata*, *Rhizaria*) showed 19.7% and LEMD267 is listed in SILVA as a nonspecific taxonomic group that refers to lobose amoebae (22.5%). However, this cannot be further determined with the available data. Excluding the unspecified eukaryotes (19%), 62.02% of the identified species OTUs (*Amoeba* 22.5%, *Archaeplastida* 0.82%, *Metazoa* 14.3%, *SAR* 19.7%, low mushrooms 4.7%) can be assigned to a biofilm in a freshwater area with a neutral pH and a lower temperature around 15 °C. The other major abundant taxa were *Intramacronucleata* (Ciliates) (16%) and *Chaetonotida* (10%). The OTU assigned to *Rhodophyceae* with 0.07% and *Echinamoeba thermarum* with 4.2% can be potentially explained by the association of the stream channel to the thermo-acidic sampling site, since both OTUs possess a thermophilic character.

The apparently qualitative identical species composition of the sample duplicates (SP4.1 and SP4.2) differs especially in terms of the fraction of nematode- and mite-related OTUs (Figure 7). The unspecified *Cyanidiaceae* and the species *Galdieria sulphuraria* belonging to the family *Cyanidiales* of phototrophic eukaryotic algae represent the most abundant taxa of the fumarole biofilm (65.3%). Unspecified *Collembola* (springtails) and *Folsomia candida* contributed to *Collembola* sp. (16.8%). *Collembola* sp. (*Arthropoda*: *Hexapoda*) represent more highly developed organisms, as they are among the typical bottom dwellers and have been described for other extreme sites, in particularly hot and particularly cold areas [33]. *Gregarina caledia* (*Apicomplexa*) contributed 14% of the biofilm consortium (Figure 7).

3.6. Structural and Morphological Description of a Eukaryotic Biofilm

The electron micrographs of a fumarole biofilm (SP5) preparation allowed inspection of individual cells, which are separated from each other by a cell wall of approx. 0.2 µm thickness. Within these cells, organelles, chloroplasts, in particular, are visible (Figure 8). Inside chloroplasts, membrane stacks of the thylakoids are prominent features. Densely packed cells (Figure 8A) are separated from each other by an extracellular matrix formed between the organisms. Remineralisation processes have caused dark precipitates to settle as small particles in the extracellular matrix. Larger particles are separated from the cells by the matrix (Figure 8A). The marked particles imitate the shape of the matrix.

Based on the number of spore cells within the mother cell, potentially either the genus *Galdieria* or *Cyanidium* could be identified. More than four spores in the cell represent a distinct structural feature of *Galdieria* (Figure 8B), while a sporangium with tetraspores could be classified as *Cyanidium* (Figure 8C). Based on the species description in Ciniglia et al. (2004) [34], the organism (Figure 8D,E) can be morphologically determined even more precisely as *Galdieria sulphuraria* under the electron microscope. The cell has a thick cell wall and contains at least seven visible chloroplast shapes in the cytosol, surrounding a central vacuole, visible as a membrane-bounded empty lumen. Ultrathin sections also reveal

the coexistence of the bacterial colony outside the extracellular matrix of the *Cyanidiaceae* biofilm (Figure 8F).
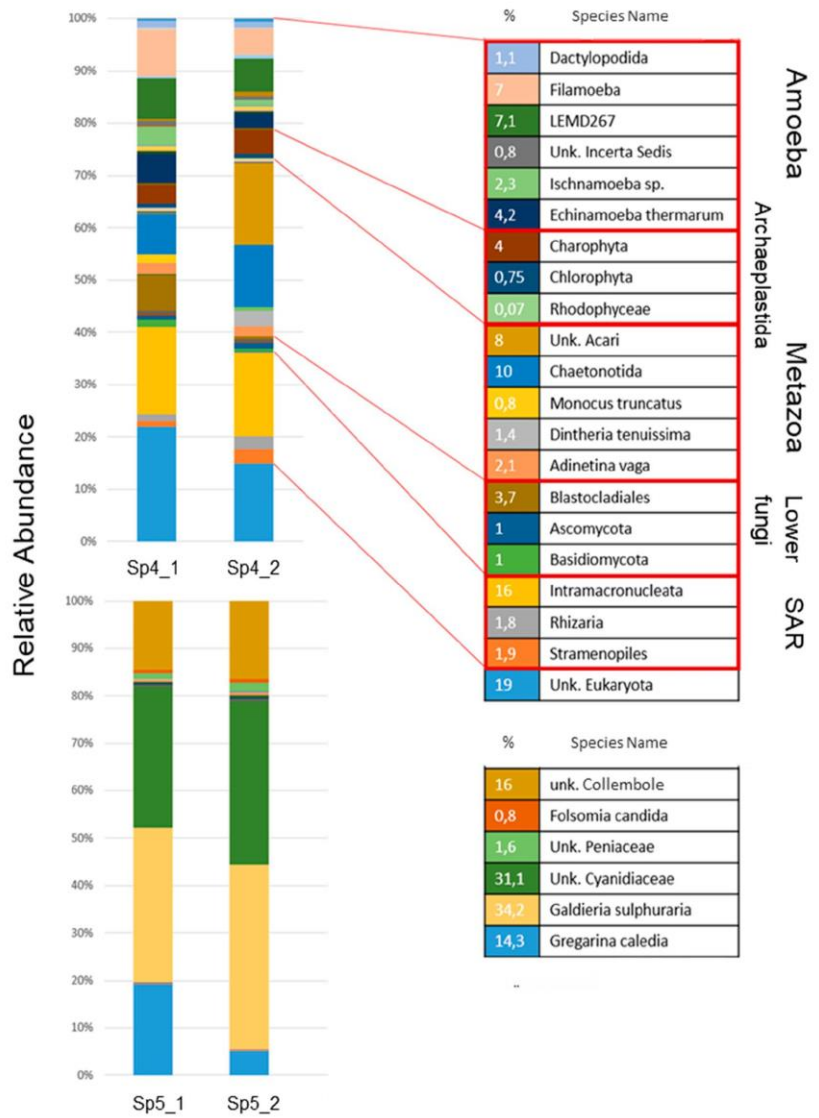
Figure 7. Eukaryotic community composition of a neutral pH (water stream biofilm SP4) and extreme (fumarole biofilm SP5) biome. Taxonomic groups are listed which show at least 0.80% of the total OTU abundance (except for the *Rhodophyceae* (0.07%) from the stream water sample to show the contrast to the fumarole and *Chlorophyta* to distinguish them from *Charophyta*). In the stream water sample, 21 species groups are listed which represent 98.02% abundance of the total OTUs, and six groups are shown in the fumarole samples, which represent 97.90% of the total OTUs (Unk, unknown).

Figure 8. Electron microscopic analysis of a section through a sample taken from the fumarole biofilm SP5. An overview of a *Cyanidiales* cell aggregate (**A**). Sporangia (mother cells; **B**,**C**) containing a varying number of autospores. In three of the five autospores (**B**), the membrane stacks of the thylakoids in the chloroplasts are clearly visible. Within the mother cell (**C**), four autospore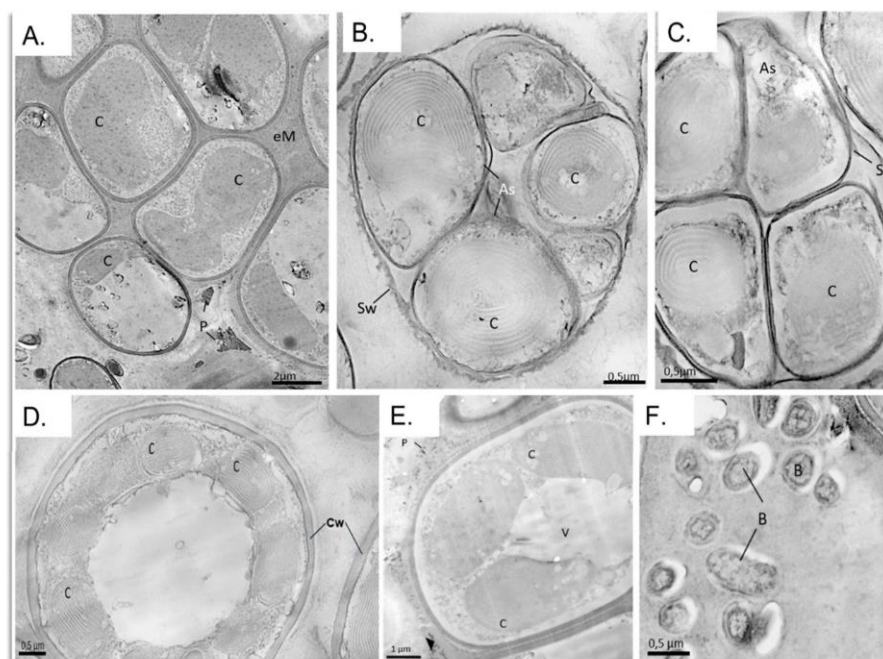s (tetraspores) already are about the same size and the sporangium wall is disintegrating. The cells in (**D**,**E**) carry multiple chloroplasts around a central vacuole. Both cells could be assigned to the genus *Galdieria*. Prokaryote morphotypes of about 0.5 μm in size attached to the *Cyanidiales* aggregate (**F**). V, vacuole; eM, extracellular matrix; Cw, cell wall; Sw, sporangium wall; B, bacteria; As, autospores; P, precipitates.

4. Discussion

Ktedonobacteria, a deeply branched bacterial class comprising of mesophilic and thermophilic representatives, is characterised by its ubiquitous prevalence in terrestrial environments, complex life cycle, and in some cases, large genome size [35]. The class is divided into two orders, *Ktedonobacterales* and *Thermogemmatisporales*. The latter order includes *Thermosporotrichaceae* as well as *Thermogemmatisporaceae* strains, whereas *Ktedonobacterales* comprise *Dictyobacteraceae*, *Ktedonobacteraceae* and *Ktedonosporobacteraceae* strains. Within these two orders, 20 proposed mesophilic and thermophilic species have been assigned to the genera *Dictyobacter*, *Tengunoibacter*, *Ktedonobacter*, *Ktedonosporobacter*, *Thermosporothrix* and *Thermogemmatispora* [36–39]. *Ktedonobacteria* isolates and related environmental DNA are derived from non-extreme sources [38,39] and extreme environments, such as an acid vapour-formed spring [40], naturally occurring CO_2 gas vents [41], a lava cave in a volcanic trench [42], volcanic fumaroles [43], steaming geothermal soil [44], and a mineral precipitating cave environment [45]. This suggests that *Ktedonobacteria* members appear to prevail in oligotrophic and extreme environments, implying these strains may have evolved versatile metabolic pathways to cope with extreme conditions. For instance, the potential to oxidise carbon monoxide (CO) has been reported for *Ktedonobacteria* members [46].

In this study, we gained insights into the abundance and diversity of *Ktedonobacteria* in the geologically diverse and environmentally extreme hydrothermal field environments of Sasso Pisano. Microbial diversity varied greatly with the type of sampling site, which included hot springs, a fumarole, and nearby soil as well as water samples. A high abundance and diversity of *Ktedonobacteria* at the hot springs and in some soil samples indicated that

the Sasso Pisano hydrothermal region offers a natural reservoir of *Ktedonobacteria* members, which became naturally enriched under stringent environmental constraints including heated surface soil, extremely low pH values, low nutritional content of nitrogen, carbon and phosphorus and high content of heavy metals as well as trace elements, especially around fumaroles [5,9,10]. The functional profile indicated that the hot spring microbiome includes a significantly higher abundance of organisms harbouring genes involved in methane, sulphur, nitrogen and aromatic compounds' metabolism, since the acidic hot springs and fumaroles emit a mixture of various gases such as ammonia, methane, carbon dioxide, hydrogen, hydrogen sulphide, hydrocarbons and aromatic compounds [2,5,47]. Moreover, transporters conferring resistance against toxic transition metals were also detected in the hot spring microbiome, which implies that *Ktedonobacteria* members inhabiting this site could have evolved high metabolic plasticity to cope with acidification, nutrient depletion, and heavy metal resistance. In our previous study of biofilm samples from Marsberg copper mine, Germany, the capability to degrade toxic aromatic compounds as well as resistance against transition metals was detected in a metagenome-assembled genome affiliated with *Ktedonobacteria* [48]. The bacterial biofilm on steam vents from Sapichu volcanoes was also mainly colonised by related similar taxa of *Chloroflexi* (*Ktedonobacteria*), *Acidobacteria,* and *Cyanobacteria*. The metabolic potential analysis predicted similar cellular metabolic pathways related to methanogenesis, sulphur respiration, nitrogen fixation, and heavy metal transport except for photosynthesis by *Cyanobacteria* [43]. The observed higher abundance of *Cyanobacteria* in low-temperature soil samples (but not in hot springs samples) was consistent with colonisation patterns of *Cyanobacteria* in lower temperature vents, as reported by Wall et al. [49], while the higher temperature vents were abundant in *Cloroflexi* [49,50].

Two important metabolic pathways in oligotrophic geothermal sites are attributed to trace gas (CO and H_2) utilisation by the type I carbon monoxide dehydrogenase (*cox* genes) and the [NiFe]-hydrogenase (*hyp* genes) as an energy source for cell growth and persistence under nutrient-limiting conditions [51]. Mounting genomic evidence suggests that three phyla, *Chloroflexi*, *Actinobacteria* and *Acidobacteria*, use CO and H_2 as substrates [52]. A variety of thermophilic bacteria belonging to the classes *Actinobacteria*, *Deinococci*, *Ktedonobacteria, Thermomicrobia* and *Clostridia* contain *cox* operons and may be capable of aerobic CO oxidation. Hot springs from different regions host different *cox* encoding communities [53]. The *Ktedonobacteria* genome also includes the reductive TCA cycle along with multiple copies of *cox* operons, conferring CO oxidation potential to this taxa, allowing it to predominate the microbial community [46] under the influence of CO-rich gas vents, hydrothermal springs, and soil environments [41,54,55]. Other high abundant taxa have evolved diverse strategies, for instance, most *Anaerolineaceae* species metabolise various organic carbon sources under anaerobic conditions through fermentative metabolism [56]. Representatives of the thermophilic *Aquificae* grow in hot springs via oxidation of dissolved ferrous iron or iron-containing minerals and conduct nitrogen fixation even at 70 °C [57,58]. *Deinococci* are resistant to several stresses due to their highly efficient DNA damage repair ability and detoxification of several toxic compounds through hydrolytic activity [59,60]. Some *Oxyphotobacteria* (photosynthetic *Cyanobacteria*) are metabolically diverse primary producers and pioneer the colonisation where light is available including ecosystems with low light, low levels of O_2 and/or sulfidic conditions, because of their ability to perform, besides oxygenic photosynthesis, anoxygenic photosynthesis [61,62].

The eukaryotic alpha diversity at the fumarole was significantly lower than in the river water sample, whereas the alpha diversity of prokaryotes did not vary significantly in the hot springs and soil samples. This indicates that thermophilic eukaryotic organisms are less diverse than thermophilic prokaryotes. *Cyanidiales* dominated the biofilm with an OTU of 65.3%, as they represent the only phototrophic organisms growing under these conditions and have been extensively studied since the 1980s [63]. The *Cyanidiales* form a monophyletic group within the red algae, which inhabits the acidic hot springs at different sites. *Cyanidiales* were identified under the electron microscope based on the chloroplasts

and reproduction pattern, as the genera *Cyanidium* and *Galdieria*. *Cyanidium* sp. propagate with tetraspores, whereas *Galdieria* is larger and forms more autospores and vacuoles than *Cyanidium* [64]. Yoon et al. (2006) described the structural composition of endolithic and interlithic populations of *Cyanidiales* at Pisciarelli in the Phlegraean fields near Naples, in Sasso Pisano (distinct from the site sampled in this study) and Monte Rotondo [65]. Overall, *Galdieria* dominated the fumaroles of Sasso Pisano, which was expected due to the mixotrophy, the resulting adaptation to fluctuating environmental factors, and availability of carbon to be exploited. *Cyanidium caldarium* was also identified.

Current descriptions of the organisms around geothermally active areas around the world usually only illuminate the photoautotrophic community of the order *Cyanidiales* without considering the heterotrophic eukaryotes. The investigation of the entire diversity is interesting as the extreme conditions can serve as models for living conditions in the early evolutionary stages of soil ecosystems and can provide further information on the development of biocoenoses in the history of the earth. On the other hand, they represent a reduced biocoenosis that can serve as a model for more complex ecosystems. *Collembola* sp. (*Hexapoda*) represent the largest and most complex species at this site. *Collembola* are among the typical bottom dwellers and have been described for other extreme sites, in particularly hot and cold areas [34]. They are known as consumers of bacteria, algae, and fungi [66–68], actively feeding and thriving on these biomass sources. *Gregarina* sp. are known to endoparasitise *Hexapoda* among invertebrates in numerous colonies [69]. Based on the detected taxa, a simple food web could be assumed in which *Cyanidium* and *Galdieria*, belonging to unicellular red algae *Cyanidiales*, represent the primary producers and build biomass which may be consumed by *Collembola* sp. as consumers, and *Gregarina* sp. endoparasitise *Collembola* sp. Symbiotic bacteria may facilitate the biofilm growth of *Cyanidiales*. *Collembola* sp. could alter the microbial communities, either directly (through selectively feeding) or indirectly (through the dissemination of microbial propagules). Ultimately, a more detailed description of the species for this site would be required to confirm the described food web as a next step. This can then serve as a model to describe the influences and interactions of the organisms involved on each other and relate these to more complex ecosystems.

5. Conclusions

The composition and morphology of microbiota around hot fluid springs, steam vent and soil samples from the geothermally active hydrothermal fields around the Sasso Pisano site (Pisa, Tuscany region) are presented. The prokaryotic community analysis indicated that *Ktedonobacteria* at the Sasso Pisano site ranged from thermophilic to mesophilic, and phylogenetically as well as metabolically diverse representatives were detected. Metabolic functional profiling of the thermophilic prokaryotic microbiota predicted a higher capability to utilise carbon (methane and aromatic compounds), sulphur and nitrogen compounds, and heavy metal resistance-conferring genes were also significantly more abundant in the hot spring microbiome as compared to the mesophilic soil microbial consortium. The diversity of eukaryotic microorganisms within biofilms at extreme conditions, with very high temperature (55 °C) and a very low pH value of 2, was evaluated. The detection of red algae *Cyanidiales*, *Arthropoda* (*Collembola* sp.) and *Apicomplexa* (*Gregarina* sp.) proposed a simplified food web at the thermophilic extremely nutrient-deprived acidic environment of a fumarole.

References

1. Barelli, A.; Bertini, G.; Buonasorte, G.; Fiordelisi, A.; Fiordelisi, C. Recent deep exploration results at the margins of the Larderello-Travale geothermal system. In Proceedings of the World Geothermal Congress 2000, Kyushu, Japan, 28 May–10 June 2000.
2. Bartoli, G.; Bottega, S.; Forino, L.M.C.; Ciccarelli, D.; Spano, C. Plant adaptation to extreme environments: The example of *Cistus salviifolius* of an active geo-thermal alteration field. *C. R. Biol.* **2014**, *337*, 101–110. [CrossRef]
3. Bertini, G.; Casini, M.; Gianelli, G.; Pandeli, E. Geological structure of a long-living geothermal system, Larderello, Italy. *Terra Nova* **2006**, *18*, 163–169. [CrossRef]
4. Narasimhan, T.; Goyal, K. Subsidence due to geothermal fluid withdrawal. *Subsid. Geotherm. Fluid Withdraw.* **1982**, *6*, 35–36. [CrossRef]
5. Bussotti, F.; Tognelli, R.; Montagni, G.; Borghini, F.; Bruschi, P.; Tani, C. Response of *Quercus pubescens* leaves exposed to geothermal pollutant input in southern Tuscany (Italy). *Environ. Pollut.* **2003**, *121*, 349–361. [CrossRef]
6. Pippucci, A.; Lorenzi, R.; Spanò, C.; Sorce, C. Stress-induced changes to the flora in a geothermal field in central Italy. *Acta Physiol. Plant* **2015**, *37*, 1–10. [CrossRef]
7. Brogi, A. The structure of the Monte Amiata volcano-geothermal area (Northern Apennines, Italy): Neogene-Quaternary compression versus extension. *Intl. J. Earth Sci.* **2007**, *97*, 677–703. [CrossRef]
8. Zelenski, M.; Taran, Y.; Galle, B. High emission rate of sulfuric acid from Bezymianny volcano, Kamchatka. *Geophys. Res. Lett.* **2015**, *42*, 7005–7013. [CrossRef]
9. D'Amore, F.; Panichi, C. Evaluation of deep temperatures of hydrothermal systems by a new gas geothermometer. *Geochim. Cosmochim. Acta* **1980**, *44*, 549–556. [CrossRef]
10. Selvi, F. *Geothermal Biotopes in Central Western Italy from a Botanical Viewpoint. Ecosystem Responses to CO2: The MAPLE Project Results*; Official Publications of the European Communities; European Communities: Luxemburg, 1999.
11. Elmarsdottir, A.; Ingimarsdottir, M.; Hansen, I.; Olafsson, J.S.; Olafsson, E. Vegetation and invertebrates in three geothermal areas in Iceland. In Proceedings of the International Geothermal Conference, Reykjavik, Iceland, 14–17 September 2003; Citeseer: Reykjavik, Iceland, 2003.
12. Rothschild, L.J.; Mancinelli, R.L. Life in extreme environments. *Nat. Cell Biol.* **2001**, *409*, 1092–1101. [CrossRef]
13. Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glöckner, F.O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, e1. [CrossRef] [PubMed]
14. Stoeck, T.; Bass, D.; Nebel, M.; Christen, R.; Jones, M.D.M.; Breiner, H.-W.; Richards, T.A. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **2010**, *19*, 21–31. [CrossRef] [PubMed]
15. Amin, N.; Schneider, D.; Hoppert, M. Bioleaching potential of bacterial communities in historic mine waste areas. *Environ. Earth Sci.* **2018**, *77*, 542. [CrossRef]
16. Dong, X.; Kleiner, M.; Sharp, C.E.; Thorson, E.; Li, C.; Liu, D.; Strous, M. Fast and simple analysis of MiSeq amplicon sequencing data with MetaAmp. *Front. Microbiol.* **2017**, *8*, 1461. [CrossRef]
17. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [CrossRef]
18. Schloss, P.D.; Westcott, S.L.; Ryabin, T.; Hall, J.R.; Hartmann, M.; Hollister, E.B.; Lesniewski, R.A.; Oakley, B.B.; Parks, D.H.; Robinson, C.J.; et al. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **2009**, *75*, 7537–7541. [CrossRef] [PubMed]
19. Edgar, R.C. UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **2013**, *10*, 996–998. [CrossRef]
20. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformation* **2014**, *30*, 614–620. [CrossRef]
21. Bushnell, B. BBMap: A fast, accurate, splice-aware aligner. In Proceedings of the 9th Annual Genomics of Energy & Environment Meeting, Walnut Creek, CA, USA, 20 March 2014; Lawrence Berkeley National Lab: Walnut Creek, CA, USA, 2014.

22. Rognes, T.; Flouri, T.; Nicholas, B.; Quince, C.; Mahe, F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **2016**, *4*, e2584. [CrossRef] [PubMed]
23. Quast, C.; Pruesse, E.; Yilmas, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glockner, F.O. The Silva ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [CrossRef] [PubMed]
24. Comeau, A.M.; Douglas, G.M.; Langille, M.G.I. Mikrobiome Helper: A custom and streamlined workflow for microbiome research. *Meth. Protoc.* **2017**, *2*, 11.
25. Iwai, S.; Weinmaier, T.; Schmidt, B.L.; Albertson, D.G.; Poloso, N.J.; Dabbagh, K.; DeSantis, T.Z. Piphillin: Improved prediction of metagenomic content by direct inference from human microbiomes. *PLoS ONE* **2016**, *11*, e0166104. [CrossRef]
26. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2016**, *44*, D457–D462. [CrossRef]
27. Dhariwal, A.; Chong, J.; Habib, S.; King, I.L.; Agellon, L.B.; Xia, J. MicrobiomeAnalyst: A web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res.* **2017**, *45*, W180–W188. [CrossRef]
28. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]
29. Stecher, G.; Tamura, K.; Kumar, S. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Mol. Biol. Evol.* **2020**, *37*, 1237–1239. [CrossRef] [PubMed]
30. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [CrossRef] [PubMed]
31. Nazem-Bokaee, H.; Gopalakrishnan, S.; Ferry, J.G.; Wood, T.K.; Maranas, C.D. Assessing methanotrophy and carbon fixation for biofuel production by *Methanosarcina acetivorans*. *Microb. Cell Factor.* **2016**, *15*, 1–13. [CrossRef]
32. Bennett, R.K.; Steinberg, L.M.; Chen, W.; Papoutsakis, E.T. Engineering the bioconversion of methane and methanol to fuels and chemicals in native and synthetic methylotrophs. *Curr. Opin. Biotechnol.* **2018**, *50*, 81–93. [CrossRef] [PubMed]
33. Hwang, I.Y.; Nguyen, A.D.; Nguyen, T.T.; Nguyen, L.T.; Lee, O.K.; Lee, E.Y. Biological conversion of methane to chemicals and fuels: Technical challenges and issues. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 3071–3080. [CrossRef]
34. Sømme, L. *Invertebrates in Hot and Cold Arid Environments*; Springer: Berlin/Heidelberg, Germany, 1995; p. 275.
35. Ciniglia, C.; Yoon, H.S.; Pollio, A.; Pinto, G.; Bhattacharya, D. Hidden biodiversity of the extremophilic Cyanidiales red algae. *Mol. Ecol.* **2004**, *13*, 1827–1838. [CrossRef]
36. Zheng, Y.; Saitou, A.; Wang, C.-M.; Toyoda, A.; Minakuchi, Y.; Sekiguchi, Y.; Ueda, K.; Takano, H.; Sakai, Y.; Abe, K.; et al. Genome Features and Secondary Metabolites Biosynthetic Potential of the Class Ktedonobacteria. *Front. Microbiol.* **2019**, *10*, 893. [CrossRef] [PubMed]
37. Yabe, S.; Aiba, Y.; Sakai, Y.; Hazaka, M.; Yokota, A. *Thermosporothrix hazakensis* gen. nov., sp. nov., isolated from compost, description of *Thermospo-rotrichaceae* fam. nov. within the class Ktedonobacteria Cavaletti et al. 2007 and emended description of the class Ktedonobacteria. *Int. J. Syst. Evol. Microbiol.* **2010**, *60*, 1794–1801. [CrossRef]
38. Yabe, S.; Aiba, Y.; Sakai, Y.; Hazaka, M.; Yokota, A. *Thermogemmatispora onikobensis* gen. nov., sp. nov. and *Thermogemmatispora foliorum* sp. nov., isolated from fallen leaves on geothermal soils, and description of Thermogemmatisporaceae fam. nov. and Thermogemmatisporales ord. nov. within the class Ktedonobacteria. *Int. J. Syst. Evol. Microbiol.* **2011**, *61*, 903–910.
39. Cavaletti, L.; Monciardini, P.; Bamonte, R.; Schumann, P.; Rohde, M.; Sosio, M.; Donadio, S. New Lineage of filamentous, spore-forming, Gram-positive bacteria from soil. *Appl. Environ. Microbiol.* **2006**, *72*, 4360–4369. [CrossRef] [PubMed]
40. Yabe, S.; Sakai, Y.; Abe, K.; Yokota, A.; Také, A.; Matsumoto, A.; Sugiharto, A.; Susilowati, D.; Hamada, M.; Nara, K.; et al. *Dictyobacter aurantiacus* gen. nov., sp. nov., a member of the family Ktedonobacteraceae, isolated from soil, and emended description of the genus *Thermosporothrix*. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 2615–2621. [CrossRef] [PubMed]
41. Jiang, Z.; Li, P.; Jiang, D.; Dai, X.; Zhang, R.; Wang, Y.; Wang, Y. Microbial community structure and arsenic biogeochemistry in an acid vapor-formed spring in Tengchong Geothermal Area, China. *PLoS ONE* **2016**, *11*, e0146331. [CrossRef]
42. De Miera, L.E.S.; Arroyo, P.; de Luis Calabuig, E.; Falagan, J.; Ansola, G. High-throughput sequencing of 16S RNA genes of soil bacterial communities from a naturally occurring CO_2 gas vent. *Int. J. Greenh. Gas. Control* **2014**, *29*, 176–184. [CrossRef]
43. Northup, D.; Melim, L.; Spilde, M.; Hathaway, J.; Garcia, M.; Moya, M.; Stone, F.; Boston, P.; Dapkevicius, M.; Riquelme, C. Lava cave microbial communities within mats and secondary mineral deposits: Implications for life detection on other planets. *Astrobiology* **2011**, *11*, 601–618. [CrossRef]
44. Brito, E.M.S.; Romero-Núñez, V.M.; Caretta, C.A.; Bertin, P.; Valerdi-Negreros, J.C.; Guyoneaud, R.; Goñi-Urriza, M. The bacterial diversity on steam vents from Paricutín and Sapichu volcanoes. *Extremophiles* **2019**, *23*, 249–263. [CrossRef]
45. Yabe, S.; Sakai, Y.; Abe, K.; Yokota, A. Diversity of *Ktedonobacteria* with Actinomycetes-like morphology in terrestrial environments. *Microbes Environ.* **2017**, *32*, 61–70. [CrossRef]
46. Ghezzi, D.; Sauro, F.; Columbu, A.; Carbone, C.; Hong, P.-Y.; Vergara, F.; De Waele, J.; Cappelletti, M. Transition from unclassified Ktedonobacterales to Actinobacteria during amorphous silica precipitation in a quartzite cave environment. *Sci. Rep.* **2021**, *11*, 1–18. [CrossRef]
47. King, C.E.; King, G.M. Description of *Thermogemmatispora carboxidivorans* sp. nov., a carbon-monoxide-oxidizing member of the class Ktedonobacteria isolated from a geothermally heated biofilm, and analysis of carbon monoxide oxidation by members of the class Ktedonobacteria. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 1244–1251. [CrossRef]

48. Arif, S.; Nacke, H.; Hoppert, M. Metagenome-assembled genome sequences of a biofilm derived from Marsberg Copper Mine. *Microbiol. Res. Announ.* **2021**, *10*, e01253-20.

49. Wall, K.; Cornell, J.; Bizzoco, R.W.; Kelley, S.T. Biodiversity hot spot on a hot spot: Novel extremophile diversity in Hawaiian fumaroles. *Microbiology* **2015**, *4*, 267–281. [CrossRef]

50. Tassi, F.; Venturi, S.; Cabassi, J.; Capecchiacci, F.; Nisi, B.; Vaselli, O. Volatile organic compounds (VOCs) in soil gases from Solfatara crater (Campi Flegrei, southern Italy): Geogenic source(s) vs. biogeochemical processes. *Appl. Geochem.* **2015**, *56*, 37–49. [CrossRef]

51. Ji, M.; Greening, C.; Vanwonterghem, I.; Carere, C.R.; Bay, S.K.; Steen, J.A.; Montgomery, K.; Lines, T.; Beardall, J.; Van Dorst, J.; et al. Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nat. Cell Biol.* **2017**, *552*, 400–403. [CrossRef] [PubMed]

52. King, G.M.; Weber, C.F.; Nanba, K.; Sato, Y.; Ohta, H. Atmospheric CO and hydrogen uptake and CO oxidizer phylogeny for Miyakejima, Japan volcanic deposits. *Microb. Environ.* **2008**, *23*, 229. [CrossRef] [PubMed]

53. Yang, J.; Zhou, E.; Jiang, H.; Li, W.; Wu, G.; Huang, L.; Hedlund, B.P.; Dong, H. Distribution and diversity of aerobic carbon monoxide-oxidizing bacteria in geothermal springs of China, the Philippines, and the United States. *Geomicrobiol. J.* **2015**, *32*, 903–913. [CrossRef]

54. Arce-Rodríguez, A.; Puente-Sánchez, F.; Avendaño, R.; Martínez-Cruz, M.; De Moor, J.M.; Pieper, D.H.; Chavarría, M. Thermo-plasmatales and sulfur-oxidizing bacteria dominate the microbial community at the surface water of a CO_2-rich hydrothermal spring located in Tenorio Volcano National Park, Costa Rica. *Extremophiles* **2019**, *23*, 177–187. [CrossRef] [PubMed]

55. De Miera, L.E.S.; Arroyo, P.; de Luis Calabuig, E.; Ansola, G. Effects of varying CO_2 flows on bacterial communities in mesocosms created from two soils. *Int. J. Greenh. Gas. Control.* **2016**, *46*, 205–214. [CrossRef]

56. McIlroy, S.J.; Kirkegaard, R.H.; Dueholm, M.S.; Fernando, E.; Karst, S.M.; Albertsen, M.; Nielsen, P.H. Culture-independent analyses reveal novel Anaerolineaceae as abundant primary fermenters in anaerobic digesters treating waste activated sludge. *Front. Microbiol.* **2017**, *8*, 1134. [CrossRef] [PubMed]

57. Kochetkova, T.V.; Toshchakov, S.V.; Zayulina, K.S.; Elcheninov, A.G.; Zavarzina, D.G.; Lavrushin, V.Y.; Bonch-Osmolovskaya, E.A.; Kublanov, I.V. Hot in Cold: Microbial life in the hottest springs in permafrost. *Microorganisms* **2020**, *8*, 1308. [CrossRef]

58. Nishihara, A.; Matsuura, K.; Tank, M.; McGlynn, S.E.; Thiel, V.; Haruta, S. Nitrogenase activity in thermophilic chemolithoautotrophic bacteria in the phylum *Aquificae* isolated under nitrogen-fixing conditions from Nakabusa Hot Springs. *Microb. Environ.* **2018**, *33*, 394–401. [CrossRef]

59. Rainey, F.A.; Nobre, M.F.; Schumann, P.; Stackebrandt, E.; Da Costa, M.S. Phylogenetic diversity of the Deinococci as determined by 16S ribosomal DNA sequence comparison. *Int. J. System. Evol. Microbiol.* **1997**, *47*, 510–514. [CrossRef] [PubMed]

60. Gerber, E.; Bernard, R.; Castang, S.; Chabot, N.; Coze, F.; Dreux-Zigha, A.; Hauser, E.; Hivin, P.; Joseph, P.; Lazarelli, C.; et al. *Deinococcus* as new chassis for industrial biotechnology: Biology, physiology and tools. *J. Appl. Microbiol.* **2015**, *119*, 1–10. [CrossRef] [PubMed]

61. Sumner, D.; Hawes, I.; Mackey, T.J.; Jungblut, A.D.; Doran, P.T. Antarctic microbial mats: A modern analog for Archean lacustrine oxygen oases. *Geology* **2015**, *43*, 887–890. [CrossRef]

62. Hamilton, T.L.; Klatt, J.M.; de Beer, D.; Macalady, J.L. Cyanobacterial photosynthesis under sulfidic conditions: Insights from the isolate *Leptolyngbya* sp. strain hensonii. *ISME J.* **2018**, *12*, 568–584. [CrossRef]

63. De Luca, P.; Taddei, R.; Varano, L. *Cyanidioschyzom merolae*: A new alga of thermal acidic enviroments. *J. Plant Taxon. Geogr.* **1978**, 33.

64. Merola, A.; Castaldo, R.; De Luca, P.; Gambardella, R.; Musacchio, A.; Taddei, R. Revision of *Cyanidium caldarium*. Three species of acidophilic algae. *G. Bot. Ital.* **1981**, *115*, 189–195. [CrossRef]

65. Yoon, H.S.; Ciniglia, C.; Wu, M.; Comeron, J.M.; Pinto, G.; Pollio, A.; Bhattacharya, D. Establishment of endolithic populations of extremophilic Cyanidiales (Rhodophyta). *BMC Evol. Biol.* **2006**, *6*, 78. [CrossRef]

66. Kühn, J.; Ruess, L. Effects of resource quality on the fitness of collembola fed single and mixed diets from the green and brown food chain. *Soil Biol. Biochem.* **2021**, *154*, 108156. [CrossRef]

67. Scheu, S.; Folger, M. Single and mixed diets in Collembola: Effects on reproduction and stable isotope fractionation. *Funct. Ecol.* **2004**, *18*, 94–102. [CrossRef]

68. Haubert, D.; Häggblom, M.M.; Langel, R.; Scheu, S.; Ruess, L. Trophic shift of stable isotopes and fatty acids in Collembola on bacterial diets. *Soil Biol. Biochem.* **2006**, *38*, 2004–2007. [CrossRef]

69. Mehlhorn, H. *Encyclopedia of Parasitology*, 4th ed.; Springer: Berlin/Heidelberg, Germany, 2016.

Supplementary Information

Sasso Pisano Geothermal Field Environment Harbors Diverse *Ktedonobacteria* Representatives and Illustrates Habitat-specific Adaptations

Sania Arif [1,*], Corinna Willenberg [1], Annika Dreyer [1], Heiko Nacke [2] and Michael Hoppert [1]

[1] Institute of Microbiology and Genetics, Department of General Microbiology, University of Göttingen, Grisebachstraße 8, 37077 Göttingen, Germany

[2] Institute of Microbiology and Genetics, Department of Department of Genomic and Applied Microbiology, University of Göttingen, Grisebachstraße 8, 37077 Göttingen, Germany

* Correspondence: Email: sarif@gwdg.de, Phone: +49-551-39-9657, Fax: +49-551-39-3808

Table S1. Complete details of the sampling sites

	Type	pH	Source 1
Sp1	Yellowish brownish Biofilm	5 – 6	Biofilm just above hot spring A vent
SP2	Yellowish-beige Precipitates on rocks	-	Nearby the hot spring A
Sp11	Grey greenish discharge	7.4	Hot spring B vent
Sp12	Grey reddish mud	-	Around the hot spring B outlet
Sp25	Black muddy water	6.7	Hot spring C
Sp26	Black muddy discharge	6.7	20cm away hot spring C
Sp27	Brownish sediment and water	6.7	60cm away downstream spring C
Sp38	Leaves mixed with black discharge	-	Downstream hot spring D
Sp310	White discharge	-	Hot Spring D outlet
Sp513-15, Sp517, Sp719, Sp819-21, Sp3 Sp122	Soil samples		All over
Sp4	Water with greenish biomass	7	Water stream
Sp5	Greenish biofilm	2	Fumarole
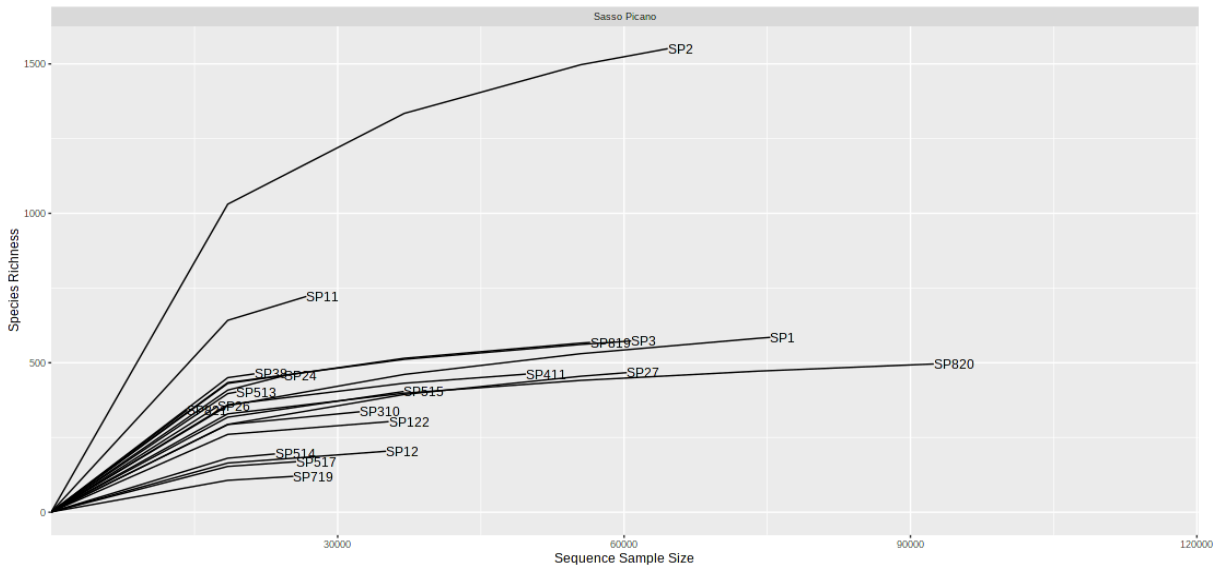
Figure S1. Rarefaction curve
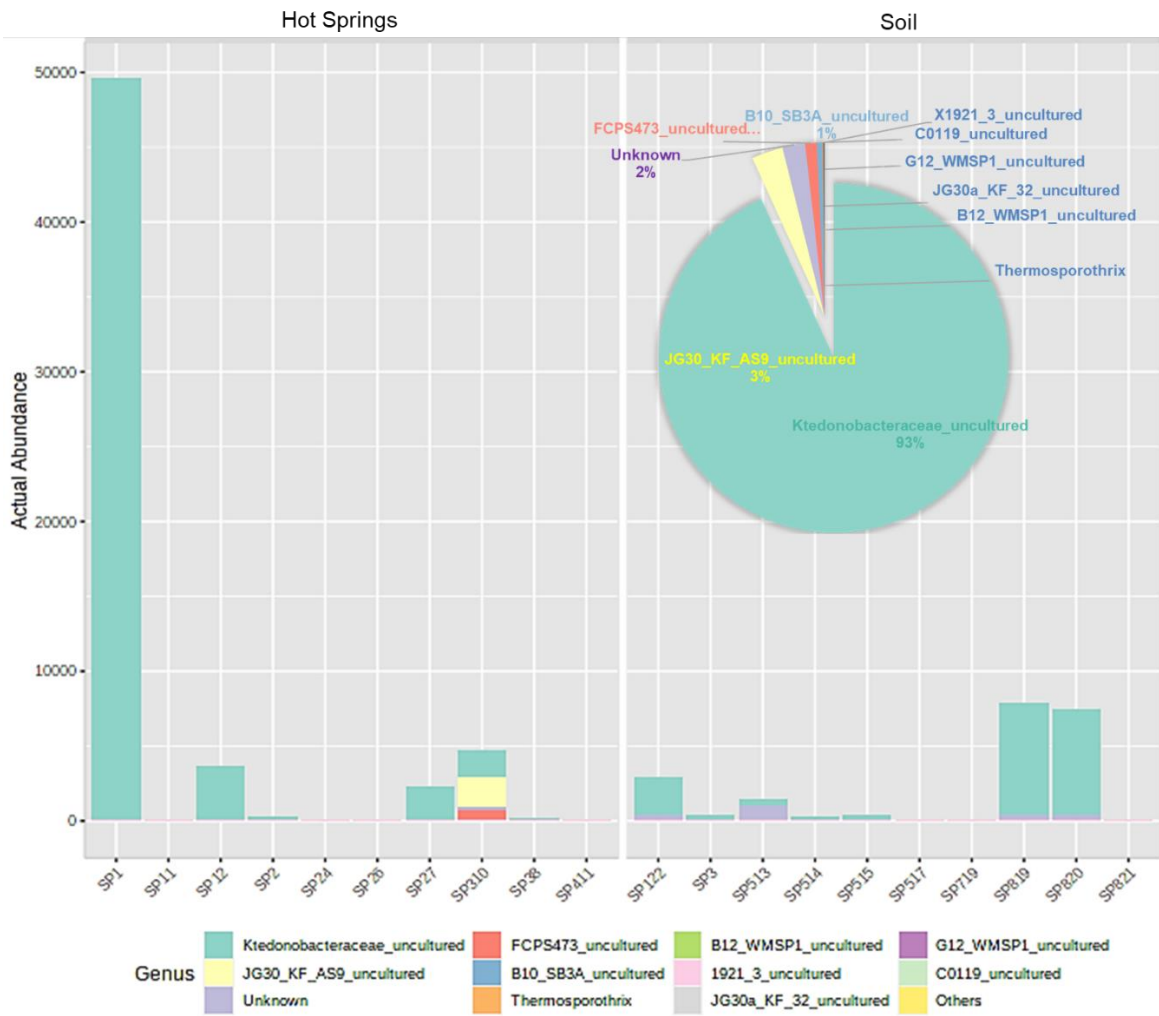
Figure S2. Number of sampled reads of the *Ktedonobacteria* members. Insert shows the relative abundance of the *Ktedonobacteria* at the genus level from all samples.
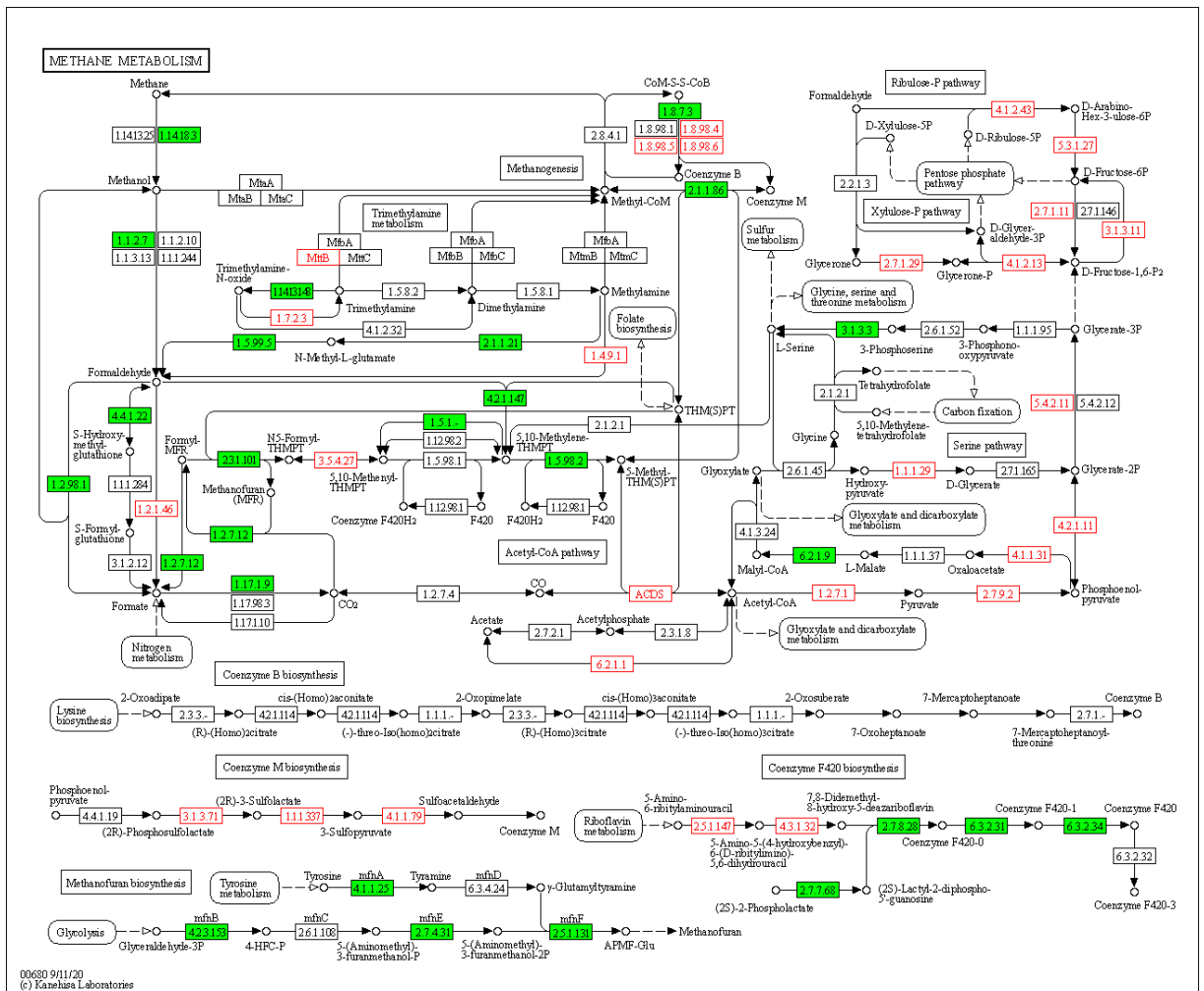
Figure S3. Differentially abundant genes involved in the methane metabolism. The differentially abundant genes KO numbers were mapped on the KEGG methane pathway map for hot springs (red) and soil (green) microbiome. The hot spring microbiome was found to be abundant in the genes involved in the anaerobic methane degradation as compared to the aerobic methane degradation in the soil microbiome.
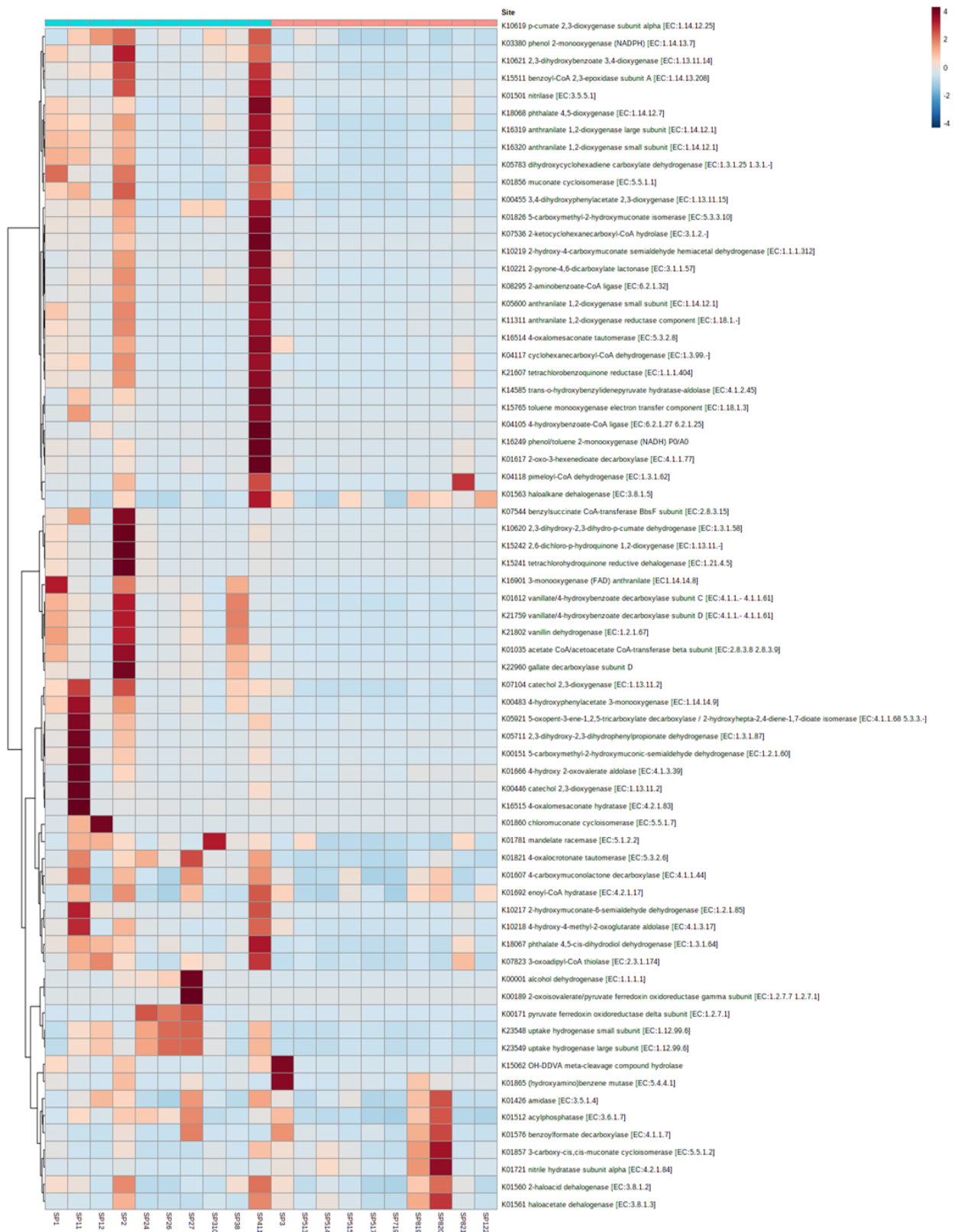
Figure S4. Relative abundance of differentially abundant genes involved in the aromatic compounds metabolism.
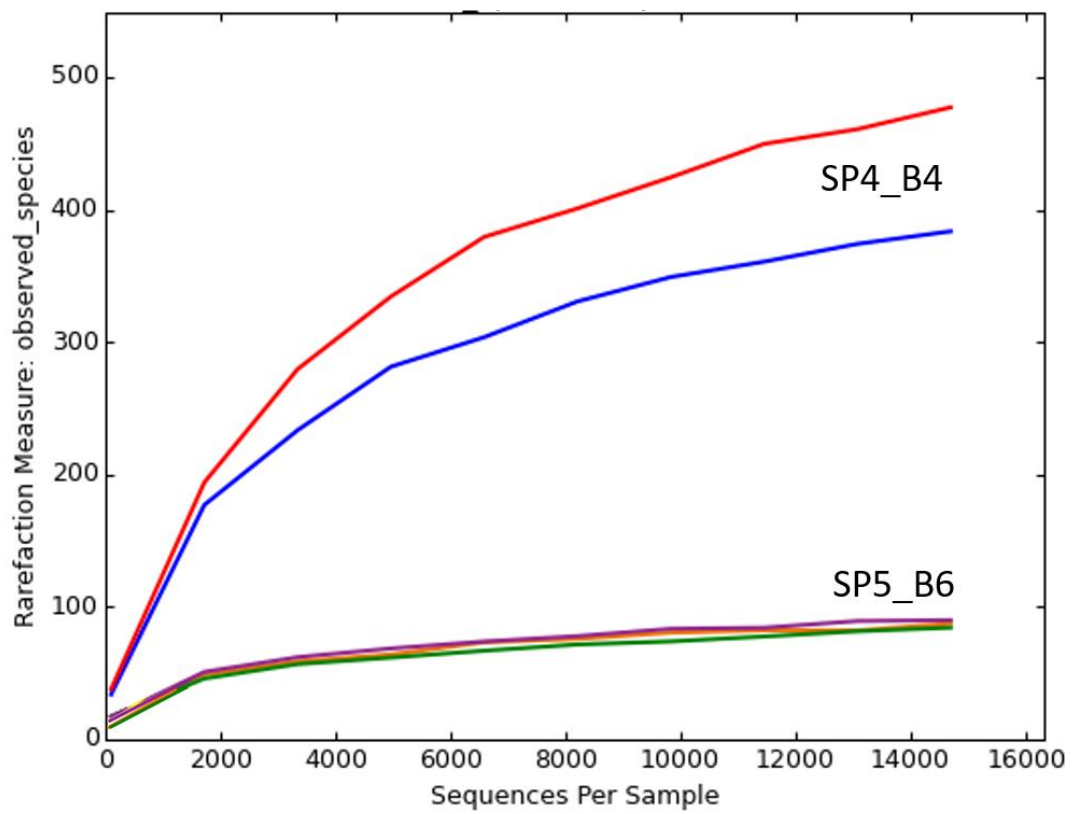
Figure S5. Eukaryotic rarefaction curve of a hot steam fumarole and neutral water stream. The number of OTUs are plotted against the reads on the x-axis. The extreme site SP5 samples OTUs have raised very quickly between 0 and 2500 reads and has almost reached saturation from 10,000 reads while the SP4 samples OTUs have a flatter rise and do not reach saturation.

5. General Discussion

5.1 Determining Abiotic Factors of Organic-rich Shales

Due to the worldwide occurrence of organic-rich shales, their geographical and abiotic circumstances vary immensely under the influence of diverse geological and/or anthropogenic activities. For instance, subsurface marine and terrestrial organic-rich shales are being exposed to microbial communities through hydraulic fracturing for oil and gas recovery (Akyon et al., 2015; Mouser et al., 2016). The abiotic factors are determining the composition of microbial communities which will colonize and prevail in the organic-rich environment. In the case of hydraulic fracturing, the temperature at the subsurface site could reach 90 °C and the saline conditions select thermophilic halotolerant bacterial groups (*Halanaerobium*, *Halomonas*, *Halolactibacillus*) related to hydrocarbon oxidation, fermentation, and sulfur-cycling metabolisms from the pumped water (Ulrich et al., 2018). In contrast, the determining abiotic factors at the exposed terrestrial organic-rich shale could be the availability of moisture, soil contamination, presence of heavy metals or pyrite, temperature, and light (Li et al., 2014). The anthropogenic activities could either influence positively or adversely these abiotic factors, as mining activities could expose organic-rich shales to atmospheric oxygen that would oxidize organic components and facilitate colonization by aerobic microorganisms. Since under the scope of this thesis, the terrestrial organic-rich shales were investigated, a comparison of geographical and abiotic factors prevailing at Monte San Giorgio and Marsberg Kilianstollen copper mine sites is required to understand their influence on microbial diversity and organic-rich shales. Under different geological circumstances, Monte San Giorgio rocks were uplifted above sea level and Marsberg Devonian and Carboniferous rocks got exposed to the lithospheric conditions, primarily due to tectonic plate movements (Emmerich, 1987). As a result, the intact anaerobic environment of deep-buried subsurface fossil and copper-bearing organic-

rich shales changed to the aerobic bare atmospheric conditions (Rieber, 2000). Monte San Giorgio has been excavated for its well preserved mid-Triassic, 245 to 230 million years old fossils (Sander & Greenwood, 1989). Marsberg carboniferous shale has been extensively mined from 1150 until the 1960s for its copper reserves (Stribrny et al., 1988). Monte San Giorgio shale is directly exposed to terrestrial weathering conditions and sunlight. Marsberg Kilian copper mine offers subsurface terrestrial dark tunnels with mine water seeping and artificial illumination. A major difference is that Monte San Giorgio organic-rich shale does not enrich metals or has lower amounts of evenly disseminated pyrite crystal (Stockar, 2010) in comparison to the metalliferous black shale of Marsberg (Emmerich, 1987). Marsberg Kilian tunnel has a constant cold temperature of 10 °C and that of Monte San Giorgio ranges from -8 to 23 °C depending on the season (data according to National Centers for Environmental Information, Asheville, NC, USA, www.ncdc.noaa.gov). The mine waters of the Kilian tunnel ranges from fresh ground water to heavy metal-containing leachate, depending upon the rocks that come in contact with the rain/ground water. Human activities, such as regular visits from tourists can bring soil microbiome into these environments as well as seeping rain/ground water could also bring microbial influx.

Based on these abiotic factors, it could be hypothesized that the native lithobionthic microbiome at both locations may be specialized to degrade and metabolize polyaromatic hydrocarbons, derived from kerogen and bitumen. In contrast to Monte San Giorgio, the Marsberg microbiome may have evolved resistance against heavy metals (especially copper) and play a major role in Fe-S cycling. The comparison of microbial communities inhabiting both organic-rich shales would help to understand whether the lithobionthic microbial communities are derived from soil/water sources or are indigenous inhabitants.

5.2 Major Abundant Microbial Communities

The major common abundant taxa at Monte San Giorgio and Marsberg organic-rich shales were *Acidobacteria*, *Actinobacteria*, *Chloroflexi*, *Proteobacteria*, *Firmicutes*, *and Bacteroidetes* (Figure 1 and 4 from chapter 2 and 3, respectively). These taxa were also found to be involved in remediating hydrocarbon contaminated water bodies (Eze et al., 2021), coalbeds (Imachi et al., 2019), oil spills (Liu & Liu, 2013), black shale weathering (Li et al., 2014), and acid mine/rock drainage (Aguinaga et al., 2018). *Acidobacteria, Actinobacteria*, and *Proteobacteria* are common soil inhibiting microorganisms (Lauber et al., 2009). In Monte San Giorgio organic-rich shale, the major proteobacterial genera (30% of the whole bacterial community) *Acidiphilium, Pedomicrobium*, *Variibacter*, *Reyranella*, *Pseudolabrys* (α-25%), *Achromobacter*, *Dechloromonas*, *Thiobacillus* (β-5%), *Halomonas*, *Shewanella*, *Chromatium*, *Legionella*, *Shigella*, *Stenotrophomonas*, *Arcobacter*, *Acinetobacter*, *Pseudomonas*, *Moraxella*, (γ-54%), *Desulfocapsa*, *Geobacter*, *Anaeromyxobacter* (δ-14%), and *Perlucidibaca* (ε-1%) are supposed to be involved in organic and sulphur compounds metabolism including polyaromatic ring catalysis (Figure S1). The biodegradative role and heavy metals resistance of *Alphaproteobacteria*, predominantly *Acidiphilium*, has been previously studied in relation to natural oil seepages and could be applied to remediate hydrocarbon-polluted acidic mine sites (Giovanella et al., 2020). *Gammaproteobacteria,* especially *Pseudomonas,* which was also the most abundant genus at the Monte San Giorgio site has been reported to contain ring-hydroxylating and ring-cleavage dioxygenases in a study of diesel-contaminated soil rhizoremediation metagenome (Garrido-Sanz et al., 2019). Moreover, n-alkane degrading genes were also found relevant to *Gammaproteobacteria* along with *Betaproteobacteria* which has been known for anaerobic aromatic hydrocarbon degradation under nitrate-reducing

conditions (Liu et al., 2014; Tan & Parales, 2019). *Delta-* and *Epsilonproteobacteria* also have been reported to oxidize organic matter and degrade crude oil.

Most of these observed *Proteobacteria* genera are commonly abundant in soil which may indicate that the Monte San Giorgio organic-rich shale microbial consortium is derived from soil. This hypothesis could be true as the beta diversity indicated that the microbial consortium at soil contaminated outside of shale showed significant similarity to that of inside organic-rich shale (Chapter 1, Figure 4).

In contrast to Monte San Giorgio organic-rich shale microbiome, the major phylum in the Marsberg leachate group was *Chloroflexi* (30%) where *Proteobacteria* represented only 10%. Interestingly, *Ktedonobacteria* and *KD4-96* represented >85% and >33% of the *Chloroflexi* at the Marsberg leachate and Monte San Giorgio organic-rich shale respectively. *KD4-96* class has been frequently detected in polycyclic aromatic compounds (Lemmel et al., 2019) and metals (iron, aluminum) contaminated soils (Wegner & Liesack, 2017) as well as mining-affected waters (Kujala et al., 2018). On the other hand, *Ktedonobacteria* were associated with di-(2-ethylhexyl) phthalate (DEHP) biodegradation in agricultural soils (Song et al., 2019) and diesel oil (over 90% C10–C28 hydrocarbons) contaminated soil (Borowik et al., 2020). Several members of *Chloroflexi* including *Ktedonobacteriales* have shown cellulolytic phenotype by gaining energy and carbon solely from crystalline cellulose (Yabe et al., 2010, 2011).

The co-abundance of *Chloroflexi* and *Cyanobacteria* at the Marsberg biofilms growing nearby leachate streams indicate a cross-feeding symbiotic relationship where autotrophic *Cyanobacteria* ferments stored carbohydrates at night into organic acids which are taken up by heterotrophic *Chloroflexi* to assimilate carbon as polyhydroxyalkanoates (Klatt et al., 2013). In return, heterotrophic *Chloroflexi* may provide protection to *Cyanobacteria* against low pH and heavy metal stress as *Oxyphotobacteria* prefer neutral to alkaline environments (Cano-Díaz et al., 2020) and

accumulate heavy metals intracellularly (Yadav et al., 2021). *Cyanobacteria* and *Chloroflexi* members are known as typical pioneering colonizers where the heterotrophic *Chloroflexi* are usually sporulating (Medrano-Santillana et al., 2017).

The low similarity indexes of *Ktedonobacteria* and *Actinobacteria* OTUs related to the Marsberg leachate group through phylogenetic tree analysis indicates that the native biofilms are indigenous and unique rather than commonly abundant soil communities. To elaborate the distinctiveness of the native microbiome, *Ktedonobacteria* MAG would be discussed later in comparison to typed strain *Ktedonobacter racemifer*, previously isolated from soil.

5.3 Comparative Functional Analysis

To test the hypothesis that the native lithobiontic microbiome at both organic-rich shales should be specialized in degrading complex hydrocarbons derived from kerogen and bitumen, the important functional pathways related to polycyclic aromatic compounds biodegradation were plotted (Figure S2). The differential abundance analysis clearly showed that the Monte San Giorgio organic-rich shales (inside layers) microbiome has evolved specialized pathways related to cleavage of the hydrocarbon rings (benzoate, biphenol, xylene, styrene, naphthalene) as compared to limestone and surface shale microbiome. In the Marsberg Kilian microbiome, the metabolism of the aromatic compound was found to be ubiquitous in three groups (leachate, unconsolidated rocks, and spring water) depending on their sampling sites. However, in the spring water group, the rising trend of aromatics metabolism correlated with the increasing heavy metal concentrations and *Actinobacteria* (*Crossiella* spp.; major abundant genus in MBS3,4) abundance. *Crossiella* spp. are characterized as rare novel *Actinobacteria* which have been detected from stone monuments causing Ca_2CO_3 precipitation (Li et al., 2018). They were also found in cave soil where they were

involved in hydrocarbon degradation (Wiseschart et al., 2019) and in Oman sea sediments among other antioxidant and cytotoxin producing isolates of *Actinobacteria* (Gozari et al., 2019). However, exact pathways and genes related to polycyclic aromatic compounds metabolism have not been investigated yet (Kämpfer, 2010). *Proteobacteria* diversity also changed from major subclass *Delta* to more specialized hydrocarbons degrading *Gamma* (MBS2,3,4) and heavy metal tolerant *Alphaproteobacteria* (MBS4). The unconsolidated rock and copper-leachate biofilms group also showed some potential of aromatic compounds metabolism in most of the samples. The *Actinobacteria* (*Acidimicrobiia* MAG 009), *Candidatus Binatota* (MAG 010 and 011), *Chloroflexi* (MAG *Ellin6529* 002, *Chloroflexia* 006, *Dehalococcoidia* 008, and *Ktedonobacteria* 019), and *Deinococcus-Thermus* (MAG 015) extracted from the biofilm MB1 as a representative of leachate group contained several genes to metabolize complex polyaromatic compounds (Table S1) which indicates KO numbers computed from 16srRNA are underrepresented.

The assumption based on diverse abiotic factors that in contrast to Monte San Giorgio, Marsberg microbiome may have evolved heavy metals resistance especially against transition metals, could be proved by exploring the type of ABC transporters detected in each microbiome (Figure S2). The differential abundance analysis of transporter genes in the Marsberg Kilian leachate group showed that the microbiome has transporter genes for transition metals and metalloids (cobalt, nickel, zinc, arsenite), and transporters related to the Monte San Giorgio organic (inside) shale microbiome were involved in the uptake of various metabolites (Table S2). Various types of specific and non-specific transporter genes involved in heavy metal haemostasis have been already detected in all MAGs (Chapter 3, Figure 8).

5.4 *Ktedonobacteria are* not only abundant at the Marsberg copper mine

Ktedonobacteria class has been detected at a quite low abundance <1% at the Monte San Giorgio limestone and shale surface that could be justified because these mycelial bacteria are present at low abundance in the soil microbiome (Yabe et al., 2017). In contrast, *Ktedonobacteria* enrichment in the Marsberg copper mine is quite significant which hints towards their heavy metal resistance and potential of aromatic compounds metabolism. Under stringent environmental conditions, a high abundance of *Ktedonobacteria* has usually been observed in various extreme environments such as fumaroles (Medrano-Santillana et al., 2017), zinc and lead polluted soil (Epelde et al., 2015), quartzite cave (Ghezzi et al., 2021) and volcano soil (Zheng et al., 2020). We have also observed *Ktedonobacteria* enriched biofilms at another completely different site, a hydrothermal field in Italy, Sasso Pisano (see chapter 4). Sasso Pisano offers boiling hot springs emitting CH_4, CO_2, H_2, sulphur compounds (H_2S, SO_4^{2-}), various hydrocarbons and aromatic compounds, heavy metals, and the acidic environment (pH 0.5–5) while Marsberg Kilian copper mine has a constant temperature (~10°C), low mine water pH (4-5), 98% humidity, copper and iron sulphides, complex hydrocarbons (2-10%) and acidic heavy metal-rich mine drainage. The high abundance of *Ktedonobacteria* at Marsberg and Sasso Pisano could be because of low pH, heavy metal stress, and radically available hydrocarbons. The extreme temperature conditions might have given these resistant extremophiles a selective advantage to proliferate and colonize. Although the inhabiting *Ktedonobacteria* strains were found to be phylogenetically distinct at Marsberg and Sasso Pisano (Chapter 3, Figure S4 and Chapter 4, Figure 4), it is hypothesized that they show functional conservation for the carbon monoxide oxidation (Cox operons), heavy metal resistance, aromatic compounds metabolism, and sporulation mechanisms. Their ubiquitous abundance in

soil and extreme environments shows their diversity to cope with several environmental stresses and ability to colonize hydrocarbon-containing sources as extremophiles.

But it raises a question that whether *Ktedonobacteria* members have evolved specialized pathways to metabolize hydrocarbons, or they are just common soil-inhabiting saprophytic bacteria that decompose organic matter. *Thermosporothrix hazakensis* SK20-1[T] was isolated from ripe compost produced from livestock excreta by a Hazaka Plant Kogyo composter (Yabe et al., 2010). *Ktedonobacter racemifer*, the first typed strain of *Ktedonobacteria* was isolated from a black locust (*Robinia pseudoacacia*) wood soil, Italy on HSA5 agar which included humic acid as a sole carbon source (Cavaletti et al., 2006). Humic acid is a random-network macromolecular product of microbially decomposed organic matter of the animal and plant remains (Li et al., 2019). Like kerogen, humic acid contains heterogeneous macromolecules, randomly aggregated to three-dimensional structures, where the metal ions may coordinate within charged groups such as carboxylates (Xing & Pignatello, 2005). Kerogen genesis is under high temperature and pressure from millions of year-old organic matter sedimented in shale, in contrast to the freshly deposited peat compost containing humic acids. The redox-active functional groups coupled with humic acid are capable of oxidizing or reducing organic molecules and ions (Eljarrat, 2012). Kerogen is more recalcitrant to chemical and thermal reactivity than humin and humic acid because of increased cross-linking, hydrophobicity, molecular weight, fused ring size, and carbon content (Xing & Pignatello, 2005).

 Due to the similarity of kerogen and humic substances macromolecules, *Ktedonobacteria* members could have adapted their specialized pathways to degrade complex ring structures of the kerogenic substrates too. Several thermophilic strains have been isolated from the hydrocarbons emitting geothermal sources where the decomposing organic matter was also accessible. *Thermogemmatispora*

carboxidivorans PM5[T] was isolated from a geothermally heated biofilm growing on a tree stump in contact with steam vents from Puhimau geothermal Kilauea Volcano, Hawaii (King & King, 2014). *Thermosporothrix narukonensis* F4[T], *Thermogemmatispora onikobensis* ONI-1T, and *Thermogemmatispora foliorum* ONI-5[T] were isolated from fallen leaves on geothermal steaming soil (Yabe et al., 2011; Yabe et al., 2016), and *Thermogemmatispora aurantia* A1-2[T] and *Thermogemmatispora argillosa* A3-2[T] were isolated from clay soil at the Onikobe Hot Springs, Japan (Zheng et al., 2019). Intriguingly, synthesis of complex secondary metabolites such as unusual branched fatty acid (Vyssotski et al., 2012), the acyloins, sattabacin and hazakacin (Park et al., 2014), two thiazole derivatives (Park et al., 2015), as well as ktedonoketone, 2'-oxosattabacin, two anthranilic derivatives and benzenoid metabolites of benzenepropanic and cinnamic acid (Igarashi et al., 2019) was reported from *T. hazakensis* SK20-1[T] (Yabe et al., 2010). Retrospectively, it could be suggested that these extremophiles also have pathways involved in the complex polycyclic aromatic ring catabolism such as *Ktedonobacteria* MAG 019 sequenced from cold Marsberg copper mine has shown benzoate, phenylacetate, carbazole degradation, biphenyl, cinnamate atrazine, diphenyl ethers and catechol degradation pathways (Table S1). *T. hazakensis* SK20-1[T] was isolated from ripe compost, which indicates that it could also metabolize polycyclic aromatic rings of humic acid or kerogenic substrates. Being a recently discovered class, *Ktedonobacteria* pathways related to hydrocarbon metabolism and colonization under extreme environmental stress conditions have not been investigated yet. Therefore, a genomic comparison between Marsberg cold copper mine MAG 019, soil-inhabiting mesophilic typed strain *K. racemifer* SOSP1-21 [T] (ASM17885v1), and thermophilic *T. hazakensis* SK20-1[T] (ASM540264v1) has been made in the following part to underscore the metabolic and functional profiles of the *Ktedonobacteria.*

5.5 Genomic Comparison of *Ktedonobacteria* Members

The shared and unique genomic features of MAG 019 as compared to *K. racemifer* and *T. hazakensis* show that *Ktedonobacteria* members have evolved multiple orthologous clusters (GO terms) related to different metabolic functions related to heterocyclic, macromolecules, and aromatic compounds. A total of 10818, 6060, and 3745 proteins were identified in *K.* racemifer, *T. hazakensis*, and MAG 019 which were collected into orthologous gene clusters 4506, 3736, and 1644, respectively. MAG019 is a relatively small genome (3.9 Mb) with 90% completion as compared to *K. racemifer* (13.66 Mb) and *T. hazakensis* (7.4 Mb). MAG 019 shared a similarity of 96.1% with *K. racemifer* and *T. hazakensis* based on orthologous shared clusters (Figure S3). More in-depth analyses based on RAST annotations showed that MAG 019 and other *Ktedonobacteria* contain genes responsible for aromatic compound metabolism (Figure S4). In contrast to other *Ktedonobacteria* genomes, MAG0 19 adaption to the cold environment of Marsberg copper mine is attributed primarily to the genes related to heat shock, copper homeostasis, bacterial chemotaxis, aromatic compounds metabolism through protocatechuate branch of beta-ketoadipate pathway, Ton and Tol transport system against heavy metal resistance, DNA repair and peptidoglycan biosynthesis (Figure S5). It could be concluded that the success of *Ktedonobacteria* extremophiles in diverse harsh and oligotrophic environments is credited to their low pH tolerance, phylogenetic diversity, aerobic carbon monoxide oxidation, heavy metal resistance, aromatic compounds metabolism, psychrophiles to thermophiles, larger genomes (5.54–13.66 Mb), genome plasticity, diverse secondary metabolite synthesis, and sporulation.

5.6 Conclusion

A large part of the organic carbon present in the lithosphere is trapped in fossil organic matter deposited in sedimentary rocks. Sedimentary formations represent 66% of Earth's surface rocks. If 10^8 gigatons of sediments (Falkowski et al., 2000) are estimated to contain just 0.3% of organics, there is potentially 300000 gigatons of potentially degradable organic carbon accrued. The exposed organic-rich shales host specialized microorganisms which are able to degrade it and eventually contribute to the return of the organic carbon and inorganic sulphur to the global cycle. It could also lead to decreased yields and recovery of gas and oil from black shales, for instance, in the process of hydraulic fracturing. The geological and abiotic factors of organic-rich shales are important decisive factors to colonize the respective microbial communities. The polycyclic aromatic rings of recalcitrant kerogen provide slow-growing specialized oligotrophs to extract energy and assimilate carbon from hydrocarbons. The extremophiles should be able to cope with environmental stresses that come along different geological settings, such as desiccation, heavy metals toxicity, temperature extremes, and lack of radically available nutrients. The geomicrobiological role of the newly discovered class of *Ktedonobacteria* and rare *Actinobacteria* members has been a neglected niche, especially in terms of their colonization as extremophiles and contribution to carbon cycle via degradation of recalcitrant kerogenic aromatic substrates which requires further research for validation. The metagenomic evaluation revealed that they have evolved multiple mechanisms against low pH, heavy metal stress, extreme temperatures, and availability of nutrients. The long-term impact of environmental microorganisms on the fossilized organic compounds and their contribution to the carbon cycle and leaching of heavy metals from shale is gaining more attention as more organic-rich shales deposits are being exploited for either gas and oil recovery or mining purposes.

5.7 References

Aguinaga, O.E., McMahon, A., White, K.N., Dean, A.P., and Pittman, J.K. (2018) Microbial community shifts in response to acid mine drainage pollution within a natural wetland ecosystem. *Frontiers in Microbiology* **9**: 1445.

Akyon, B., Stachler, E., Wei, N., and Bibby, K. (2015) Microbial mats as a biological treatment approach for saline wastewaters: The case of produced water from hydraulic fracturing. *Environmental Science & Technology* **49**: 6172-6180.

Borowik, A., Wyszkowska, J., Kucharski, M., and Kucharski, J. (2020) The role of dactylis glomerata and diesel oil in the formation of microbiome and soil enzyme activity. *Sensors* **20**: 3362.

Cano-Díaz, C., Maestre, F.T., Eldridge, D.J., Singh, B.K., Bardgett, R.D., Fierer, N., and Delgado-Baquerizo, M. (2020) Contrasting environmental preferences of photosynthetic and non-photosynthetic soil cyanobacteria across the globe. *Global Ecology and Biogeography* **29**: 2025-2038.

Cavaletti, L., Monciardini, P., Bamonte, R., Schumann, P., Rohde, M., Sosio, M., and Donadio, S. (2006) New lineage of filamentous, spore-forming, gram-positive bacteria from soil. *Applied and Environmental Microbiology* **72**: 4360-4369.

Eljarrat, E. (2012) 1.02 - Methodologies for Sample Preservation and Stabilization. In *Comprehensive Sampling and Sample Preparation*. Pawliszyn, J. (ed). Academic Press, Oxford, pp. 31-49.

Emmerich, H.P. Heydemann, A. (1987) Sekundärmineralbildung aus Grubenwässern im Kupferbergwerk Niedermarsberg [Secondary mineral formation from mine water in the Niedermarsberg copper mine]. *Der Aufschluss* 38, 149-156.

Epelde, L., Lanzen, A., Blanco, F., Urich, T., and Garbisu, C. (2015) Adaptation of soil microbial community structure and function to chronic metal contamination at an abandoned Pb-Zn mine. *FEMS Microbiology Ecology* **91**: 1-11.

Eze, M.O., Hose, G.C., George, S.C., and Daniel, R. (2021) Diversity and metagenome analysis of a hydrocarbon-degrading bacterial consortium from asphalt lakes located in Wietze, Germany. *AMB Express* **11**: 89.

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.

Garrido-Sanz, D., Redondo-Nieto, M., Guirado, M., Pindado Jiménez, O., Millán, R., Martin, M., and Rivilla, R. (2019) Metagenomic insights into the bacterial functions of

a Diesel-degrading consortium for the rhizoremediation of Diesel-polluted soil. *Genes* **10**: 456.

Ghezzi, D., Sauro, F., Columbu, A., Carbone, C., Hong, P.Y., Vergara, F. et al. (2021) Transition from unclassified Ktedonobacterales to Actinobacteria during amorphous silica precipitation in a quartzite cave environment. *Scientific Reports* **11**: 1-18.

Giovanella, P., Vieira, G.A., Otero, I.V.R., Pellizzer, E.P., de Jesus Fontes, B., and Sette, L.D. (2020) Metal and organic pollutants bioremediation by extremophile microorganisms. *Journal of Hazardous Materials* **382**: 121024.

Gozari, M., Bahador, N., Jassbi, A.R., Mortazavi, M.S., Hamzehei, S., and Eftekhar, E. (2019) Isolation, distribution and evaluation of cytotoxic and antioxidant activity of cultivable actinobacteria from the Oman Sea sediments. *Acta Oceanologica Sinica* **38**: 84-90.

Igarashi, Y., Yamamoto, K., Ueno, C., Yamada, N., Saito, K., Takahashi, K. et al. (2019) Ktedonoketone and 2'-oxosattabacin, benzenoid metabolites from a thermophilic bacterium Thermosporothrix hazakensis in the phylum Chloroflexi. *The Journal of Antibiotics* **72**: 653-660.

Imachi, H., Tasumi, E., Takaki, Y., Hoshino, T., Schubotz, F., Gan, S. et al. (2019) Cultivable microbial community in 2-km-deep, 20-million-year-old subseafloor coalbeds through ~1000 days anaerobic bioreactor cultivation. *Scientific Reports* **9**: 2305.

King, C.E., and King, G.M. (2014) Description of *Thermogemmatispora carboxidivorans* sp. nov., a carbon-monoxide-oxidizing member of the class Ktedonobacteria isolated from a geothermally heated biofilm, and analysis of carbon monoxide oxidation by members of the class Ktedonobacteria. *International Journal of Systematic and Evolutionary Microbiology* **64**: 1244-1251.

Klatt, C.G., Liu, Z., Ludwig, M., Kühl, M., Jensen, S.I., Bryant, D.A., and Ward, D.M. (2013) Temporal metatranscriptomic patterning in phototrophic Chloroflexi inhabiting a microbial mat in a geothermal spring. *The ISME Journal* **7**: 1775-1789.

Kujala, K., Mikkonen, A., Saravesi, K., Ronkanen, A.-K., and Tiirola, M. (2018) Microbial diversity along a gradient in peatlands treating mining-affected waters. *FEMS Microbiology Ecology* **94**.

Kämpfer, P. (2010) Actinobacteria. In *Handbook of Hydrocarbon and Lipid Microbiology*. Timmis, K.N. (ed). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1819-1838.

Lauber, C.L., Hamady, M., Knight, R., and Fierer, N. (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and Environmental Microbiology* **75**: 5111-5120.

Lemmel, F., Maunoury-Danger, F., Fanesi, A., Leyval, C., and Cébron, A. (2019) Soil properties and multi-pollution affect taxonomic and functional bacterial diversity in a range of French soils displaying an anthropisation gradient. *Microbial Ecology* **77**: 993-1013.

Li, J., Sun, W., Wang, S., Sun, Z., Lin, S., and Peng, X. (2014) Bacteria diversity, distribution and insight into their role in S and Fe biogeochemical cycling during black shale weathering. *Environmental Microbiology* **16**: 3533-3547.

Li, Q., Zhang, B., Yang, X., and Ge, Q. (2018) Deterioration-associated microbiome of stone monuments: structure, variation, and assembly. *Applied and Environmental Microbiology* **84**: e02680-02617.

Li, Y., Fang, F., Wei, J., Wu, X., Cui, R., Li, G. et al. (2019) Humic Acid Fertilizer Improved Soil Properties and Soil Microbial Diversity of Continuous Cropping Peanut: A Three-Year Experiment. *Scientific Reports* **9**: 12014.

Liu, H., Xu, J., Liang, R., and Liu, J. (2014) Characterization of the medium-and long-chain n-alkanes degrading *Pseudomonas aeruginosa* strain SJTD-1 and its alkane hydroxylase genes. *PloS One* **9**: e105506.

Liu, Z., and Liu, J. (2013) Evaluating bacterial community structures in oil collected from the sea surface and sediment in the northern Gulf of Mexico after the Deepwater Horizon oil spill. *MicrobiologyOpen* **2**: 492-504.

Medrano-Santillana, M., Souza-Brito, E.M., Duran, R., Gutierrez-Corona, F., and Reyna-López, G.E. (2017) Bacterial diversity in fumarole environments of the Paricutín volcano, Michoacán (Mexico). *Extremophiles* **21**: 499-511.

Mouser, P.J., Borton, M., Darrah, T.H., Hartsock, A., and Wrighton, K.C. (2016) Hydraulic fracturing offers view of microbial life in the deep terrestrial subsurface. *FEMS Microbiology Ecology* **92**: fiw166.

Park, J.-S., Yabe, S., Shin-Ya, K., Nishiyama, M., and Kuzuyama, T. (2015) New 2-(1′ H-indole-3′-carbonyl)-thiazoles derived from the thermophilic bacterium Thermosporothrix hazakensis SK20-1 T. *The Journal of Antibiotics* **68**: 60-62.

Park, J.S., Kagaya, N., Hashimoto, J., Izumikawa, M., Yabe, S., Shin-Ya, K. et al. (2014) Identification and Biosynthesis of New Acyloins from the Thermophilic Bacterium Thermosporothrix hazakensis SK20-1T. *Chembiochem* **15**: 527-532.

Rieber, H.P. (2000) Monte San Giorgio und Besano, mittlere Trias, Schweiz und Italien. Monte San Giorgio and Besano, Middle Triassic, Switzerland and Italy. *Europäische Fossillagerstätten* **12**: 83-90.

Sander, P.M., and Greenwood, P.H. (1989) The pachypleurosaurids (Reptilia: Nothosauria) from the Middle Triassic of Monte San Giorgio (Switzerland) with the description of a new species. *Philosophical Transactions of the Royal Society of London B, Biological Sciences* **325**: 561-666.

Song, M., Wang, Y., Jiang, L., Peng, K., Wei, Z., Zhang, D. et al. (2019) The complex interactions between novel DEHP-metabolising bacteria and the microbes in agricultural soils. *Science of the Total Environment* **660**: 733-740.

Stockar, R. (2010) Facies, depositional environment, and palaeoecology of the Middle Triassic Cassina beds (Meride Limestone, Monte San Giorgio, Switzerland). *Swiss Journal of Geosciences* **103**: 101-119.

Stribrny, B., Urban, H., and Weber, H. (1988) The Lower Carboniferous black shale formation, a possible source for noble and base metal deposits in the NE Rhenish Massif, Federal Republic of Germany. *Mineralogy and Petrology* **39**: 129-143.

Tan, W.A., and Parales, R.E. (2019) Hydrocarbon degradation by Betaproteobacteria. In *Taxonomy, genomics and ecophysiology of hydrocarbon-degrading microbes*. McGenity, T.J. (ed.). Springer, Cham (Switzerland), pp. 125-141.

Ulrich, N., Kirchner, V., Drucker, R., Wright, J.R., McLimans, C.J., Hazen, T.C. et al. (2018) Response of aquatic bacterial communities to hydraulic fracturing in Northwestern Pennsylvania: A five-year study. *Scientific Reports* **8**: 5683.

Vyssotski, M., Ryan, J., Lagutin, K., Wong, H., Morgan, X., and Stott, M. (2012) A novel fatty acid, 12, 17-dimethyloctadecanoic acid, from the extremophile *Thermogemmatispora sp.* (strain T81). *Lipids* **47**: 601-611.

Wegner, C.-E., and Liesack, W. (2017) Unexpected dominance of elusive Acidobacteria in early industrial soft coal slags. *Frontiers in Microbiology* **8**: 1023.

Wiseschart, A., Mhuantong, W., Tangphatsornruang, S., Chantasingh, D., and Pootanakit, K. (2019) Shotgun metagenomic sequencing from Manao-Pee cave, Thailand, reveals insight into the microbial community structure and its metabolic potential. *BMC Microbiology* **19**: 144.

Xing, B., and Pignatello, J.J. (2005) SORPTION | Organic Chemicals. In *Encyclopedia of Soils in the Environment*. Hillel, D. (ed). Oxford: Elsevier, pp. 537-548.

Yabe, S., Sakai, Y., and Yokota, A. (2016) *Thermosporothrix narukonensis* sp. nov., belonging to the class Ktedonobacteria, isolated from fallen leaves on geothermal soil, and emended description of the genus *Thermosporothrix*. *International Journal of Systematic and Evolutionary Microbiology* **66**: 2152-2157.

Yabe, S., Sakai, Y., Abe, K., and Yokota, A. (2017) Diversity of Ktedonobacteria with actinomycetes-like morphology in terrestrial environments. *Microbes and Environments* **32**: ME16144.

Yabe, S., Aiba, Y., Sakai, Y., Hazaka, M., and Yokota, A. (2010) *Thermosporothrix hazakensis* gen. nov., sp. nov., isolated from compost, description of Thermosporotrichaceae fam. nov. within the class Ktedonobacteria Cavaletti et al. 2007 and emended description of the class Ktedonobacteria. *International Journal of Systematic and Evolutionary Microbiology* **60**: 1794-1801.

Yabe, S., Aiba, Y., Sakai, Y., Hazaka, M., and Yokota, A. (2011) *Thermogemmatispora onikobensis* gen. nov., sp. nov. and *Thermogemmatispora foliorum* sp. nov., isolated from fallen leaves on geothermal soils, and description of Thermogemmatisporaceae fam. nov. and Thermogemmatisporales ord. nov. within the class Ktedonobacteria. *International Journal of Systematic and Evolutionary Microbiology* **61**: 903-910.

Yadav, A.P.S., Dwivedi, V., Kumar, S., Kushwaha, A., Goswami, L., and Reddy, B.S. (2021) Cyanobacterial extracellular polymeric substances for heavy metal removal: A mini review. *Journal of Composites Science* **5**: 1.

Zheng, Y., Wang, C.-M., Sakai, Y., Abe, K., Yokota, A., and Yabe, S. (2019) *Thermogemmatispora aurantia* sp. nov. and *Thermogemmatispora argillosa* sp. nov., within the class Ktedonobacteria, and emended description of the genus Thermogemmatispora. *International Journal of Systematic and Evolutionary Microbiology* **69**: 1744-1750.

Zheng, Y., Wang, C.-m., Sakai, Y., Abe, K., Yokota, A., and Yabe, S. (2020) *Dictyobacter vulcani* sp. nov., belonging to the class Ktedonobacteria, isolated from soil of the Mt Zao volcano. *International Journal of Systematic and Evolutionary Microbiology* **70**: 1805-1813.

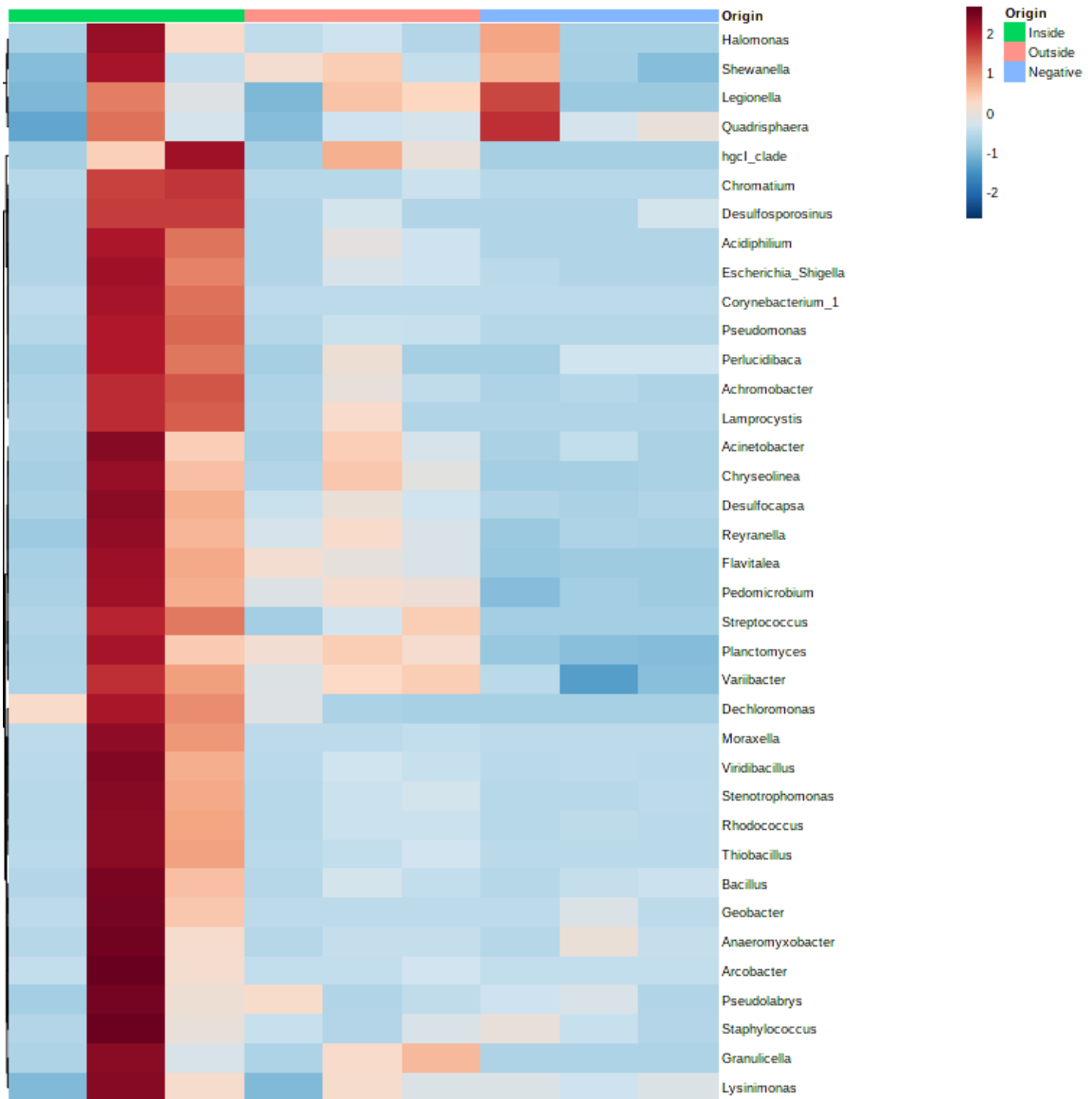**5.8 Supplementary Information**

Figure S1. Relative abundance of genera in the Monte San Giorgio organic-rich shale microbiome.
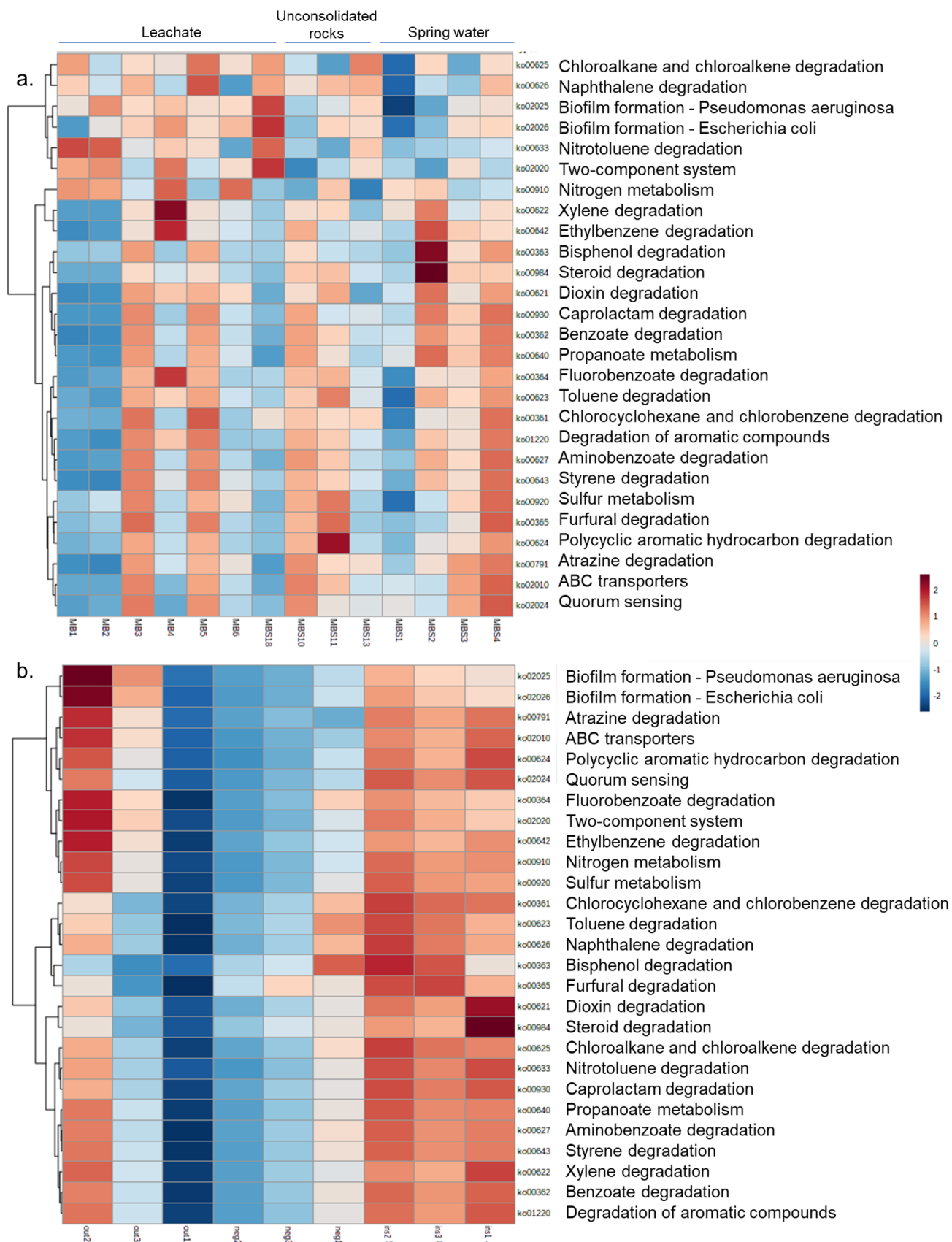
Figure S2. Distribution of aromatic compounds degradation pathway among Marsberg and Monte San Giorgio samples.

Table S1. Aromatic compounds metabolism pathways related to MAGs

MAGs	Aromatic compounds metabolism pathways
002 *Ellin6529*	Phenylacetate degradation I (aerobic), 5-nitroanthranilate degradation, 4-sulfocatechol degradation, Urate conversion to allantoin I, and chlorinated phenol degradation
006 *Chlorolflexia*	Phenylacetate degradation I (aerobic), naphthalene degradation (aerobic), cinnamate, and 3-hydroxycinnamate degradation to 2-hydroxypentadienoate
008 *Dehalococcoidia*	Phenylacetate degradation I (aerobic), naphthalene degradation (aerobic), cinnamate and 3-hydroxycinnamate degradation to 2-hydroxypentadienoate, and catechol degradation to 2-hydroxypentadienoate, 4-hydroxyacetophenone degradation
009 *Acidimicrobiia*	catechol degradation to 2-hydroxypentadienoate, 4-hydroxyacetophenone degradation, cinnamate and 3-hydroxycinnamate degradation to 2-hydroxypentadienoate, and superpathway of phenylethylamine degradation
010 *Candidatus Binatota*	Naphthalene degradation (aerobic), phenylacetate degradation I (aerobic), biphenyl degradation, carbazole degradation, 4-nitrotoluene degradation, 4-nitrophenol degradation, 4-chloronitrobenzene degradation, benzoyl-CoA degradation I (aerobic), 3,5-dichlorocatechol degradation, cinnamate and 3-hydroxycinnamate degradation to 2-hydroxypentadienoate, 4-sulfocatechol degradation, 4-hydroxymandelate degradation, 2-aminophenol degradation, dibenzothiophene desulfurization, diphenyl ethers degradation, protocatechuate degradation II (ortho-cleavage pathway), benzene degradation, anthranilate degradation I (aerobic), 2-chlorobenzoate degradation, and salicylate degradation
011 *Candidatus Binatota*	Naphthalene degradation (aerobic), diphenyl ethers degradation, protocatechuate degradation II (ortho-cleavage pathway), carbazole degradation, cinnamate and 3-hydroxycinnamate degradation to 2-hydroxypentadienoate, benzoyl-CoA degradation I (aerobic), 3,5-dichlorocatechol degradation, 3,5-dichlorocatechol degradation, 3,4,6-trichlorocatechol degradation, anthranilate degradation I (aerobic), 3-chlorocatechol degradation I (*ortho*), 4-chlorocatechol degradation, benzoate degradation II (aerobic and anaerobic), 5-chloro-3-methyl-catechol degradation, 4-chloro-2-methylphenoxyacetate degradation, and 2,4 dichlorophenoxyacetate degradation
015 *Deinococcus-Thermus*	Phenylacetate degradation I (aerobic), carbazole degradation, biphenyl degradation, cinnamate, and 3-hydroxycinnamate degradation to 2-hydroxypentadienoate, and catechol degradation to 2-hydroxypentadienoate.
019 *Ktedonobacteria*	Phenylacetate degradation I (aerobic), carbazole degradation, biphenyl degradation, cinnamate, and 3-hydroxycinnamate degradation to 2-hydroxypentadienoate, benzoate degradation II (aerobic and anaerobic), atrazine degradation I (aerobic), diphenyl ethers degradation, and catechol degradation to 2-hydroxypentadienoate.

Table S2. Differentially abundant genes related to transporters

Monte San Giorgio organic-rich shale (inside)	Marsberg (Leachate group)
K10009 tcyB; L-cystine transport system permease protein	K02009 cbiN; cobalt/nickel transport protein
K08173 ydfJ; MFS transporter, MHS family, metabolite:H+	K02008 cbiQ; cobalt/nickel transport system permease

symporter

K05820 hcaT; MFS transporter, PPP family, 3-phenylpropionic acid transporter
K02768 fruB; fructose PTS system EIIA component [EC:2.7.1.202]
K02770 fruA; fructose PTS system EIIBC or EIIC component [EC:2.7.1.202]
K02769 fruAb; fructose PTS system EIIB component [EC:2.7.1.202]
K02821 ulaC; ascorbate PTS system EIIA or EIIAB component [EC:2.7.1.194]
K02822 ulaB; ascorbate PTS system EIIB component [EC:2.7.1.194]
K12516 bigA; putative surface-exposed virulence protein
K16088 TC.FEV.OM1; outer-membrane receptor for ferric coprogen and ferric-rhodotorulic acid
K11735 gabP; GABA permease
K14052 puuP; putrescine importer
K03758 arcD; arginine:ornithine antiporter / lysine permease
K16263 yjeH; amino acid efflux transporter
K07788 mdtB; multidrug efflux pump
K11688 dctP; C4-dicarboxylate-binding protein DctP
K11689 dctQ; C4-dicarboxylate transporter, DctQ subunit
K11690 dctM; C4-dicarboxylate transporter, DctM subunit
K21395 yiaO; TRAP-type transport system periplasmic protein
K05834 rhtB; homoserine/homoserine lactone efflux protein
K07794 tctB; putative tricarboxylic transport membrane protein
K12536 hasD; ATP-binding cassette, subfamily C, type I secretion system permease/ATPase
K02424 fliY; L-cystine transport system substrate-binding protein
K10022 aotJ; arginine/ornithine transport system substrate-binding protein
K10023 aotM; arginine/ornithine transport system permease protein
K10024 aotQ; arginine/ornithine transport system permease protein
K10025 aotP; arginine/ornithine transport system ATP-binding protein [EC:7.4.2.1]
K08195 pcaK; MFS transporter, AAHS family, 4-hydroxybenzoate transporter
K08368 yaaU; MFS transporter, putative metabolite transport protein
K02798 cmtB; mannitol PTS system EIIA component [EC:2.7.1.197]
K02800 mtlA; mannitol PTS system EIICBA or EIICB component [EC:2.7.1.197]
K16081 algE; alginate production protein
K07789 mdtC; multidrug efflux pump
K07791 dcuA; anaerobic C4-dicarboxylate transporter DcuA
K16345 xanP; xanthine permease XanP
K21393 yiaN; TRAP-type transport system large permease protein
K21394 yiaM; TRAP-type transport system small permease protein
K13929 mdcA; malonate decarboxylase alpha subunit [EC:2.3.1.187]
K13931 mdcC; malonate decarboxylase delta subunit
K02048 cysP; sulfate/thiosulfate transport system substrate-binding protein
K11073 potF; putrescine transport system substrate-binding protein
K10002 gltK; glutamate/aspartate transport system permease protein
K10003 gltJ; glutamate/aspartate transport system permease protein
K10004 gltL; glutamate/aspartate transport system ATP-binding protein [EC:7.4.2.1]

protein

K02007 cbiM; cobalt/nickel transport system permease protein
K02195 ccmC; heme exporter protein C
K02004 ABC.CD.P; putative ABC transport system permease protein
K02429 fucP; MFS transporter, FHS family, L-fucose permease
K02557 motB; chemotaxis protein MotB
K03286 TC.OOP; OmpA-OmpF porin, OOP family
K03304 tehA; tellurite resistance protein
K03313 nhaA; Na+:H+ antiporter, NhaA family
K01551 arsA; arsenite/tail-anchored protein-transporting ATPase [EC:7.3.2.7 7.3.-.-]
K00184 dmsB; dimethyl sulfoxide reductase iron-sulfur subunit
K00185 dmsC; dimethyl sulfoxide reductase membrane subunit
K06160 pvdE; putative pyoverdin transport system ATP-binding/permease protein
K06073 btuC; vitamin B12 transport system permease protein
K07091 lptF; lipopolysaccharide export system permease protein
K07114 yfbK; Ca-activated chloride channel homolog
K03442 mscS; small conductance mechanosensitive channel
K03561 exbB; biopolymer transport protein ExbB
K03559 exbD; biopolymer transport protein ExbD
K03562 tolQ; biopolymer transport protein TolQ
K03560 tolR; biopolymer transport protein TolR
K06076 fadL; long-chain fatty acid transport protein
K05807 bamD; outer membrane protein assembly factor BamD
K07238 TC.ZIP; zinc transporter, ZIP family
K07301 yrbG; cation:H+ antiporter
K03832 tonB; periplasmic protein TonB
K03640 pal; peptidoglycan-associated lipoprotein
K03646 tolA; colicin import membrane protein
K03641 tolB; TolB protein
K07307 dmsB; anaerobic dimethyl sulfoxide reductase subunit B
K04758 feoA; ferrous iron transport protein A
K04759 feoB; ferrous iron transport protein B
K12530 rtxB; ATP-binding cassette, subfamily B, bacterial RtxB
K11603 mntA; manganese transport system ATP-binding protein [EC:7.2.2.5]
K11720 lptG; lipopolysaccharide export system permease protein
K09697 natA; sodium transport system ATP-binding protein [EC:7.2.2.4]
K09808 lolC_E; lipoprotein-releasing system permease protein
K16092 btuB; vitamin B12 transporter
K16089 TC.FEV.OM2; outer membrane receptor for ferrienterochelin and colicins
K16087 TC.FEV.OM3; hemoglobin/transferrin/lactoferrin receptor protein
K12543 lapE; outer membrane protein, adhesin transport system
K18904 nodT; outer membrane protein, multidrug efflux system
K15987 hppA; K(+)-stimulated pyrophosphate-energized sodium pump
K08355 aoxA; arsenite oxidase small subunit [EC:1.20.2.1 1.20.9.1]
K08356 aoxB; arsenite oxidase large subunit

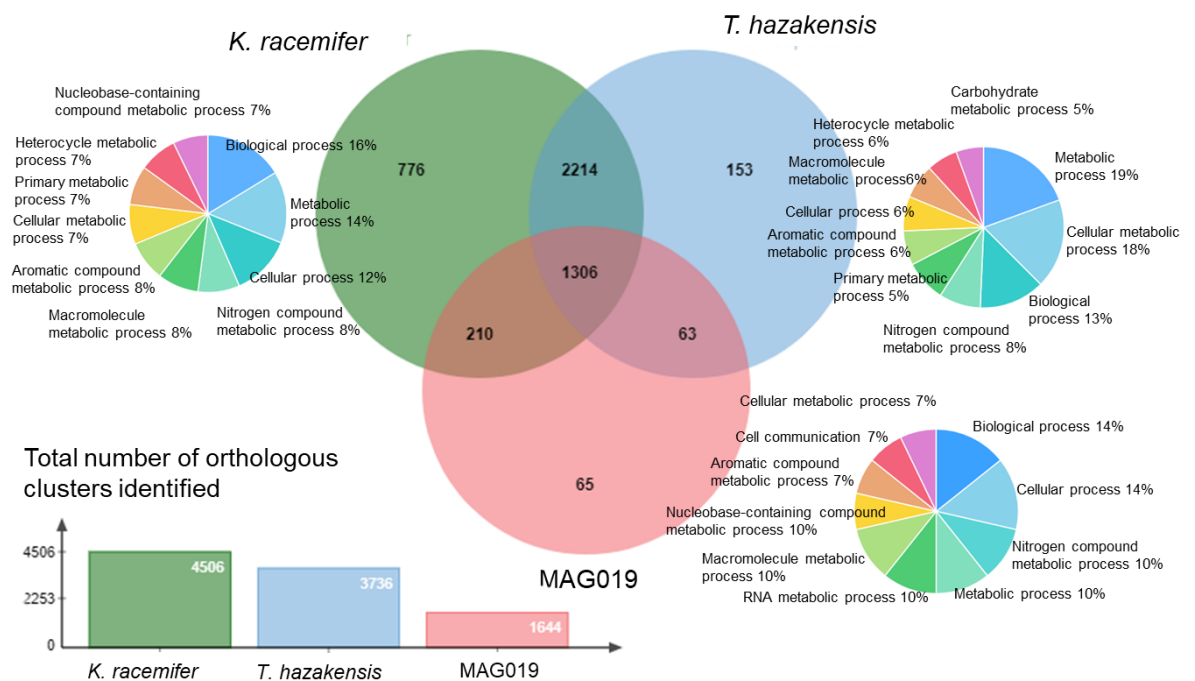|	
K13893 yejA; microcin C transport system substrate-binding protein K13894 yejB; microcin C transport system permease protein K13895 yejE; microcin C transport system permease protein K09817 znuC; zinc transport system ATP-binding protein [EC:7.2.2.20] K03535 gudP; MFS transporter, ACS family, glucarate transporter K05819 mhpT; MFS transporter, AAHS family, 3-hydroxyphenylpropionic acid transporter K05548 benK; MFS transporter, AAHS family, benzoate transport protein K08159 sotB; MFS transporter, DHA1 family, L-arabinose/isopropyl-beta-D-thiogalactopyranoside export protein K03475 ulaA; ascorbate PTS system EIIC component K02024 lamB; maltoporin K18300 oprN; outer membrane protein, multidrug efflux system K20543 bcsC; cellulose synthase operon protein C K11734 aroP; aromatic amino acid transport protein K18299 mexF; multidrug efflux pump K07084 yuiF; putative amino acid transporter K11745 kefC; glutathione-regulated potassium-efflux system ancillary protein K02824 uraA; uracil permease	[EC:1.20.2.1 1.20.9.1] K13888 macA; membrane fusion protein, macrolide-specific efflux system K16554 exoP; polysaccharide biosynthesis transport protein

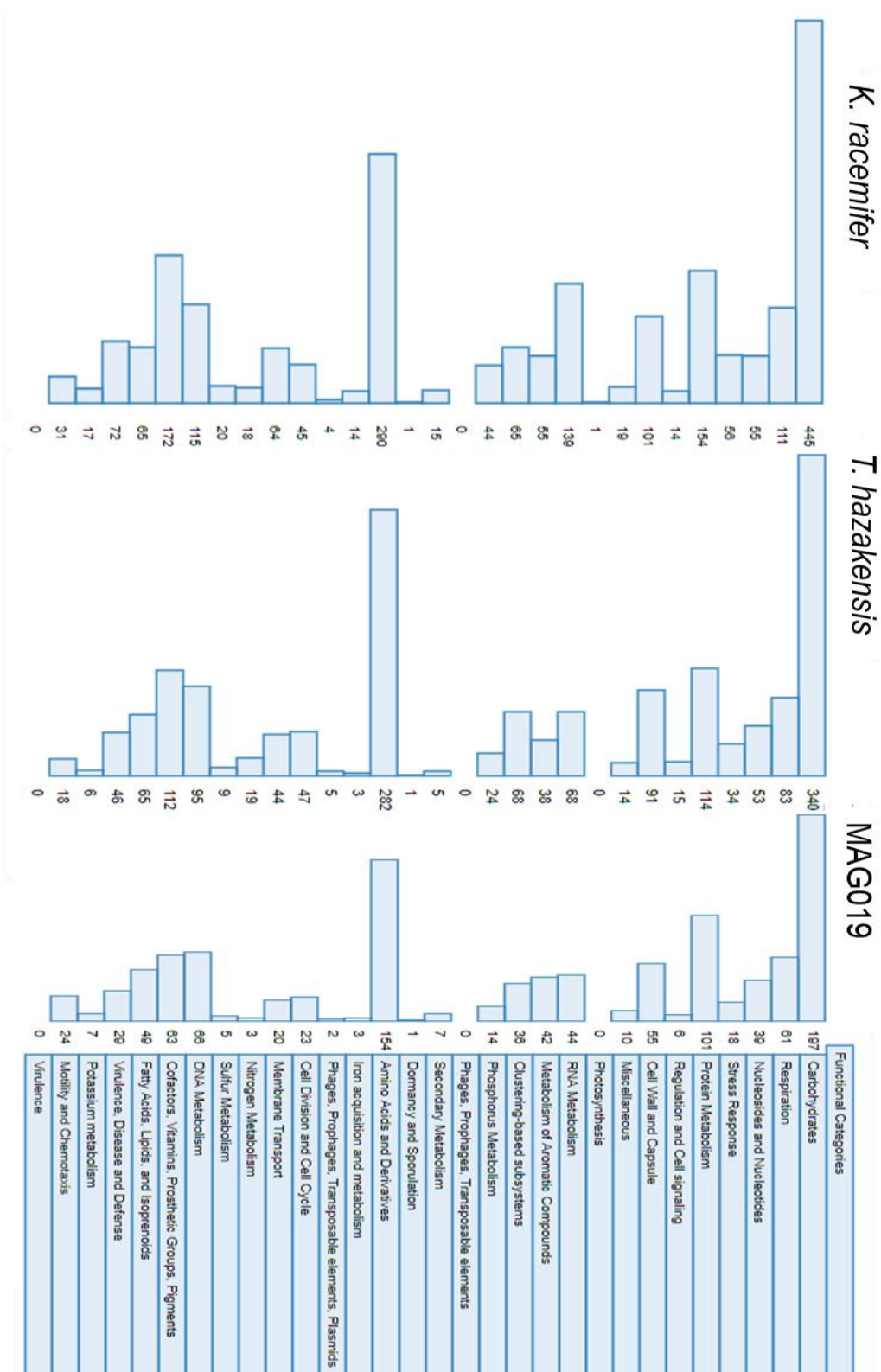**Figure S3. Shared and unique orthologous protein clusters**

Functional Categories	K. racemifer	T. hazakensis	MAG019
Carbohydrates	445	340	197
Respiration	111	83	81
Nucleosides and Nucleotides	55	53	39
Stress Response	58	34	18
Protein Metabolism	154	114	101
Regulation and Cell signaling	14	15	6
Cell Wall and Capsule	101	91	55
Miscellaneous	19	14	10
Photosynthesis	1	0	0
RNA Metabolism	139	68	44
Metabolism of Aromatic Compounds	55	38	42
Clustering-based subsystems	65	68	38
Phosphorous Metabolism	44	24	14
Phages, Prophages, Transposable elements	0	0	0
Secondary Metabolism	15	5	7
Dormancy and Sporulation	1	1	1
Amino Acids and Derivatives	290	282	154
Iron acquisition and metabolism	14	3	3
Phages, Prophages, Transposable elements, Plasmids	4	5	2
Cell Division and Cell Cycle	45	47	23
Membrane Transport	64	44	20
Nitrogen Metabolism	18	19	3
Sulfur Metabolism	20	9	5
DNA Metabolism	115	95	66
Cofactors, Vitamins, Prosthetic Groups, Pigments	172	112	83
Fatty Acids, Lipids, and Isoprenoids	65	65	49
Virulence, Disease and Defense	72	46	29
Potassium metabolism	17	6	7
Motility and Chemotaxis	31	18	24
Virulence	0	0	0

Figure S4. Comparison of genome features annotated with RAST.

Figure S5. SEED based functional profile analysis based on RAST

Appendix

Declaration

I hereby declare that this thesis has been written independently, without the use of other sources and aids than those cited.

Sania Arif

Göttingen, July 23 2021

Acknowledgment

I am very much thankful to my supervisor PD. Dr. Michael Hoppert. He believed in me that I could win the DAAD scholarship when I myself was not sure and gave me the acceptance letter to apply for it. I cannot thank him enough for his constant support and guidance throughout my Ph.D. Next, I would like to extend my thanks to Dr. Heiko Nacke that he performed and guided me through the major whole metagenome analysis. I sincerely appreciate Prof. Rolf Daniel and Prof. Joachim Reitner for their suggestions and contributions as my thesis committee members. Without their constructive remarks, this research work could not be made better. I would also like to thank Dr. Anja Poelhein and Dr. Dominik Schneider from Prof. Dr. Rolf Daniel's group for performing the sequencing and analysis.

The emotional support from my parents and siblings held me strong and focused to execute my Ph.D. successfully. Despite my sisters Hira and Sara were busing with their studies, they were always there to talk to me. Without them, life has no meaning and happiness. I cannot pay them back for their care and sincere intentions. Most importantly, I am grateful to God that He gave me a family in Germany in the shape of my German granny Dr. Christel Eilers who was there during my Ph.D. as strong support and gave me the best advice. We talked like best friends like nothing could stand between us as if we were connected by blood. In the end, I wish to thank my friends Fania and Alaa who became my close friends in the last year of my Ph.D. when I needed it the most. Thank you for encouraging and cooking food for me so that I can finish writing this thesis and bringing laughter and fun moments to my life that I will cherish for the rest of my life.