Exposing the Dark Microbial Biosphere

VALERIE HUBALEK
Abstract

Dark biosphere research has been widely neglected, although by volume this biome comprises the lion’s share of habitats on our planet. In these systems the main metabolic strategies are of chemotrophic nature, leading to gradual depletion of redox gradients essential for sustaining life. Thus these environments are regarded more or less close to chemical equilibrium.

Here, we use sequence data of whole community metagenomes and taxonomic marker approaches to study the ecology of environments close to the thermodynamic limit: deep terrestrial aquifers and aphotic systems impacted by petroleum-derived products. We show that these systems select for individuals with reduced genomes and cell sizes, likely as a mode to save energy. Due to genome reduction, these so called “streamlined” cells are reduced in the number of genes and metabolic pathways. This loss has led to community members sharing the metabolic burden of synthesizing in particular energy costly metabolites, creating tight interdependencies between the community members, as a consequence. In addition, we propose that cells scavenging anabolic products derived from detrital biomass and intermediate fermentation products are equally important in these systems. Hence, life at the thermodynamic limit involves a much more complex biological system than previously shown, that goes beyond traditionally described electron- and intermediate metabolite-transfer dependencies.

This thesis furthermore includes ecological implications, demonstrating how species diversity and community metabolism are shaped by redox gradients and dispersal potential in the deep biosphere and contaminated sediments. This research is also relevant from a practical point of view, as it pinpoints new opportunities for enhanced bioremediation through metabolite additions in order to raise the efficiency of degradation processes.

Valerie Hubalek, Department of Ecology and Genetics, Norbyvägen 18 D, Uppsala University, SE-752 36 Uppsala, Sweden.

© Valerie Hubalek 2015

ISSN 1651-6214
urn:nbn:se:uu:diva-264115 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-264115)
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I  Xiaofen Wu, Karin Holmfeldt, Valerie Hubalek, Daniel Lundin, Mats Åström, Stefan Bertilsson and Mark Dopson. The microbial interactome of the terrestrial deep biosphere reveals metabolic partitioning among populations. ISME Journal, In Press

II Valerie Hubalek, Xiaofen Wu, Alexander Eiler, Moritz Buck, Christine Heim, Mark Dopson, Stefan Bertilsson and Danny Ionescu. Connectivity driven bacterial diversity patterns and functional potential in three deep aquifers of the Fennoscandian shield. Submitted.

III Valerie Hubalek & Moritz Buck, BoonFei Tan, Julia Foght, Annelie Wendeberg, David Berry, Stefan Bertilsson and Alexander Eiler. Metabolic partitioning in an alkane degrading bioreactor operating under methanogenic condition. Manuscript.

“Because the history of evolution is that life escapes all barriers. Life breaks free. Life expands to new territories. Painfully, perhaps even dangerously. But life finds a way.”

*Michael Crichton - Jurassic Park (1990)*

*Dedicated to Darkness*
List of additional Papers and Manuscripts


## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>16Sr RNA</td>
<td>16S Ribosomal Ribonucleic Acid</td>
</tr>
<tr>
<td>Äspö HRL</td>
<td>Äspö Hard Rock Laboratory</td>
</tr>
<tr>
<td>ASS</td>
<td>Alkylsuccinate synthase</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Tool</td>
</tr>
<tr>
<td>BSS</td>
<td>Benzylsuccinate Synthase</td>
</tr>
<tr>
<td>BTEX</td>
<td>Benzene, Toluene, Ethylbenzene and Xylene</td>
</tr>
<tr>
<td>CH₄</td>
<td>Methane</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>Fe</td>
<td>Iron</td>
</tr>
<tr>
<td>H₂ase</td>
<td>Hydrogenase</td>
</tr>
<tr>
<td>HS⁻</td>
<td>Hydrogensulfide</td>
</tr>
<tr>
<td>H₂SO₄</td>
<td>Sulfate</td>
</tr>
<tr>
<td>MAS</td>
<td>Methyl-Alkylsuccinate Synthase</td>
</tr>
<tr>
<td>Mn</td>
<td>Manganese</td>
</tr>
<tr>
<td>NGS</td>
<td>Next Generation Sequencing</td>
</tr>
<tr>
<td>N₂</td>
<td>Nitrogen gas</td>
</tr>
<tr>
<td>NH₄⁺</td>
<td>Ammonium</td>
</tr>
<tr>
<td>NMS</td>
<td>Napthyl-methylsuccinate synthase</td>
</tr>
<tr>
<td>NO₃⁻</td>
<td>Nitrate</td>
</tr>
<tr>
<td>NO₂⁻</td>
<td>Nitrite</td>
</tr>
<tr>
<td>PAH</td>
<td>Polycyclic aromatic hydrocarbons</td>
</tr>
<tr>
<td>SLiMEs</td>
<td>Subsurface Lithoautotrophic Microbial Ecosystems</td>
</tr>
<tr>
<td>SO₃⁻</td>
<td>Sulfite</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1. The redox tower of energy gradients

Figure 2. Alkane degradation and recycling of propionyl-CoA to fumarate

Figure 3. The geo-gas driven deep biosphere

Figure 4. The Äspö Hard Rock Laboratory (HRL) sampling site

Figure 5. The Syncrude Mildred Lake plant

Figure 6. Model of potential metabolic pathways in deep aquifers

Figure 7. Alpha diversity measures in deep aquifers

Figure 8. Model of a syntrophic network and associated secondary partners

Figure 9. The microbial community composition in Lake Grötingen
1 Introduction

1.1 The Unseen World
Microorganisms represent a central component of our planet. Historically they were the first life forms to emerge and quantitatively they are dominant by populating every single habitat on the planet, starting in the gut of the tiniest insect (or on the in- and outside of other organisms) and deep sea vents with boiling temperatures, to oil contaminated environments and rocks many kilometers beneath the surface of the Earth.

They are central to the food web and mediate processes ranging from photosynthesis to decomposition of organic and inorganic matter, and play central roles in all major elemental cycles. Needless to say, all life on Earth quintessentially depends on these microscopic entities as primary source of nutrients, recyclers of dead matter and by their ability to metabolize even toxic compounds.

Studying the microbial world does not only offer us endless potential for advancing technology in the medical or industrial area, but also gives us insight into the beginning of life and its limitations. Limits that have been and will likely also continue to be redefined.

1.2 The Dark Biosphere
The dark biosphere constitutes the largest collection of habitats by volume on our planet, mainly comprising the aphotic zones of all major water bodies, the oceanic sub-seafloor sediments and the continental subsurface.

Although being such a big fraction of the biosphere, the lion’s share of microbial studies conducted thus far has been directed towards light-exposed environments. The boundaries for the truly dark environment which is supposedly independent of sunlight driven production, are difficult to define as the products of photosynthesis are pervasive and will likely reach, at least to some degree, most biotic habitats on our planet. Perhaps the only clear distinction between the photosphere and the dark biosphere is that the prevailing metabolic processes in dark systems are of chemotrophic nature and more or less moving towards the thermodynamic limit. Underlying chemical reactions that sustain life in these environments, are restricted by low substrate concentration, the availability of catalysts, as well as available reduct-
ants and oxidants yielding a relatively low free Gibb’s energy. Thus we consider these systems to move towards the thermodynamic and kinetic limits for sustaining life.

1.2.1 Terrestrial Deep Aquifers

Major technical advances in the last few years have enabled a more extensive exploration of the terrestrial deep subsurface. In contrast to the first studies of sub-seafloor microbiology, this research was initially facilitated by the quest of finding oil or natural gas, minerals and water depositories. Another reason was the design and validation of safe long-term storage facilities for nuclear waste as for the Åspö Hard Rock Laboratory (HRL) in Southern Sweden (more details in following chapter). The aspect of studying resource and contaminant degradation as well as waste container stability from a microbiological point of view were the next major motivation for the initiation of terrestrial deep biosphere research.

The need for drilling or tunnel excavation to reach these environments pose an obvious risk for introducing contamination by adding potential electron-acceptors from the drilling equipment and drilling fluids and by altering the salinity or other chemical components by draining the aquifer. Various strategies to minimize these effects are being applied, for example by adding tracer substances to the drill fluids in order to test if contaminating drill water has intruded the aquifer. Triple tube drilling can to some extent protect aquifer surfaces from the drill water, but the risk for washout of fracture materials and contamination still remains. Without the availability of undisturbed controls, the significance of these potential effects are difficult to evaluate.

However, the finding of 3.5 billion-year-old fossilized biofilm at 450 m depth at the Åspö HRL provides encouraging evidence that microorganisms found in granitic aquifers are not merely contamination artifacts. Furthermore, this could suggest that hard rock aquifers might be one of the oldest habitats for microbial life on the planet. It is plausible that before the atmosphere of the Earth was formed, the subsurface of our early planet sheltered the first life forms from high levels of radiation. This concept inspired NASA’s scientists to look for life on other planets in deep underground settings in their future explorations.

The definition of what is regarded as “deep” and “shallow” is most rationally defined based on the degree of connectivity to the surface rather than actual depth.

The chemical properties of deep aquifers are determined by the origin and mixing of the groundwater, which is essentially the result of the surrounding geology. Groundwater-conducive fractures can range in size from micrometers to crush zones in faults that can hold large quantities of water. Addition-
ally, these systems can vary in the degree of isolation, not only from the surface, but also from other aquifers that can replenish electron-donors and more favorable electron acceptors\(^5\).

Undoubtedly, the abundance and activity of microbes in the deep subsurface are orders of magnitude lower than in surface sediments, as the energy yield of photosynthetic derived processes when compared to chemosynthesis is higher and thus produces more biomass at a faster rate. Despite the slow growth, microbial biomass from the subseafloor has been estimated to \(2.9 \times 10^{29}\) cells, corresponding to 4.1 petagram (Pg) of carbon, which is \(~0.6\%\) of Earth’s total living biomass\(^9\).

It has been proposed that microorganisms residing at these great depths use hydrogen and methane of primarily abiotic and to a lesser extent biogenic origin as energy sources to reduce ubiquitous carbon dioxide for autotrophic growth. Several studies have documented the general presence of hydrogen and methane in the Fennoscandian Shield, strengthening the hypothesis of a “geo-gas” driven deep biosphere\(^{10-12}\). These microorganisms have been termed subsurface lithoautotrophic microbial ecosystems (SLiMEs) in the 1990s\(^{13}\).

A very recent finding and likely another important factor shaping the SLiMEs, which is also a strong indicator for their activity and growth, is the relatively high abundance and diversity of viruses, which were found in the Äspö HRL groundwater\(^{14}\). These viruses, aside from controlling the numbers and activity of their prey, might cause a release of biomolecules from induced lysis of cells. This might provide an essential and highly available carbon and nutrient source for the rest of the community. Moreover, phages are known to be important vehicles for gene transfer in other ecosystems and may well play a similar role in the deep biosphere\(^{14}\).

### 1.2.2 Petroleum Environments

Petroleum is organic matter (biomass) that once got buried and embedded in the Earth’s crust. These organic carbon pools are thus also part of the deep biosphere. Protected from the biotic carbon cycle of the surface, detritus was subsequently transformed by long-term geological processes to its present form\(^{15}\). This diverse mixture of compounds may reach the surface as a result of geological events (e.g. earthquakes) or anthropogenic activities (e.g. drilling). Accordingly they are among the world’s most problematic and widespread contaminants. Global estimates suggest that about 47% of the crude oil that enters the marine environment is from natural seeps, whereas 53% results from leaks and spills during the extraction, transportation, refining, storage, and utilization of petroleum\(^{16}\). Given the impact that recent oil spill accidents have on the environment combined with their carcinogenic and neurotoxic potential for humans, research on the distribution, degradation, and persistence of these compounds is of utmost importance.
Petroleum is one of the most complex mixtures of organic compounds with more than 17,000 distinct chemical components. Crude oils feature unique and variable chemical and physical characteristics, making it difficult to predict their specific environmental fate. In broad terms, there are four classes of constituents: the saturates, the aromatics, the asphaltenes (phenols, fatty acids, ketones, esters, and porphyrins), and the resins (pyridines, quinolines, carbazoles, sulfoxides, and amidess). Resin and asphaltene fractions contain trace amounts of nitrogen, sulfur and/or oxygen in addition to the bulk elements carbon and hydrogen. These compounds in turn often form complexes with heavy metals.

The good news is that the highly reduced hydrocarbons in petroleum represent a rich source of energy for some microbes. Therefore, petroleum constituents are biodegradable when suitable electron acceptors are present. Nevertheless, and this is the bad news, aromatic hydrocarbons and polar fractions, which are very toxic, represent the most persistent fraction because of high molecular stability and physical properties making them less susceptible to enzyme attack. This can lead to decade- or century-long residence times in the environment with far-reaching consequences for human health and the environment.

Monoaromatic hydrocarbons referred to as BTEX (benzene, toluene, ethylbenzene and xylene) are highly volatile and soluble substances commonly found in gasoline. Polycyclic aromatic hydrocarbons (PAHs) are compounds built on fused benzene rings and the most difficult to biodegrade. The range of environmental significance for biodegradation lies between naphthalene (C_{10}H_{8}) and coronene (C_{24}H_{12}). Anthropogenic sources include combustion of fossil fuels and waste or other industrial activities and is known to significantly affect coastal and inland surface waters as a result of their ability to be transported over long distances as gases or aerosols.

The first steps of hydrocarbon biodegradation result in the loss of the saturated compounds, while aromatic hydrocarbons and the polar fractions are more resistant to biodegradation. Saturated hydrocarbons constitute the largest fraction of crude oil by mass, thus the biodegradation of saturated hydrocarbons is quantitatively the most important process in the removal of crude oil from the environment. Overall, the susceptibility of hydrocarbons to microbial degradation is ranked as follows: linear alkanes > branched alkanes > small aromatics > cyclic alkanes.

Aside from hydrocarbons, petroleum reservoirs contain a wide range of other organic substrates, that can serve as carbon and energy sources for microorganisms. They include fatty acids and especially its main synthesis and degradation intermediate acetate, formate, propionate and butyrate, and more complex organic acids such as napthenic acids with concentrations in crude oil up to 100 mM. All these substrates may be oxidized under anaerobiosis in the presence of terminal electron acceptors possibly present in oilfield waters (e.g. sulfate).
1.3 Microbial Metabolism in the Darkness

The base of any food web starts with primary producers, autotrophs ("self-feeding", from the Greek autos "self" and trophe "nourishing") which per definition can either use light or inorganic substances as energy sources and a reductant (usually water or just hydrogen) to reduce CO$_2$ from the surroundings in order to build up complex carbohydrates (such as glucose). These compounds are then being further oxidized to release energy for synthesis of proteins, lipids and nucleic acids that also derive from primary production. Primary producers are then being consumed by heterotrophs, subsequently distributing these organic molecules across the food web.

Organisms that are cut-off from both photons and oxygen, such as organisms residing in the bottom layers of open water systems and deep aquifers, will have to rely on electron acceptors with lower thermodynamic yield than oxygen. In a sequential order, according to the free energy yielded by their reduction, that will be NO$_3^-$ which can be reduced to N$_2$ or NH$_4^+$, Mn(IV) to Mn(II), Fe(III) to Fe(II), SO$_4^{2-}$ to HS$^-$ or S$^2-$ and at the end of the spectrum CO$_2$ being reduced to CH$_4$ (Fig. 1) $^{22}$. Thus, in terms of reactions, these are ordered in the following manner: Denitrification, Mn(IV) or Fe(III) reduction, sulfate reduction and finally methanogenesis. The further down in the redox cascade, the smaller the energy yield of the respective reaction, which typically results in a slower growth and low cell division rates.

![Figure 1. The redox tower showing the spectrum of energy gradients. EA= Electron acceptor, EY= energy yield, RS= reduced species form of the electron acceptor.](image)

1.3.1 Anaerobic metabolic cooperation: Syntrophy

Syntrophy is regarded as a version of symbiosis that is beneficial for one or both of the interacting organisms, but is restricted to metabolic interactions$^{23}$. In the context of anoxic carbon cycling, the central syntrophic relationship consists of methanogens and fermenting organisms. The syntrophic fermentor provides acetate, hydrogen and carbon dioxide, while the meth-
anogenic partner is responsible for efficient removal of hydrogen and formate, which would otherwise hinder fermentation because of thermodynamic constraints. As methane production is energetically the least favorable process, it requires efficient metabolic coordination with its partners. Many syntrophs are known to rely on the capacity to perform reverse electron transport-driven energy-conserving H₂ production through electron-confurcating hydrogenase (H₂ase) in combination with reduced ferredoxin.

1.3.2 Petroleum hydrocarbons as food and energy source

Petroleum hydrocarbons, being highly reduced organic molecules, can be a good substrate for microbial growth under the right conditions. With oxygen being available, most hydrocarbons are susceptible to enzymatic attack using monoxygenases (for alipathic and some aromatic hydrocarbons) and dioxygenases (for aromatic hydrocarbons), in processes where O₂ acts both as an electron acceptor and co-substrate.

Under anaerobic conditions, the breakdown of hydrocarbons can be accomplished by anaerobic respiration with nitrate, ferric iron and sulfate. Under methanogenic conditions, hydrocarbon degradation is thermodynamically feasible only in syntrophic association with hydrogen-consuming microorganisms, such as methanogens or sulfate reducers. To make the compound chemically reactive, several activation reactions have been proposed: carboxylation, hydroxylation and radicalization by fumarate addition. The latter is the only proven mechanism so far, in which a family of glycol radical enzymes are involved in the activation of structurally different hydrocarbons: Alkylsuccinate synthase (ASS) mainly found under sulfate- and nitrate-reducing conditions, benzylsuccinate synthase (BSS) reported under methanogenic, sulfate-, nitrate- and iron-reducing conditions and methylalkylsuccinate synthase (MAS) reported under nitrate-reducing conditions, in addition to the non-homolog naphthyl-methylsuccinate synthase (NMS).

The primary substrates of ASS are n-alkanes, while BSS binds to toluene and NMS on 2-methynaphthalene. These enzymes add the respective substrate to the double bond of a fumarate molecule, forming a succinylated product. The following reduction to acetyl-CoA is analogous for aromatic hydrocarbons after ring-cleavage (Fig. 2). After alkane activation with fumarate, the carbon skeleton is rearranged in order to allow CO₂ demerging via decarboxylation during which a free coenzyme A is being attached (likely produced by methanogens). After beta-oxidation the resulting acetyl-CoA is demerged from the chain resulting in the removal of two carbons. Beta-oxidation is then repeated until the endproducts are one acetyl-CoA or propionyl-CoA. The recycling of propionyl-CoA to fumarate is further demonstrated in Fig. 2.
Acetyl-CoA can further feed into acetotrophic methanogenesis, and the hydrogen that is produced in parallel, can be incorporated by hydrogenotrophic methanogens (see section 1.3.4).

![Diagram of alkane activation and formation to acetyl-CoA or acyl-CoA (steps 1-9) and subsequent recycling of propionyl-CoA to fumarate (11-15). Red indicating newly formed bonds, grey signifying dissociating molecules and coloured boxes are for products that feed into other reactions or pathways.](image)

**Figure 2.** Alkane activation and formation to acetyl-CoA or acyl-CoA (steps 1-9) and subsequent recycling of propionyl-CoA to fumarate (11-15). Red indicating newly formed bonds, grey signifying dissociating molecules and coloured boxes are for products that feed into other reactions or pathways.

1.3.3 Energy- and nutrient cycling in the deep biosphere

The deep subsurface is mainly anaerobic and the supply of organic carbon is scarce. Hence it has been proposed that “geo-gases” of abiotic origin will be of higher importance. Highly abundant hydrogen and methane would in this case function as electron donors and carbon dioxide as an oxidizing complement. Part of the methane can be of biogenic origin, whereas H₂ can be formed abiotically mainly through radiolysis of water and to a minor extent through rock transformation. Migration of methane through faults and fracture zones in the Fennoscandian Shield to the surface has been reported, showing that there can be a flow of gases from the deep that travels...
upwards. This provides strong evidence that the electron donors H₂ and CH₄ are broadly present in the deep subsurface.

Thus, with water available and under moderate temperature conditions, microbes can in theory survive independent of light-driven processes (Fig. 3)³⁴. Methanogens would use H₂ to reduce CO₂ to CH₄, and acetogenic bacteria would also require these two components to produce acetate. This acetate can then be metabolized by acetoclastic methanogens, iron and sulfate reducing bacteria and heterotrophs. This would position them as primary producers in these systems³⁴.

![Diagram](image)

**Figure 3.** The geo-gas driven deep biosphere. Modified after Karsten Pedersen *et al.*³⁴

Under certain circumstances, it has been hypothesized that this process can be reversed. If a shortage in hydrogen is being created, for example when methanogens co-occur with more fast-growing chemoautotrophs, such as sulfate reducers, the opposite reaction would become more favorable and methane or acetate would be stepwise oxidized anaerobically to CO₂³⁵. Not only energy and carbon are required for microbial growth, but also other inorganic nutrients and cofactors. For the build-up of proteins and nucleic acids, phosphorus is available in minerals like apatite, while nitrogen predominates most ground waters and is bioavailable via nitrogen fixations³⁴.
1.3.4 The start- and end products of the carbon cycle: Methane and Carbon Dioxide

The greenhouse gases methane and carbon dioxide are of global importance, as they represent the end products of organic matter decay and are of significance in global warming\(^36\). In the scope of the here presented work, they are especially relevant as these are the central components for chemolithoautotrophic primary production. Methanogenic archaea can synthesize biomass by using CO\(_2\) and hydrogen as carbon and energy sources. In anaerobic hydrocarbon degradation networks, methanogens are obligate partners for the syntrophic primary degraders, by removing hydrogen that would otherwise hamper this fermentative process. Biotically, methane is produced enzymatically by methanogenic archaea from a variety of substrates. These are the three pathways simplified to the chemical reaction\(^37-39\):

(i) *Hydrogenotrophic methanogenesis*

\[
\text{CO}_2 + 4 \text{H}_2 \rightarrow \text{CH}_4 + 2 \text{H}_2\text{O}
\]
Carbon dioxide is being reduced to methane using hydrogen as the electron donor.

\[
4 \text{HCOOH} \rightarrow 3 \text{CO}_2 + \text{CH}_4 + 2 \text{H}_2\text{O}
\]
Hydrogenotrophs can also utilize formate or secondary alcohols, which substitute for hydrogen.

(ii) *Acetoclastic methanogenesis from acetate*

\[
\text{CH}_3\text{COO}^- + \text{H}^+ \rightarrow \text{CO}_2 + \text{CH}_4
\]
This reaction occurs in combination with acetate fermentation, in which the acetate is simultaneously oxidized and reduced (disproportionation reaction) in order to form carbon dioxide and methane.

(iii) *Methylotrophic methanogenesis*

\[
4 \text{CH}_3\text{OH} \rightarrow 3 \text{CH}_4 + 1 \text{CO}_2 + 2 \text{H}_2\text{O}
\]
Similarly to the acetoclastic pathway, methylated compounds such as methanol, methylamines and methyIsulfides undergo a disproportionation reaction in order to produce methane and carbon dioxide.

Methane oxidation is the reverse reaction in which methane is anaerobically oxidized to carbon dioxide. This still rather enigmatic process has predominantly been previously described for anoxic marine and freshwater systems. Analogous to methanogenesis, this involves a syntrophic partnership\(^40\) using the terminal electron acceptors sulfate, nitrite or metal oxides such as manganese or iron\(^41\).
1.4 Metagenomics and the Uncultured World

Microbes organized in communities are strongly interconnected to the surrounding environment and other organisms. In order to understand this complex interplay, cultivation based techniques are not sufficient, as it is well established that laboratory settings fail miserably in reproducing the myriads of factors that microorganisms are subjected to\textsuperscript{42}. A consequence of this is that only a fraction of the microbes that are found in the natural sample can be cultured, a phenomenon termed “the great plate count anomaly”\textsuperscript{43}.

In the mid 1980s microbiologists came to the realization that they had to drop the concept that only culturable microorganisms exist. At this point it was realized that a more systemic approach was needed to understand the microbial world\textsuperscript{32}. In 1985, Norman Pace was the first to announce the idea of cloning marker genes directly from the environment\textsuperscript{44}, but it was not until the 1990s that the first study on cloning the universal taxonomical marker gene of the 16S rRNA derived directly from marine samples was published\textsuperscript{45}. DNA taken directly from the environment without prior cultivation of the organisms has been termed the “metagenome” and the study thereof “metagenomics”\textsuperscript{46}.

The development of “next generation sequencing” (NGS) and the drastic reduction of sequencing costs, enabled scientists to go beyond the question of “who is there” when sequencing phylogenetic marker genes, to “what can they do” by collecting the total genomic information from a given environment. By applying large scale sequencing, it became possible to correlate genomic content with environmental conditions, to disentangle the relationship between symbiosis partners and to dig deeper into the diversity of a gene families\textsuperscript{47}, just to name a few possibilities.

By sequencing single cells or pure cultures, the extracted DNA will represent the genomic content of a clonal population of cells (providing there is no contamination). This will not be the case with a metagenome, where cells with distanced evolutionary history will be sequenced simultaneously. However, gaining complete coverage of all organisms within a complex sample is in most cases unattainable, and this derives both from incomplete DNA recovery as well as the difficulty in assembling reads from different but closely related organisms which create interspecies chimeras\textsuperscript{47}.

The identification of genes is mostly based on annotations to known homologs in databases. This approach is more reliable if a closely related organism has been sequenced and the function of the encoding genes has been experimentally verified. BLAST (Basic Local Alignment Search Tool)\textsuperscript{48,49} is commonly used to identify gene family members within a metagenome. This approach can be combined with \textit{ab initio} gene prediction tools which are mostly based on supervised learning and statistical pattern recognition methods, using Markov or hidden Markov models\textsuperscript{47,50}. 
In order to assign functions found in the metagenome to the respective organism, short sequence reads are first assembled to longer fragments (contigs) and subsequently “binned”. One strategy to accomplish this is either by grouping contigs according to their coverage, GC content, called compositional binning, or phylogenetic binning based on homology searches to reference sequences\textsuperscript{47}. 
2 The Present Thesis

2.1 Aims

The different chapters in this thesis compare and contrast the effects of distinct environmental gradients (level of connectivity to the surface; hydrocarbon contamination gradient) on microbial communities living in energy-limited systems. Using metagenomic approaches, patterns of community structure, linkages and dependencies could be revealed.

The following main objectives were addressed:

• Chapter I: Reconstructing microbial food webs on the metagenome level: identifying co-dependencies and partnerships at different degrees of isolation from the surface. What replaces the sun as a main energy source in deep aquifers?

• Chapter II: The effects of varying connectivity levels to the sunlit surface on microbial diversity and community structure. What happens to a microbial community in different water masses over six years?

• Chapter III: Assigning identity to specific metabolic roles of primary alkane degraders and their syntrophic partners beyond transfer of electrons and intermediate fermentation products. What interactions can we predict among community members in hydrocarbon degradation under methanogenic conditions?

• Chapter IV: Is there a higher remediation capacity in heavily contaminated sites when compared to more pristine locations? Why do aromatic pollutants resist degradation on biologically active sediments?
2.2 Methods

2.2.1 Study Sites

Äspö
The Äspö Hard Rock Laboratory (Äspö HRL), operated by the Swedish Nuclear Fuel and Waste Management Co. (SKB), is an underground research facility located in the southeast of Sweden (Fig. 4). It consists of a 3,600 meter long tunnel going down to a depth of 450 m under the Äspö island.

The geological formation of this area is referred to as the Fennoscandian Shield, which is dated to be between 1.6 – 3.1 billion years old. The bedrock mainly consists of granite and quartz-monzodiorite, which has been fractured by glaciations leading to the formation of various aquifers, which have been studied extensively in terms of geology, hydrology, and chemistry. Altogether six water bodies from depths between 170 to 450 m, differing in water type and age were sampled for the different projects of this thesis.

Alberta Oil Sand Tailings
The Athabasca oil sands (also called the Athabasca tar sands or Alberta tar sands) are large deposits of bitumen or extremely heavy crude oil, located in northeastern Alberta, Canada.
Oil sands tailings are byproducts of bitumen extraction from surface mining and processing of oil sands ores, consisting of a mixture of slightly alkaline water, sand, clay and residual hydrocarbons. Oil sands operations in northern Alberta, Canada produce \( \sim 1.3 \) million barrels of bitumen and \( \sim 262,000 \) m\(^3\) of tailings per day that are deposited into settling basins (tailings ponds). Tailings accumulate since the producing companies operate under a zero discharge policy.

![Figure 5. The Syncrude Mildred Lake plant.](https://commons.wikimedia.org/wiki/File:Syncrude_mildred_lake_plant.jpg#/media/File:Syncrude_mildred_lake_plant.jpg) (Licensed under Public Domain via Commons)

Challenges associated with tailings ponds include the presence of inorganic and organic contaminants such as metals, salts, petroleum hydrocarbons and the emission of the greenhouse gases \( \text{CH}_4 \) and \( \text{CO}_2 \). The fine tailings settle by gravity most rapidly during the first 3-4 years after deposition, forming mature fine tailings (MFT) which then require decades for significant incremental settling. The samples for this study were taken from the Mildred Lake Settling Basin (Fig.5).

**Lake Grötingen, Jämtland**

Lake Grötingen (62°51’ N, 15°29’ E) is located in Grötingen, along the river Gimän in central Sweden, and part of the Natura 2000 area of Gimän. During industrial coal and tar production between 1890-1930, the sewage from industrial processes was directly discharged into the lake. Consequently, a large area of the lake bottom close to the former factory is still covered by high amounts of tar residue. Three sampling sites were chosen for this
study. The first located at the discharge of the factory, the second 2.5 km upstream and the third 860 m downstream.

2.2.2 Analytical Methods

Sampling
For chapter I and II, ground waters from six separate boreholes differing in water age and depth (171-455 m) were sampled from the Åspö Hard Rock Laboratory, in the South of Sweden.

The study in chapter III was on short-chain alkane degrading cultures (SCADC) isolated from mature fine tailings from the Mildred Lake Settling Basin in Alberta, Canada.

For chapter IV sediment cores were taken from the tar-contaminated lake Grötingen located in Jämtland, Central Sweden.

Molecular analysis
The foundation for all studies in this thesis is in depth analysis of recovered environmental DNA sequences that were interpreted within the respective environmental context. Each dataset included chemical analyses of the corresponding habitat to examine genetic adaptation to the prevailing conditions.

Single cell picking & whole genome amplification.
To study populations of filamentous Archaea from an alkane degrading consortium (chapter III), single cells were picked under the microscope using a micromanipulator (Eppendorf TransferMan® NK 2). This mechanical device enables the uptake and transfer of single cells with a microcapillary and hence enable recovery of a simplified sub-community from initially complex samples.

In order to attain a sufficient amount for DNA sequencing, the filaments were subjected to genome amplification, termed multiple displacement amplification (MDA)\(^1\). Random primers bind to denatured DNA followed by strand displacement synthesis of the Phi 29 polymerase. Each displaced strand serves additionally as a template, generating high yields of amplified DNA. The phage derived Phi 29 polymerase, furthermore exhibits proof-reading activity that delivers up to 1000-fold higher fidelity compared to the Taq DNA polymerase\(^2\).

Barcoded amplicon sequencing.
Partial 16S rRNA gene sequences were obtained for the studies in chapter II and III, by using a two-step PCR procedure with primers targeting the V3 and V4 regions of the 16S rRNA gene. The first amplification was to acquire more material while minimizing amplification bias and amplicons from this step were subsequently used as template for a second PCR with identical
primers, except that both forward and reverse primers included 7 bp DNA barcodes that were unique for each amplified sample. The obtained 16SrRNA amplicons were submitted to the SciLifeLab SNP/SEQ sequencing facility (Uppsala University) where Illumina MiSeq technology was used. An in-house sequence analysis and annotation pipeline was then applied on the obtained 2 x 300 bp reads, containing following steps: barcode demultiplexing, read pairing, quality control, clustering and taxonomic assignment against the SilvaMod database.

Whole genome shotgun sequencing.
DNA from the studies conducted in chapters I-IV were submitted to the SciLifeLab SNP/SEQ sequencing facility (Uppsala University) where metagenome library construction and Illumina HiSeq sequencing was performed. In short, genomic DNA was sheared using a focused-ultrasonicator, followed by preparation of sequencing libraries generated for clustered flowcell sequencing, which was performed on the Illumina HiSeq sequencer operated in 2 x 100 cycle mode.

Sequence analysis involved trimming, assembly, binning (only for chapters I & III) and gene prediction/annotation. For assembly, MEGAHIT (in chapter II and III) and RAY (chapter I) were used. Contigs were scaffolded with NEWBLET (chapter I) or the BESST software (chapter III). Coverage was computed by mapping back the reads to the scaffolds using bowtie2, and for computing coverage, bedtools was used. After filtering out short contigs and scaffolds, Concoct was used for coverage and nucleotide composition-based binning. The criteria for high quality bins, so called metagenome-assembled genomes (MAGs), are based on completeness and contamination. An additional criteria used in chapter III was an overlap by more than 95% in contig content as defined by concoct and taxonomic binning. We used PHYLOPYTHIA+S, which uses a marker gene approach to distribute contigs and scaffolds into taxonomical clades.

2.3 Results

2.3.1 Exploring the Deep Terrestrial Biosphere
The finding that life is possible independent of light energy, resulted in a paradigm shift in the 1980s. In fact, it has been estimated that the terrestrial deep biosphere hosts up to 20% of Earth’s biomass. Difficulties in accessing these deep terrestrial systems still make them one of the least understood ecosystems on earth.
It has been hypothesized that these ecosystems are sustained by more or less omnipresent H₂ and CO₂ and to some extent CH₄. All three are viewed as geo-gases and can be derived from their geological surroundings.

Deprived of high yielding energy sources, we expect to find microbial populations with extremely long generation times stretching from hundred to thousands of years.

**Paper I. The microbial interactome of the terrestrial deep biosphere reveals metabolic partitioning among populations.**

**Aim:** To infer the metabolic potential and traits of dominant community members in three deep aquifers with contrasting connectivity to the surface. Here, we hypothesize that with decreasing connectivity and according energy limitation, cells will be selected to become streamlined.

**Study:** Three aquifers located at depths between 171- 450 m were sampled for two different cell fractions, one for cells bigger than 0.22 μm and one for cells and viruses smaller than 0.22 μm. Apart from viruses, the smaller size
fraction also contained an abundance of reads assigned to bacteria, therefore it was included in the study.

The recovered Illumina HiSeq sequences allowed the metabolic reconstruction of 69 distinct microbial genomes, each with > 86% coverage.

Based on metabolic key characteristics, the most abundant populations from each water mass were classified into broad functional groups (I-VI). Traits considered were e.g. fermentation pathways, respiration pathways, C and N fixation. Significant metabolic pathways represented by these functional groups were mapped for each aquifer and a conceptual model of putative metabolic interactions and dependencies were combined into an interactome (Fig. 6).

In the small cell fraction, a phylogenetically distinct subset with reduced estimated genome size was observed in all three aquifers. This is analogous to other studies carried out in oligotrophic environments, showing that in order to save energy for replicating unnecessary genes, they are lost over time. A smaller cell size is another adaptation to higher efficiency, by minimizing nutrient needs and increasing the volume to surface ratio and hence also the capacity to transport solutes.

With longer isolation from the surface, the communities are depleted in organic matter that otherwise was washed in by recharge water. This depletion lead to a tendency to depend more on chemosynthesis, which was reflected by the observation of a more heterotrophic or mixotrophic lifestyle for the upper aquifers, while the deepest and most disconnected site was dominated by chemolithoautotrophic processes, especially sulfate reduction.

In general, this lifestyle generated by isolation from the surface, created tight interdependencies between the community members.

**Paper II. Connectivity driven bacterial diversity patterns and functional potential in three deep aquifers of the Fennoscandian shield.**

**Aim:** To characterize the general effects of isolation on the community dynamics and diversity patterns of three deep aquifers over a time period of six years. Here, we speculate that with increased isolation (disconnection to the sunlit surface) chemolithoautotrophic processes will become more important in the community and resulting slow growth in combination with long (or strong) selection has lead to species extinction. This species extinction driven by dispersal and energy limitation should be reflected by a decrease in overall richness with depth.

**Study:** In this study, microbial communities of three deep terrestrial subsurface aquifers of depths between 180-455 m were investigated by sequencing of 16S rRNA amplicons and shotgun metagenomes. The amplicon dataset revealed the phylogenetic structure, diversity and community dynamics over
six years. Two samples of the deeper aquifers resulted in the successful recovery of whole shotgun metagenomes that were used to detect homologs to key genes involved in nitrogen and carbon fixation, sulfate reduction, sulfide oxidation and fermentation.

Figure 7. Bacterial community alpha diversity measures of the three aquifers, showing A) permuted beta-dispersion to test on homogeneity of the samples within the water masses; B) chao1 alpha diversity estimation; C) Simpson diversity index (1/D); and D) Simpson evenness (E1/D).

The community at the shallow site, that has in-flow of Baltic Sea water every couple of weeks, showed a highly dynamic population, dominated by putatively sulfur oxidizing *Sulfurovum* and *Sulfurimonas* genera at all time points. The intermediate aquifer (with a water turnover of approximately five years) had a less variable and dynamic microbial community and was strongly dominated by candidate phylum OD1. The deepest water mass (5000 years old saline glacial melt and meteoric waters) had the lowest taxon-
richness and surprisingly contained *Cyanobacteria*, that likely were trapped since the last water in-flow thousands of years ago.

In this study, the degree of connectivity to the surface seems to be the main factor in shaping community dynamics and microbial diversity.

In accordance to paper I, this study showed that the occurrence of key genes for nitrogen and carbon fixation, sulfate reduction, sulfide oxidation and fermentation, point towards a gradient in the ratio between chemoorganohetero- and chemolithautotrophic processes in deep aquifers that is related to connectivity to the sunlit surface.

We furthermore observed that species richness (Fig. 7 B) declines with isolation of the environment, which is due to a decreasing number of ecological niches from depleting resources and gradients. Another reason is that immigration events that would bring new organisms, are more rare.

In contrast, the evenness (Fig. 7 D) increases with depth. We speculate that, as species invasions become extremely rare, interactions between remaining community members become tighter and more indispensable under such a condition. Moreover, preceding succession events could have already taken place through competition and/or predation leading to an overall lower richness and to higher evenness, by eliminating dominant species.

2.3.2 Studies on Petroleum impacted environments

Oil derived hydrocarbons are some of the world’s most widespread contaminants and therefore prioritized in environmental management. Environmental-friendly decontamination by microorganisms has in many cases been demonstrated to be the last resort for removal of these toxic compounds. Especially biodegradation under anaerobic conditions is a most challenging task and studying microbial biodegradation of pollutants is important for several reasons:
1. Finding ways to catalyze and improve bioremediation.
2. The adverse or beneficial effects of microbial activity on hydrocarbon profiles for the oil industry.
3. The production of the greenhouse gases methane and carbon dioxide at the end of hydrocarbon degradation. This has an implication from an environmental point of view, but also for methane as a commercial energy resource.

**Paper III. Metabolic partitioning in an alkane degrading bioreactor operating under methanogenic condition.**

**Aim:** To reconstruct the interdependencies within syntrophic microbial assemblages beyond the first level of electron and intermediate metabolite transfer. Here, we hypothesize that enhanced cross-feeding is needed for metabolic processes to be optimized for reduced individual metabolic burden when living in conditions of shortage of bioavailable energy.
**Study:** In order to study the underlining processes of alkane degradation, a microbial community indigenous to oil sands tailing ponds in Alberta, Canada, was isolated. These cultures were grown under strictly methanogenic conditions on short-chain alkanes in bioreactors. The dataset included the full sample of this mixed culture, as well as Archaeal-Bacterial consortia that were isolated using micromanipulation. Combined with metagenomic approaches, we obtained sequence data of these synergistic consortia.

Figure 8. A proposed model for the syntrophic network and associated secondary partners.

Three types of syntrophic consortia were identified, each dominated by different types of methanogens, represented by the genera *Methanoregula*, *Methanoseta* and *Methanolinea*. Metabolic reconstruction revealed that the bacteria specifically associated with these methanogens, perform fermentation, as well as a syntrophy and acetogenesis facilitated by energy conservation revolving around $\text{H}_2$ metabolism. Other taxa were shown to scavenge on anabolic products (proteins and lipids) derived from detritus.

The genomic reconstruction of this study points towards more complex forms of metabolic dependencies, beyond the classic $\text{H}_2$-producing, syntrophic alkane degrader and the methanogenic partner merely scavenging $\text{H}_2$. We rather detected a complex network of strict labor division (Fig. 8).
Essential, but metabolically costly biosynthesis functions were lost in various taxonomic groups, indicating that energy and nutrient limitations select for a streamlined genome (link to paper I). Beyond that, sharing the metabolic burden is an optimization strategy for a community living under thermodynamic conditions that are close to hinder cellular metabolism.

**Paper IV. Characterization of microbial communities along a tar-contamination gradient in lake Grötingen.**

**Aim:** Assessing the factors causing the slow bioremediation and to determine differences in functional gene pools at varying pollutant concentrations. Here, we expected tar contaminated sites to have a higher genetic potential for hydrocarbon degradation than pristine sites.

**Study:** Lake Grötingen became heavily polluted when sewage from industrial production of coal and tar was discharged directly into the lake. Hundred years after the dismantling of this factory, a large area of lake sediment is still covered with high amounts of tar residue in form of polycyclic aromatic hydrocarbons (PAHs), which are the most problematic to degrade.

We studied differences of microbial bioremediation capacity within three sites of varying distance to the pollution source. Community composition (Fig. 9) and genetic potential of microbial tar degraders were studied using whole metagenome shotgun sequencing. Broadly defined metabolic functions were assessed with MEGAN, while a separate blast search was carried out for specific marker genes involved in hydrocarbon degradation. Combined with chemical analyses on the lake sediment cores, we furthermore aimed to assess factors causing the long bioremediation time, especially at the heavy contaminated site.
The comparison of functional genes connected to hydrocarbon degradation, although not statistically significant, indeed pointed towards a slight accumulation of these genes at the site closest to the factory. Overall community and functional composition did not indicate specific differences between the sites, although the heavily polluted site did show a distinct community pattern compared to the rest.

The main factor for the slow degradation rate was likely low bioavailability of the pollutants. The fact that we did not observe major differences in the gene pool implies that the potential to degrade a wide variety of compounds is likely a general feature of microorganisms in lake sediment. Thus the presence of certain compounds may not have an effect on the inherent potential, but instead on the transcription rate of the genes involved in this process.

### 2.4 Conclusions and Outlook

Any system that is devoid of photons and feature few steep chemical gradients can be seen as close to chemical equilibrium. Without redox reactions, life would seize to exist. In the context of the research presented in my thesis, various barriers (soil, rock, water or the vessel of a bioreactor) limit or prevent photon and matter entry at least temporarily. This causes these systems to move towards the thermodynamic limits for life.
As a rule of thumb, redox reactions will happen in the order from the energetically most favorable to the reaction with the least energy yield (Fig. 1). For the deep aquifers this implies, that as long as recharge water supplies the system with reduced organic matter, nitrates and sulfates, sulfate reduction and denitrification will dominate. In deeper and more isolated environments where reduced organic matter is likely more depleted and recalcitrant, chemolithotrophic processes will instead drive the system. If geo-gases such as hydrogen and carbon dioxide are available, organisms may depend on chemolithoautotrophy for primary production. In general these strategies are energetically less favorable and will result in an overall slower growth rates within these systems. Consequently generation times of hundred to thousands of years have been suggested for the SLiMES (subsurface lithoautotrophic microbial ecosystems).

In general anaerobic environments, with a rather limited amount of available free energy, force microbes into tight syntrophic interactions. These complex syntrophic partnerships require a highly adapted genome streamlined to maximize energy conservation, a trait in agreement with the observed cell-size minimization. We observed this effect especially in the deep biosphere and the short chain alkane degrading cultures. In addition, we propose that cells scavenging anabolic products derived from detrital biomass and intermediate fermentation products are also important in these systems.

As a contrast, the tar contaminated Lake Grötingen serves as an example of a complex system with substrate alternatives that are more favorable than the organic pollutants under scrutiny. Tar degrading specialists are either being outcompeted, or their metabolic plasticity allows them to switch to more favorable carbon sources, which presumably selects for cells with broad metabolic capabilities in these communities.

Low energy availability furthermore has an effect on biodiversity. We could show, that species richness decreased with the grade of isolation from energy and nutrient sources. Another interesting aspect worth mentioning is related to species migration and the role of stochasticity. While nutrient and energy availability are usually the strongest environmental filters, microbes in the deep biosphere also experience extensive filtering through long time isolation, where predation and the geological formation will hinder major species invasions. It is perhaps a provocative thought to consider that communities in isolated biomes are not always capable of exploiting the full spectrum of available energy and nutrient cycling, simply because functionally essential members never arrived.

We are still at the beginning of understanding the underlying chemical reactions that enable life on Earth. Furthermore it is likely that each of these
reactions are found in various alternative pathways that yet need to be uncovered and described. Metagenomic research is strongly limited by the incomplete knowledge on metabolic pathways and by the amount of trustworthy assigned functions in databases. Thus research on biochemistry followed by gene expression verification will most certainly remain the foundation for future research in the field.

From a microbiologist’s point of view, I believe it is important to study both the genome complexity at the level of communities and individual cell (or filament) if we want to disentangle interactions. A simplified setting, such as an enrichment culture, allows us to study processes under controlled laboratory conditions, in order to assign function to identity and to understand the underlying process without cross-signals from peripheral reactions. Studying genomic potential, especially of streamlined genomes, already tells us a lot about adaptation to a specific lifestyle. Nevertheless we need to add activity levels and turnover rates to the equation, which includes a need for transcriptomic and proteomic data.

The deep biosphere still leaves a lot of open questions that need to be asked and answered. The already mentioned activity measurements need to be done more extensively on a transcriptome level and linked to actual process rates, in order to understand the genetic underlying mechanisms of life in “slow-motion”. Furthermore, studying the trophic effects of viral predation for shaping the host community and their influence on carbon turnover, as well as their influence on evolution by facilitating gene-transfer needs to be put in focus.

In a practical sense, the work on alkane degradation could also help finding strategies for making recovery of oil sands more efficient by adding essential metabolites in to the system, that are otherwise energetically expensive for the microbes.
3 Summary in Swedish

Den mörka biosfären breder ut sig under våra fötter! Djupa akviferer i berggrunden, sediment under sjöar och hav och en rad andra liknande miljöer nås aldrig av det direkta solljus som normalt utgör den huvudsakliga direkta eller indirekta energikällan för det liv som vi möter i vårt dagliga liv på jordytan. Av förklarliga, främst praktiska, skäl vet vi fortfarande väldigt lite om de organismer och processer som kännetecknar dessa svårtillgängliga ekosystem, även om de i såväl volym som biomassa överskuggar de organismer som lever i solljus. Vad vi vet är att pågående metabola processer och långa uppehållstider leder till gradvis förbrukning och utarmning av de redoxaktiva ämnen och redox-gradierter som behövs för att driva kemotrof tillväxt och ny produktion av biomassa. Man kan därför betrakta många av dessa system som vore de i termodynamisk jämvikt där kemiskt bunden energi för att driva metabola processer är en bristvara.

I mitt avhandlingsarbete har jag utnyttjat storskalig sekvenseringsteknik för att beskriva artsammansättningen i komplexa mikrobiella samhällen samt deras samlade och individuella genetiskt kodade funktionella potential. Genom att jämföra mikrobiella samhällen i olika typer av miljöer och deras dynamik över tid har det även varit möjligt att använda denna information för att beskriva och förstå ekologin hos de organismer och samhällen som återfinns i dessa miljöer. Det har också varit möjligt att beskriva strategier och anpassningar för att överleva och tillväxa i miljöer där metabola energiflöden är mycket låga.

Mina huvudsakliga modellsystem har för detta ändamål varit akviferer i den djupa terrestra biosfären; vattenmassor som rör sig i berggrunden hundratals meter under våra fötter, och mer ytliga sediment och slaggmassor som förorenats med aromatiska föreningar på grund av mänsklig påverkan. Mina analyser visar att dessa miljöer gynnar bakterier och arkéer som har reducerade genom och även är små till storleken, vilket skulle kunna vara anpassningar för att spara energi. Som konsekvens saknar dessa organismer ofta vissa betydelsefulla eller rent av livsnödvändiga gener och funktioner. Att sakna livsnödvändiga funktioner kan tyckas svår förstålig, men en uppenbar fördel är att kostsamma metabola processer kan delas av ett flertal populations inom ett samhälle. Detta skapar auxotrofier och starka beroendeförhållanden inom samhället. Mina resultat visar också att ytterligare beroendeförhållanden inom den mörka biosfärens mikrobiella samhällen skapas genom
att många populationer inte främst utnyttjar primära substrat såsom geo-gas eller aromatiska kolväten som huvudsaklig energikälla, utan istället lever på att bryta ner och tillkansa sig sekundära metaboliter eller proteiner, kolhydrater, lipider och andra biomolekyler som producerats vid primärproducenternas biomassatillväxt. Detta pekar på att mikrobiella samhällen som växer fram i de energibegränsade miljöer som jag studerat är mycket mer komplext sammansatta än vad som är den gångse uppfattningen och att interaktionerna går bortom traditionellt definierade anaeroba näringskedjor där gradvis mer oxiderade metaboliter överförs mellan olika mikrobiella populationer.

I framförallt den djupa berggrundens mikrobiom genomfördes mer detaljerade studier över hur mikrobiella funktioner och populationer varierade mellan vattennivåer och över tid. Även om den huvudsakliga uppbyggnaden av dessa samhällen med avseende på funktioner var liknande så fanns det också anmärkningsvärd skillnader. Vattennivåer som under tusentals år är varit isolerade från de solbelysta biosfären uppvisade en mer tydlig hydrogenotrof och autotrof metabolism, medan mikroorganismer i yngre vattennivåer som är i mer direkta kontakt med ytvatten är övervägande heterotrofa och sannolikt förlitar sig åtminstone delvis på tillförsel av substrat därifrån. Det är kanske inte heller förvånande att dessa mer tytliga bakteriesamhällen uppvisar en betydligt större dynamisk variation över tid, medan bakteriesamhällen som har en omsättningstid på tusentals år uppvisar betydligt högre tidsmässig stabilitet och mindre artikedom. Detta kopplas naturligt till förväntad förbrukning och utarmning av såväl energirika kolföreningar som termodynamiskt fördelaktiga elektronacceptorer över tid, och den sänkta tillväxthastighet som detta i teorin bör leda till.


Min forskning har således kastat nytt ljus på den hittills dåligt utforskade mörka biosfären i berggrunden under våra fötter och i sedimenten under våra ytvatten och förutom att påvisa värdet av funktionell samverkan har faktorer som styr mikrobiella samhällens sammansättning och funktion sats under
lupp. Resultaten kan också användas för att hitta nya vägar för att, genom exempelvis tillförsel av nyckelm metaboliter eller främjandet av samarbetande mikrobiella konsortier, styra naturligt förekommande mikroorganismer till att utföra önskvärda ekosystemtjänster såsom nedbrytning av organiska föroreningar.
4 Acknowledgements

These past years in Uppsala and Leipzig have been an incredible journey, not only in a professional way but especially in making me grow personally.

All this would not have been possible without the support of many people starting with the great triumvirate of brilliant and unconventional minds: Annelie, Stefan and Alex. Annelie and Stefan, I’m forever grateful for all the opportunities to work in interesting fields, to cooperate and work internationally and most of all, for your kindness and mentorship over the years. Alex, thank you so much for guiding and challenging me and at the same time providing an atmosphere where I could ask stupid questions.

Lars, you are an intimidatingly cool personality and an inspiration. By the way: sorry for being a nuisance with the study plan every year!

To my cooperation partners Xiaofen & Mark thank you for fruitful discussions and for being a true pleasure to work with!

My past and recent office mates: Birgit, Janne, Tom, Claudia, Hannah & Jingying. Thank you all for inspiring chats about science and life, especially Hannah for making me practice my German on a daily basis and helping me to finish all chocolate within a radius of 30 km 😊.

Alina (my symbiosis partner in hedonistic-guilty-pleasure-indulgence) and Frieda (although it took us some time to discover the Amore!)... I have no words that would give justice to our special friendship! The same goes for you Inga, you good soul... I will never forget our good times in Germany and Sweden, and for sure more to come!

Monica, thank you for sharing your home with me for more than three years. It was a pleasure and thank you for being a good friend and listener.

Moritz and Lucas, you did not only provide crucial groundwork for this thesis, but I especially appreciate you guys for being fun outside of work!

A shout-out to my other fellow PhD’s: Torsten, my favorite rival! Thanks for the red bull and fixing the graph! Anastasija, the best free-diving and sporadic-climbing partner! Blaize, Jovana, Anne & Roger (ok you guys are a post-
docs by now), Karolina, Lorena, Maria, Yinghua, Martin, Annika, Leyden, ... with all of you I share special memories! Martha and Sari, you incredibly hardworking and kind people, I can never repay in chocolate or whatever, in how much you were of help!

Pilar, Omneya, Anna, Andrea, Nuria, Kristin, Dolly, Cristian, Silke, Eva, Sebastian, Gesa, Peter, (and the ones I forgot to mention)... you all deserve more than a brief mention; thank you for providing this exceptionally fantastic working environment!

To the Leipzig crew: Karolin, thank you for showing me the beauty of Leipzig and also for your tips and tricks in the lab, which of course goes to Martha and Ute as well!

To those of you who have been with me from the beginning: Die “Biohendl” Angela, Lisi, Leni, Matsch, Verena, Vero, Mette and Alex. “Die Gruppe” Lukas, David, Thomas, Andi², Christoph, Stefan and Bernadette. Gloria & Hannes, my second home in Berlin. Alice & Martin my Austrian base in Sweden! I’m incredibly lucky to have you guys as my friends!

Deep gratitude goes to my family, my foundation. My brother and his family for their moral support during the past years. To my late father who implanted my interest in biology by getting me the first book I can remember, which was about dinosaurs (although I eventually ended up studying smaller organisms)! And last but not least I thank my mother, I owe you so much more than just my existence, but especially for always encouraging me to pursue my studies.
5 References

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1301

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)