Julia Kranyecz, BSc

Understanding terrestrial microbial communities and their antibiotic resistance pattern from the Aral Sea and Alpine glacier regions

MASTERARBEIT

zur Erlangung des akademischen Grades

Master of Science

Masterstudium Biochemie und Molekulare Biomedizin

eingereicht an der

Technischen Universität Graz

Betreuerin

Univ.-Prof. Dipl.-Biol. Dr.rer.nat, Gabriele Berg

Institut für Umweltbiotechnologie

Graz, Mai 2020

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

Datum

Unterschrift

Acknowledgements

I would like to express my deep gratitude to my supervisor Professor Dr. Gabriele Berg, for giving me the opportunity to work at her institute, to be part of her team and for trusting me with this interesting topic.

I would also like to thank Dr. Maximilian Mora for sharing his vast knowledge and insight in all aspects of my thesis, for his patient guidance, enthusiastic encouragement, useful critiques of this thesis and for his assistance in keeping my progress on schedule. My grateful thanks are also extended to Dr. Tomislav Cernava for his valuable and constructive suggestions during the planning and development of this research work, to Dr. Wisnu Wicaksono for his help with data analysis and providing positive controls for the PCRs, to Ms Birgit Wassermann for her support with data analysis in anvi'o and to Dr. Melanie Obermeier for her useful and constructive recommendations for data visualization.

I would also like to extend my thanks to the technicians Ing. Monika Schneider-Trampitsch, Ing. Angelika Schaefer, Ing. Barbara Fetz and to the apprentices for offering me help in all thesis-relevant questions and giving emotional support whenever needed.

Special thanks should be given to my fellow Master students Tamara Reisenhofer-Graber and Matthias Schweitzer for scientific discourse and emotional support on a daily basis.

Finally, I must express my very profound gratitude to my family and to my friends for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.
Thank you.
Julia Kranyecz

Abstract

Increasing antibiotic resistances (AR) in bacteria mark one of the biggest rising obstacles for health care and medicine. A commonly accepted concept is, that increasing AR are caused by bacterial adaption to extensive use of antibiotics by humans in healthcare, agriculture, etc. However, a less well studied question in this regard is to which extent antibiotic resistances occur already naturally without human influence. To tackle this issue, this thesis focuses on the AR potential in two regions: First, the Aral Sea region in Uzbekistan, an altered ecosystem, as example for AR occurring in an environment which is heavily influenced by humans and second, the Austrian glacier region, a region untouched by civilization to set a baseline of natural AR occurring in an untouched environment. By using 16S rRNA gene amplification sequencing, antibiotic resistance gene (ARG) targeted qPCR analysis and metagenomic analysis, the abundance of taxa of both areas could be shown. The microbiome of terrestrial habitats in the Dachstein region showed a balanced, healthy structure, while the Aral Sea's microbiome depicted overabundances of taxa and an overall unhealthy state. Both locations possess antibiotic resistance potential, however, the ARG potential located on plasmids in both cases is low, implying a low risk of ARG dissemination.

Kurzfassung

Die steigenden Resistenzen gegen Antibiotika in Bakterien stellen mehr und mehr ein ernst zu nehmendes Problem in der Medizin dar. Die vorherrschende Meinung ist, dass Bakterien Resistenzen als Antwort auf den hohen Einsatz von Antibiotika in Medizin, Landwirtschaft etc. entwickeln. Doch auch in der vom Menschen wenig beeinflussten Natur kommen Antibiotikaresistenzen vor. Um dieses bisher wenig erforschte Phänomen näher zu beleuchten, beschäftigt sich diese Masterarbeit mit den Antibiotika Resistenzen in zwei Gebieten: 1) Die Aral See Region in Usbekistan, eine stark durch den Menschen veränderte Region und 2) die Dachstein-Gletscherregion, wo nur wenig Einfluss durch den Menschen besteht. Durch 16S rRNA Genamplifikation, Metagenomanalyse und

durch qPCR zeigt diese Arbeit, dass die bakterielle Zusammensetzung im Mikrobiom am Dachstein über die Jahre stabil und ausgewogen bleibt, während am Aral See wenige Taxa vorherrschen und das Mikrobiom dominieren. An beiden Orten konnten Antibiotika-Resistenz-Gene gefunden werden, von denen jedoch ein nur sehr geringer Anteil auf Plasmiden liegt und somit ein Problem für das Resistenzproblem in der Medizin darstellt.

Table of Content

I. Introduction

1. About the current situation of antibiotic resistances

1.1. Resistances in health care and agriculture

Antimicrobial resistances (AMR), especially antibiotic resistances (AR), is getting more and more a serious threat to global public health. Currently, AR are on the rise, because microorganisms are getting less sensitive to formerly effective drugs. As a consequence, established treatments against microbial infections work less efficient or show no effects at all which prolongs a patient's hospital stays. A spread of infections is more likely to take place and is harder to control, while economic and social costs explode.[1]

Antibiotics have been a highly efficient weapon against bacterial infections for decades, being able to cure medical conditions that often led to death in the pre-antibiotic era. But soon after the discovery of antibiotics and the medical use thereof, the existence of ARs but also their increase, which is correlated with a higher use of antibiotics, came into focus.[2] Overuse and inappropriate administration of antibiotics in the medical field, e.g. for common cold or other viral infections, was – and sadly still is- a major driving factor leading to today's dramatic situation: multi resistant bacteria, highly enriched in hospitals, causing incurable diseases and up to 700 000 deaths per year worldwide.[3,4]

The global rise of antibiotic resistant bacteria is not only associated with the increasing morbidity and mortality, it also threatens the achievements of modern medicine and it could, as discussed in the scientific community and announced by the WHO, potentially lead to a post-antibiotic era.[1,5] Up to 10 million deaths per year worldwide after 2050 could be possible if no new antimicrobial drugs or antiretroviral therapies hit the market and if the current trends continue.[4] In 2019, extended-spectrum β-lactamase – producing *Enterobacteriaceae* were already resistant to the majority of third- and fourth-generation cephalosporins.[6] Bacteria resistant to Colistin, a last resort antibiotic, were found in China in 2015 in humans and pigs alike, spreading across Europe and the USA as well.[4] According to the WHO, very high rates of AR in common bacteria such as *Escherichia coli, Klebsiella pneumonia* and

Staphylococcus aureus were found in all WHO regions, carrying resistance against various antibiotics such as 3rd gen. cephalosporins, fluoroquinolones, 3rd gen. carbapenems and methicillin. However, as they state, there are significant gaps in surveillance of pathogens which again pose a major threat to public health.[1] Yet, while the development of new antibiotics should be top priority in order to prevent a post-antibiotic era, private and public investments keeps decreasing.[7]

But not only the medical field struggles with multi resistances, the wasteful use of antibiotics in agriculture, especially for non-therapeutic administration, could cause resistance in plants. This, on the other hand, could subsequently become a problem for humans as well, when these AR bacteria are ingested and their resistance spread in human environment. Since the 1950s, it became common practice to administer subtherapeutic quantities of antibiotics to animals such as poultry, beef cattle and swine, in order to enhance the feed-to-weight ratio[3]. The downside of this common procedure, however, is, that these long exposure periods create perfect conditions for bacteria to develop and spread resistances and contribute to increased rates of antibiotic resistance in human pathogens.[4,8] Via horizontal gene transfer of genetic elements, e.g. plasmids, resistances can be shared with other microbes via conjugation, leading to multi resistant bacteria.[3]

The transmission of AR bacteria from livestock to humans can occur in various ways, including direct contact of farmworkers, butchers, veterinarians and others involved in the process of meat and dairy production as well as through the food chain after ingestion.[3,9] Studies show tetracycline-resistant *E. coli* strains in the gut microbiome of chicken farmers after administration of tetracycline feed to their poultry as well as a direct transmission of the AR bacteria within chicken populations through feces.[10] But not only food-borne pathogens, e.g. *E.coli* and *Salmonella spp.*, were found to have a link between AR genes in livestock and humans, also species of *Enterococcus* and methicillin-resistant *Staphylococcus aureus* (MRSA) were found.[3] In the Netherlands, MRSA, a bacterium that in its nonpathogenic form usually colonizes the nasal and intestinal tract, was found to be present in 12% of meat products freely available in retail.[3] The meat most likely got contaminated during the slaughtering process, as the contents within the intestinal tract got in contact

with the meat during that process. Interestingly, however, meat of chicken, which did not get treated with antibiotics prior to slaughter, showed lower values of AR bacteria.[3,11] In Italy, MRSA strains were found in bovine milk and cheese products.[12] Studies also show that ingested glycopeptide and streptogramin-resistant *Enterococci* are able to survive gastric passage and can be found in feces for up to 14 days after consumption of contaminated meat. In Denmark in 1997, 50% of chicken meat from retail stores contained *E. faecium* and 10% of these were vancomycin resistant. This leads to the assumption that enterococcal contamination of meat and meat products is widespread.[9]

The term One Health is often found in literature in this context and represents the idea of interconnected, geographically close ecosystems that share microorganisms and their antimicrobial resistances, host organisms and environmental factors that contribute to the spread of resistances.[13] On the other hand, Global Health's view is in a worldwide context attending to global conditions that lead to a spread of resistances by regarding political and socioeconomic actions of countries on a global stage. Also, the microbiome of one area, even if it may be located rather remotely, is not sealed off from influences, e.g. through wastewater treatment, pollution, farming, a steady exchange of resistances occurs. On a worldwide view, these local communities are interconnected through weather, travel, trade or animal migration leading to an interchange among geographical areas[13].

1.2. Resistances in natural habitats

The mechanisms of antibiotic resistance formation and spread in bacteria were described in the passage above. Some bacteria, however, show intrinsically low susceptibility to antibiotics in their natural environment as well.[14] In order for an antibiotic substance to show effectiveness, the substance must pass the cell envelope and reach its target at a sufficient concentration. There are certain defense mechanisms to protect a cell from antibiotic substances: Target modification, target protection, or a general reduction of antibiotic concentrations. Efflux pumps, originally known to be involved in signal trafficking and offering resistance to toxic compounds from various sources, show great antibiotic resistance potential as well.[14–16]

The soil also contains natural antibiotic-producing bacteria, which are

part of complex communities where they interact with other species.[15] Following a relocation of such resistance genes from the chromosomal DNA to mobile elements, these resistances can become transferable to clinical pathogens. Around 50% of *Actinomycetes* isolated from soil are capable of synthesizing antibiotics[8], therefore the presence of ARG in bacteria is a not surprising defense mechanism. Soil samples from all over the world besides Antarctica contained AR bacteria, highly implying that natural resistances already existed before humans started producing and using antibiotics on a grand scale.[17]

It is questionable to which extend a seemingly natural environment really is not anthropogenically influenced, since common antimicrobial resistance genes (ARG) found in humans and livestock can also be detected in ground water and soil. This could be due to a distribution thereof by physical factors, such as rainfall, water flow, wind and animal migration, especially bird migration.[6,18] This antibiotic gene leakage to natural environments drastically impacts the local bacterial communities, as it is believed to alter eco-evolutionary feedback loops present in microbial communities.[15]

2. The Dachstein glacier – an endangered environment

The Dachstein glacier, particularly the Hallstadt glacier, is part of the Eastern Alps and located at the border of Styria and Upper Austria. With a size of 3.04 km^2 in 2007, it is the biggest glacier of the northern Limestone Alps. The Hallstadt Glacier was formed during the last ice age called Wyrm approximately 18 000 years ago and reached its largest documented size in 1865 during the so-called small ice age in the 18^{th} and 19^{th} century. Since then, the size has been impacted by climate change, endangering the permafrost bodies of the glacier, which lead up tp the loss of approximately 63% of its volume since 1865.[19,20]

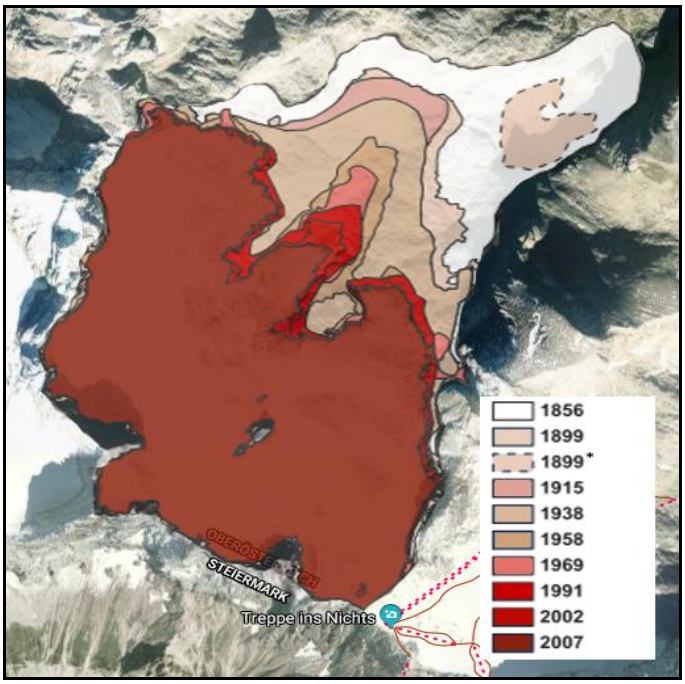


Figure 1: Decrease of the Dachstein glacier size since 1865, modified from Helfricht, K., 2009[19] and Google maps[21]. *dead ice, that got separated from the glacier main body

Warmer temperatures and a lower total annual precipitation, especially during winter and summer, led to a tremendous decrease in size, which is shown in Figure 1. Since 2006, researchers annually measured the mass balance and change in mass of the Hallstadt glacier. Even in 2012/13, the year with the least negative mass balance since start of the measurement, the glacier lost total mass while its size stagnated around 3,016km^2.[22] Future scenarios predict a split of the glacier in an eastern and western part and a dramatic reduction of ice thickness.[19]

Regarding the distribution of microorganisms after a retreat of glaciers, it has to be noted that not much is known so far. Best studied are polar and Antarctic glacier fore fields, while the Dachstein lacks information thereof. One study featuring the Damma glacier fore field located in Switzerland's Alps suggests that shifts in microbial communities can happen because of increasing Carbon and nutrient availability, plant colonization and associated changes in pH over the years after receding of the glacier. [23]

3. The Aral Sea – an anthropogenic disaster

3.1. A once flourishing region – impacted by the Soviet Union's agriculture system

The Aral Sea is located at the border of Kazakhstan and Uzbekistan in central Asia and was once the world's fourth largest lake with a total size of 67,500 km^2 in 1960, only behind the Caspian Sea, Lake Superior and Lake Victoria. [24] Two rivers, the Amu Darya and Syr Darya, fed the Aral Sea with a volume of 56 km^3 water per year and sustained a diverse flora and fauna in the river deltas. Irrigated agriculture was a common practice, as well as animal husbandry, hunting, fishing and reed harvesting. [24,25]

3.2. Shrinking of the Aral Sea and its impact on the environment and the local population

Agricultural diversions and the unsustainable use of water from the two water suppling rivers in the early 1960s (especially for cotton production in the Soviet Union) led to the drying of the lake as the water balance of the Aral Sea could not be sustained and the volume decreased rapidly[25]. The maximum water level declined more than 26 m from 1960 until 2009.[21] The shrinking water levels can be seen in Figure 2 a – f. In 1989, as a consequence of the water loss, the Aral Sea got separated into the Small Aral Sea in the northern part and the Large Aral Sea in the south[24–26]

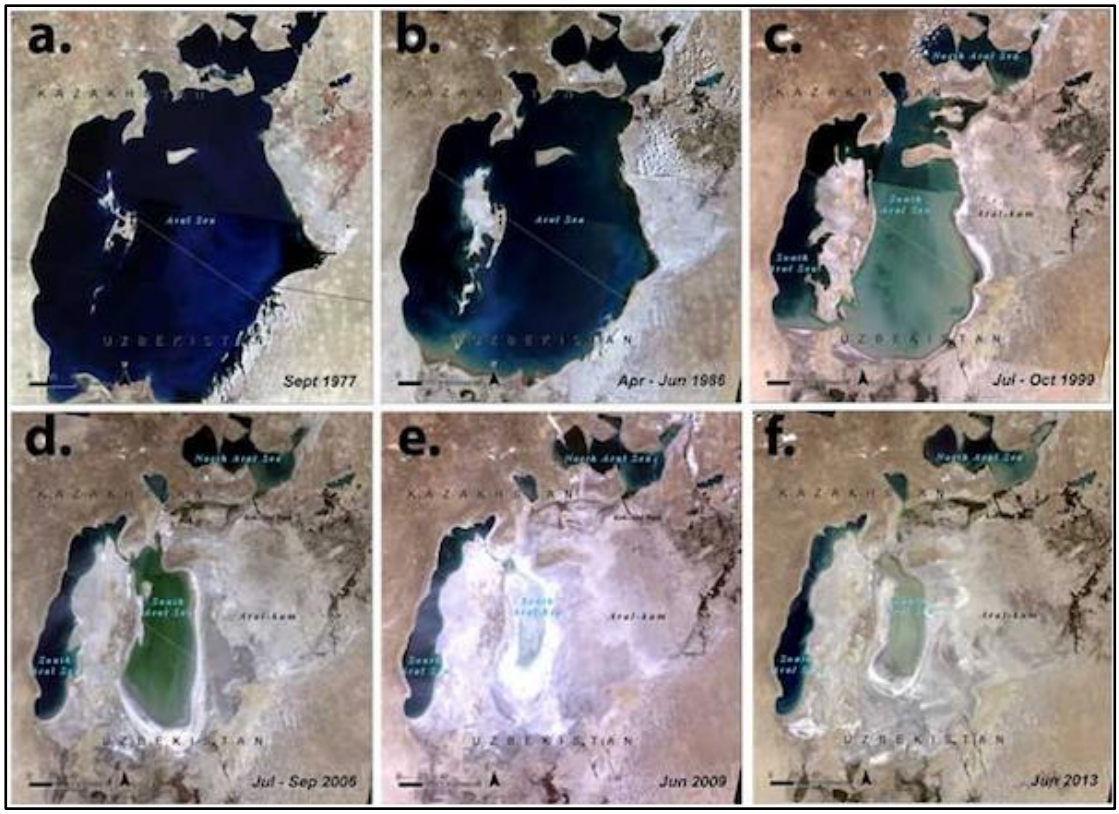


Figure 2: Satellite images monitoring the shrinkage of the Aral Sea from 1977 – 2013. a. 1977, b. 1985, c. 1999, d. 2005, e. 2009, f. 2013. [27]

Due to the desiccation of the Aral Sea, traditional and influential industries such as fishery vanished completely, but it also impacted the climate of the whole region.[28] Maritime conditions have been replaced by a more continental and desert-like climate with warmer summers, cooler winters and an overall reduced humidity, leading to a shortened growing season for crops.[29] The increasingly dry climate also impacted agriculture: animal husbandries decreased in number, areas of cultivatable land have been reduced and so has the quality of the soil. Crops such as cereal, rice, potatoes and vegetables are grown on irrigated land, the yield is affected by the deterioration of the soil and excessive salinization. The soil in the Syr Darya river delta has been found to contain alkali- (HCO_3) and carbon-trioxide (CO_3), which are both known to be harmful for environment and vegetation. The salinity in the Eastern Basin of the Large Aral Sea exceeds 200 g/L.[24,28]

Not only the high salinity poses a threat to the environment and human health, but also the high value of industrial pollutants such as PCB-compounds, DDT, HCH, heavy metals and pesticides for parasite control. They first accumulated in water, but after desiccation, the remaining soil was also

contaminated and through atmospheric transport these toxic compounds enter the food chain leading to humans.[30]

The local population living in the ecological disaster zone around the Aral Sea suffers acute health problems such as respiratory and digestive diseases, or even cancer caused by toxic compounds in salt and dust as well as a loss of diversity in nutrition. Combined with poor sanitation, this led to a collapse of the healthcare system and the infant mortality rate rose from an average of 45/1000 live births in 1965 to more than 80/1000 in districts near the former seashore. [25] Potable water is often contaminated with human pathogens, promoting the prevalence of viral hepatitis, typhoid, paratyphoid and dysentery diseases. Liver and kidney ailments are widespread, probably due to the high salt content of drinking water. [29,30]

3.3. The current situation and outlook for the future

Since the 1980s, there have been measures of revegetation and reforestation of parts of the dried bottom of the Aral Sea basin in order to stabilize and lower the probability of wind erosion in this area. These measures as well as the creation of lakes and wetlands are one first step to battle further desertification of the Aral Sea basin.[24]

In 2005, a dam was built in order to stabilize the water level of the Small Aral Sea by accumulating the water of the Syr Darya. Since then, the ecosystem of the Small Aral Sea has improved, reaching a stable water level of 42 m and an average salinity of 10-14 $g*L^{-1}$, which is close to the natural salinity the Aral Sea had back in 1960 (~ 10 $g* L^{-1}$). Also, the dissolved oxygen levels seem to be high and contributed to a comeback of fish life in the Small Aral Sea. Species such as carps and pike-perch fish started to thrive again after the salinity levels have decreased.[24,25,29]

This, however, only applies to the Small Aral Sea, as the Large Aral Sea still faces desiccation and environmental problems. In order to avert further desiccation or to even reverse the current crisis, a complex approach is needed: The old irrigation system must be reconstructed, although this would significantly harm the economic and social state of the local population due to its importance in the food production for the growing population. Development and introduction of effective irrigation technologies such as drip irrigation should

be installed, while a reconstruction of water supply systems and optimization of the drainage runoffs need to take place.[24,27] Since these are costly investments, it is unlikely for the Aral Sea to completely recover in the near future. However, there is hope for a partial restoration like it is already happening for the Small Aral Sea.[24]

4. Objectives of this study

This thesis aims to show the bacterial community composition of rhizosphere and soil at the Dachstein and at the Aral Sea, antibiotic resistance potential and AR gene abundance by using 16S rRNA gene amplification sequencing, AR gene targeted qPCR analysis and metagenomic analysis. My hypothesis features a less diverse microbial community and a higher ARG abundance in the Aral Sea samples, due to a higher anthropogenic impact, while the Dachstein samples are expected to feature more microbial diversity and less ARG abundance. Furthermore, this thesis aims to bring more insight to the underdeveloped research on the subject of naturally occurring antibiotic resistances and how anthropogenic influence can alter natural environments.

II. Materials and Methods

To generate data for this thesis, rhizosphere samples as well as soil samples were taken at multiple locations of the sampling areas (Aral Sea and Dachstein glacier). Afterwards, DNA was extracted and used for 16S rRNA gene amplicon sequencing, metagenome sequencing, and PCR as well as qPCR. Negative controls were treated the same way as the samples throughout all procedures. The data from amplicon sequencing and metagenome sequencing was processed with bioinformatic tools further explained later. In
Figure *3* the workflow of the procedures is illustrated.

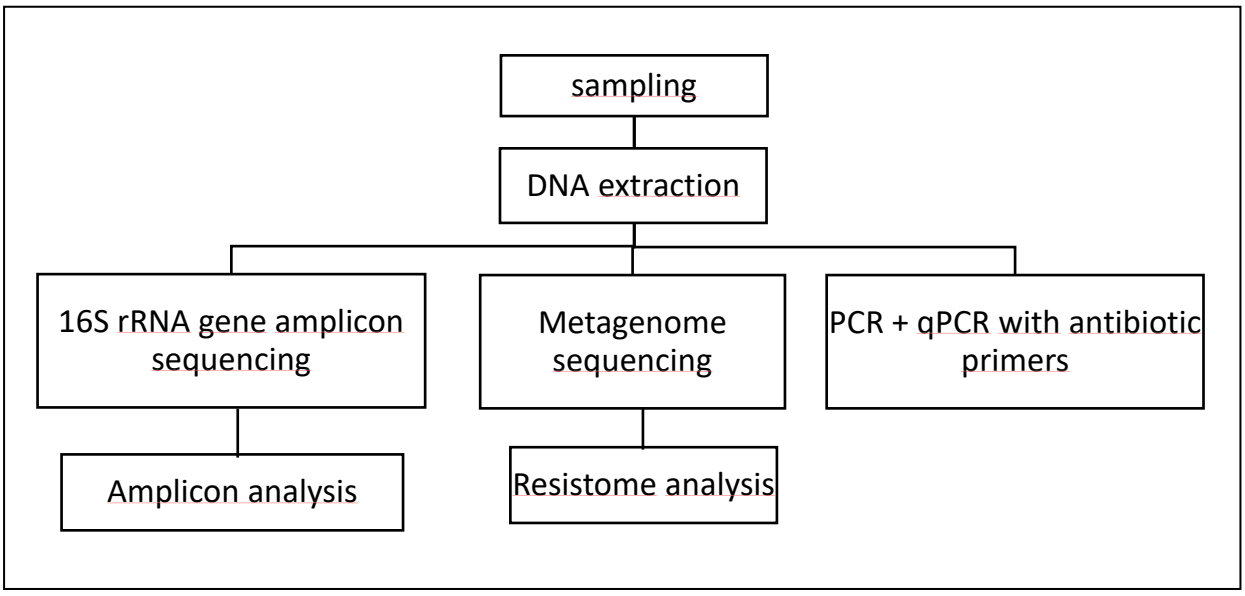


Figure 3: Visual overview of the workflow for this work

1. Sampling

1.1 Dachstein samples

Rhizosphere samples of alpine plant species *Papaver alpinum* (L.) (engl. Burser-Alpine poppy, ger. Alpen-Mohn), further referred to as *Papaver alpinum, Hornungia alpina* (L.) O. Appel (engl. chamois cress, ger. Gämskresse), further referred to as *Hornungia alpina* and *Sedum atratum* L. (engl. dark stonecrop, ger. Dunkler Mauerpfeffer), further referred to as *Sedum atratum,* were taken near the Hallstatt glacier. Additionally, corresponding soil samples were taken, for exact GPS data and visualization thereof, please see Figure 4 . The plants

were sampled in areas where the glacier receded 10 years, 70 years and 150 years ago; further referred to as Dachstein 10a, 70a and 150a samples; the plant associated rhizosphere samples will be referred to as plant samples.

Figure 4: Sampling points of the Dachstein samples. Blue: 10 years, yellow: 70 years, red: 150 years without ice. GPS data visualized by Latlong[31], card material courtesy of OpenStreetMap® [32]

Table 1: GPS information of the Dachstein samples: Latitudinal and longitudinal data of the sampling spots

Sample name	years without ice	Lat	Long	Altitude	chosen for metagenome analysis
Ho1	10a	47,4888988	13,618165	2270	x
Ho2	10a	47,4873205	13,6202097	2211,14136	
Ho3	10a	47,486995	13,620387	2274	
Ho4	10a	47,4873205	13,6202097	2211,14136	
Ho70	70a	47,4939661	13,6243246	2054,44751	
Ho150	150a	47,495682	13,6229381	2103,71484	
Pa1	10a	47,4888988	13,618165	2270	
Pa2	10a	47,4884432	13,6215702	2173,89063	
Pa3	10a	47,4889409	13,6222171	2156,34644	
Pa70	70a	47,4939661	13,6243246	2054,44751	
Pa150	150a	47,495682	13,6229381	2103,71484	
Se1	10a	47,4888988	13,618165	2270	
Se2	10a	47,4884432	13,6215702	2173,89063	
Se3	10a	47,4889409	13,6222171	2156,34644	
Se70	70a	47,4939661	13,6243246	2054,44751	
Se150	150a	47,495682	13,6229381	2103,71484	
So1	10a	47,4888988	13,618165	2270	
So2	10a	47,4873205	13,6202097	2211,14136	
So3	10a	47,486995	13,620387	2274	
2So1	10a	47,4873205	13,6202097	2211,14136	
2So2	10a	47,4873205	13,6202097	2211,14136	
2So3	10a	47,4873205	13,6202097	2211,14136	

Plants and soil were sampled for 16S rRNA gene amplicon sequencing and metagenome sequencing. All samples were stored in portable cooling bags and processed within 48 hours after sampling.

1.2 Aral Sea samples

The Aral Sea plant samples from the species *Suaeda acuminata* (C.A.Mey.) Moq. (no english or german name available), further referred to as *Suaeda acuminata* as well as their corresponding soil were taken in the Aral Sea basin and near the Large Aral Sea's west shore line, for exact GPS data and visualization thereof, please see Figure 5 and Table 2. The plants were sampled in areas where the water receded 5 years, 10 years and 40 years ago; since the 5 year and 10 year spots were closely together, they appear as one spot in the map. The samples where the water receded 5 years, 10 years and 40 years ago will be referred to as Aral Sea 5a, 10a and 40a in this work; the plant associated rhizosphere samples will be referred to as plant samples. Again, all samples were transported at ambient temperature and were processed within three (40a samples) or four days (5a and 10a samples).

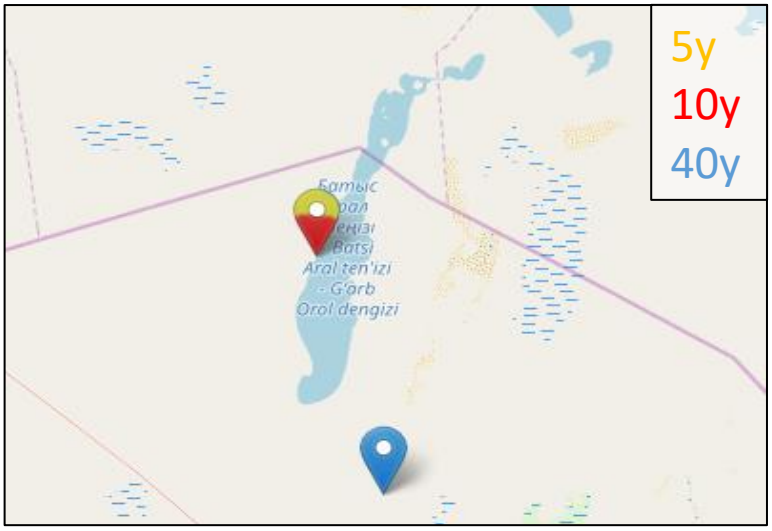


Figure 5: Sampling points of the Aral Sea samples. Blue: 40 years, red: 10 years, yellow: 5 years without water. GPS data visualized by Latlong [31], card material courtesy of OpenStreetMap® [32]

Table 2: GPS information of the Aral Sea samples: Latitudinal and longitudinal data of the sampling spots

Sample name	years without water	Latitude	Longitude	Altitude	chosen for metagenome analysis
A-51	5a	45,093333	58,340556	10	
A-52	5a	45,093333	58,340556	10	x
A-53	5a	45,093333	58,340556	10	
A-101	10a	45,093333	58,340556	10	
A-102	10a	45,093333	58,340556	10	
A-103	10a	45,093333	58,340556	10	
A-401	40a	44,014722	58,717778	20	
A-402	40a	44,014722	58,717778	20	x
A-403	40a	44,014722	58,717778	20	
AS-51	5a	45,093333	58,340556	10	x
AS-52	5a	45,093333	58,340556	10	x
AS-53	5a	45,093333	58,340556	10	x
AS-101	10a	45,093333	58,340556	10	
AS-102	10a	45,093333	58,340556	10	
AS-103	10a	45,093333	58,340556	10	
AS-401	40a	44,014722	58,717778	20	x
AS-402	40a	44,014722	58,717778	20	x
AS-403	40a	44,014722	58,717778	20	x

2. Sample processing

2.1 Sample preparation

At least three independent biological replicates, each consisting of roots of several individual plants to reach at least 5 g sample mass per biological replicate, were taken at the Dachstein 10a sampling spots and for all Aral Sea samples. One independent biological replicate, again consisting of several individual plants, was taken at the 70a and 150a sampling spots at the Dachstein. Soil samples were taken in triplicates at each Aral Sea sampling spot and in the Dachstein 10a zone.

To isolate the DNA from soil and rhizosphere, 5 g of soil or rhizosphere were added to 20 mL sterile 0.85% NaCl, shaken by hand and vortexed for 3 min. Fractions of 2 mL of the resulting suspensions were centrifuged at 16.000 x g and 4°C with a Sorvall RC-5B Refrigerated Superspeed Centrifuge (DuPont Instruments, USA) for 20 min until four technical replicates, consisting each of a pellet of about 0.1 g, were obtained. The pellets were weighed and stored at -20°C until extraction.

2.2 DNA extraction

In a next step, the DNA was extracted using the FastDNA SPIN kit for soil (MP Biomedicals, Solon, OH, USA). All of the pellets were treated accordingly to manufacturer's protocol. In short, 978 µL sodium phosphate buffer and 122 µL MT buffer were added to the pellets. The cells were homogenized by FastPrep FP120 instrument (Qbiogene, BIO101, Carlsbad, CA, USA) for 40 seconds at a speed of 6.0 m/s. After 15 min centrifugation at 14,000 × g, 250 µL protein precipitation solution (PPS) were added to the supernatant. After five min centrifugation at 14,000 × g, one mL binding matrix solution was added to the supernatant. After inverting the tube for two minutes by hand the tubes were incubated at room temperature for at least three min to allow settling of silica matrix. After that, the supernatant was discarded, while the resuspended binding matrix was transferred to a spin filter with centrifugation at 14,000 × g for one min. Afterwards, the pellet got resuspended in 500 µL SEWS-M, the suspension was again centrifuged at 14,000 × g for one min. After air drying the spin filter for five min, the binding matrix was resuspended in 60 µL DNase free water. For better binding, the matrix was incubated for five minutes at 55 °C. After another centrifugation step at 14,000 × g for one min, DNA was ready. DNA concentrations of all samples were measured via a NanoDrop™ 2000/2000c Spectrophotometer and afterwards all samples were stored at -20 °C until use.

3. Amplification and sequencing of 16S rRNA fragments

In order to identify the unknown bacteria in each sample, the V4 region of the 16S rRNA gene was amplified by a polymerase chain reaction (PCR) using universal bacterial primers (Caporaso[33,34]) featuring an inhouse barcode system (see Table 3 and Table 4 for details). To ensure that the amplification was successful, the PCR products were loaded onto a 1% agarose gel at 140 V for 60 min (genXpress ; 1xTAE buffer) and purified using the Wizard SVGel and PCR Clean-Up System (Promega, Madison, USA). Afterwards, the DNA concentration of the purified barcoded samples were measured via Qubit dsDNA BR assay (Thermo Fischer Scientific) and pooled (equimolar, 500 ng). Sequencing was done on Illumina MiSeq (2 × 300 bp paired end) at Genewiz

(Leipzig, Germany). This was done for three of the four technical replicates from each sample.

Table 3: 16S rRNA gene PCR reaction mix per sample

Reagent	**Volume (µL)**
H_2O (nuclease free PCR water)	20.6
515f_BC Primer (5 µM)	1.2
806r_BC Primer (5 µM)	1.2
6x Taq-&GO™ mastermix (MP Biomedicals, Solon, OH, USA)	6.0
DNA template	1.0
Total	30.0

Table 4: Nucleotide sequence of the Caporaso[33] amplicon primers used for PCR (515F Illumina tag primer + 806R Illumina tag primer (Microsynth AG, Balgach, Switzerland) in Table 3.

Primer Label	Nucleotide sequence (5'-3')
515F (Caporaso)	GTGCCAGCMGCCGCGGTAA
806R (Caporaso)	GGACTACHVGGGTWTCTAAT

Table 5: 16S rRNA gene PCR program settings for the thermocycler (Whatman Biometra Tpersonal (Biometra 141 GmbH Göttingen, Germany))

Initial denaturation	96°C/5 min	
DNA denaturation	96°C/1 min	30 cycles
Primer annealing	54°C/1 min	
Elongation	74°C/1 min	
Final extension	74°C/10 min	
Finish	15°/∞	

4. DNA preparation for metagenome sequencing

For metagenome sequencing, the 4 technical replicates of each sample were pooled; the Aral Sea soil triplicates at 5a, 10a and 40a sampling spots were pooled due to a very low DNA content in the individual replicates. DNA concentration of each individual sample was measured via Qubit dsDNA BR and HS assay (Thermo Fisher Scientific). Afterwards, total community DNA shotgun sequencing was performed on an Illumina HiSeq2000 system (2 × 150 bp) by Genewiz (Leipzig, Germany).

5. Bioinformatic procedures

5.1 Amplicon sequence processing

The sequences were analyzed with QIIME 2[35] (version 2019.10). Primer and adapter sequences were removed, demultiplexing was done via cutadapt[36] with a minimal length of 50 bp. Dada2[37] was used (through Dada2 plugin in QIIME2 pipeline) to denoise and dereplicate as well as to remove chimeras of paired end sequences, resulting in amplicon sequence variants (ASVs). Classification of the taxonomy was done via QIIME 2 vsearch[38] and remaining sequences of chloroplastidal and mitochondrial origin were removed. As reference database SILVA[39–41] (version 132) was used. Output sequences were classified as kingdom, phylum, class, order, family and genus depending on the depth of reliable classifier assignments. Further statistical analyses and visualizations were conducted in QIIME 2 and in Calypso[42] (version 8.72).

For alpha diversity, Shannon metric was used to measure the similarity, while beta diversity was shown as a Bray-Curtis dissimilarity matrix, also visualized via Emperor[43] plot. With the help of Calypso, LEfSe[44] (Linear discriminant analysis Effect Size) plots were generated. To evaluate effects of plant species and sampling site on the richness, Analysis of variance (ANOVA) was used. Permutational analyses of variance tests (PERMANOVA) were performed with 999 permutations.

QIIME 2 commands and the taxa contained in the negative controls, which were removed from the sample datasets, can be seen in the appendix VII.1 and VII.2.

5.2 Bioinformatic data analysis of the metagenome

The metagenome data assembly was quality checked by FASTQC[45,46], trimmomatic[47] (version 0.39) was used for filtering, assembly was done by megahit[48] (version 1.2.9); further data processing and downstream analysis were done via bowtie2[49] (version 2.3.5.1), samtools[50,51], anvi'o[52] (version 6.1) + the annotation via COG database[53], kaiju[54] for taxonomic classification (version 1.7.3 + NCBI non redundant database (nr)) and RGI[55] (version 5.1.0) against CARD[55] (version 3.0.7) and wildCARD[55] (Prevalence, Genomes, & Variants data, version 3.0.6) databases. The graphic output of the metagenomic data

was created with the help of krona[56], Circos[57] (Circos Table Viewer v0.63-9), RAWGraphs[58] and Calypso.

Commands for the metagenome analysis can be seen in the appendix VII.3.

Due to time reasons, all of the metagenomics analysis in anvi'o were done with the samples A-52, A-402, A-Soil5 and A-Soil40 to show the difference between the metagenome of very dry environment and recently dried out areas. The Dachstein analysis was left out due to time reasons but will be a project for the future. The data of viral taxa was removed from the datasets for a better comparison with the 16S data.

For the resistome analysis in RGI, several approaches were evaluated, but decided to use RGI bwt approach due to time reasons. In order to obtain information about the resistome of the Dachstein as well, sample Ho1 were included in the RGI bwt approach, as well as the Aral Sea samples listed above.

The resistome data acquired through RGI bwt contained all perfect and strict hits above 40% identity and was used for analysis without further filtering. Thresholds for resistome data are currently discussed in literature, but the "ideal" threshold is yet to be found. For analysis of the whole resistome datasets, an identity threshold of 87%[59] for strict hits was found to be useful in sewage impacted environments and is used in the analysis of the whole metagenome samples. RGI reads with two or more hits in Drug classes were combined into "multi resistant" for reasons of simplicity.

6. Antibiotic resistance screening with primers

6.1 PCR Screening

In order to screen the samples for antibiotic resistances and quantify them via qPCR, established primers used mainly in studies about environmental samples or glacier regions (Table 6) were ordered.

Table 6: Sequences of the qPCR primers, [b]Y= T or C; R = A or G; S= G or C

antibiotic	Gene name	Primer forward (5'-3')	Primer reverse (5'-3')	length (bp)	source
Tetracycline	tetA	CCGCGCTTTGGGTCATT	TGGTCGCGTCCCAGTGA	51	[60]
	tetB	AGGCGCATCGCTGGATT	CAGCATCCAAAGCGCACTT	55	[60]
Sulfonamide	sul1	CCGTTGGCCTTCCTGTAAAG	TTGCCGATCGCGTGAAGT	67	[61,62]
Beta-lactamase	CTX-Mg8/25	AACRCRCAGACGCTCTAC[b]	TCGAGCCGGAASGTGTYAT[b]	326	[63]

	blaKPC	TTGTTGATTGGCTAAAGGG	CCATACACTCCGCAGGTT	106	[64]
Vancomycine	**vanR**	**ATGTTATCGTCCACTCCGGC**	**AACTCGGTGGGAGTAAGGGA**	**91**	[65]
	vanH	**GACAGTTGGTGTGGTGGGAA**	**CGGCTGCGACTATAAGCCAA**	**96**	[65]
Streptomycine	str(A)	TCAATCCCGACTTCTTACC	CACCATGGCAAACAACCATA	126	[8]
	str(B)	ATCGCTTTGCAGCTTTGTTT	ATGATGCAGATCGCCATGTA	143	[8]
	aad(A)	CAGCGCAATGACATTCTTGC	GTCGGCAGCGACAYCCTTCG[b]	295	[66]
Erythromycine	ermA	AAGCGGTAAACCCCTCTGA	TTCGCAAATCCCTTCTCAAC	190	[67]

20-30 ng DNA of the Dachstein samples as well as the Aral Sea samples were tested for antibiotic resistance genes through PCR with the primers listed above. The reagents used for the master mix can be seen in Table 7, the total volume of each PCR reaction was 30 µL. The PCRs run with the primers tetA, tetB, sul1, vanR, vanH, str(B) , aad(A) and ermA were tested with the high melting temperature program in Table 8, while the primers CTX-Mg8/25, blaKPC and str(A) were run with the low melting program in Table 8, because of the lower melting temperatures of these primers.

Table 7: AB PCR reaction mix per sample

Reagent	**Volume (µL)**
H_2O (nuclease free PCR water)	21.8
Primer forward (10 µM)	0.6
Primer reverse (10 µM)	0.6
6x Taq-&GO™ mastermix (MP Biomedicals, Solon, OH, USA)	6.0
DNA template	1.0
Total	30.0

Table 8: AB PCR program settings for the thermocycler – high/low melting (Whatman Biometra Tpersonal (Biometra 141 GmbH Göttingen, Germany))

	high melting	**low melting**	
Initial denaturation	94°C/10 min	94°C/10 min	
DNA denaturation	94°C/40 sec	94°C/40 sec	35 cycles
Primer annealing	60°C/40 sec	50°C/40 sec	
Elongation	72°C/40 sec	72°C/40 sec	
Final extension	72°C/7 min	72°C/7 min	
Finish	4°/∞	4°/∞	

6.2 Real time qPCR

In addition to the PCR screening with all 11 antibiotic primers, a qPCR (quantitative PCR) was done with vanH and vanR primers to determine the abundance of vancomycin resistance genes in the bacterial community of the Dachstein and Aral Sea samples. To run the qPCR a Rotor-Gene 6000 real-

time rotary analyzer (Corbett Research, Sydney, Australia) and Rotor-Gene 6000 Series Software 1.7 were used.

The qPCR reaction mix components can be seen in Table 9, the qPCR program ran accordingly to Table 10. Gene copy number determinations were made by comparing Ct-values of the samples to Ct-values of a standard curve. The standard curve was created by using DNA with given concentrations of the Vancomycin gene, dilutions of 10^{-4} to 10^{-7} (concentrations of 10^{3} -10^{6} ng/µl) were used in this case. With the help of the Ct-values of the samples and the standard curve, the copy number of Vancomycin genes in each sample could be calculated.

Table 9: vanH/ vanR qPCR reaction mix per sample

Reagent	**Volume (µL)**
H_2O (nuclease free PCR water)	3.0
KAPA SYBR Green 2x MM (KAPA Biosystems Boston, Massachusetts, United States)	5.0
vanR_f [10 pmol/µl] /vanH_f [10 pmol/µl	0.5
vanR_r [10 pmol/µl] /vanH_r[10 pmol/µl	0.5
DNA template	1.0
Total	10.0

Table 10: vanH/ vanR qPCR Program - Three step with melt

Pre-incubation	95 °C	5 min	
Amplification	95 °C	20 s	40 cycles
	60 °C	15 s	
	72 °C	20 s	
Final melt curve	72°C-96°C		

7. Antibiotic primer screening in metagenome

To test the binding capacity of the ARG primers in the metagenome dataset, bioinformatical screening with the software CLC main workbench 8.1.2, Qiagen Bioinformatics, was done in silico with 4 Aral Sea samples (A-52, A-402, A-Soil5, A-Soil40).

III. Results

1. Composition of bacterial communities in Dachstein and Aral Sea assessed by 16S rRNA gene amplicon sequencing

1.1 General information

The general diversity of all samples was assessed by 16S rRNA gene amplicon sequencing. In total these were 37 samples (biological replicates and 3 technical replicates of each sample). Initially, not all samples contained equal amounts of DNA, which was taken into consideration for PCR and an equimolar amount of PCR product was sent for sequencing. Figure 6 shows the amount of DNA that was extracted per gram pellet and differences between Dachstein and Aral Sea samples, but also between plant associated and soil samples can be analyzed.

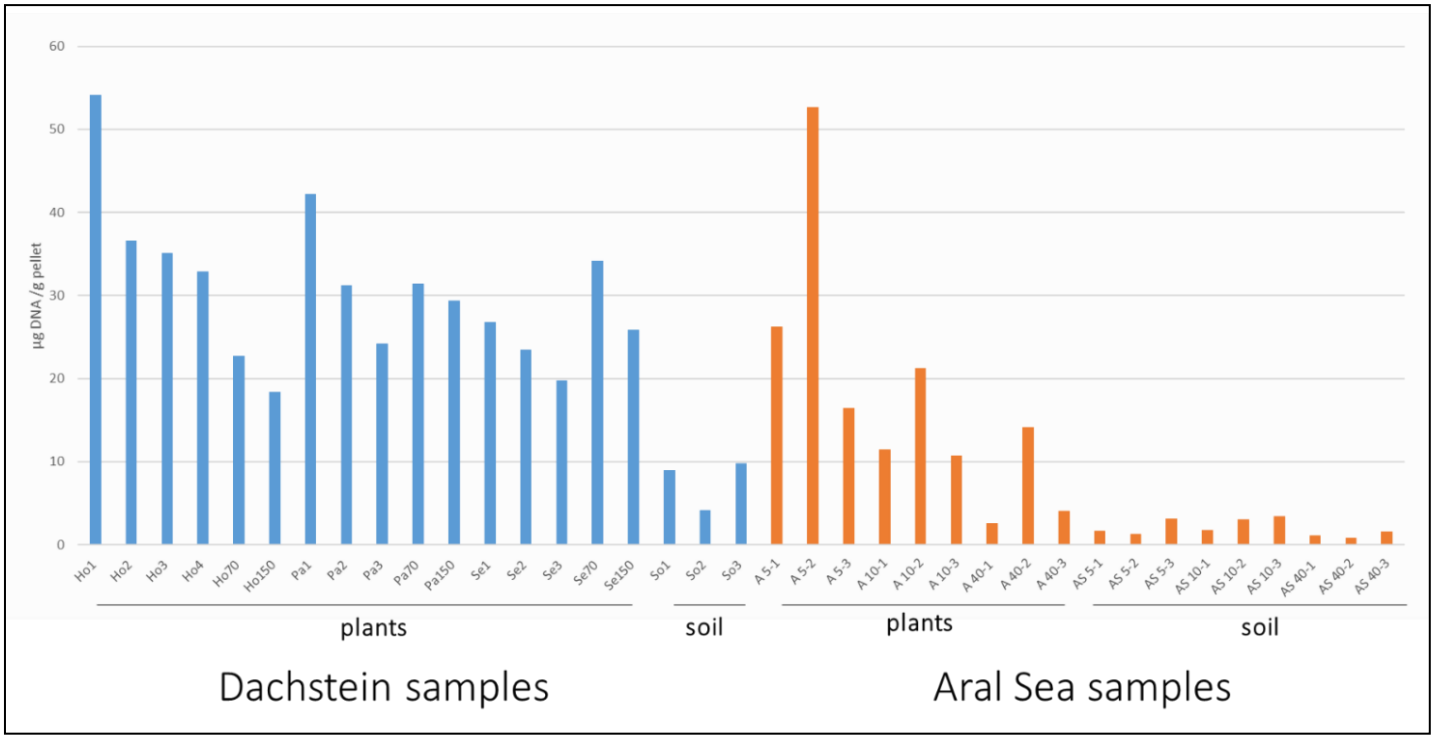


Figure 6: DNA content extracted out of 1 g pellet of each sample; Ho: *Hornungia alpina*, Pa: *Papaver alpinum*, Se: *Sedum atratum*, So: Dachstein soil, A: Aral Sea plant *Suaeda acuminata* , AS: Aral Sea soil; 1,2,3,4: 10 years without ice; 70: 70 years without ice, 150: 150 years without ice; ,5-1,5-2,5-3: 5 years without water; 10-1, 10-2, 10-3: 10 years without water, 40-1, 40-2, 40-3 : 40 years without water.

Three technical replicates per sample were sequenced, but we merged them into one dataset because these did not differ substantially from each other. These similarities within the technical replicates served as an additional internal control for reproducibility.

The total number of ASVs that were retrieved by QIIME2 was 31,349, out of these 23,882 originated from the Dachstein samples and 8,988 from the Aral Sea; 1,521 ASVs were shared in both data sets and mostly contained typical plant associated bacteria like *Burkholderiacaea*.

The Alpha diversity (Shannon index) of the Dachstein and Aral Sea samples shows a statistically significant (ANOVA, $p<5x10^{-7}$) diversity when comparing the Dachstein ASVs to the Aral Sea's ASVs, while the diversity of the Aral Sea samples differ greatly from each other; they can be seen in Figure 7.

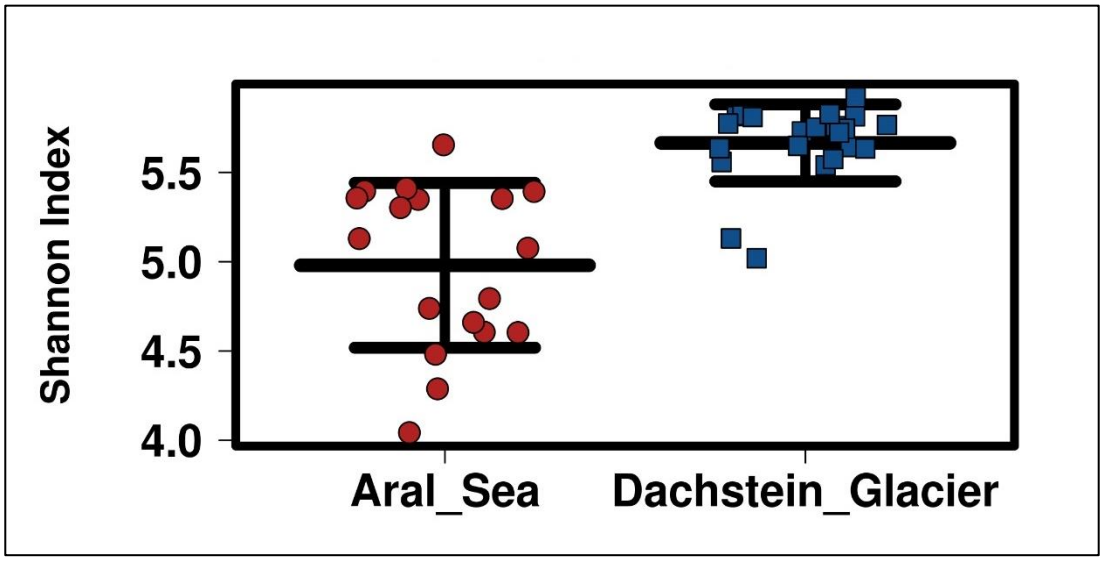


Figure 7: Alpha diversity Shannon index distribution of Dachstein and Aral Sea samples on ASV level; p-value (ANOVA): 3x10^-7; F=38; each symbol represents one sample.

1.2 Bacterial community structure of the Dachstein

The Dachstein samples all together contained 1,587 different taxa after collapsing the 23,882 ASVs on genus level (based on SILVA database release 132). The beta diversity of these taxa is shown as a barplot featuring the top 20 most abundant taxa in the biological Dachstein samples on genus level (Figure 8). No clear trend regarding an increase or decrease of taxa is visible.

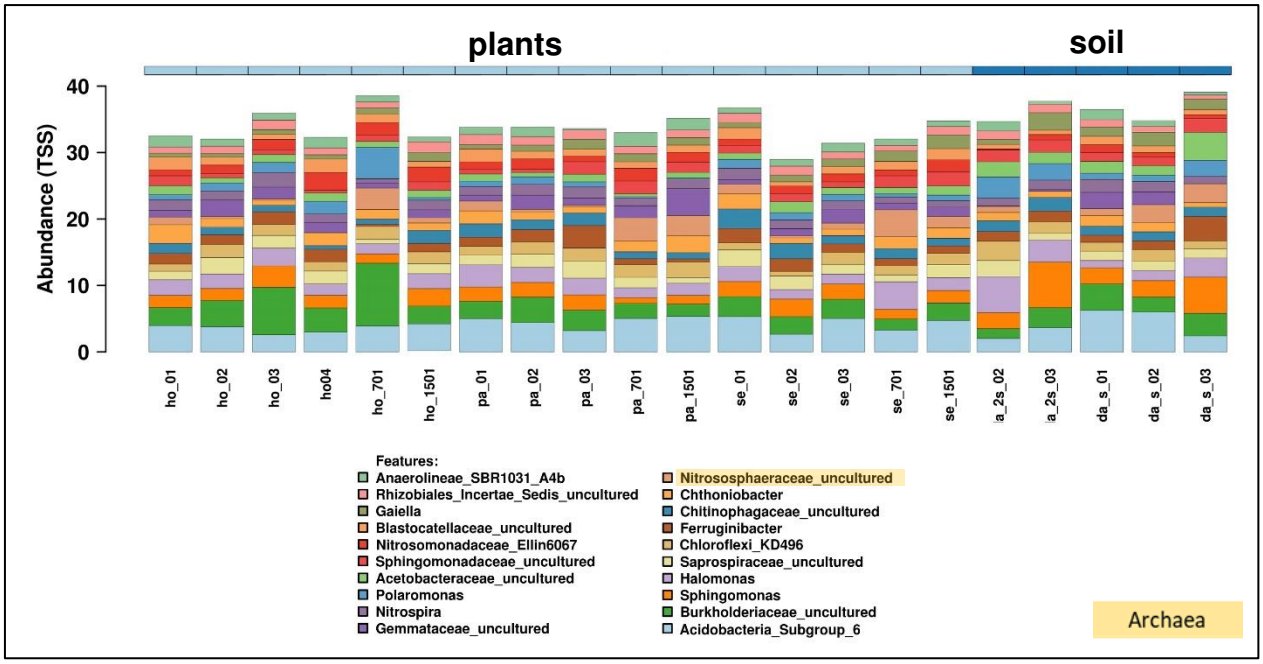


Figure 8: Bacterial community structure of the Dachstein – top 20 taxa on genus level; ; Ho: *Hornungia alpina*, Pa: *Papaver alpinum*, Se: *Sedum atratum*, ds_s: Dachstein soil; technical replicates merged; Archaea marked with yellow bar; 01,02,03,04: 10 years without ice; 70: 70 years without ice, 150: 150 years without ice; axis features values up to 40%, values up to 100% would be classified as "other" and were excluded here for visual clarity.

In order to further explore the similarities and dissimilarities of the found taxa, Figure 9 shows two PCoA (Principal Coordinates Analysis) plots of the Dachstein ASVs' beta diversity colored according to a) the years without ice and b) according to plant species *Hornungia alpina, Papaver alpinum, Sedum atratum* and soil. Similarities within the 70a and 150a samples can be observed and three of the soil samples show high similarity with each other, while portraying dissimilarity to the other three soil samples as well as the plant associated rhizosphere samples. Out of the three soil samples with high similarities (So3, 2So2, 2So3), two were sampled at the same location, while So3 was sampled at a different location. The high similarity can therefore not be explained by geographical proximity.

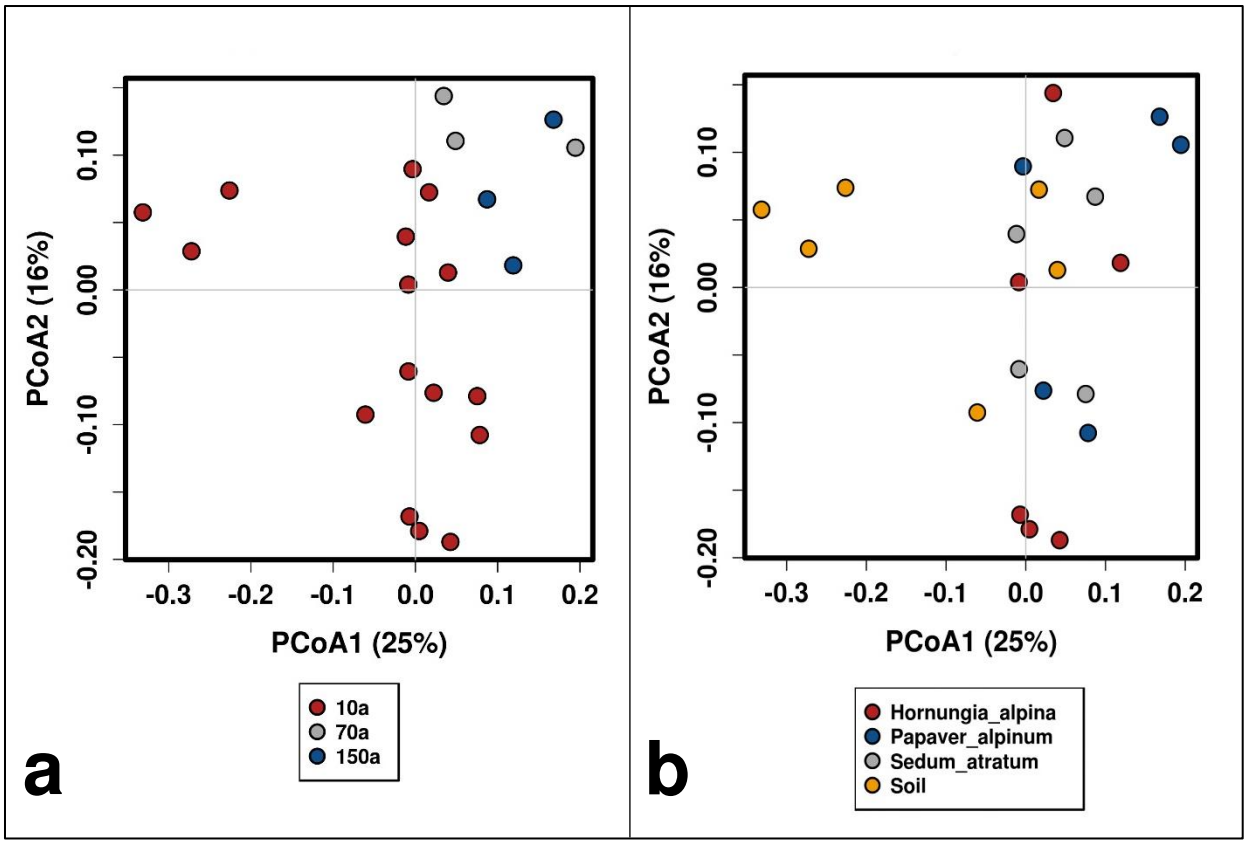


Figure 9: Principal coordinates analysis (PcoA) beta diversity plots of the bacterial communities of the Dachstein analysed by 16S rRNA amplicon sequencing on ASV level based on Bray-Curtis dissimilarity. a: points in the coordinate system colored according to years without ice, red: 10 years, grey: 70 years, blue: 150 years. b. points in the coordinate system colored according to plant species, red: *Hornungia alpina*, blue: *Papaver alpinum*, grey: *Sedum atratum*, yellow: soil; each dot represents one biological replicate.

In order to get a complete picture of the diversity present in the Dachstein glacier samples, alpha diversity showing the mean species diversity is crucial. The alpha diversity distribution of the Dachstein samples can be seen in Figure 10. Here we can see that the ASVs of the plant samples show in average a statistically relevant higher Shannon Index (ANOVA, p=0.039) and a standard deviation lower than in the soil samples. In Figure 11 the plant samples are further divided into the three plant species and show a high alpha diversity within all plant species with the highest diversity in *Sedum atratum*; also statistically relevant. The soil samples, like before, show a lower alpha diversity in general and a wide distribution of Shannon index values.

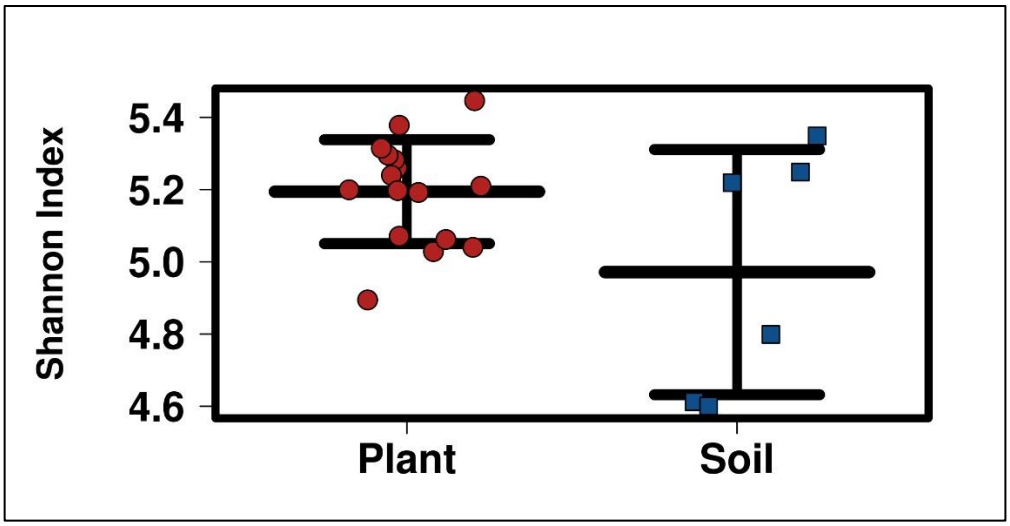


Figure 10: Dachstein alpha diversity Shannon index distribution of plant and soil grouped samples on ASV level. Biological samples clustered in plant and soil samples; each symbol represents one sample; p-value (ANOVA)= 0.039, F=4.9.

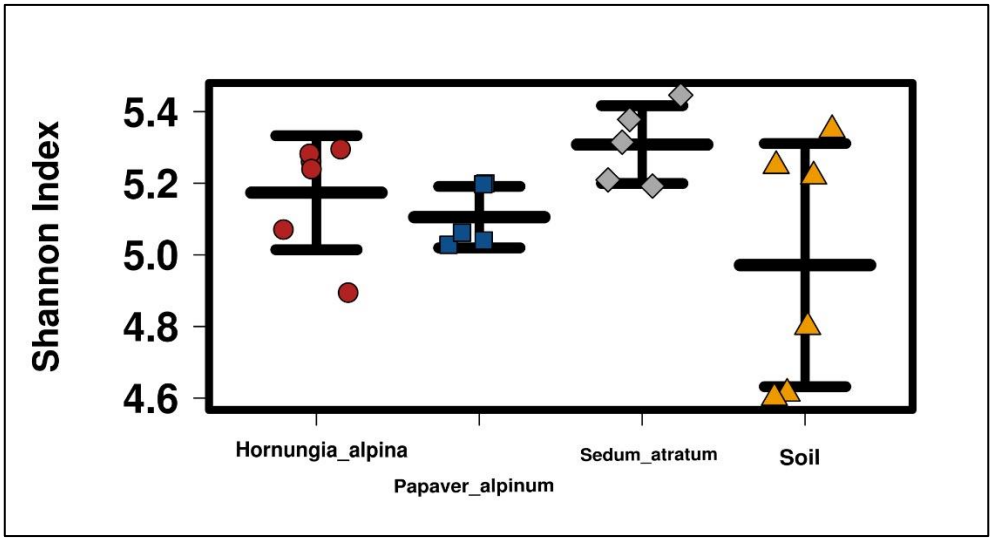


Figure 11: Dachstein alpha diversity Shannon index distribution of plant and soil grouped samples on ASV level. Samples clustered in plant species *Hornungia alpina, Papaver alpinum, Sedum atratum* and soil; each symbol represents one sample; p-value (ANOVA) = 0.093, F= 2.5.

To decipher which taxa cause observed differences and to show species that are highly associated with features such as plant species or years, LEfSe method was applied. Ten taxa appear to be associated with soil samples, four taxa with *Sedum atratum* samples, six taxa with *Papaver alpinum* and five taxa with *Hornungia alpina* (Figure 12). Additionally, five genera are significantly more abundant, reaching an LDA score of slightly under 4. *Blastocatella* is significantly enriched in soil samples of the Dachstein, while *Rhodanobacteraceae* are found more often in *Sedum atratum* samples, *Pseudomonas* appears to be enriched in *Papaver alpinum* samples, and

Nitrosomonadaceae (Ellin6067) is more likely to be found in *Hornungia alpina* samples. The significance of these correlations can be seen in the appendix 4.1.

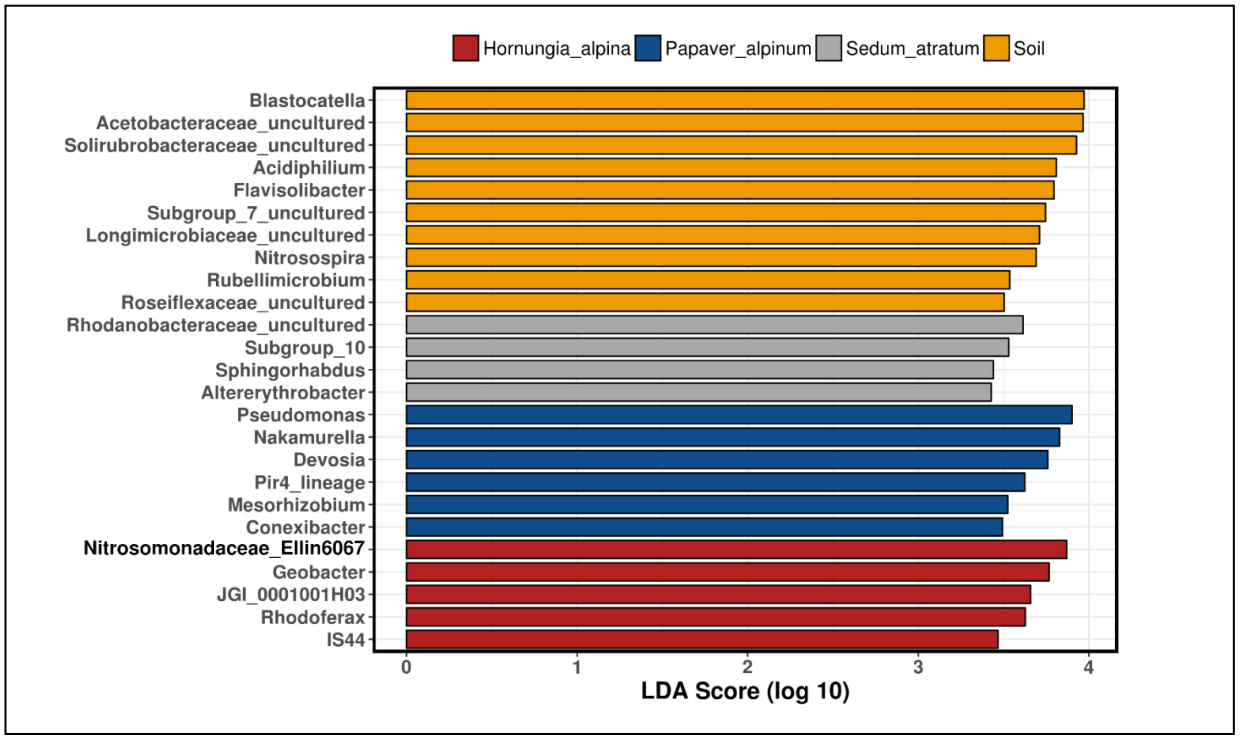


Figure 12: Linear discriminant analysis effect size (LEfSe) plot of the Dachstein samples showing LDA score values clustered in plant species and soil; taxa on genus level.

To get a better picture of taxa associated with the rhizosphere samples and because the soil samples were taken at the 10 years sampling spot only, the soil samples were omitted for the next analysis and the remaining data was used for LEfSe method. The following Figure 13 depicts the Dachstein dataset of the plant associated samples clustered in 10 years, 70 years and 150 years without ice. The bacterial taxa in the rhizosphere of the plants change over the years and every time point has different highly associated bacterial taxa. Nine different taxa appear to be associated with 10a samples, six with 70a samples and five with 150a samples. Interestingly, *Nitrososphaeraceae* is found to be enriched in the plants 70 years without ice, is an Archaea. The top 20 most abundant taxa in the Dachstein samples feature *Nitrososphaeraceae* as well, as the only Archaeal taxon occurring in such high numbers. According to LEfsE, *Nitrososphaeraceae* appear to be highly associated with the 70a plant samples, a finding that can also be seen in the barplot. Again, the significance of these correlations can be seen in the appendix 4.1.

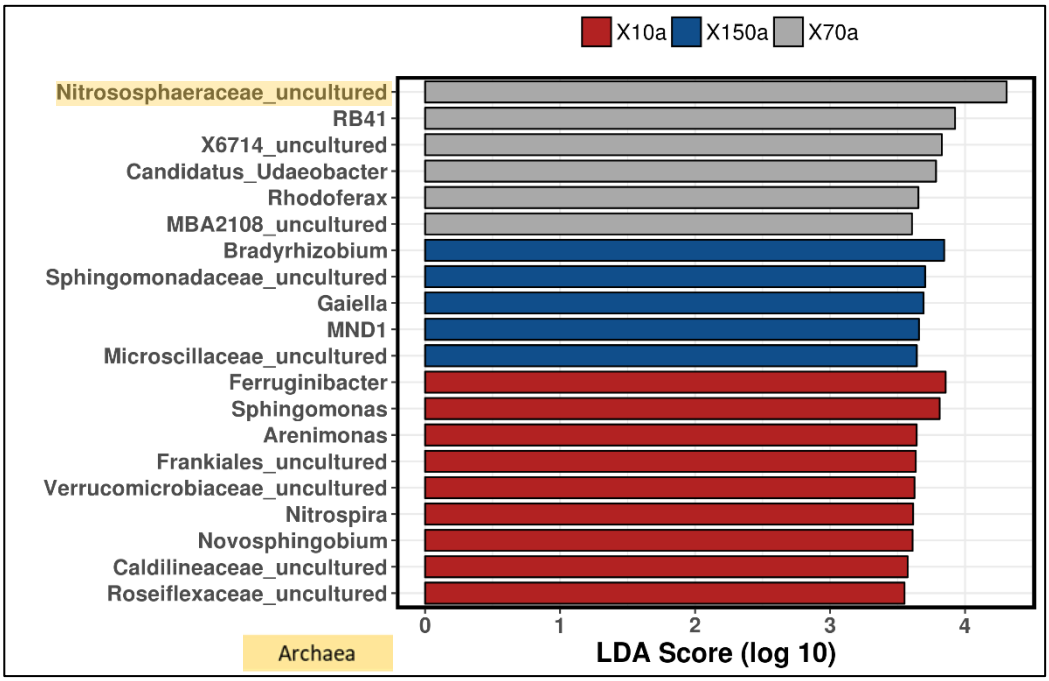


Figure 13: Linear discriminant analysis effect size (LEfSe) plot of the Dachstein plant only dataset– comparing the samples based on the time when the specific sampling sites emerged from the glacier 10 years (X10a, red), 70 years(X70a, blue), 150 years(X150a, grey); taxa on genus level; Archaea marked with orange bars.

PERMANOVA compares the diversity of each sample against each other and provides a value according to their similarity. In addition to that, permutations are used to provide further statistical relevance. This method was used to show the clustering of all plant and soil distance values (Figure 14. a) and clustering in years without ice (Figure 14. b). In a) an overlap of the interquartile range as well as a mean variation of the values can be observed. In b. we see smaller variations especially int the 70a and 150a samples, but these occur due to the low sample number and lack statistical relevance.

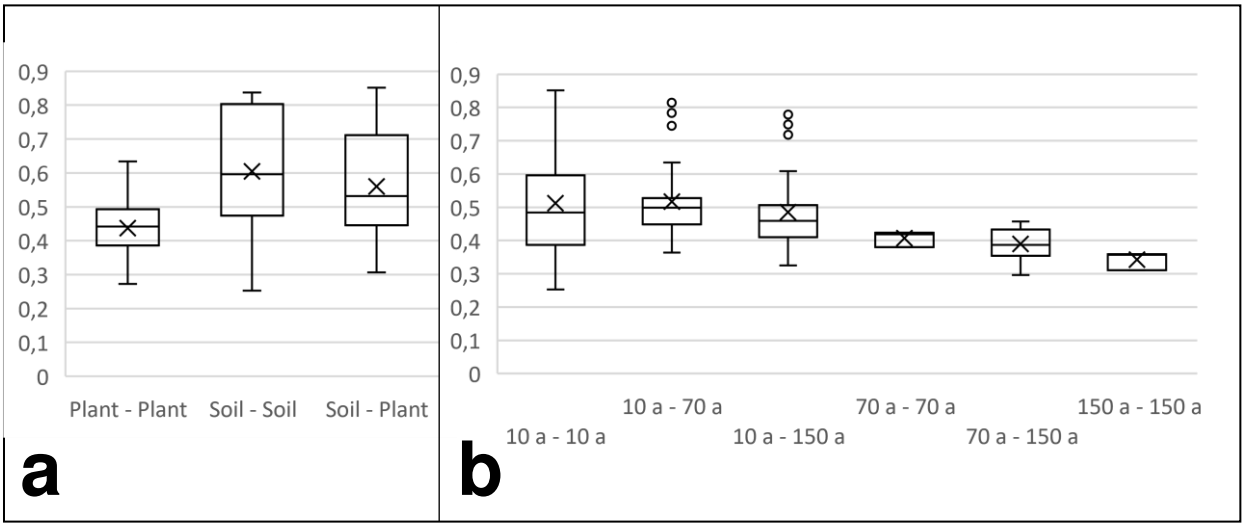


Figure 14: Dachstein PERMANOVA paired with 999 permutations; a: plant : soil; b: 10 years, 70 years, 150 years without ice; box represents second and third quartile; whiskers mark the minimum and maximum excluding outliers; outliers marked with dots.

1.3 Composition of bacterial communities in the Aral Sea assessed by 16S rRNA gene amplicon sequencing

The Aral Sea samples all together contained 827 different taxa after collapsing the 8,988 ASVs on genus level (based on SILVA database release 132). The top 20 must abundant taxa of these samples can be seen in a barplot featuring the beta diversity in the biological replicates of the Aral Sea samples on genus level (Figure 15). Here compositional differences between plant associated and soil samples as well as between the years without water can be observed.

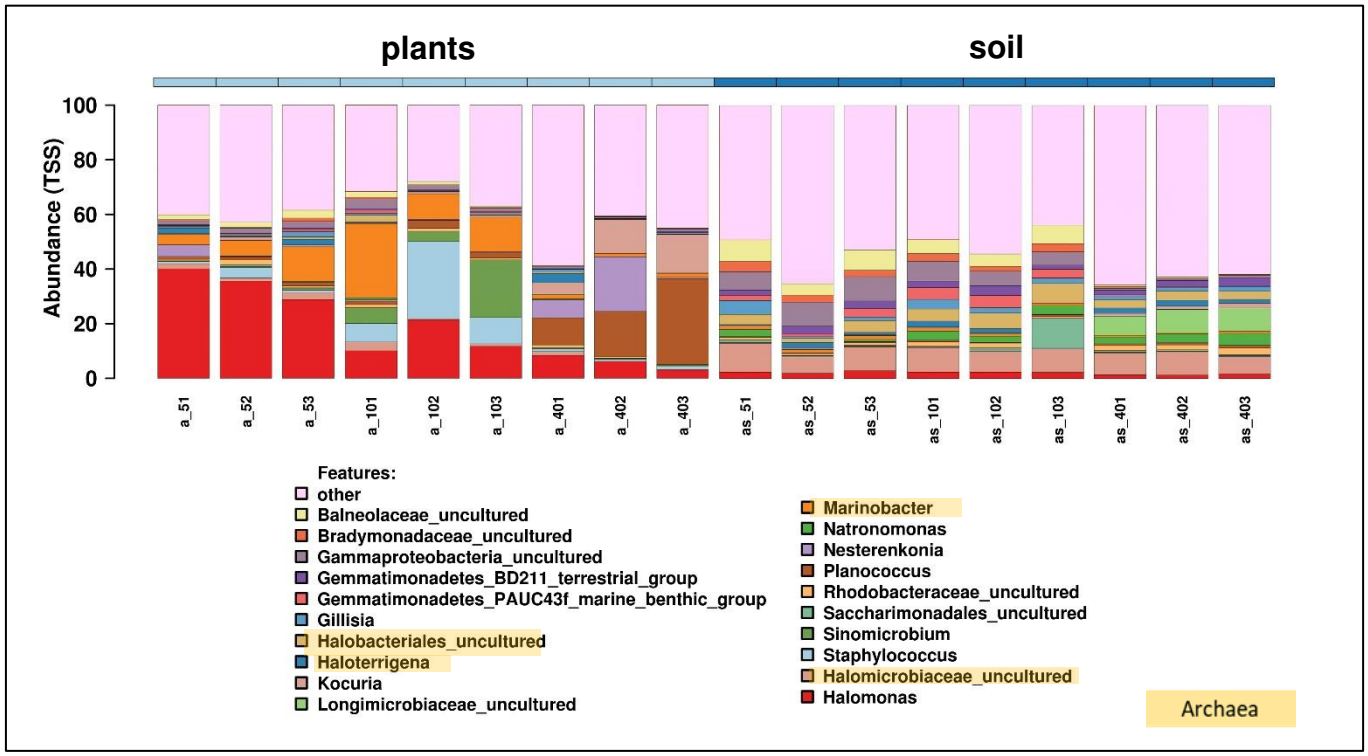


Figure 15: : Bacterial community structure of the Aral Sea– top 20 taxa on genus level; a*: Suaeda acuminata*, as: Aral Sea soil,51,52,53: 5 years without water; 101, 102, 103: 10 years without water, 401, 402, 403 : 40 years without water; other: remaining taxa (minus top 20); technical replicates merged; Archaea marked with yellow bars.

To investigate the distribution of taxa present within the Aral Sea samples, the alpha diversity showing the mean species diversity is one important part of biodiversity studies. The alpha diversity distribution (Shannon index) of the Aral Sea samples can be seen in

Figure *16*, where a higher Shannon Index can be observed within the soil samples when comparing to the plant samples with lower Shannon Index values. The diversity differences between the plant and soil samples is statistically relevant (ANOVA p-value = 5.6x10^-7). Although seeing striking

differences in the top 20 most abundant taxa when comparing 5a, 10a and 40a samples in Figure 15, clustering in 5a+10a vs 40a samples shows insignificant p-values consulting multiple statistical methods (Shannon p-value: 0.45; richness/Chao1 p-value: 0.69; evenness p-value: 0.5).

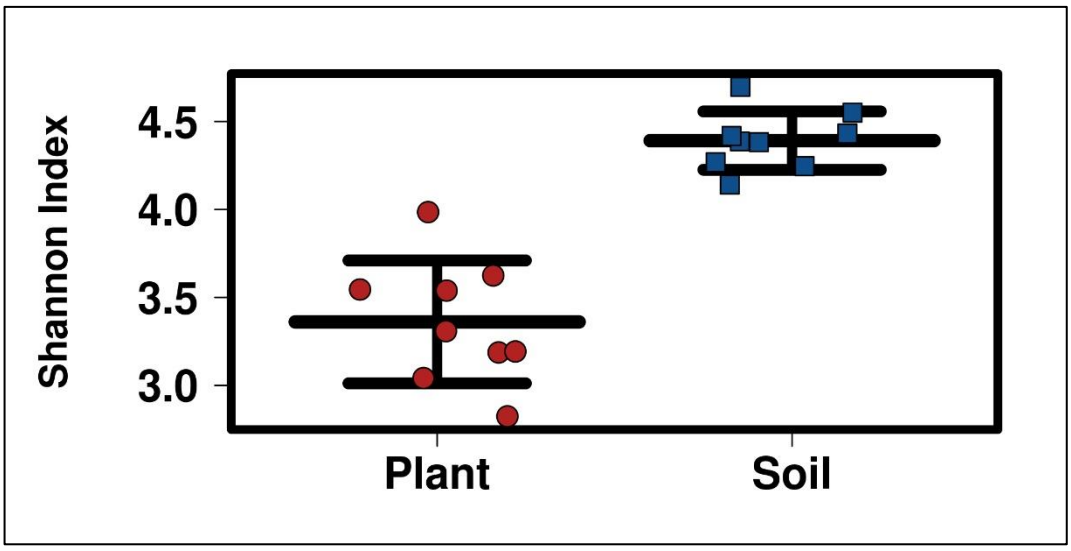


Figure 16: Aral Sea alpha diversity Shannon index distribution of plant and soil grouped samples on ASV level. Samples clustered in plant and soil samples; each symbol represents one sample; p-value (ANOVA) = 5.6x10^-7, F=64.

To visualize similarities and dissimilarities between the Aral Sea data, Principal Coordinates Analysis (PCoA) was used. The clustering of the beta diversity of the Aral Sea ASVs can be seen as PCoA plots in Figure 17 colored according to a) origin of the sample (rhizosphere of *Suaeda acuminata* or soil) and b) the years without water. Here, a more distinct clustering of the samples' diversity and similarity takes place than in the Dachstein samples. The diversity clusters of 40a samples are clearly distinguishable from the other cluster formed by samples which are 5-10 years without water based on the PCoA2 axis. Additionally, rhizosphere and soil samples can be clearly distinguished from each other based on the PCoA1 axis. The similarity of the 5a and 10a samples is stronger impacted if they originated from plant or from soil, leading to mixed clusters of 5a and 10a samples in Figure 17 b.

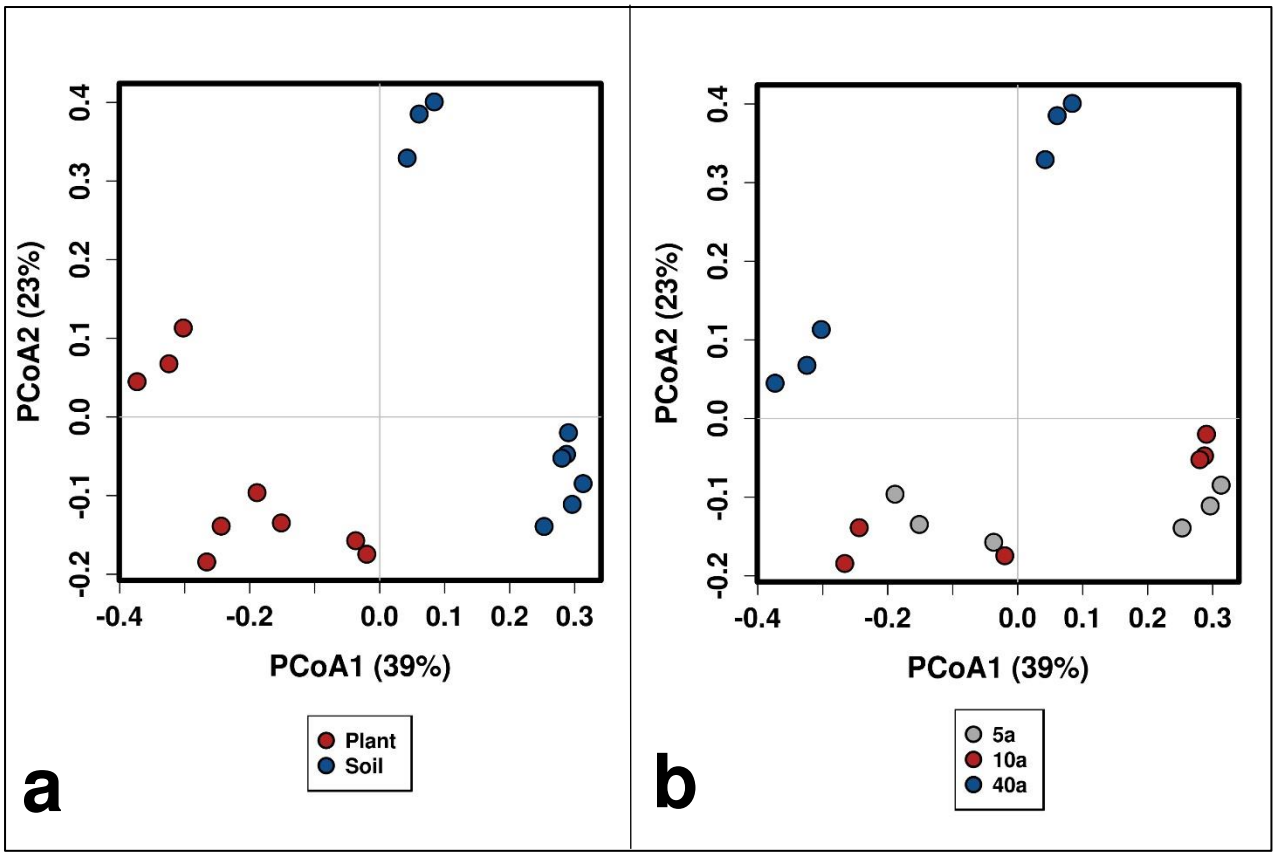


Figure 17: Principal coordinates analysis (PcoA) beta diversity plots of the Aral Sea on ASV level based on Bray-Curtis dissimilarity. a: points in the coordinate system colored according to plant or soil b: points in the coordinate system colored according to years without water, grey: 5 years, red: 10 years, blue: 40 years; each dot represents one biological replicate.

To identify taxa which are highly associated with features such as plant species or years without water, LEfSe method was applied. Regarding the LEfSe plot clustering the samples in plant and soil (Figure 18), 20 taxa appear to be associated with the soil samples, while eleven taxa are shown to be highly associated with the plant associated samples. Furthermore, it can be observed that *Halomicrobiaceae* are significantly enriched in soil samples of the Aral Sea, while *Halomonas* are found more often in the plant associated samples. The significance of these correlations can be seen in appendix 4.2.

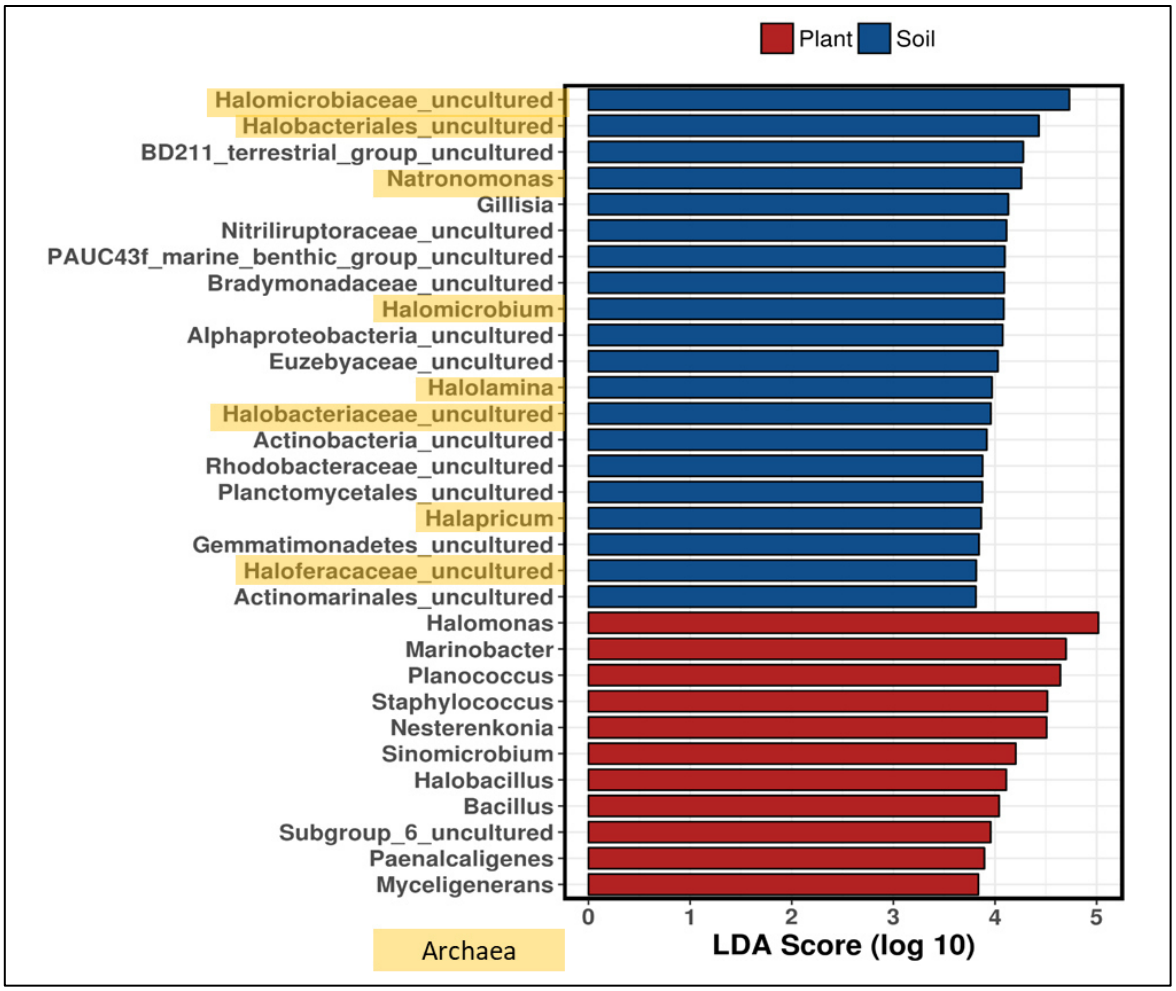


Figure 18: Linear discriminant analysis effect size (LEfSe) plot of the Aral Sea samples, clustered in plant and soil; plant associated samples (red), soil samples (blue); taxa on genus level; Archaea marked with yellow bars.

Because of the striking differences between plant associated samples and soil samples in the beta diversity, the datasets were split and separately analyzed as LEfSe plots to examine which genera were more abundant within each dataset. Figure 19 shows five taxa from the plant only dataset highly associated with the 5a feature, one taxon appears to be associated with 10a samples and three taxa are enriched in the 40a samples. Furthermore, we see, that *Halomonas* shows the highest association with the 5a samples, while *Myceligenerans* is highly enriched in the 10a samples and *Solirubrobacter* in the 40a samples.

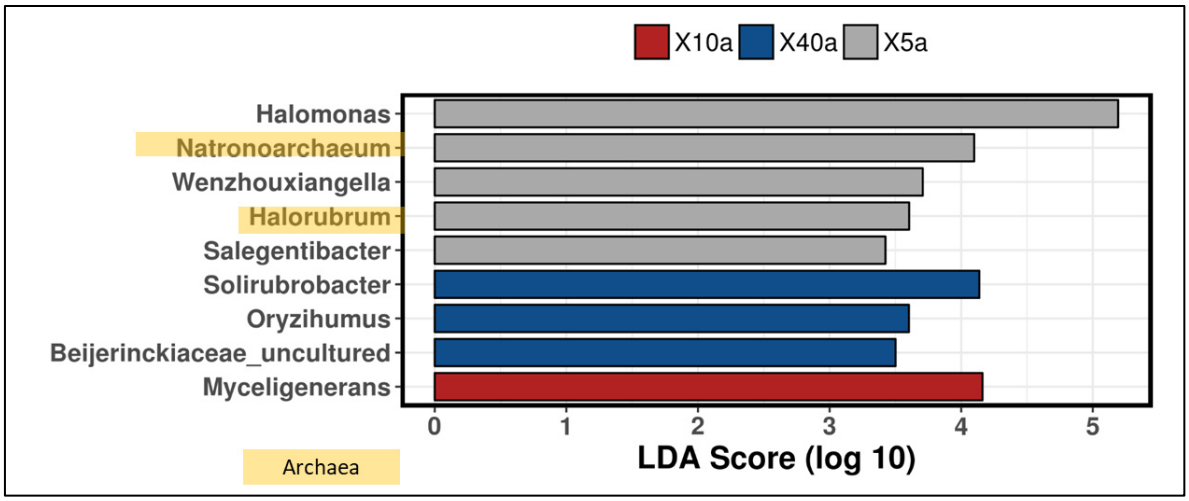


Figure 19: Linear discriminant analysis effect size (LEfSe) plot of the Aral Sea plant only associated to the time, the respective sampling sites have been without water: 5 years (X5a, grey), 10 years(X10a, red), 40 years(X40a, blue); taxa on genus level; Archaea marked with yellow bars.

To complete the picture, the highest associated taxa from the soil only dataset depicted as a LEfSe plot feature six different taxa for 5a samples, two taxa for 10a and six taxa for the 40a samples. *Gammaproteobacteria* showing the highest association with the feature 5a, *PAUC43f,* a subgroup of *Gemmatimonadetes,* for the 10a samples and *Longimicrobiaceae* for the 40a samples (see Figure 20).

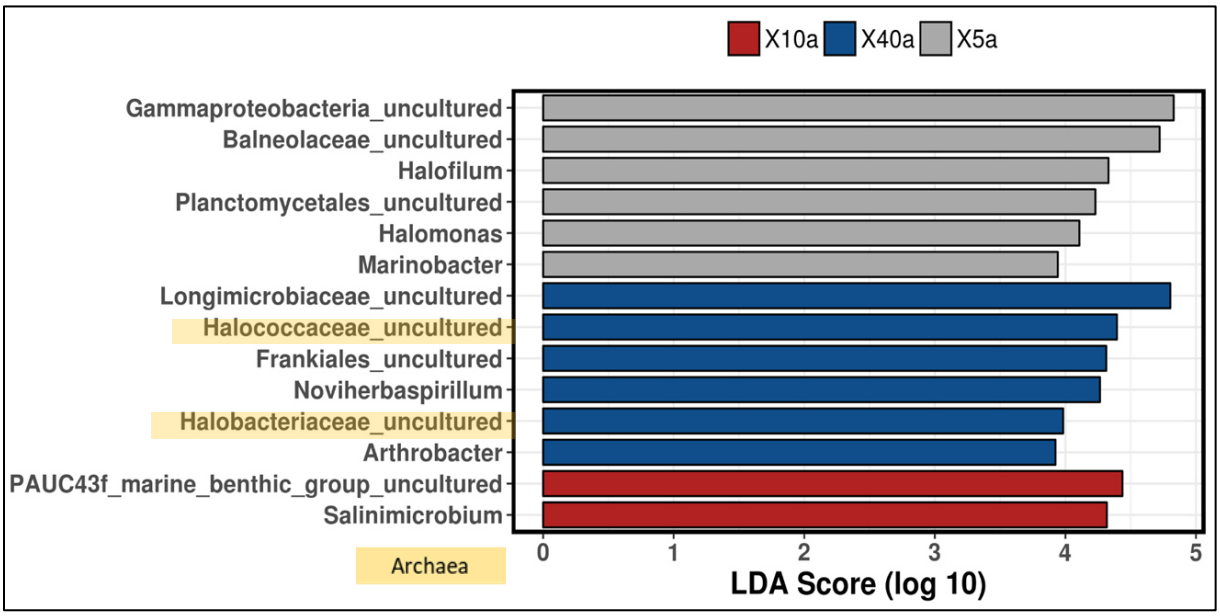


Figure 20: Linear discriminant analysis effect size (LEfSe) plot of the Aral Sea soil only associated to the time the respective sampling sites have been without water: 5 years (X5a, grey), 10 years(X10a, red), 40 years(X40a, blue); taxa on genus level; Archaea marked with yellow bars.

PERMANOVA was used to compare the diversity of each sample against each other and to obtain a value according to their similarity. In addition to that, permutations are used to provide further statistical relevance. This method was used to show the Aral Sea data clustered according to plant and soil (Figure 21 a) as well as clustered according to the years without water (Figure 21 b). In Figure 21 a as well as in Figure 21 b, overlaps of the interquartile range as well as the mean variation can be observed. The diversity does not depict statistically relevant dissimilarity.

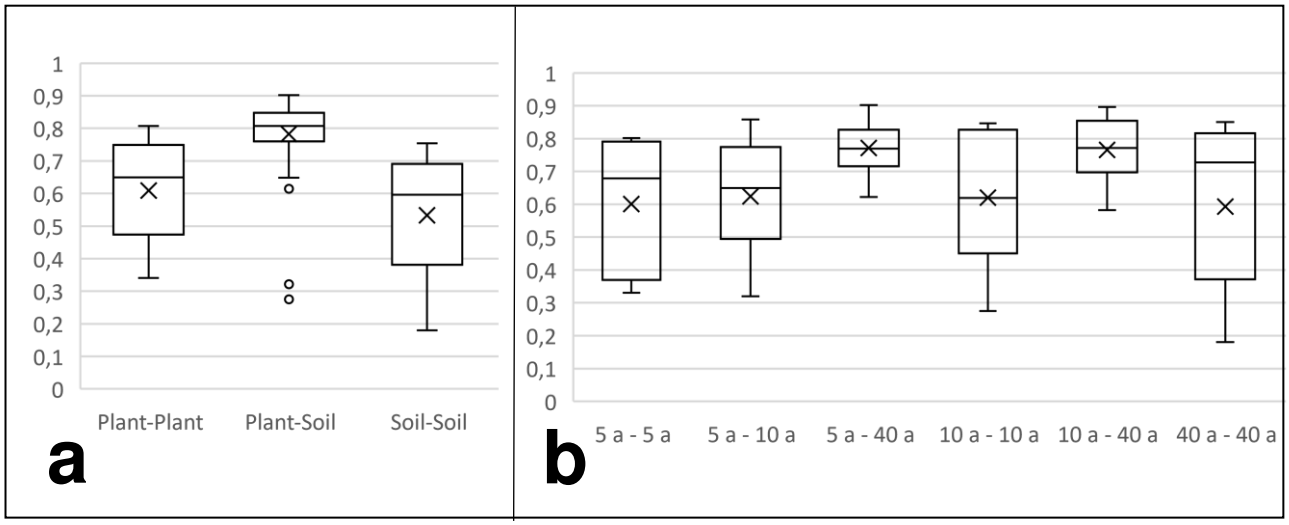


Figure 21: Aral Sea PERMANOVA paired with 999 permutations; a: plant : soil; b: 5 years, 10 years, 40 years without water; box represents second and third quartile; whiskers mark the minimum and maximum excluding outliers; outliers marked with dots.

2. Antibiotic resistance screening with primers

2.1 PCR based detection of antibiotic resistance genes

In order to detect antibiotic resistance genes in both the Dachstein and Aral Sea samples, PCRs with primers targeting antibiotic resistance genes in agricultural and environmental samples were conducted. After screening the Dachstein and the Aral Sea samples, the two primer pairs vanH and vanR displayed positive results. VanH produced clear bands at expected length in all of the Dachstein and Aral Sea samples, while vanR produced clear bands in all Aral Sea samples and in the Dachstein samples Pa3, Se1, Se2, Se3, Se70, Se150, So1, So2 and So3. All of the other primers showed either unspecific bands on agarose gel or no bands at all.

2.2 Copy number of vanR gene assessed via qPCR

After the positive PCR results of vanH and vanR, qPCRs were done with both primer pairs. In both cases, however, differences between the melting curves of the standard and the samples appeared. While the standards melted around 80° C, the samples melted between 85-92° C. A check of the qPCR amplicons on agarose gel showed multiple and unspecific bands, in contrast to the single bands we saw with normal PCR. However, the theoretical abundances of Vancomycin resistance genes (assessed with vanR primers) in the Dachstein samples were calculated to show a understanding of the procedure; the results can be seen in Figure 22.

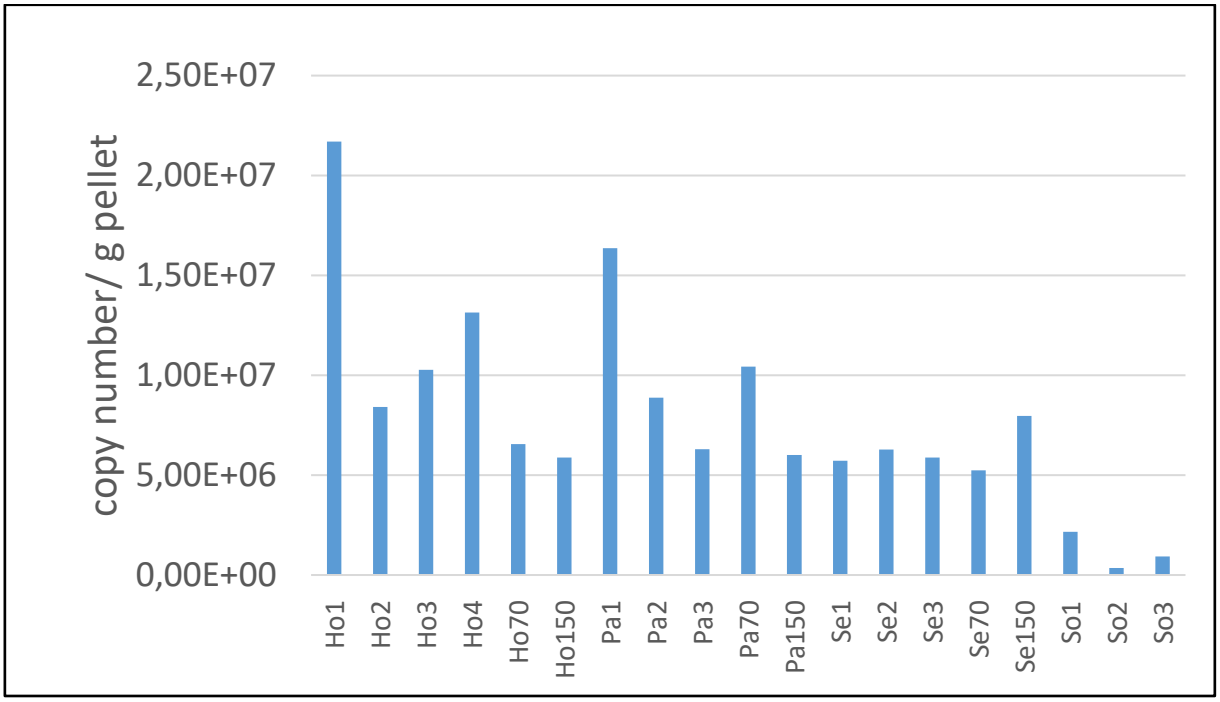


Figure 22: Theoretical vanR gene abundance in the Dachstein samples calculated from the Ct values of the standard curve in copy number per g of the pellet; Ho: *Hornungia alpina*, Pa: *Papaver alpinum*, Se: *Sedum atratum*, So: soil; 1,2,3,4: 10 years without ice; 70: 70 years without ice, 150: 150 years without ice.

3. Bacterial composition and resistomal information assessed by metagenomic analysis

3.1 Alpha and beta diversity of the metagenomes

Figure 23a shows the top 20 taxa of the 4 Aral Sea samples obtained through the metagenome workflow; as a comparison; b.) features the 16S rRNA amplicon data already shown in Figure 15 for the 4 samples used for metagenome sequencing. As the soil samples of one sample spot were merged

for the metagenome analysis, but not for the amplicon sequencing, the soil samples appear merged in a.), but as triplicates in b.).

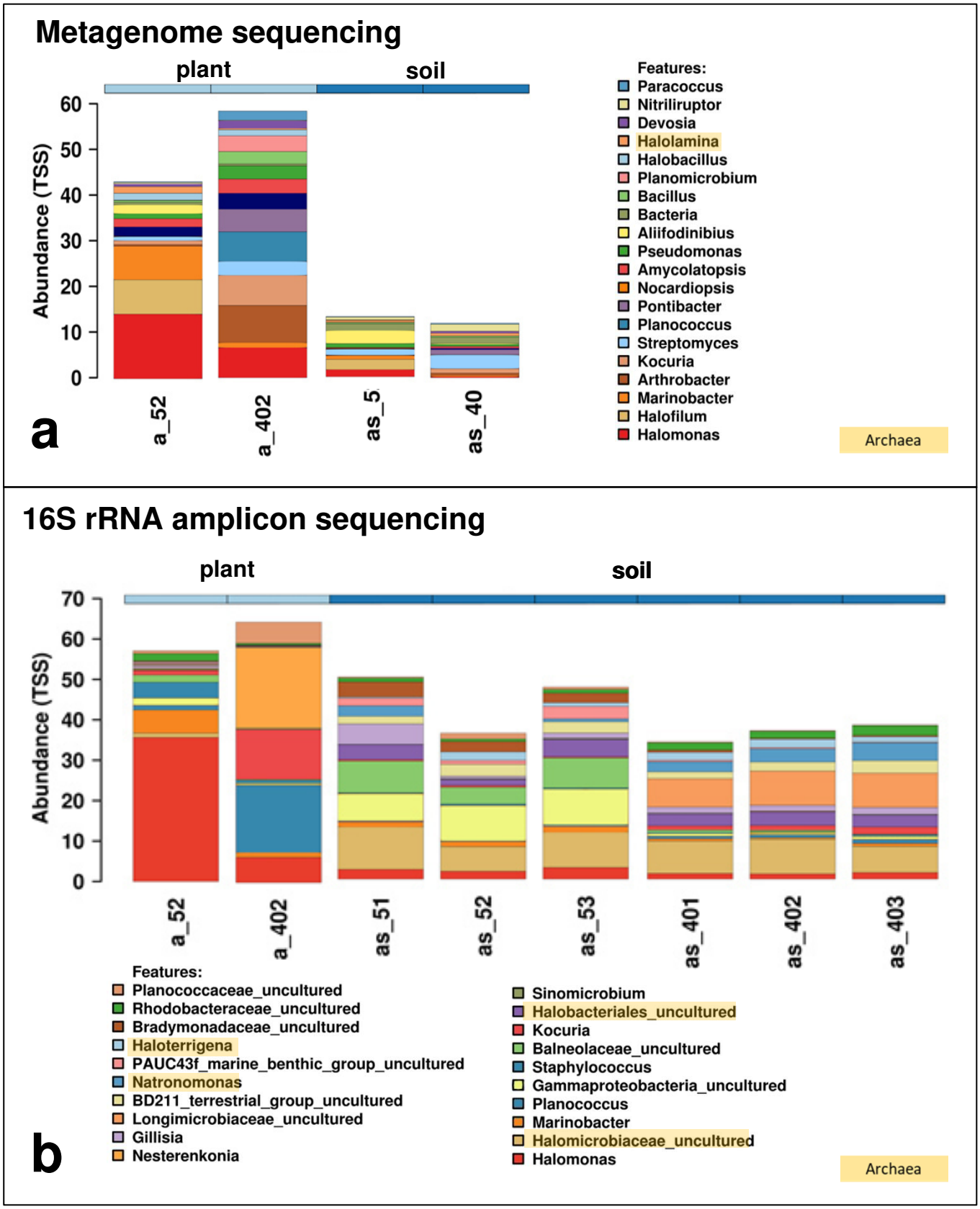


Figure 23: Bacterial composition of the Aral Sea - top 20 taxa barplot of metagenome sequencing (a) and 16S rRNA amplicon sequencing (b) in comparison – top 20 taxa on genus level; a: Aral Sea plant sample; as: Aral Sea soil sample; 51,52,53: 5 years without water, 401, 402, 403 : 40 years without water; Archaea marked with orange bars.

In order to complete the picture of the diversity present in the four metagenome samples, a glimpse at the alpha diversity is needed. The alpha diversity of the metagenomic data clustered in plant associated and soil samples can be seen in Figure 24, while the clustering in 5a and 40a samples is displayed in Figure 25. The sample size of four in total is quite low for statistical relevance, however, the clustering in plant and soil does show significance with a p-value below 0.05.

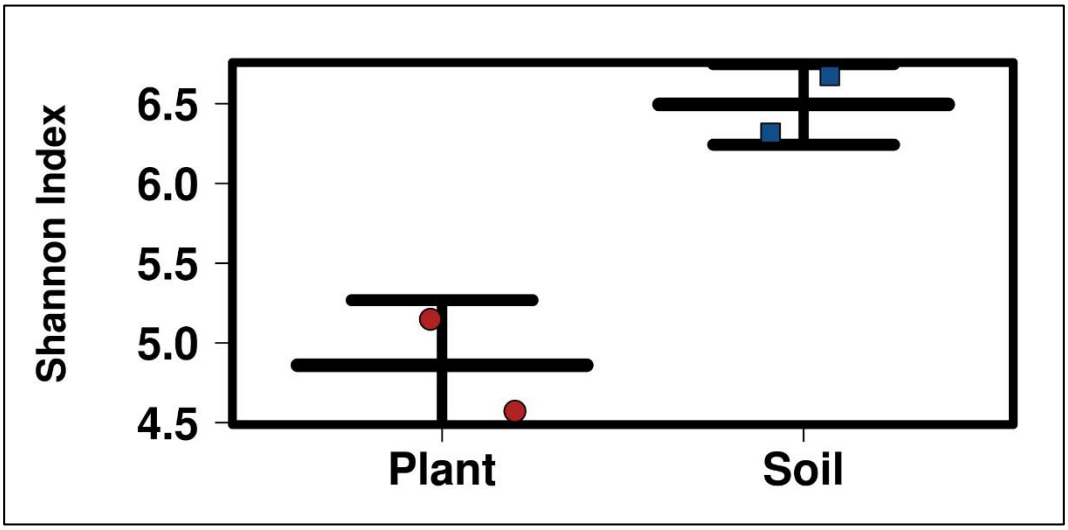


Figure 24: : Aral Sea alpha diversity Shannon index distribution of the metagenome data, grouped in plant and soil samples on genus level; p-value = 0.04, F=23 (ANOVA); each symbol represents one biological sample.

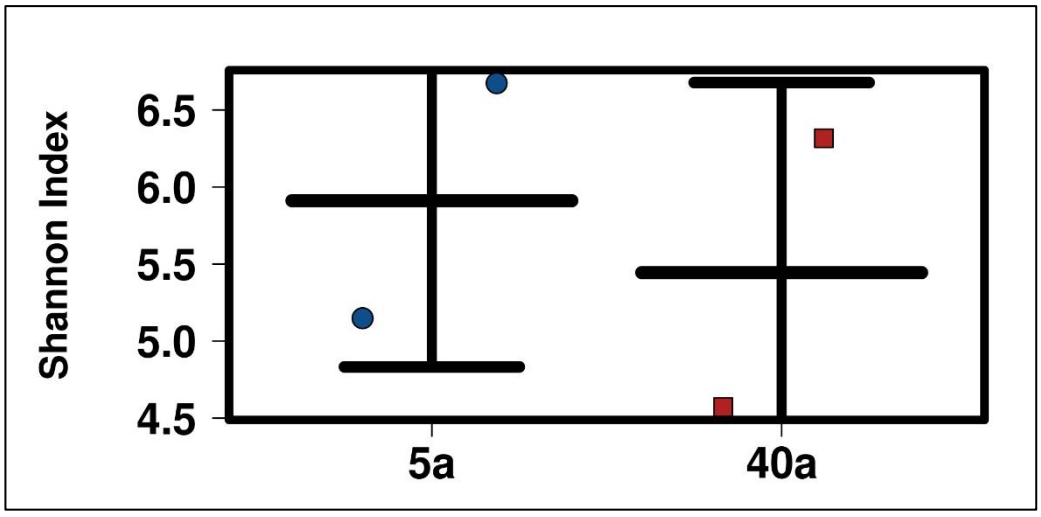


Figure 25: : Aral Sea alpha diversity Shannon index distribution of the metagenome data associated to the time the respective sampling sites have been without water: 5 years (blue) and 40 years(red); samples on genus level; p-value = 0.73, F=0.16 (ANOVA); each symbol represents one biological sample.

3.2 Antibiotic resistance discovered in metagenomes

The screening of the metagenomes against the CARD (version 3.0.7) and wildCARD (version 3.0.6) databases provided information about the antibiotic resistance potential of the samples, the abundance of the genes as well as the different resistance mechanisms. The samples used for the metagenomic analysis are A-52 (5a Aral Sea plant sample), A-402 (40a Aral Sea plant sample), A-Soil5 (5a Aral Sea soil sample), A-Soil40 (40a Aral Sea soil sample) and Ho1 (10a Dachstein plant sample *Hornungia alpina*). Table 11 gives information about the total quality filtered reads of the metagenome samples as well as the RGI reads, while showing percentage values. It can be observed that some samples contained more antibiotic gene reads in relation to the total read count, e.g. A-52 features the highest total reads according to RGI and the highest percentage when comparing RGI and total filtered reads. Interestingly, both soil samples show lower percentage and total RGI reads than the plant associated samples.

Table 11: Total RGI hits (unfiltered) in relation to quality filtered reads.

	A-52	*A-402*	*A-Soil5*	*A-Soil40*	*Ho1*
total reads RGI	86,002	71,776	30,950	26,074	54,022
total filtered reads FASTQC	49,626,559	44,031,630	50,770,435	37,697,305	55,560,928
Percentage RGI to FASTQC reads	0.17%	0.16%	0.06%	0.07%	0.10%

In order get a closer look at the antibiotic resistance genes abundant in the metagenome samples, the top five most abundant resistance genes for each sample in absolute values are shown in Table 12. For better comparison, the percentage values of top five most abundant resistance genes can be seen in Figure 26.

Table 12: Top five most abundant resistance genes of each sample (unfiltered) in absolute values.

sample	*ARO term*	*reads*
A-52	adeF	38,928
	sul1	4,919
	Bifidobacterium adolescentis rpoB conferring resistance to rifampicin	4,904
	AAC(3)-IIb	3,758
	msbA	3,123
	total	**86,002**

A-402	Bifidobacterium adolescentis rpoB conferring resistance to rifampicin	10,234
	Streptomyces rishiriensis parY mutant conferring resistance to aminocoumarin	9,537
	Nocardia rifampicin resistant beta-subunit of RNA polymerase (rpoB2)	8,644
	Bifidobacteria intrinsic ileS conferring resistance to mupirocin	7,018
	adeF	6,846
	total	**71,776**
A-Soil5	kdpE	7,192
	sul1	4,176
	adeF	4,089
	Bifidobacterium adolescentis rpoB conferring resistance to rifampicin	2,609
	AAC(3)-IIb	2,455
	total	**30,950**
A-Soil40	kdpE	6,184
	Nocardia rifampicin resistant beta-subunit of RNA polymerase (rpoB2)	5,604
	Bifidobacterium adolescentis rpoB conferring resistance to rifampicin	2,992
	adeF	1,541
	Streptomyces rishiriensis parY mutant conferring resistance to aminocoumarin	1,494
	total	**26,074**
Ho1	adeF	18,583
	Bifidobacterium adolescentis rpoB conferring resistance to rifampicin	3,946
	Nocardia rifampiciin resistant beta-subunit of RNA polymerase (rpoB2)	3,330
	kdpE	3,120
	AAC(3)-IIb	2,970
	total	**54,022**

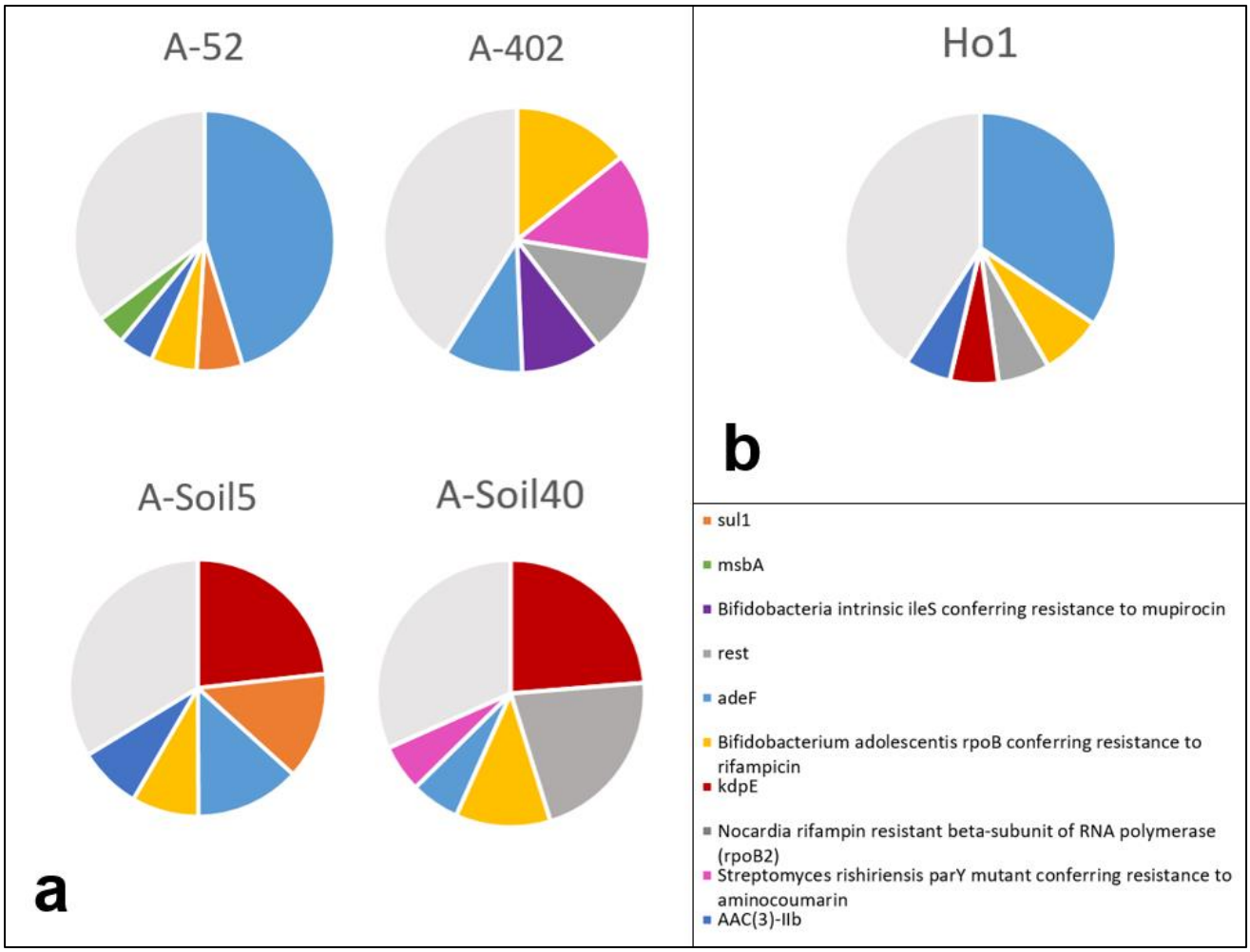


Figure 26: Top five most abundant antibiotic resistance genes in each sample presented by their relative abundance; a: resistance genes of the Aral Sea samples; b: resistance genes of Dachstein sample *Hornungia alpina* rhizoshpere 1.

Regarding the treatment of antibiotic resistant bacteria, knowledge of resistance mechanisms is crucial. Also, for this study, we took interest in the abundance of antibiotic resistance mechanisms present in the resistome of the samples. In Figure 27, the focus lies on the different mechanisms occurring in the metagenomes of the four Aral Sea samples (a) and the Dachstein sample (b). A high abundance of efflux pumps can be observed in A5-2, A-Soil5 and Ho1, while both 40a samples feature high proportion of target alteration and target replacement.

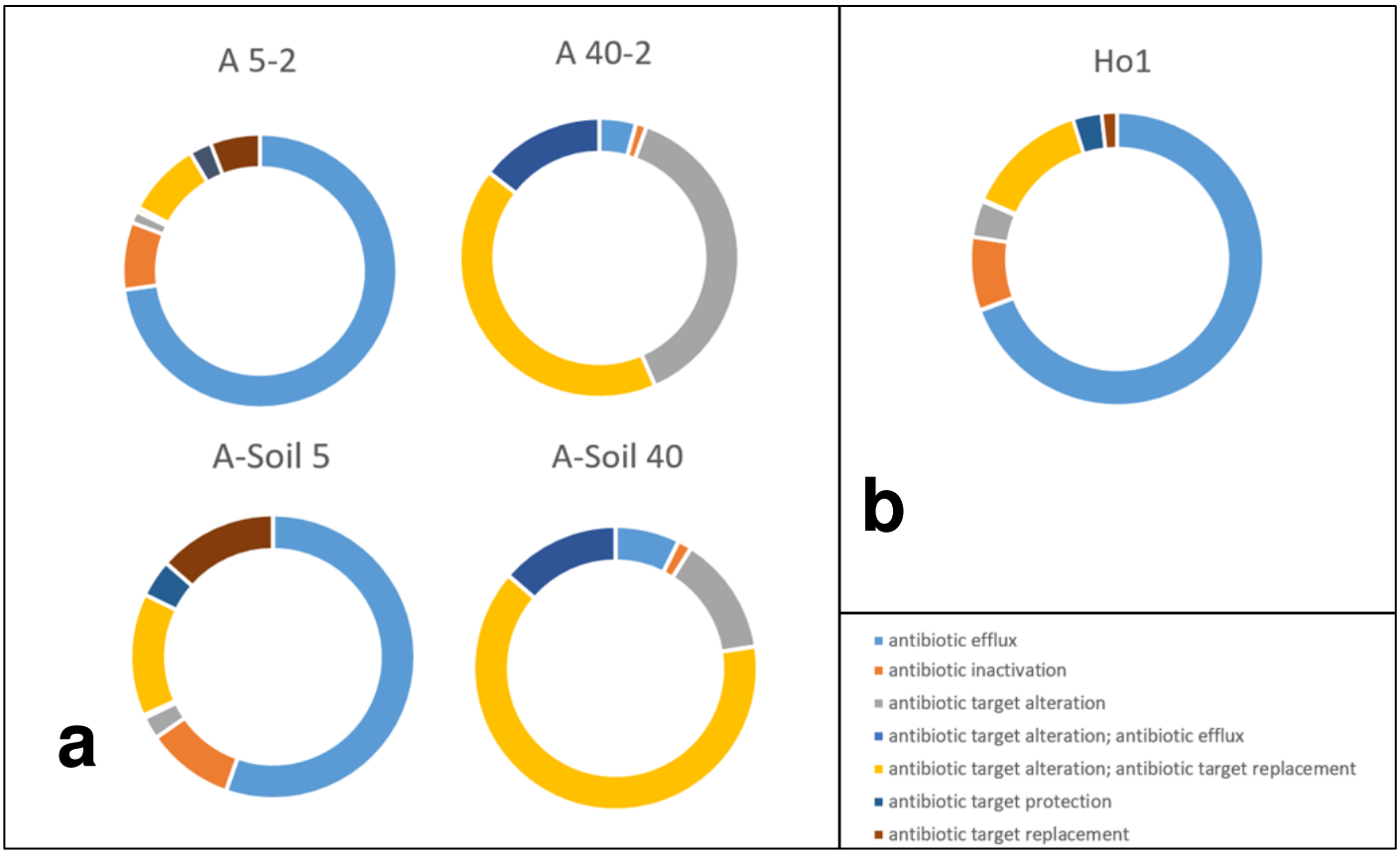


Figure 27: Mechanisms of antibiotic resistance presented by their relative abundance; a: mechanisms of the Aral Sea samples; b: mechanisms of Dachstein sample *Hornungia alpina* rhizoshpere 1.

Since only two of the eleven primers used for PCR were able to produce clear bands on agarose gel, we wanted to further investigate the abundance of the genes targeted with the eleven primers in the RGI hits. Therefore, the metagenomes were screened for the genes targeted by the antibiotic primers and the results can be seen in Table 13. Each sample contained at least five of the eleven targeted genes. Interestingly, only A-402 contained vanH, while vanR could be detected in every sample. In contrast to that, according to PCR, vanH was able to produce clear bands at expected length in all of the Dachstein and Aral Sea samples, while vanR produced clear bands in all Aral Sea samples and in the Dachstein samples Pa3, Se1, Se2, Se3, Se70, Se150, So1, So2 and So3.

Table 13: Antibiotic resistance primers found in the metagenomes of the Aral Sea samples (A-52, A-402, A-Soil5, A-Soil40) and Dachstein sample Ho1.

sample	*Gene name*	*ARO term*	*reads*	*database*
A-52	tetA	tetA(48)	14	CARD
	tetB	tetB(P)	2	wildCARD

	sul1	sul1	4,919	wildCARD
	CTXM	CTX-M-76	498	wildCARD
	blaKPC	PC1 beta-lactamase (blaZ)	44	wildCARD
	vanR	vanRM/vanRO	8	CARD
	aad(A)	aada4/6/8b/11/16	644	CARD/ wildCARD
	ermA	erm31/39/46	20	CARD
A-402	tetA	tetA(48)/Corynebacterium striatum tetA	1,074	CARD
	tetB	tetB(P)	16	CARD/ wildCARD
	sul1	sul1	6,358	wildCARD
	CTXM	CTX-M-76	536	wildCARD
	blaKPC	PC1 beta-lactamase (blaZ)	36	wildCARD
	vanR	vanRF/vanRN/vanRO	128	CARD/ wildCARD
	vanH	vanHO	8	CARD
	aad(A)	aada6/8b/10/11/16/17	912	CARD/ wildCARD
	ermA	ermA	24	wildCARD
A-Soil5	tetA	Corynebacterium striatum tetA	68	CARD
	sul1	sul1	4,176	wildCARD
	CTXM	CTX-M-76	129	wildCARD
	blaKPC	PC1 beta-lactamase (blaZ)	4	wildCARD
	vanR	vanRO	24	CARD
	aad(A)	aada6/8/8b/10/11/16/24	228	CARD/ wildCARD
A-Soil40	tetA	tet(A)/tetA(48)/Corynebacterium striatum tetA	98	CARD/ wildCARD
	tetB	tetB(48)	6	CARD
	sul1	sul1	255	wildCARD
	vanR	vanRO/RF	32	CARD
	aad(A)	aada /2/4/6/8/8b/11/16	212	CARD/ wildCARD
Ho1	tetA	tetA(48)/Corynebacterium striatum tetA	69	CARD
	sul1	sul1	848	wildCARD
	CTXM	CTXM 21/50/75/125	65	wildCARD
	vanR	vanRO	202	CARD
	aad(A)	aada /2/4/6/8b/10/11/16/24	315	CARD/ wildCARD

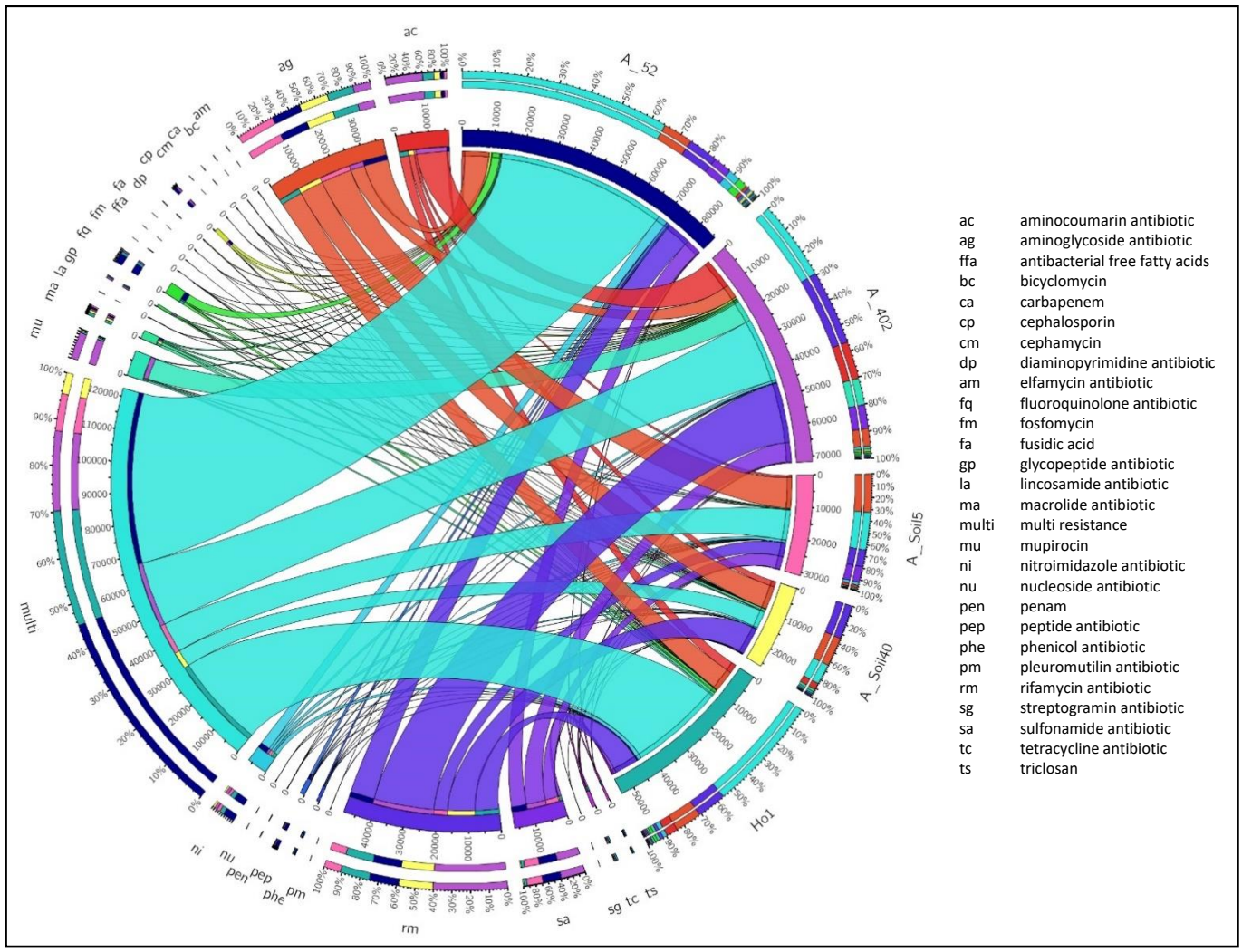


Figure 28: Antibiotic drug class abundance of the samples A-52, A-402, A-Soil5, A-Soil40 and Dachstein's *Hornungia alpina* associated sample Ho1 visualized as Circos plot; resistance to two or more antibiotic drugs appear as "multi" ; A-52: Aral Sea plant *Suaeda acuminata* sample from 5 years without water sampling spot, biological replicate number 2; A-402: Aral Sea plant *Suaeda acuminata* sample from 40 years without water sampling spot, biological replicate number 2; A-Soil5: Aral Sea soil sample from 5 years without water sampling spot; A-Soil40: Aral Sea soil sample from 40 years without water sampling spot; Ho1: Dachstein plant *Hornungia alpina* sample from 10 years without ice sampling spot.

Circos plots can be used to get a good overview over complex data and, in this thesis, feature the abundance of resistances against certain antibiotic types according to their absolute values in each sample. In Figure 28 the resistance of the samples to drug classes feature a high abundance of multi resistances, the majority resulting from highly abundant multi resistant efflux pumps. Because the abundance of resistances of lower numbers was masked by the highly abundant multi resistances, Figure 29 features the same data but with efflux pumps excluded and allows a better overview. A high abundance of resistance to rifamycin antibiotic can be observed, featuring 30% to over 50% of total resistances to drug classes, depending on the sample.

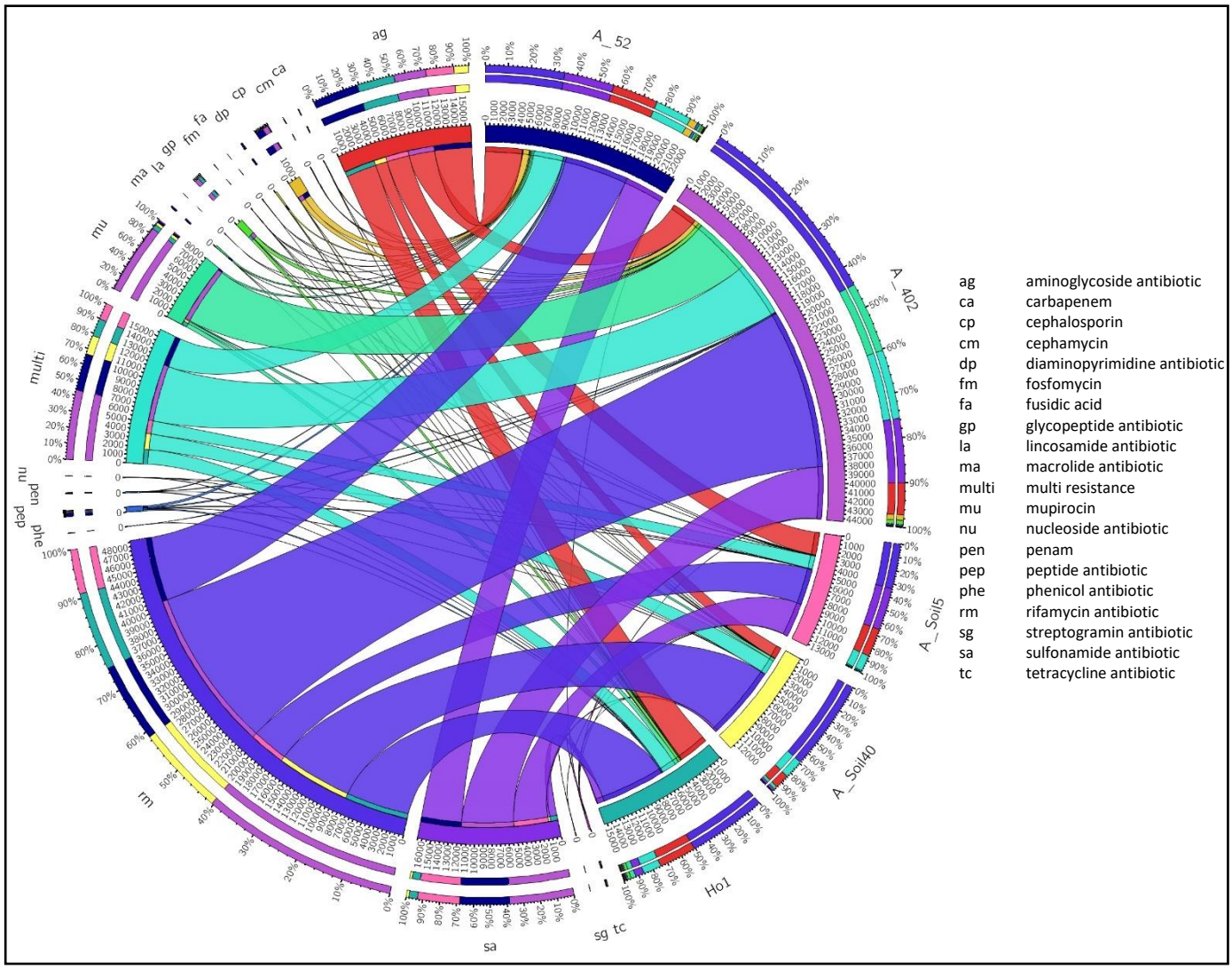


Figure 29: Drug classes with efflux pumps excluded; drug class abundance of the samples A-52, A-402, A-Soil5, A-Soil40 and Dachstein's *Hornungia alpina* associated sample Ho1 visualized as Circos plot; resistance to two or more antibiotic drugs appear as "multi" ; A-52: Aral Sea plant *Suaeda acuminata* sample from 5 years without water sampling spot, biological replicate number 2; A-402: Aral Sea plant *Suaeda acuminata* sample from 40 years without water sampling spot, biological replicate number 2; A-Soil5: Aral Sea soil sample from 5 years without water sampling spot; A-Soil40: Aral Sea soil sample from 40 years without water sampling spot; Ho1: Dachstein plant *Hornungia alpina* sample from 10 years without ice sampling spot.

To visualize all ARG-like genes (since no filtering took place) in each sample according to their respective abundance, the data of each sample is displayed as a bubble plot featuring antibiotic resistance gene abundance, drug class and resistance mechanism (Figure 30-34). As the data is not normalized, a size comparison is included in each figure. Although every sample features a unique ARG abundance profile, similarities within the most abundant genes can be observed, e.g. adeF, sul1 and rpoB/2, which code for resistance to sulfonamide antibiotics, rifampicin antibiotics, fluoroquinolone antibiotics and tetracycline antibiotics, respectively appear in high abundance in all samples.

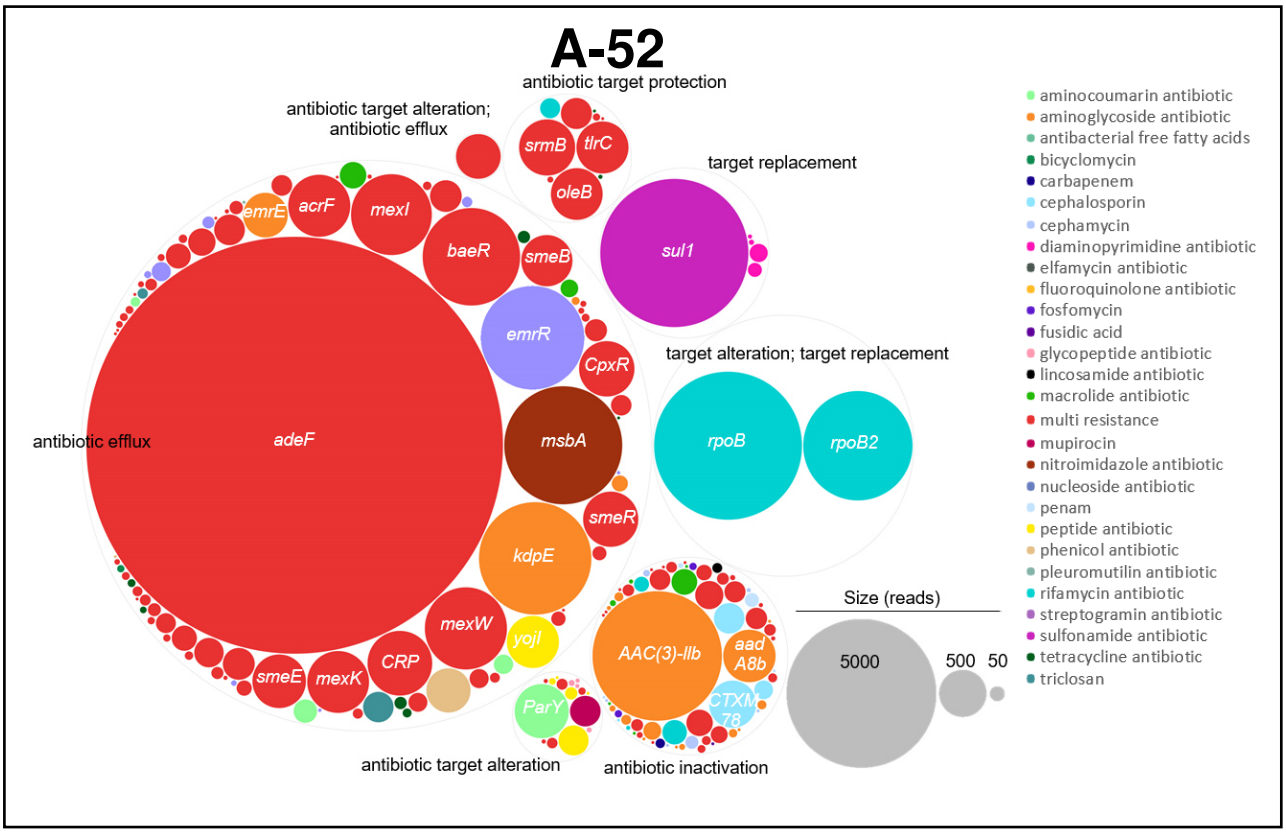


Figure 30: Antibiotic resistance genes of Aral Sea rhizoshpere sample A-52; size according to absolute abundance of ARG, color represents drug class, grouped according to general resistance mechanisms of detected genes; 52: 5 years without water, biological replicate number 2.

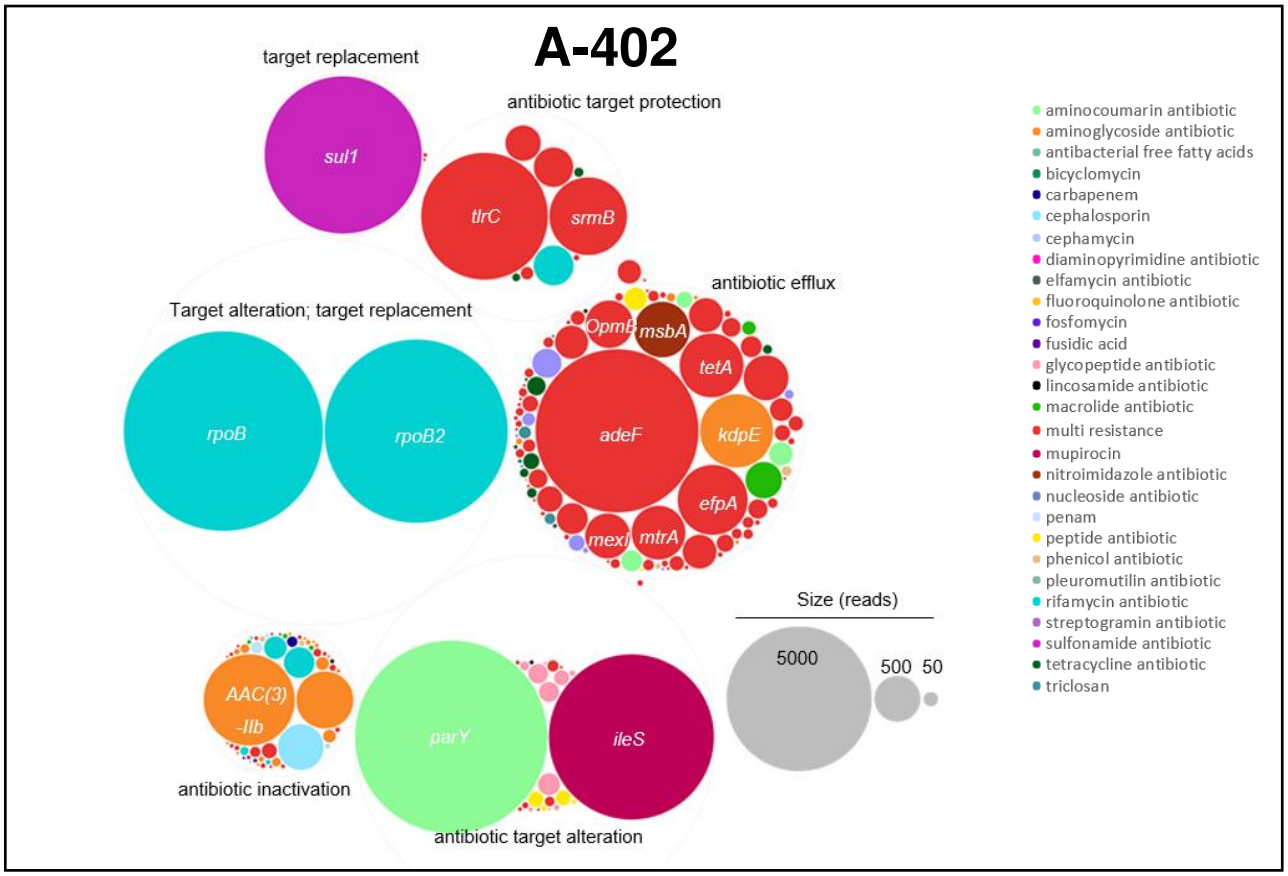


Figure 31: Antibiotic resistance genes of of Aral Sea rhizoshpere sample A-402; size according to absolute abundance of ARG, color represents drug class, grouped according to general resistance mechanisms of detected genes; 402: 40 years without ice, biological replicate number 2.

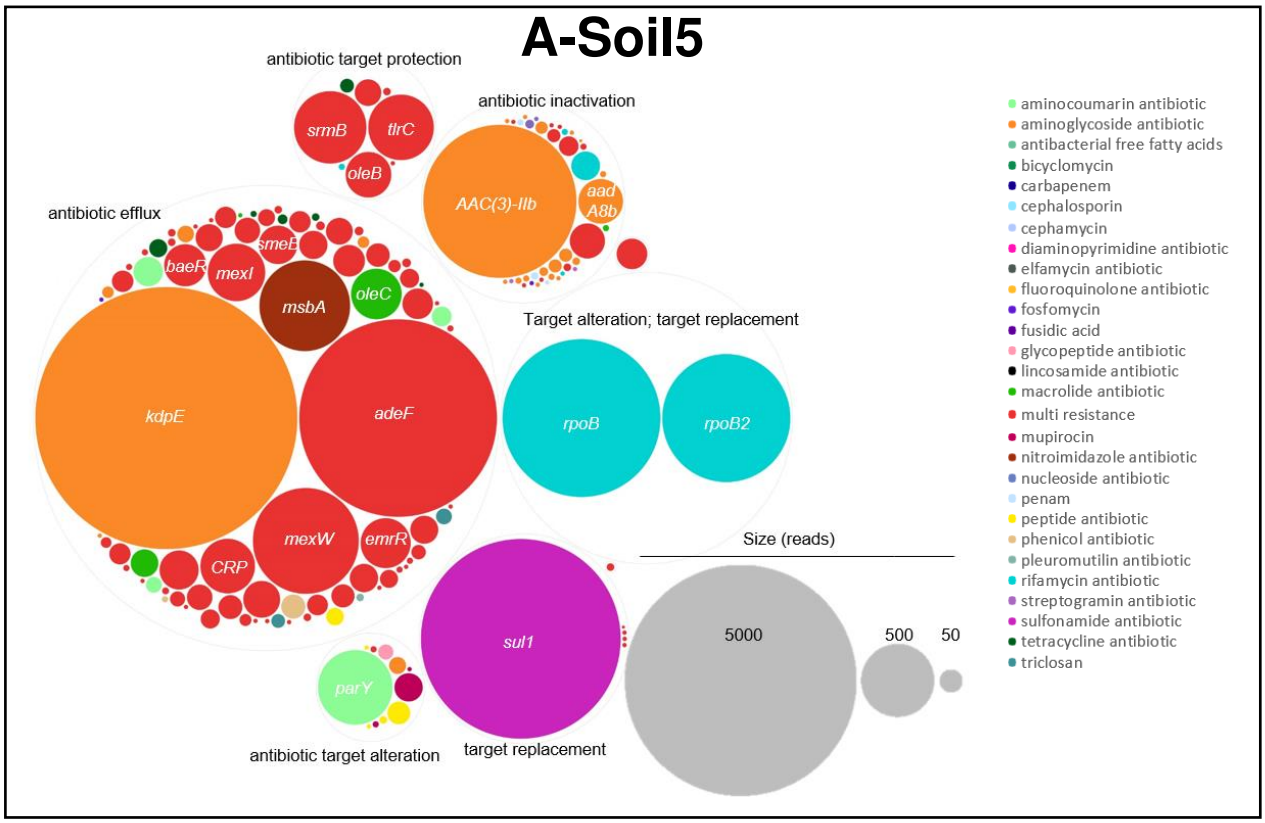


Figure 32: Antibiotic resistance genes of of Aral Sea soil sample A-Soil5; size according to absolute abundance of ARG, color represents drug class, grouped according to general resistance mechanisms of detected genes; 5: 5 years without water.

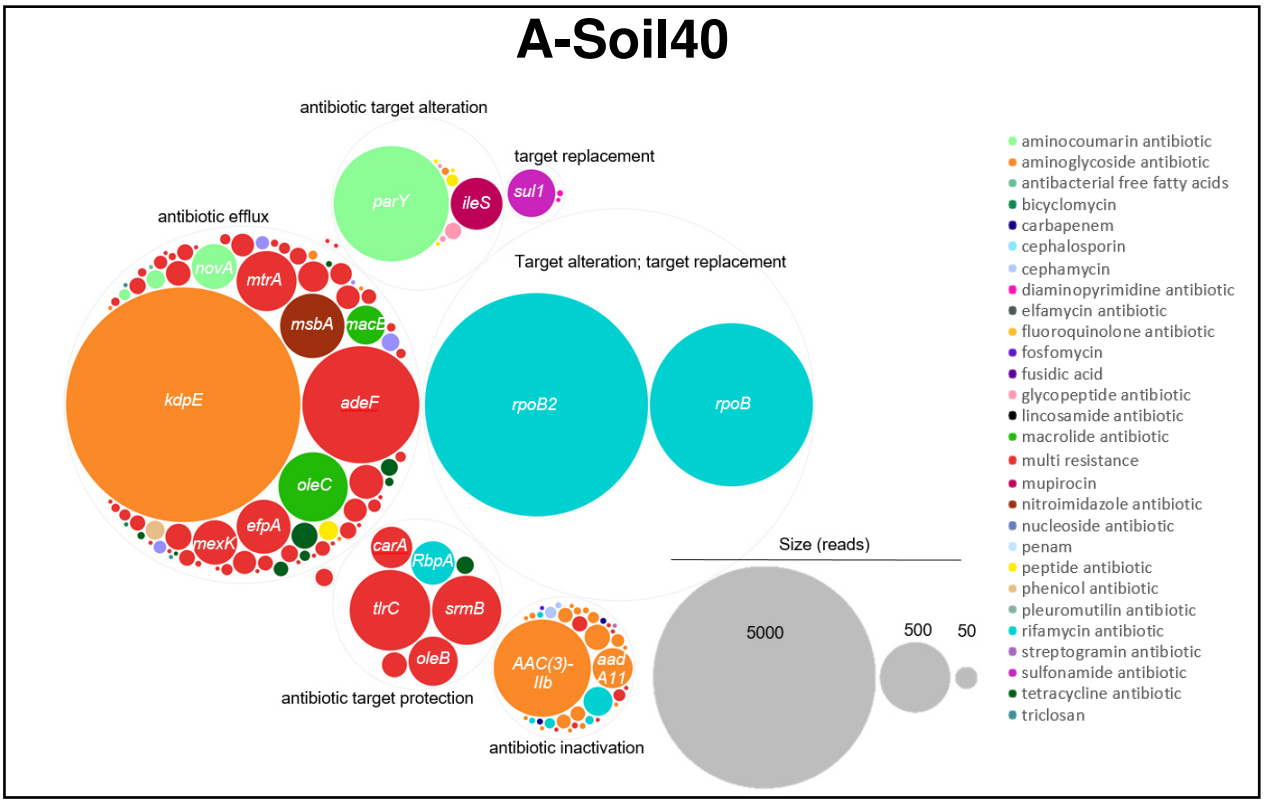


Figure 33: Antibiotic resistance genes of Aral Sea soil sample A-Soil40; size according to absolute abundance of ARG, color represents drug class, grouped according to general resistance mechanisms of detected genes; 40: 40 years without water.

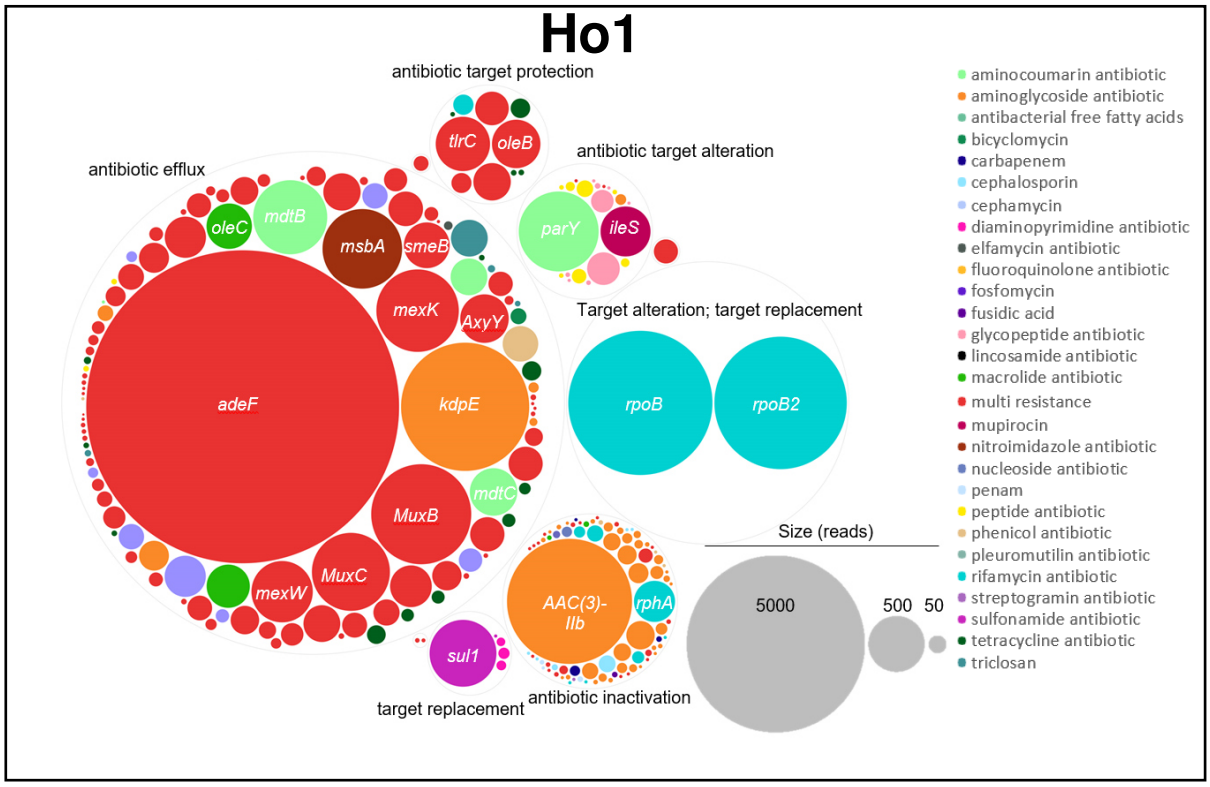


Figure 34: Antibiotic resistance genes of Dachstein rhizosphere sample Ho1; size according to absolute abundance of ARG, color represents drug class, grouped according to general resistance mechanisms of detected genes; Ho1: rhizosphere of *Hornungia alpina* 10 years without ice.

After analyzing the whole resistome data in each sample, the ARGs only located on plasmids were further investigated, as these can potentially spread within the bacterial community via horizontal gene transfer and cause sharing and accumulation of multi resistances. Since RGI bwt method was used to analyze the resistome and works with raw reads (no contigs), these results heavily rely on the quality of information contained in CARD and wildCARD database. In order to increase quality of the results, only perfect and strict hits > 87%[59] identity were included. Figure 35 shows the ARG data of all five samples that was found to be located on plasmids. For comparison, the ARG are grouped into sample and sub grouped into resistance mechanism, while the color shows the ARG. 19 different ARGs were detected in all of the samples.

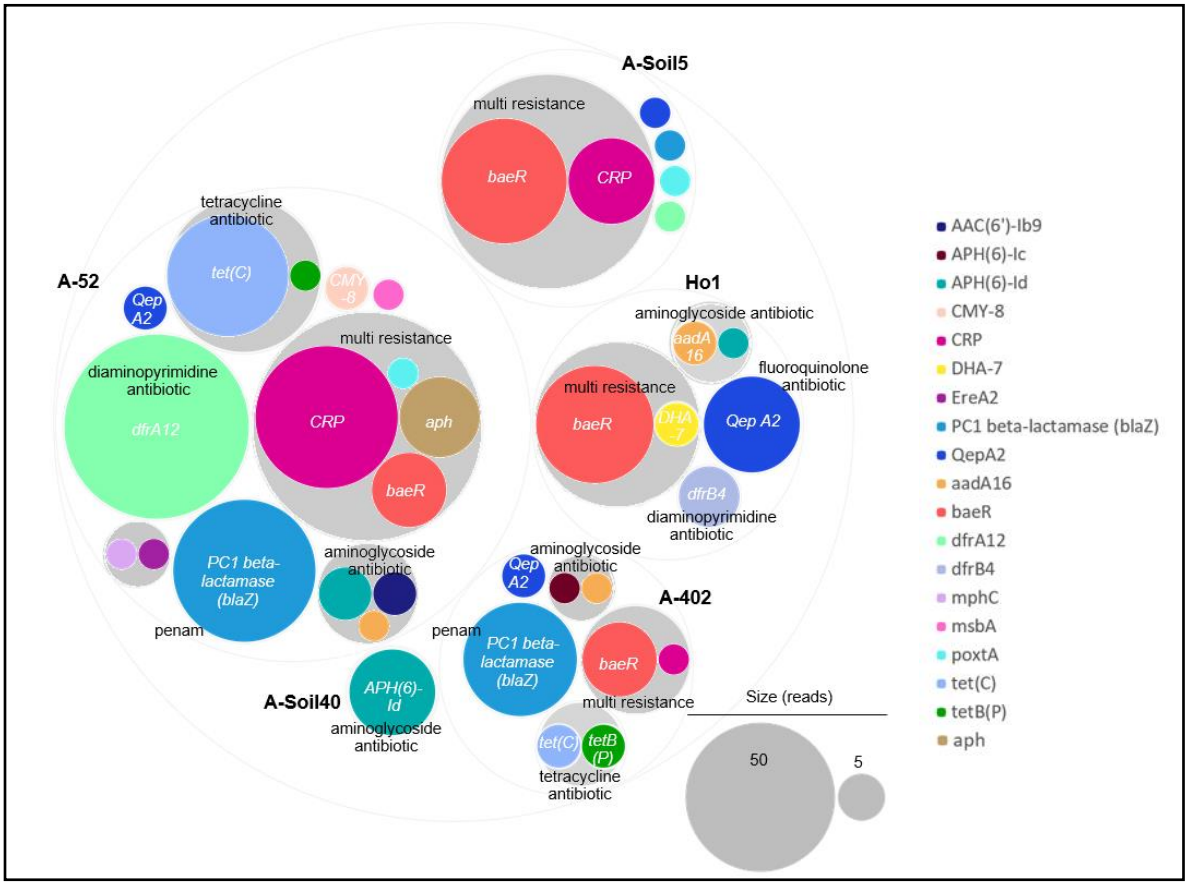


Figure 35: Antibiotic resistance genes located on plasmids; bubble size according to abundance of ARG, color represents ARG, grouped according to samples and sub grouped according to resistance mechanisms; A-52: Aral Sea plant *Suaeda acuminata* sample from 5 years without water sampling spot, biological replicate number 2; A-402: Aral Sea plant *Suaeda acuminata* sample from 40 years without water sampling spot, biological replicate number 2; A-Soil5: Aral Sea soil sample from 5 years without water sampling spot; A-Soil40: Aral Sea soil sample from 40 years without water sampling spot; Ho1: Dachstein plant *Hornungia alpina* sample from 10 years without ice sampling spot.

IV. Discussion

In order to confirm or reject the hypothesis of less microbial diversity in the Aral Sea samples due to a higher anthropogenic impact, and to unravel both region's bacterial community composition, 16S rRNA gene amplification sequencing, metagenomic analysis and qPCR were carried out. It could be shown, that the Dachstein and the Aral Sea, two unique ecosystems, both possess antibiotic resistance potential and that the Dachstein samples feature a more diverse microbial community.

The Dachstein samples generally contained higher DNA amounts than the Aral Sea samples (Figure 6), and the microbial diversity was enriched in the glacier fore field samples. However, for both datasets the plant samples were significantly higher in DNA content compared to the soil. Especially the Aral Sea soil samples contained very low amounts of DNA. One reason for this is, the rhizosphere, one hotspot for microbial activity, and the fact that microorganisms can extend functional repertoire of plants and enhance nutrient availability, while also protecting against stress or diseases.[68] The high cell density around the plant roots makes them a perfect sampling spot and great for comparing to the bacteria active in pure soil. The diversity of the samples, however, feature a higher alpha diversity (Figure 7) in the Aral Sea soil samples compared to the plants. The high amount of DNA especially in the plant samples could also result from plant DNA that is present in the rhizosphere.

It is quite well understood that the physio-chemical properties of the soil or the climate can influence the composition of microbiota and that an unbalanced composition of microbial taxa is associated with diseases in plants [69].The rock type in the study area of the Dachstein is very compact Dachstein-specific limestone [20] and the growth of one of the sampling plants *Hornungia alpina* is restricted to limestone[70]. *Papaver alpinum* has a preference for sparsely vegetated, mostly calcareous scree[71] and *Sedum atratum* grows preferably on limestone in subalpine to alpine climate conditions[72]. Therefore, all three plants are perfectly adapted to the living conditions at the glacier fore field.

The top 20 taxa barplot of the Dachstein according to 16S rRNA analysis depicts high diversity within all samples, with low differences. The LEfSe results for plant to soil (Figure 12), however, show a significant difference, featuring *Blastocatella* among others as enriched taxon in the soil samples. Since the microbial abundance and the composition only changes slightly over the timespan of 10 to 150 years without ice with only small alterations, the high abundance and diversity within the Dachstein samples hint towards an overall healthy microbiome, as there are no overabundances of taxa. However, plant diversity increased with the years without ice of the sampling spots (Mora M, personal communication). The decision to sample *Hornungia alpina, Papaver alpinum* and *Sedum atratum* was made due to their confirmed appearance within all sampling areas. This, however, could have influenced and biased the results, as these plans might always prefer similar soil types and since no soil samples free of vegetation could be taken at 70a and 150a sampling spots. Thus, a comparison between pure soil at different locations cannot be made. Nevertheless, considering that three different plants were sampled and that the differences between them are not big, the bias of sticking to the same species in all sampling areas is expected to be negligibly small. Furthermore, if a different plant species growing exclusively at another sampling spot would show a different microbial composition, it would not be possible to determine if that change would be dependent on location or host organism.

One other 16S rRNA gene amplicon study done in the Swiss Alps reported an increase in alpha diversity during soil development on glacier fore fields[23]. This was not detectable in our study. However, the soil composition of the Swiss samples was mainly coarse-grained metamorphic granite along the fore field, which differs from the Dachstein's Limestone.

All in all, the microbial community at the Dachstein glacier only changed slightly over the years without ice and no overabundances of certain taxa occurred. This indicates a stable and overall healthy microbiome at this location with little anthropogenic impact. The stable microbial community is most likely formed within the first years after glacier recession and while plants do statistically increase the diversity of the soil microbiome, a statistically significant difference between different plant species growing in this region could not be measured.

The Aral Sea samples, in contrast to that, changed a lot throughout the years without water, not only in their microbial composition. *Suaeda acuminata*, the plant sampled at the Aral Sea for this study, is an adaptive member of the family *Chenopodiaceae*, can grow on soil high in salinity and is therefore one of very few plants that survive the Aralkum desert conditions. The growth of salt-resistant plants like *Suaeda acuminata* on saline and alkaline soils can encourage desalinization and improvement of the soil[73], which could explain why the abundance of *Halomicrobiaceae* was significantly higher in soil samples than in the plant samples. Soil at the area where the water receded 40 years ago, however, is known to be less high in salinity than 5a or 10a soil, which could have an influence on the growth of this plant as well as the rhizosphere bacteria, as the plant might be restricted in growth on drier soil with less salinity[74]. Regarding the former abundant animal species at the Aral Sea, it is evident that the number of different species reaching from insects, birds or mammals to water organisms dropped significantly. In 1991, the number of bird and mammal species were already cut in half [75], until now, the situation has worsened; according to Tomislav Cernava and Gabriele Berg in personal discussion with the author, there was no visual contact with any living being during sampling.

Regarding the dire climatic environment in the Aral Sea region while keeping the abundance of toxic components and salt in the Aral Sea sediment in mind, it is to be expected that the microorganisms which are not only surviving but even thriving on this kind of soil have to be survivalists and somehow accordingly genetically equipped. Archaea are known to be able to colonize and live at extreme environments[76], therefore it is interesting to see four out of the top 20 most abundant taxa to be of archaeal origin. *Halomicrobiaceae* even mark the highest abundant taxon in the soil, while being significantly less abundant in the plant samples, according to LEfSe. The used amplicon primers were universal bacterial primers and recommended for Archaea as a good choice in terms of avoiding bias and allowing for good representation of known bacterial and archaeal groups[77], However, a newer study with extensive archaeal primer comparison questions this, stating that the universal primers do not depict the true abundance of Archaea, but only giving a distorted, underrepresented and biased view on the archaeal community present in the

samples[78]. Nevertheless, the primers that were used for this thesis are still the best suited universal primers due to the lack of a better alternative, which targets bacterial as well as archaeal taxa. When keeping that in mind, it is especially interesting, that many archaeal taxa could be found in the samples. All 4 archaeal taxa out of the top 20, *Halomicrobiaceae*, *Haloterrigena*, *Natronomonas* and *Halobacteriales,* are halophilic Archaea of the class *Halobacteria,* which are known to colonize environments with high salinity[79] and have been mentioned before as high abundant taxa in saline soils[76].

The top 20 taxa barplot (Figure 15) of the Aral Sea samples suggests huge differences not only between plant and soil, but also between the years without water. The number of *Halomonas*, taking up about one third of the taxal abundance (and showing up as the taxa with the highest association with the feature 5a in LEfSE plot Figure 18) in the 5a samples, significantly reduced over the years, yet stayed very abundant in all soil samples. The Dachstein samples on the other hand just contained around 2% *Halomonas* of the total taxa abundance, a value that is only changing slightly over the years without ice. This highly suggests that the microbiome of the plants undergoes changes, probably due to the different environmental conditions the plants are facing on more arid and less saline soil at the Aral Sea. *Halomonas*, extremophile bacteria and known to tolerate desiccation and salinity quite well, were found to promote plant growth on high saline soil when associated with the rhizosphere. [80,81] However, the drier the soil gets over the years, the more the *Halomonas* abundance drops. On the other hand, a sudden rise in *Planococcus,* a halotolerant *Bacilli* subgroup, that is able to promote plant growth in the presence of growth inhibitory levels of salt[82], can be seen in the 40a plant samples. This leads to the assumption that the plants most likely need some growth promoting bacteria to survive in this harsh environment.

The top 20 taxa abundance barplot (Figure 15) as well as the PCoA plots (Figure 17) also show strong taxal and diversity differences between the 40a samples and the rest. While the 5a and 10a samples share a lot of common taxa such as *Halomonas*, *Marinobacter, Balneolaceae, Bradymonadaceae* and not further classifiable *Gammaproteobacteria*, the taxa and diversity of the 40a samples differ greatly, featuring different species as well as a generally higher

diversity. The locations of the 5a and 10a sampling spots, however, were very close to each other (Figure 5) and this proximity could also have influenced and biased our findings, as soil composition or environmental factors such as dust migration through winds do not differ much.

Sinomicrobium, a *Flavobacteriaceae* subgroup, can be found in water or soil, depending on the species, and was found to be able to thrive on high saline and alkaline soil [83]. A high abundance of *Sinomicrobium* was found in the plant 10a samples, but only low values appeared in the other samples. This could hint that the soil at 10a sampling spot was contaminated with additionally potential saline or alkaline substances. A further analysis of the soil composition would be of great interest for answering questions related to salinity and alkalinity.

Marinobacter, depending on the species, can degrade petroleum[84], be halotolerant[85] or halophilic and are widely distributed throughout the world's oceans. Therefore, it is not a surprising hit in water with high salinity[86]. Since the water of the Aral Sea, especially after shrinkage of its size, accumulated high amounts of salt, findings of *Marinobacter* in the soil after the water receded is not unexpected, and the bacteria that were living in the saline water can now be found in the soil. This also explains the sudden drop of this genus in the 40a samples as the water 40 years ago was not that high in salinity as it was later in the 10a or 5a sampling spots.

The metagenome analysis of the abundant taxa was used to provide a more profound insight in the abundant taxa, as the metagenome sequencing uses all present DNA of the sample to identify the taxa instead of only the 16S rRNA gene that is used for 16S amplicon sequencing and omits the targeted amplification step, which can introduce primer bias. The major players, however, stay the same: *Halomonas, Halofilum* and *Marinobacter* appear as the most abundant top three taxa. As the 16S rRNA data could not offer taxa on species level and often struggled with genus level as well, some bacteria like uncultured *Balneolaceae* or uncultured *Halobacteriales* are taxa on higher level. The metagenome sequencing, however, provides more reliable information, which led to a slightly different naming of the taxa. The uncultured *Balneolaceae* most likely show up as halotolerant *Aliifodinibius* in the metagenome bar plot. On the other hand, taxa like *Halomonas* or *Planococcus* appear in similar percentages as they did in the amplicon data, suggesting a high consistency and credibility

of the data. According to the amplicon data, a trend can be seen and clear clustering of the alpha diversity when comparing plant and soil samples takes place. The samples differ more greatly when taken from plant or soil than at a different sampling spot.

When comparing the abundant taxa of Dachstein and Aral Sea, great differences between the top 20 most abundant taxa at both locations can be observed. The only genus both areas have in common is the former mentioned *Halomonas*. However, when looking at the data on phylum level, we see that both share similar abundance of *Proteobacteria, Bacteriodetes, Actinobacteria* and *Planctomycetes* as top four most abundant phyla. On class level and levels further below, these similarities decrease. This could be heavily influenced by the different environmental conditions; the Dachstein features limestone soil with annual high precipitation, low temperatures and alpine conditions, while the Aral Sea offers desert like conditions with low precipitation, high salinity and overall continental weather conditions. Regarding these differences, it is highly unlikely to find many taxa shared in both areas. However, the observed compositional changes within the Aral Sea samples as well as the low numbers of extracted DNA and predominant abundance of certain taxa such as *Halomonas* suggest an imbalance in the microbial composition, which is connected with being less healthy, whereas the Dachstein samples feature a balanced and overall healthy microbiome.

It can be said, that the microbiome at the Aral Sea underwent changes in microbial composition and overabundances of certain taxa, stating that the microbiome at the Aral Sea is far from being in a healthy and stable condition. As much as humans have altered the environmental conditions at the Aral Sea, from a once flourishing lake environment to a desert high in salinity and toxic compounds, the microbiome also depicts this imbalance.

In order to get insight in the antibiotic resistance gene abundance of our samples, the metagenomes were screened against CARD and wildCARD databases. The total RGI reads compared to the total reads according to FASTQC show a higher ratio of RGI to FASTQC reads for some samples than for the rest. Especially the soil samples appear to a have a lower relative abundance of AR gene hits in their total reads.

The analysis of resistance mechanisms (Figure 27) suggests a prevalence of efflux pumps conferring antibiotic resistance for the 5a samples as well as the Dachstein sample Ho1. The 40a samples, however, feature target alteration as the most abundant resistance mechanism of antibiotic resistance. This is further interesting since this difference is presumably based on the different environments. When comparing the environment at the 5a sampling spot with the 40a one, the 40a soil appears to be much drier, even sand-like, while the 5a soil was rather wet and loamy. Furthermore, the Aral Sea water was not that high in salt and heavy metal pollution 40 years ago when the sampling spot dried out, than when the spot at 5 years dried out. 5 years ago, the shrunken lake's salt and poison content accumulated in the remaining water, leading to a higher salinity in the 5a soil. It could potentially be the case, that the bacterial community at 40a sampling spot faces situations where efflux pumps are not necessary in high numbers, e.g. less salt to pump out of the cell. Also, due to a drought, only limited metabolism can take place and therefore energy-cost-intensive multi efflux pumps do not offer increased evolutionary fitness anymore.

The high abundance of antibiotic target alteration and antibiotic target replacement genes in A-402 and A-Soil40 level could occur due to the high abundance of rpoB and rpoB2, which are both coding for rifamycin resistant RNA polymerase subunits[87]. Rifampicin is a widely used antibiotic in healthcare, especially for treatment of tuberculosis[88], where rifampicin is a key component and patients with resistance to this drug have a poor prospect of healing as rifampicin resistance is often associated with resistance to other antituberculosis drugs[89]. The mechanism of rpoB conferring resistance to rifampicin is well understood in *Mycobacterium tuberculosis*[90] and marks a serious threat in tuberculosis treatment. Not so much is known about their occurrence in environmental samples and a high abundance of resistance against a critical last resort antibiotic in environmental samples is a troublesome finding. One explanation, however, could be, that *E. coli*, as studies show, is able to develop rpoB conferring resistance to rifampicin in an antibiotic-free environment during thermal stress due to hot conditions (temperature around 42° C)[91]. Since the climate conditions in the Aralkum desert are dry and hot, this could have led to a development and accumulation of these resistances in local bacteria, besides natural abundance of the AR gene.

The top five abundant antibiotic resistance genes in each sample (Figure 26) feature AdeF and KdpE genes as highly abundant in all tested samples. AdeF codes for resistance nodulation cell division type efflux pumps as part of the adeFGH operon[92]. It equips the cell with resistances to fluoroquinolone antibiotic and tetracycline antibiotic, but has not been found on plasmids yet, therefore it is not a potential candidate for transferring resistances. This, however, does not exclude translocations to plasmids and spread of resistances in the future, as bacteria are under constant selection pressure.

KdpE gene is a part of the KdpFABC complex coding a K^+ transporting efflux pump. As a transcriptional regulator that binds to a specific 22-bp sequence in the promoter region of kdpFABC operon, it positively regulates the kdpFABC expression, while also being necessary for K^+ homeostasis. [93,94] Cells depending on this system might possess a high abundance of the KdpE gene. According to wildCARD database, the KdpE gene has not been found on plasmids yet either.

An additional filtering of the RGI data with a threshold of 87%[59] identity would obviously reduce the size of the datasets, but not at equal amounts. A-52 would reduce its total reads from 86002 to 42753, a loss of reads of approx. 50%. The other samples A-402 (13%), A-Soil5 (22%), A-Soil40 (10%) and Ho1 (41%) would also lose significant numbers of hits, but a higher threshold ensures higher reliability of the output data and will therefore be integrated in future data analysis.

Regarding the antibiotic resistance screening with primers, no useful information from the PCR and qPCR could be obtained, since all of the antibiotic primers showed unspecific bands, multiple bands or no bands at all. This could be due to the fact that the majority of the primers was originally designed for agricultural environments e.g. manure (heuer smalla 2007, walsh 2011). As the metagenomes of glacier or lake environments greatly differ from these, it is possible that the primers were not suited for the environmental samples of this study. On the other hand, it is also highly likely that extracted plant DNA interfered with the PCR and caused unspecific bands or an inhibition of the reaction. A quick search in the metagenome contigs of AS-52, AS-402, Soil-5, Soil-40 with the sequence of the antibiotic primers revealed one binding spot for CTX-Mg8/25 fwd in Soil5, implying potentially binding capacity at various

locations and the possibility of bands of undefined length via PCR. This *in silico* search, however, could not make allowance for base mismatches or hairpin structures. According to the PCRs, many primers were able to produce bands with varying lengths, implying the existence of multiple binding spots. In addition to that, the variation in the melting temperature of the qPCR with primer vanR hints towards the existence of more than one amplified DNA fragment. Because of this, the values shown in Figure 22 do not reflect the true abundance of vancomycin genes within the samples, as the binding of the primers took place at multiple loci and is therefore not suited for calculating the abundance of one certain gene. Also, the results do not match with the first PCR done with vanR on the Dachstein samples, which showed a higher abundance within the *Papaver alpinum* and *Sedum atratum* samples and almost no occurrence in *Hornungia alpina*.

The computational screening of the metagenomes for the 11 genes targeted by the ARG PCR primers (Table 13) revealed the presence of at least 5 of the 11 targeted genes in every sample. VanH was, according to PCR, abundant in every sample, but appeared only in A-402 in computational screening, while on the other hand, vanR could be detected virtually in every sample, in contrast to the prior PCR results. However, sul1, the most abundant one of the 11 tested genes according to RGI, did not give any positive results in the PCR. This divergence most likely results from the PCR primers, which are not suitable for our environments, because genes, even though they were as present in high abundance as sul1, could not be visualized by PCR.

When comparing the outcome of the *in silico* screening of the primer sequences used for PCR with the ARGs found via resistome analysis, it is interesting to see, that even though primers for abundant genes such as sul1 or aad(A) were used, the primer sequences in the metagenome could not be found. However, when comparing the metagenome to CARD and wildCARD databases, some of the genes were found in high abundance. This further hints towards the fact, that both *in silico* methods have their limitations and that screening of genes in metagenomes allows a deeper insight than screening for short and specific primers.

Regarding the multi resistances in the resistome of the 5 samples, the Circos plot (Figure 28) features a high abundance, especially A-52 and Ho1 with

more than 50% of their resistance genes featured as multi resistant. The majority of these multi resistances originate from efflux pumps, which originally confer outstanding resistance to the action of toxic compounds, but also confer resistance against antibiotics[95]. To get a better graphical overview of the drug classes, efflux pumps in Figure 29 were excluded and a high abundance of rifamycin, sulfonamide antibiotics and other non-efflux pump mediated multi resistances can be seen. Furthermore, sul1, the one sulfonamide antibiotic with the highest abundance, occurs in every sample. It is a well understood resistance gene and is often part of the class 1 integron (intI1), which can be located on either plasmids or chromosomes, especially on chromosomes of *Betaproteobacteria* as mobile elements[96]. Mobile elements can be shared within a bacterial community and therefore lead to a spread of resistances. As intI1 genes occur in environmental samples and are linked to antibiotic, heavy metal and disinfectant resistances, they can be used to measure anthropogenic pollution in the environmental samples[97]. However, sul1 was also found outside of the intI1 cassette for certain bacterial species[97] and the abundance of sul-ARGs was found to be decreasing as salinity increases[98]. Unfortunately, CARD and wildCARD databases do not provide information about intI1 abundance in our samples and the abundance of sul1 cannot be used to measure anthropogenic pollution in our samples as not every sul1 gene is part of intI1. According to wildCARD, the sul1 genes found in our samples are not located on plasmids and therefore do not appear in the bubble plot featuring plasmid data only. However, for confirming the exact location of the genes, a main RGI approach with data split into chromosome and plasmid DNA by a tool like plasflow[99] might produce additional information.

As only the ARG located on plasmids or mobile elements can be transferred between microorganisms, these are of great importance regarding the rise of antibiotic resistant bacteria in health care. When taking a look at the quality filtered plasmid only data (Figure 35), we see, that the most abundant genes throughout the samples are dfrA12, baeR, CRP and PC1 betalactamase. DfrA12 is frequently found to be associated with class 1 integron cassettes (mobile elements), which makes it a potential risk for health as it confers resistance to diaminopyrimidine antibiotic between species[100]. BaeR is known as an activator for transcription of the RND-type efflux pumps AcrD and

MdtABC[101]. BaeR is part of the resistome of every sample, besides AS-40. However, as an efflux pump activator located on plasmids, it might not confer antibiotic resistance without the respective efflux pump system. CRP is known to be a global regulator that represses MdtEF multidrug efflux pump expression and the control of the multidrug efflux pump seems to be connected to sugar metabolism.[102]. However, only small percentages of total CRP reads in the samples are said to be located on plasmids, while the majority is on chromosomes. The PC1 beta-lactamase (blaZ) is found in *Bacillus subtilis* and *Staphylococcus aureus* and is a plasmid encoded enzyme, which is able to confer antibiotic resistance by catalyzing the hydrolysis of β-lactams such as penicillin G and ampicillin[103]. With both 16S rRNA gene amplicon sequencing and metagenome analysis, *Staphylococci* and *Bacilli* were detected in the samples.

To confirm, that the rather low amount of ARG hits in this graph do not result from the strict filtering of the plasmid located genes (>87% identity), the plasmid-ARG below 87% identity were also checked. However, these genes with identity below 87%, adeF, ErmA and KpnH, turn up with negligibly low read counts compared to the most abundant plasmid-ARG. Sample A-Soil40, interestingly, only features one ARG on plasmids, APH(6)-Id, a streptomycin-inactivating enzyme.[104]

However, all reads located on plasmids occur in very low numbers (<75 reads per sample) and are not very abundant within the bacterial communities compared to the total RGI reads (depending on the sample between 0.04% (A-402) and 0.11% (AS-5)). Therefore, potential risk to health based on mobile ARG abundance can be considered low.

V. Conclusions

To sum up, both areas possess antibiotic resistance potential, the anthropogenically via tourism only slightly affected Dachstein glacier as well as the agriculturally heavily influenced Aral Sea basin. In contrast to the hypothesis, the resistome of the Dachstein shows resistance mechanisms and involved genes in similar amounts as one sample of the Aral Sea. According to these results, the anthropogenic impact does not alter the ARG abundance. To back up these findings, a comparison of the metagenomes of all samples, is needed and will further be a project in the future. However, due to the environmental changes in the Aral Sea region, a shift throughout the years without water can be seen, as the bacterial composition of the Aral Sea changes and adapts to the new conditions facing high salinity and toxic compounds. Overall, the microbiome is imbalanced and clearly portrays the unhealthy state the whole Aral Sea's environment is currently facing. On the other hand, the metagenome of the Dachstein only undergoes subtle changes throughout the 150 years without ice, while slightly adapting to changes of soil composition. The microbiome is stable and no overabundances occur, which are all factors that define a healthy microbiome. Interestingly, the resistance mechanisms of Aral Sea's soil and plant sample at 40 years without water depict a very distinct picture that deviates from the other samples, and further investigation will be needed to reach full understanding.

The resistome data shows that clinically relevant resistances also occur in high abundance in environmental samples and that the problem of antibiotic resistance genes is not solely the problem of the health sector. Regarding the One Health hypothesis, natural occurring ARG migrated towards hospital environments, where the volutionary selection and accumulation thereof causes problems. However, the risk of spreading ARG in our study is fortunately low as only a low number of ARG reads were found to be located on plasmids, while the majority of ARG are located on the chromosomes and are therefore less likely to spread within the bacterial community. These are unlikely to contribute to the antibiotic crisis in health care. Additionally, due to the dire situation

regarding animal and bird life near the Aral Sea, the possibility of ARG transmission via migration birds is minimized. [6,18]

Outlook

As a future project, all of the metagenome samples will be analyzed to complete the picture of natural antibiotic resistance at the Dachstein and the Aral Sea. Regarding the rise of antibiotic resistances in healthcare, it would be important to invest more money into finding and developing new drugs. Since natural antibiotics and ARG occur in anthropogenically uninfluenced and influenced regions alike, both could be promising starting points to search for new drugs. As this thesis focused on ARG at the Aral Sea and Dachstein, taking samples at these locations in order to cultivate microbial species and to isolate potential antibiotic producers could be the next step. The wildCARD database that was used for the resistome analysis of this thesis allows to search for sequences with predicted ARG potential that have not been found so far. With the help of this prediction tool, the discovery of new antibiotics or an improved efficacy of an already known antibiotic might be possible. Furthermore, a reduction of antibiotic usage in agriculture is inevitable to secure the effectivity of antibiotics currently in use in healthcare.

VI. References

1. World Health Organization. Antimicorbial Resistence Global Report on Surveillance. *Physiol. Genomics* **1**, 1–8 (2014).
2. Levy, S. B. Antibiotic and antiseptic resistance: Impact on public health. *Pediatr. Infect. Dis. J.* **19**, 120–122 (2000).
3. Marshall, Bonnie M. & Levy, Stuart B. Food animals and antimicrobials: Impacts on human health. *Clin. Microbiol. Rev.* **24**, 718–733 (2011).
4. Robinson, T. P., Bu, D. P., Carrique-Mas, J., Fèvre, E. M., Gilbert, M., Grace, D., Hay, S. I., Jiwakanon, J., Kakkar, M., Kariuki, S., Laxminarayan, R., Lubroth, J., Magnusson, U., Thi Ngoc, P., Van Boeckel, T. P. & Woolhouse, M. E. J. Antibiotic resistance is the quintessential One Health issue. *Trans. R. Soc. Trop. Med. Hyg.* **110**, 377–380 (2016).
5. Obermeier, Melanie-Maria, Taffner, Julian & Bergna, Alessandro. Unravelling native plant resistimes - The Sphagnum microbiome harbours versatile and novel antimicrobial resistance genes. *bioRxiv* (2019) doi:http://dx.doi.org/10.1101/695973.
6. Teng, Lin, Lee, Shinyoung, Ginn, Amber, Markland, Sarah M., Mir, Raies A. & DiLorenzo, Nicolas. Genomic Comparison Reveals Natural Occurrence of Clinically Multidrug-Resistant Extended-Spectrum- □-Lactamase-Producing Escherichia coli Strains. *Am. Soc. Microbiol.* **85**, 1–16 (2019).
7. Lack of new antibiotics threatens global efforts to contain drug-resistant infections. https://www.who.int/news-room/detail/17-01-2020-lack-of-new-antibiotics-threatens-global-efforts-to-contain-drug-resistant-infections, date accessed: 2020-01-28.
8. Popowska, Magdalena, Rzeczycka, Marzenna, Miernik, Antoni, Krawczyk-Balska, Agata, Walsh, Fiona & Duffy, Brion. Influence of soil use on prevalence of tetracycline,streptomycin, and erythromycin resistance and associated resistance genes. *Antimicrob. Agents Chemother.* **56**, 1434–1443 (2012).
9. Sørensen, Thomas Lund, Blom, Marianne, Monnet, Dominique L., Frimodt-Møller, Niels, Poulsen, Rikke Lykke & Espersen, Frank. Transient Intestinal Carriage after Ingestion of Antibiotic-Resistant Enterococcus faecium from Chicken and Pork. *N. Engl. J. Med.* **345**, 1161–1166 (2001).
10. Levy, Stuart B., Fitzgerald, George B. & Macone, Ann B. Spread of antibiotic-resistant plasmids from chicken to chicken and from chicken to man. *Nature* **260**, 40–42 (1976).
11. de Boer, E., Zwartkruis-Nahuis, J. T. M., Wit, B., Huijsdens, X. W., de Neeling, A. J., Bosch, T., van Oosterom, R. A. A., Vila, A. & Heuvelink, A. E. Prevalence of methicillin-resistant Staphylococcus aureus in meat. *Int. J. Food Microbiol.* **134**, 52–56 (2009).
12. Normanno, G., Corrente, M., La Salandra, G., Dambrosio, A., Quaglia, N. C., Parisi, A., Greco, G., Bellacicco, A. L., Virgilio, S. & Celano, G. V. Methicillin-resistant Staphylococcus aureus (MRSA) in foods of animal origin product in Italy. *Int. J. Food Microbiol.* **117**, 219–222 (2007).
13. Hernando-Amado, Sara, Coque, Teresa M., Baquero, Fernando & Martínez, José L. Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nat. Microbiol.* **4**, 1432–1442 (2019).
14. Martínez, José L. Antibiotics and Antibiotic Resistance Genes in Natural Environments. *Science (80-.).* **321**, 365–367 (2008).

15. Hiltunen, Teppo, Virta, Marko & Laine, Anna-Liisa. Antibiotic resistance in the wild: An ecoevolutionary perspective. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, (2017).
16. Martinez, Jose L. General principles of antibiotic resistance in bacteria. *Drug Discov. Today Technol.* **11**, 33–39 (2014).
17. Segawa, Takahiro, Takeuchi, Nozomu, Rivera, Andres, Yamada, Akinori, Yoshimura, Yoshitaka, Barcaza, Gonzalo, Shinbori, Kunio, Motoyama, Hideaki, Kohshima, Shiro & Ushida, Kazunari. Distribution of antibiotic resistance genes in glacier environments. *Environ. Microbiol. Rep.* **5**, 127–134 (2013).
18. Su, Sicong, Li, Chenyu, Yang, Jiping, Xu, Qunying, Qiu, Zhigang, Xue, Bin, Wang, Shang, Zhao, Chen, Xiao, Zhonghai, Wang, Jingfeng & Shen, Zhiqiang. Distribution of Antibiotic Resistance Genes in Three Different Natural Water Bodies-A Lake, River and Sea. *Int. J. Environ. Res. Public Health* **17**, 1–12 (2020).
19. Helfricht, K. Veränderungen des Massenhaushaltes am Hallstätter Gletscher seit 1856. (Leopold-Franzens-Universität Innsbruck, 2009).
20. Rode, Matthias, Schnepfleitner, Harald, Sass, Oliver, Kellerer-Pirklbauer, Andreas & Gitschthaler, Christoph. Permafrost distribution and conditions at the headwalls of two receding glaciers (Schladminger and Hallstadt glaciers) in the Dachstein Massif, Northern Calcareous Alps, Austria. *Cryosph. Discuss.* **i**, 1–20 (2019).
21. Google maps. https://www.google.at/maps, date accessed: 2020-01-26.
22. Hartl, Lea, Stocker-Waldhuber, Martin Reingruber, Klaus & Fischer, Andrea. Hallstätter Gletscher. (2013).
23. Rime, Thomas, Hartmann, Martin, Brunner, Ivano, Widmer, Franco, Zeyer, Josef & Frey, Beat. Vertical distribution of the soil microbiota along a successional gradient in a glacier forefield. *Mol. Ecol.* **24**, 1091–1108 (2015).
24. Micklin, Philip. The past, present, and future Aral Sea. *Lakes Reserv. Res. Manag.* **15**, 193–213 (2010).
25. Izhitskiy, A. S., Zavialov, P. O., Sapozhnikov, P. V., Kirillin, G. B., Grossart, H. P., Kalinina, O. Y., Zalota, A. K., Goncharenko, I. V. & Kurbaniyazov, A. K. Present state of the Aral Sea: Diverging physical and biological characteristics of the residual basins. *Sci. Rep.* **6**, 1–9 (2016).
26. Bekchanov, Maksud, Ringler, Claudia & Bhaduri, Anik. A Water Rights Trading Approach to Increasing Inflows to the Aral Sea. *L. Degrad. Dev.* **29**, 952–961 (2018).
27. Krapivin, Vladimir, Mkrtchyan, Ferdenant & Rochon, Gilbert. Hydrological Model for Sustainable Development in the Aral Sea Region. *Hydrology* **6**, 91 (2019).
28. Karlykhanov, Orazkhan K. & Toktaganova, Gulzhaz B. The Assessment of Irrigated Land Salinization in the Aral Sea Region. *Int. J. Environ. Sci. Educ.* **11**, 7946–7960 (2016).
29. Micklin, Philip. The Aral Sea Disaster. *Annu. Rev. Earth Planet. Sci.* **35**, 47–72 (2007).
30. Jensen, S., Mazhitova, Z. & Zetterström, R. Environmental pollution and child health in the Aral Sea region in Kazakhstan. *Sci. Total Environ.* **206**, 187–193 (1997).
31. Latlong. www.latlong.net/Show-Latitude-Longitude.html, date accessed: 2020-02-10.
32. Openstreetmap. www.openstreetmap.org/copyright, date accessed: 2020-02-10.
33. Caporaso, J. Gregory, Lauber, Christian L., Walters, William A., Berg-Lyons,

Donna, Lozupone, Catherine A., Turnbaugh, Peter J., Fierer, Noah & Knight, Rob. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4516–4522 (2011).

34. Caporaso, J. Gregory, Lauber, Christian L., Walters, William A., Berg-Lyons, Donna, Huntley, James, Fierer, Noah, Owens, Sarah M., Betley, Jason, Fraser, Louise, Bauer, Markus, Gormley, Niall, Gilbert, Jack A., Smith, Geoff & Knight, Rob. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624 (2012).
35. Bolyen, Evan, Rideout, Jai Ram, Dillon, Matthew R., Bokulich, Nicholas A., Abnet, Christian C., Al-Ghalith, Gabriel A., Alexander, Harriet, Alm, Eric J., Arumugam, Manimozhiyan, Asnicar, Francesco, Bai, Yang, Bisanz, Jordan E., Bittinger, Kyle, Brejnrod, Asker, Brislawn, Colin J., Brown, C. Titus, Callahan, Benjamin J., Caraballo-Rodríguez, Andrés Mauricio, Chase, John *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
36. Martin, Marcel. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* (2011).
37. Callahan, Benjamin J., Mcmurdie, Paul J., Rosen, Michael J., Han, Andrew W., A, Amy Jo & Holmes, Susan P. Dada2: High resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581–583 (2016).
38. Rognes, Torbjørn, Flouri, Tomáš, Nichols, Ben, Quince, Christopher & Mahé, Frédéric. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **2016**, 1–22 (2016).
39. Glöckner, Frank Oliver, Yilmaz, Pelin, Quast, Christian, Gerken, Jan, Beccati, Alan, Ciuprina, Andreea, Bruns, Gerrit, Yarza, Pablo, Peplies, Jörg, Westram, Ralf & Ludwig, Wolfgang. 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J. Biotechnol.* **261**, 169–176 (2017).
40. Quast, Christian, Pruesse, Elmar, Yilmaz, Pelin, Gerken, Jan, Schweer, Timmy, Yarza, Pablo, Peplies, Jörg & Glöckner, Frank Oliver. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2013).
41. Yilmaz, Pelin, Parfrey, Laura Wegener, Yarza, Pablo, Gerken, Jan, Pruesse, Elmar, Quast, Christian, Schweer, Timmy, Peplies, Jörg, Ludwig, Wolfgang & Glöckner, Frank Oliver. The SILVA and 'all-species Living Tree Project (LTP)' taxonomic frameworks. *Nucleic Acids Res.* **42**, 643–648 (2014).
42. Zakrzewski, Martha, Proietti, Carla, Ellis, Jonathan J., Hasan, Shihab, Brion, Marie Jo, Berger, Bernard & Krause, Lutz. Calypso: A user-friendly web-server for mining and visualizing microbiome-environment interactions. *Bioinformatics* **33**, 782–783 (2017).
43. Vázquez-Baeza, Yoshiki, Gonzalez, Antonio, Smarr, Larry, McDonald, Daniel, Morton, James T., Navas-Molina, Jose A. & Knight, Rob. Bringing the Dynamic Microbiome to Life with Animations. *Cell Host Microbe* **21**, 7–10 (2017).
44. Segata, Nicola, Izard, Jacques, Waldron, Levi, Gevers, Dirk, Miropolsky, Larisa, Garrett, Wendy S. & Huttenhower, Curtis. Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011).
45. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, date accessed: 2020-05-10 (2010).
46. FastQC. https://qubeshub.org/resources/fastqc, date accessed: 2020-05-06 (2015).
47. Bolger, Anthony M., Lohse, Marc & Usadel, Bjoern. Trimmomatic: A flexible

trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
48. Li, Dinghua, Liu, Chi Man, Luo, Ruibang, Sadakane, Kunihiko & Lam, Tak Wah. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
49. Langmead, B. & Salzberg, S. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
50. Li, Heng, Handsaker, Bob, Wysoker, Alec, Fennell, Tim, Ruan, Jue, Homer, Nils, Marth, Gabor, Abecasis, Goncalo & Durbin, Richard. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
51. Li, Heng. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
52. Eren, A. Murat, Esen, Ozcan C., Quince, Christopher, Vineis, Joseph H., Morrison, Hilary G., Sogin, Mitchell L. & Delmont, Tom O. Anvi'o: An advanced analysis and visualization platformfor 'omics data. *PeerJ* **2015**, 1–29 (2015).
53. Galperin, Michael Y., Makarova, Kira S., Wolf, Yuri I. & Koonin, Eugene V. Expanded Microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* **43**, D261–D269 (2015).
54. Menzel, Peter, Ng, Kim Lee & Krogh, Anders. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* **7**, (2016).
55. Alcock, Brian P., Raphenya, Amogelang R., Lau, Tammy T. Y., Tsang, Kara K., Bouchard, Mégane, Edalatmand, Arman, Huynh, William, Nguyen, Anna Lisa V., Cheng, Annie A., Liu, Sihan, Min, Sally Y., Miroshnichenko, Anatoly, Tran, Hiu Ki, Werfalli, Rafik E., Nasir, Jalees A., Oloni, Martins, Speicher, David J., Florescu, Alexandra, Singh, Bhavya *et al.* CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **48**, D517–D525 (2020).
56. Ondov, Brian D., Bergman, Nicholas H. & Phillippy, Adam M. Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* **12**, 385 (2011).
57. Krzywinski, Martin, Schein, Jacqueline, Birol, Inanç, Connors, Joseph, Gascoyne, Randy, Horsman, Doug, Jones, Steven J. & Marra, Marco A. Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).
58. Mauri, Michele, Elli, Tommaso, Caviglia, Giorgio, Uboldi, Giorgio & Azzi, Matteo. RAWGraphs: A visualisation platform to create open outputs. *ACM Int. Conf. Proceeding Ser.* **Part F1313**, (2017).
59. Naidoo, Yashini, Valverde, Angel, Cason, Errol D., Pierneef, Rian E. & Cowan, Don A. A clinically important, plasmid-borne antibiotic resistance gene (β-lactamase TEM-116) present in desert soils. *Sci. Total Environ.* **719**, 137497 (2020).
60. Guarddon, M., Miranda, J. M., Rodríguez, J. A., Vázquez, B. I., Cepeda, A. & Franco, C. M. Real-time polymerase chain reaction for the quantitative detection of tetA and tetB bacterial tetracycline resistance genes in food. *Int. J. Food Microbiol.* **146**, 284–289 (2011).
61. Heuer, Holger & Smalla, Kornelia. Manure and sulfadiazine synergistically increased bacterial antibiotic resistance in soil over at least two months. *Environ. Microbiol.* **9**, 657–666 (2007).
62. Knapp, C. W., Yang, R. K., Hanfelt, M. M., Smith, M. S., Aga, D. S., Graham, D. W., S, Van Overbeek L., Wellington, E. M. H., Egan, S., Smalla, K., Heuer, H.,

Collard, J., Karagouni, A. D., Nikolakopoulou, T. L. & Prevalence, J. D. Van Elsas. Abundance of six tetracycline resistance genes in wastewater lagoons at cattle feedlots with different antibiotic use strategies- supplementary data. *Gene* (2007).

63. Dallenne, Caroline, da Costa, Anaelle, Decré, Dominique, Favier, Christine & Arlet, Guillaume. Development of a set of multiplex PCR assays for the detection of genes encoding important β-lactamases in Enterobacteriaceae. *J. Antimicrob. Chemother.* **65**, 490–495 (2010).
64. Wang, Lijun, Gu, Haitong & Lu, Xinxin. Rapid low-cost detection of Klebsiella pneumoniae carbapenemase genes by internally controlled real-time PCR. *J. Microbiol. Methods* **91**, 361–363 (2012).
65. Liu, Yuan, Jia, Yuqian, Yang, Kangni, Li, Ruichao, Xiao, Xia & Wang, Zhiqiang. Antagonizing Vancomycin Resistance in Enterococcus by Surface Localized Antimicrobial Display-Derived Peptides. *ACS Infect. Dis.* 1–9 (2019) doi:10.1021/acsinfecdis.9b00164.
66. Walsh, F., Ingenfeld, A., Zampicolli, M., Hilber-Bodmer, M., Frey, J. E. & Duffy, B. Real-time PCR methods for quantitative monitoring of streptomycin and tetracycline resistance genes in agricultural ecosystems. *J. Microbiol. Methods* **86**, 150–155 (2011).
67. Stoll, C., Sidhu, J. P. S., Tiehm, A. & Toze, S. Prevalence of clinically relevant antibiotic resistance genes in surface water samples collected from Germany and Australia. *Environ. Sci. Technol.* **46**, 9716–9726 (2012).
68. Brugman, Sylvia, Ikeda-Ohtsubo, Wakako, Braber, Saskia, Folkerts, Gert, Pieterse, Corné M. J. & Bakker, Peter A. H. M. A Comparative Review on Microbiota Manipulation: Lessons From Fish, Plants, Livestock, and Human Research. *Front. Nutr.* **5**, 1–15 (2018).
69. Flandroy, Lucette, Poutahidis, Theofilos, Berg, Gabriele, Clarke, Gerard, Dao, Maria Carlota, Decaestecker, Ellen, Furman, Eeva, Haahtela, Tari, Massart, Sébastien, Plovier, Hubert, Sanz, Yolanda & Rook, Graham. The impact of human activities and lifestyles on the interlinked microbiota and health of humans and of ecosystems. *Sci. Total Environ.* **627**, 1018–1038 (2018).
70. Winkler, Manuela, Tribsch, Andreas, Paun, Ovidiu, Englisch, Thorsten & Schönswetter, Peter. Pleistocene distribution range shifts were accompanied by breeding system divergence within Hornungia alpina (Brassicaceae) in the Alps. *Mol. Phylogenet. Evol.* **54**, 571–582 (2010).
71. Schönswetter, Peter, Solstad, Heidi, García, Pedro Escobar & Elven, Reidar. A combined molecular and morphological approach to the taxonomically intricate European mountain plant Papaver alpinum s.l. (Papaveraceae) - Taxa or informal phylogeographical groups? *Taxon* **58**, 1326–1343 (2009).
72. Infoflora. https://www.infoflora.ch/de/flora/sedum-atratum.html, date accessed: 2020-04-28.
73. FAO-Food and Agriculture Organization of the United Nations. *Handbook for saline soil management*. *Fao* (2018).
74. Breckle, SW, Scheffer, A. & Wucherer, W. Halophytes on the Dry Sea Floor of the Aral Sea. in *Sustainable Land Use in Deserts* (2001).
75. Williams, W. D. & Aladin, N. V. The Aral Sea: Recent limnological changes and their conservation significance. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **1**, 3–23 (1991).
76. Chaban, Bonnie, Ng, Sandy Y. M. & Jarrell, Ken F. Archaeal habitats - From the extreme to the ordinary. *Can. J. Microbiol.* **52**, 73–116 (2006).
77. Kuczynski, Justin, Lauber, Christian L., Walters, William A., Parfrey, Laura

Wegener, Clemente, Josã C., Gevers, Dirk & Knight, Rob. Experimental and analytical tools for studying the human microbiome. *Nat. Rev. Genet.* **13**, 47–58 (2012).
78. Pausan, Manuela R., Csorba, Cintia, Singer, Georg, Till, Holger, Schöpf, Veronika, Santigli, Elisabeth, Klug, Barbara, Högenauer, Christoph, Blohs, Marcus & Moissl-Eichinger, Christine. Exploring the Archaeome: Detection of Archaeal Signatures in the Human Body. *Front. Microbiol.* **10**, 1–13 (2019).
79. Najjari, Afef, Elshahed, Mostafa S., Cherif, Ameur & Youssef, Noha H. Patterns and determinants of halophilic archaea (Class halobacteria) diversity in tunisian endorheic salt lakes and sebkhet systems. *Appl. Environ. Microbiol.* **81**, 4432–4441 (2015).
80. Musilova, Michaela, Wright, Gary, Ward, John M. & Dartnell, Lewis R. Isolation of Radiation-Resistant Bacteria from Mars Analog Antarctic Dry Valleys by Preselection, and the Correlation between Radiation and Desiccation Resistance. *Astrobiology* **15**, 1076–1090 (2015).
81. Kearl, Jennifer, McNary, Caitlyn, Lowman, J. Scott, Mei, Chuansheng, Aanderud, Zachary T., Smith, Steven T., West, Jason, Colton, Emily, Hamson, Michelle & Nielsen, Brent L. Salt-tolerant halophyte rhizosphere bacteria stimulate growth of alfalfa in salty soil. *Front. Microbiol.* **10**, (2019).
82. Siddikee, M. A., Chauhan, P. S., Anandham, R., Han, Gwang Hyun & Sa, Tongmin. Isolation, characterization, and use for plant growth promotion under salt stress, of ACC deaminase-producing halotolerant bacteria derived from coastal soil. *J. Microbiol. Biotechnol.* **20**, 1577–1584 (2010).
83. Cheng, Bin, Li, Chunfang, Lai, Qiliang, Du, Miaofen, Shao, Zongze, Xu, Ping & Yang, Chunyu. Sinomicrobium pectinilyticum sp. nov., A pectinase-producing bacterium isolated from alkaline and saline soil, And emended description of the genus Sinomicrobium. *Int. J. Syst. Evol. Microbiol.* **64**, 2939–2943 (2014).
84. Woźniak, Małgorzata & Gałązka, Anna. The Rhizosphere Microbiome and Its Beneficial Effects on Plants - Current Knowledge and Perspectives. *Adv. Microbiol.* **58**, 59–69 (2019).
85. Gauthier, M. J., Lafay, B., Christen, R., Fernandez, L., Acquaviva, M., Bonin, P. & Bertrand, J. C. Marinobacter hydrocarbonoclasticus gen. nov,, sp. nov,,a New , Extremely Halotolerant , Hydrocarbon-Degrading Marine Bacterium. *Int. J. Syst. Bacteriol.* **42**, 568–576 (1992).
86. Amin, Shady A., Green, David H., Waheeb, Dhuha Al, Gärdes, Astrid & Carrano, Carl J. Iron transport in the genus Marinobacter. *BioMetals* **25**, 135–147 (2012).
87. Miller, L. P., Crawford, J. T. & Shinnick, T. M. The rpoB gene of Mycobacterium tuberculosis. *Antimicrob. Agents Chemother.* **38**, 805–811 (1994).
88. Decroo, Tom, de Jong, Bouke C., Piubello, Alberto, Lynen, Lutgarde & Van Deun, Armand. Tuberculosis treatment: one-shot approach or cascade of regimens? *Lancet Respir. Med.* **8**, e4–e5 (2020).
89. Telenti, A., Imboden, P., Marchesi, F., Matter, L., Schopfer, K., Bodmer, T., Lowrie, D., Colston, M. J. & Cole, S. Detection of rifampicin-resistance mutations in Mycobacterium tuberculosis. *Lancet* **341**, 647–651 (1993).
90. Mariam, Deneke H., Mengistu, Yohannes, Hoffner, Sven E. & Andersson, Dan I. Effect of rpoB mutations on fitness of Mycobacterium tuberculosis. *Antimicrob Agents Chemother* **48**, 1289–1294 (2004).
91. Rodríguez-Verdugo, Alejandra, Gaut, Brandon S. & Tenaillon, Olivier. Evolution of escherichia coli rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evol. Biol.* **13**, (2013).
92. Coyne, Sébastien, Rosenfeld, Nicolas, Lambert, Thierry, Courvalin, Patrice &

Périchon, Bruno. Overexpression of resistance-nodulation-cell division pump AdeFGH confers multidrug resistance in Acinetobacter baumannii. *Antimicrob. Agents Chemother.* **54**, 4389–4393 (2010).
93. Pedersen, Bjørn P., Stokes, David L. & Apell, Hans Jürgen. The KdpFABC complex - K+ transport against all odds. *Mol. Membr. Biol.* **35**, 21–38 (2019).
94. Ali, Maria K., Li, Xinfeng, Tang, Qing, Liu, Xiaoyu, Chen, Fang, Xiao, Jinfeng, Ali, Muhammad, Chou, Shan Ho & He, Jin. Regulation of inducible potassium transporter KdpFABC by the KdpD/KdpE two-component system in Mycobacterium smegmatis. *Front. Microbiol.* **8**, (2017).
95. Fernandes, Pedro, Sommer Ferreira, Bruno & Sampaio Cabral, Joaquim Manuel. Solvent tolerance in bacteria: Role of efflux pumps and cross-resistance with antibiotics. *Int. J. Antimicrob. Agents* **22**, 211–216 (2003).
96. Gillings, Michael, Boucher, Yan, Labbate, Maurizio, Holmes, Andrew, Krishnan, Samyuktha, Holley, Marita & Stokes, H. W. The evolution of class 1 integrons and the rise of antibiotic resistance. *J. Bacteriol.* **190**, 5095–5100 (2008).
97. Koczura, Ryszard, Mokracka, Joanna, Taraszewska, Agata & Łopacinska, Natalia. Abundance of Class 1 Integron-Integrase and Sulfonamide Resistance Genes in River Water and Sediment Is Affected by Anthropogenic Pressure and Environmental Factors. *Microb. Ecol.* **72**, 909–916 (2016).
98. Lu, Zihao, Na, Guangshui, Gao, Hui, Wang, Lijun, Bao, Chenguang & Yao, Ziwei. Fate of sulfonamide resistance genes in estuary environment and effect of anthropogenic activities. *Sci. Total Environ.* **527–528**, 429–438 (2015).
99. Krawczyk, Pawel S., Lipinski, Leszek & Dziembowski, Andrzej. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res.* **46**, e35 (2018).
100. Padungtod, Pawin, Tribuddharat, Chanwit & Chuanchuen, Rungtip. Widespread presence of dfrA12 and its association with dfrA12-aadA2 cassette in Salmonella enterica isolates from swine. *Southeast Asian J. Trop. Med. Public Health* **42**, 1471–1476 (2011).
101. Pletzer, Daniel, Stahl, Antje, Oja, Anna Elisabeth & Weingart, Helge. Role of the cell envelope stress regulators BaeR and CpxR in control of RND-type multidrug efflux pumps and transcriptional cross talk with exopolysaccharide synthesis in Erwinia amylovora. *Arch. Microbiol.* **197**, 761–772 (2015).
102. Nishino, Kunihiko, Senda, Yasuko & Yamaguchi, Akihito. CRP Regulator Modulates Multidrug Resistance of Escherichia coli by Repressing the mdtEF Multidrug Efflux Genes. *J. Antibiot. (Tokyo).* **61**, 120–127 (2008).
103. Yuan, Ji, Chow, Dar Chone, Huang, Wanzhi & Palzkill, Timothy. Identification of a β-lactamase inhibitory protein variant that is a potent inhibitor of Staphylococcus PC1 β-lactamase. *J. Mol. Biol.* **406**, 730–744 (2011).
104. Ashenafi, Meseret, Ammosova, Tatiana, Nekhai, Sergei & Byrnes, W. Malcolm. Purification and characterization of aminoglycoside phosphotransferase APH(6)-Id, a streptomycin-inactivating enzyme. *Mol. Cell. Biochem.* **387**, 207–216 (2014).

VII. Appendix

1. QIIME 2 protocol:

#0) demultiplex

cutadapt -e 0.17 --no-indels -g file:barcodes.fasta -G file:barcodes.fasta --minimum-length 50 -o trimmed-{name}.1.fastq.gz -p trimmed-{name}.2.fastq.gz AD.1.fastq.gz AD.2.fastq.gz

#1) import into a QIIME Artifact file

qiime tools import --type 'SampleData[PairedEndSequencesWithQuality]' --input-path /media/sf_Mintfolder/manifest.txt --output-path /media/sf_Mintfolder/AD_analysis/AD_demux.qza --input-format PairedEndFastqManifestPhred33V2

#2) denoise, dereplicate, remove chimera with dada2

qiime dada2 denoise-paired --i-demultiplexed-seqs /media/sf_Mintfolder/AD_analysis/AD_demux.qza --p-trunc-len-f 220 --p-trunc-len-r 130 --p-n-threads 2 --o-table /media/sf_Mintfolder/AD_analysis/AD_denoisrepchim_table.qza --o-representative-sequences /media/sf_Mintfolder/AD_analysis/AD_denoisrepchim_RSV.qza --o-denoising-stats /media/sf_Mintfolder/AD_analysis/AD_denoisrepchim_log.qza

#3) merge

cat '/media/sf_sf_mint/Qiime2/data_neu_18.11/unknown_anchored_sorted_to_0mismatch_unanchored/trimmeduk-A_51_A.1.fastq.gz' '/media/sf_sf_mint/Qiime2/data_neu_18.11/original_1mismatch_anchored/trimmed-A_51_A.1.fastq.gz' > trimmedcat-A_51_A.1.fastq.gz'

#3.2 import reference database

qiime tools import –input-path media/sf_sf_mint/Qiime2/database/silva_132_99_16S.fna --output-/media/sf_sf_mint/Qiime2/database/silva_132_99_16S.fna_featuresequence.qza --type FeatureData[Sequence]

qiime tools import --input-format HeaderlessTSVTaxonomyFormat --input-path /media/sf_Mintfolder/databases/SILVA_132_QIIME_release/taxonomy/16S_only/99/consensus_taxonomy_7_levels.txt --output-path /media/sf_Mintfolder/databases/silva_132_99_16S_consensus_taxonomy_7_levels_featuretaxonomy.qza --type FeatureData[Taxonomy]

#4) classify vsearch

qiime feature-classifier classify-consensus-vsearch --i-query '/media/sf_sf_mint/Qiime2/data_neu_18.11/featuredata_rep-seqs_dachstein_only.qza' --i-reference-reads '/media/sf_sf_mint/Qiime2/database/silva_132_99_16S_featuresequence.qza' --i-reference-taxonomy '/media/sf_sf_mint/Qiime2/database/silva_132_99_16S_consensus_taxonomy_7_levels_featuretaxonomy.qza' --o-classification '/media/sf_sf_mint/Qiime2/data_neu_18.11/classification_dachstein_only.qza' --verbose

#5 Chloroplasts and mitochondria filtering

qiime taxa filter-table --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/featuretable_dachstein_only.qza' --i-taxonomy '/media/sf_sf_mint/Qiime2/data_neu_18.11/classification_dachstein_only.qza' --p-

exclude mitochondria,chloroplast --o-filtered-table featuretable_dachstein_only_filtered.qza

#5.1 unassigned filtering

qiime taxa filter-table --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/featuretable_triplicates_grouped_aral_only.qza' --i-taxonomy '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/classification_aral_only.qza' --p-exclude unassigned --o-filtered-table featuretable_triplicates_grouped_aral_only_no_unassigned.qza

#6 merging of triplicates

qiime feature-table group --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/featuretable_aral_only_filtered.qza' --p-axis sample --m-metadata-file '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/meta_column_triplicate_merch_ARAL_2.txt' --m-metadata-column "sampleidnew" --p-mode sum --o-grouped-table featuretable_triplicates_grouped_aral_only

#7 Collaps

qiime taxa collapse --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/featuretable_aral_only_ftriplicates_grouped_no_unasigned.qza' --i-taxonomy '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/classification_aral_only.qza' --p-level 7 --o-collapsed-table featuretable_aral_only_ftriplicates_grouped_no_unasigned_collapsed --verbose

#8 subsampling

qiime feature-table rarefy --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/ohne_NCs/featuretable_triplicates_grouped_aralsee_no_unassigned_NCs_taxaabgezogen_collapsed.qza' --p-sampling-depth 17263 --o-rarefied-table featuretable_triplicates_grouped_aralsee_no_unassigned_NCs_taxaabgezogen_collapsed_subsampled

qiime feature-table rarefy --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/dachstein_only/no_ncs/featuretable_nachNC_taxa_Abzug_collapsed.qza' --p-sampling-depth 13603 --o-rarefied-table featuretable_nachNC_taxa_Abzug_collapsed_subsampled

#9 Visualization

#9.1 alpha diversity

qiime taxa barplot --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/featuretable.qza' --i-taxonomy '/media/sf_sf_mint/Qiime2/data_neu_18.11/classification.qza' --m-metadata-file '/media/sf_sf_mint/Qiime2/meta.txt' --o-visualization visualisation_taxa_barplot

qiime diversity alpha --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/featuretable.qza' --p-metric 'shannon' --o-alpha-diversity alpha_diversity.qza

#9.2 beta diversity

qiime diversity beta-group-significance --i-distance-matrix '/media/sf_sf_mint/Qiime2/data_neu_18.11/dachstein_only/no_ncs/featuretable_nachNC_taxa_Abzug_distance_matrix.qza' --m-metadata-file '/media/sf_sf_mint/Qiime2/data_neu_18.11/dachstein_only/no_ncs/calypso/meta_after_merging_for_barplot_dachstein_only_yearsafterice.txt' --m-metadata-column years-without-water-or-ice1 --p-pairwise --o-visualization featuretable_triplicates_grouped_aralsee_no_unassigned_NCs_taxaabgezogen_betadiv_years

qiime diversity beta --i-table '/media/sf_sf_mint/Qiime2/data_neu_18.11/featuretable.qza' --p-metric 'braycurtis' --o-distance-matrix DA_distance_matrix

qiime diversity pcoa --i-distance-matrix '/media/sf_sf_mint/Qiime2/data_neu_18.11/DA_distance_matrix.qza' --o-pcoa DA_PCoA_matrix

qiime emperor plot --i-pcoa '/media/sf_sf_mint/Qiime2/data_neu_18.11/DA_PCoA_matrix.qza' --m-metadata-file '/media/sf_sf_mint/Qiime2/meta.txt' --o-visualization DA_emperor_visualisation

#convertions

biom to tsv

biom convert -i '/media/sf_sf_mint/Qiime2/data_neu_18.11/DA/test.biom' -o test.txt --to-tsv

featuretable to biom

biom convert -i '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/ohne_NCs/Taxonomy+Data_ohne_NCs.txt' -o featuretable_triplicates_grouped_aralsee_no_unassigned_NCs_taxaabgezogen --to-hdf5 --table-type="OTU table"

qiime tools import --input-path '/media/sf_sf_mint/Qiime2/data_neu_18.11/aral_only/ohne_NCs/featuretable_triplicates_grouped_aralsee_no_unassigned_NCs_taxaabgezogen' --type 'FeatureTable[Frequency]' --input-format BIOMV210Format --output-path featuretable_triplicates_grouped_aralsee_no_unassigned_NCs_taxaabgezogen.qza

classification txt to qza

qiime tools import --type FeatureData[Taxonomy] --input-path '/media/sf_sf_mint/Qiime2/data_neu_18.11/DA/classification_taxonomy_with_other.txt' --output-path classification_mergedID.qza

2. Removed Taxa

Aral Sea dataset:

ASV	Kingdom	Phylum	Class	Order	Family	Genus	Reason for removal
0ff9a3e11785afae02b80dbb3d7bb218	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Halobacillus	taxa in AS_NC
2b5e64b305728566bbecd9c285ac2b5d	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Nesterenkonia	taxa in AS_NC
9cc08a7fe344d950abd0f2734266c4b2	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	taxa in AS_NC
b391dfffdcc75c247cc529e82e55652d	Bacteria	Chloroflexi	KD4-96	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in AS_NC
163e1e5c9a2635d3f3bd7d4922f0e544	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Burkholderia-Caballeronia-Paraburkholderia	taxa in AS_NC
ed13da9ea6575ac2bc815586d064d762	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	uncultured	taxa in AS_NC
5f107787e615e69b05383d7a1e2538fb	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium	taxa in AS_NC
3cd7ec22048bf733dce99ab861eb5988	Bacteria	Firmicutes	Bacilli	Bacillales			taxa in AS_NC
5dc6dc7ce2f6fcf52c919645d9ccce60	Bacteria	Proteobacteria	Deltaproteobacteria	Oligoflexales	Oligoflexaceae	Oligoflexus	taxa in AS_NC
403abb56bd1f13c8b26a7a7f7b52770d	Archaea	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	uncultured archaeon	taxa in AS_NC and PCR_NC
2ec6eb3de57fc12eaef4cbfc6c3314a7	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Halobacillus	taxa in PCR_NC
97af0565c040d8c28eb495a49fa73dad	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in PCR_NC
b92c12801a462abb993cf52f6c3d9406	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
8cd8b6bfed8d925d71a7510a58ddb32c	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Halobacillus	taxa in PCR_NC
e97198f6f59eab9c5bf70f3c64be5712	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
790be0c6974b2385365b3b16a7bf5322	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
7a8f286934bf73d26f00d06fad86acc8	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
c338cee355e9b1189c86476c426f8eb9	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
272b57ad42945ade456620056fc3ce80	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
2c2bea982ae5be8246700e96458987bb	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Haloterrigena	taxa in PCR_NC
1aef8bb9aac2c996c860b54d8a88442c	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	Ferruginibacter	taxa in PCR_NC
5fe39e8826654b0887c4a2e18f3bbc69	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
cdebed842c7ea83145a4f2258d959a50	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Pontibacillus	taxa in PCR_NC
daad403a3c8593cb7e0fc9ae64507503	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Pontibacillus	taxa in PCR_NC
5164626fccbddfab353196d453bb5e3e	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Haloarcula	taxa in PCR_NC
98d9ee0239c0c750dfd4d6879a6a8b53	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
5a77d84c7594224338252b20d0e5006e	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
94bafe6e7b0fd5ece4aa353081b6d798	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in PCR_NC
c3d0796e602cdda8475948e700f2c1fe	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
cd32bf961dd25e8f25abd7037f744d9d	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
853219b52702aa68d9c6f5c8160380ea	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
032e7ad856fefbc54db50dab1b785863	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Haloarcula	taxa in PCR_NC
21a3a87a0bf673cfeb467e9343c01363	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Haloarcula	taxa in PCR_NC
9b9bb090113836e28bce146bdedb82c2	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	taxa in PCR_NC
8124867736ddf5b9c60e14e300b0505f	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Haloarcula	taxa in PCR_NC
25007237c14777654a01b65df4f24696	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	JGI 0001001-H03	taxa in PCR_NC
02cb0b0f8685661d3f3788663ac43597	Archaea	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	uncultured archaeon	taxa in PCR_NC
2a5719b4cb3d67f798b055bb0bfebd18	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Nesterenkonia	taxa in PCR_NC
055bf88af1e8cc581183ccef5e690640	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Nesterenkonia	taxa in PCR_NC
029d7672a259e3d1c6f182afb9d4130f	Bacteria	Bacteroidetes	Rhodothermia	Balneolales	Balneolaceae	uncultured	taxa in PCR_NC
12a727cf3709b3d6724099bf6999489e	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rubellimicrobium	taxa in PCR_NC

4f238b241370a477cd4a5f2cf5324059	Bacteria	Acidobacteria	Subgroup 6				taxa in PCR_NC
e9484318818cee93323164de8a2fe4a5	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC
12df2e0dd541423665509febfdcf4567	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	uncultured	taxa in PCR_NC
db01b0746f06c12a4bb88e819fc4d638	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halolamina	taxa in PCR_NC
4ad1ba0447bab76ae30ac2c5f9ebf638	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halococcaceae		taxa in PCR_NC
d5d62cf8a16e1144e82e3614e6c2395b	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	JGI 0001001-H03	taxa in PCR_NC
46c4acd1b5b9c278af45975425ca19c1	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC
9d41cbca8b37373a79f2cb16d7b60458	Bacteria	Chloroflexi	KD4-96	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC
c5327762d3860f1066a63aa28ffcbb82	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in PCR_NC
ec3ff0b35da56b8930611cb70223dc66	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in PCR_NC
3fc0ee6ef16d4977e3811396cb14dbea	Bacteria	Chloroflexi	KD4-96	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC
59c267e2d6fd1ca6477bed8aedcf016c	Bacteria	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Rubritaleaceae	Luteolibacter	taxa in PCR_NC
1584bda4b06f5b21588ce273096d0303	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Nesterenkonia	taxa in PCR_NC
569b7dc34beb98ab4d511f126dfcaee5	Bacteria	Bacteroidetes	Rhodothermia	Balneolales	Balneolaceae	uncultured	taxa in PCR_NC
80c9b35bcf4ac13dc1101386a779067f	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in PCR_NC
eb7e701b11ff0ad37c95d0a32d70cad3	Bacteria	Chloroflexi	Anaerolineae	Caldilineales	Caldilineaceae	uncultured	taxa in PCR_NC
13ea357817dd4276afdc01a860d20945	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Kocuria	taxa in PCR_NC
6168edba7206a2b4220a7d6d8648330a	Bacteria	Proteobacteria	Deltaproteobacteria	Desulfuromonadales	Geobacteraceae	Geobacter	taxa in PCR_NC
fa1620623cfa2a2030f277b1695e01bf	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC
0c261cb300d4e1cfc5c59ae3099692cc	Bacteria	Chloroflexi	OLB14	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC
dc9b0496419e8f4ef082527e462207eb	Bacteria	Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	taxa in PCR_NC
9a7b31b73a05776daf4e9d92dd34956d	Bacteria	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Rubritaleaceae	Luteolibacter	taxa in PCR_NC
b2a8d4f63fddc9918a507789c9119502	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	JG30-KF-CM45	uncultured bacterium	taxa in PCR_NC
c5227d674bb2db2fab607c7861a2ef02	Bacteria	Chloroflexi	TK10				taxa in PCR_NC
2cc404bd46f9bdd940ddce792b2c2ebb	Bacteria	Planctomycetes	Planctomycetacia	Pirellulales	Pirellulaceae	Blastopirellula	taxa in PCR_NC
bc9fc445f23e0b544dc9f98339927cd4	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Paenalcaligenes	taxa in PCR_NC
d36b9db25cf20829de50a8c61cfbd3ed	Bacteria	Verrucomicrobia	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter	taxa in PCR_NC
93389ea87016a24a185fa28c2eb93ce0	Bacteria	Chloroflexi	KD4-96	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC
648435c5d2edad98e12ff25746c28c21	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	Pontibacter	taxa in AS_NC
44d73fe225852240b3e2913f8aed0bf3	Archaea	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	uncultured archaeon	taxa in AS_NC
7026a5ed099e71c4013b4c37caac6bc1	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	taxa in AS_NC
ae6caa27684b7e09602d46e7b7ac00e4	Bacteria	Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	taxa in AS_NC
23f664a018bdb4c1b74ad90198c26b45	Bacteria	Verrucomicrobia	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Chthoniobacter	taxa in AS_NC
cd1e9fbf42b741fed5f677fbb0df2690	Bacteria	Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	taxa in AS_NC
e18781dd62c91a4f9ea20f073e654b8c	Bacteria	Planctomycetes	Planctomycetacia	Pirellulales	Pirellulaceae	Pirellula	taxa in AS_NC

Dachstein data set:

ASV	**Kingdom**	**Phylum**	**Class**	**Order**	**Family**	**Genus**	**Reason for removal**
c7c556276d982788d084d7d8b580d928	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Halomicrobium	taxa in Da_NC_01_A
d9f52cbd0d38f85d9cb8815313ee12cc	Bacteria	Actinobacteria	Nitriliruptoria	Euzebyales	Euzebyaceae	uncultured	taxa in Da_NC_01_A
0165a10a6c2567ada79f045f6a2f8dee	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	Pontibacter	taxa in Da_NC_01_A
c4011a2aa45c9f36a936de3208a0d11a	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Noviherbaspirillum	taxa in Da_NC_01_A
a3f2933386b409572e70993d80f4b593	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rubellimicrobium	taxa in Da_NC_01_A
af1275d7fd665dccfa1d2469a3d6daa1	Archaea	Euryarchaeota	Halobacteria	Halobacteriales			taxa in Da_NC_01_A

213016b2c63eae181931d1d8aaa622a9	Bacteria	Actinobacteria					taxa in Da_NC_01_A
91f4531477bbe74afe7e3c1ac74d4235	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	JG30-KF-CM45	uncultured bacterium	taxa in Da_NC_01_A
d661638f1a825b7f4ec90e8790b09145	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
ac93fb65d979109135ff85f51d5226e5	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
4ca63434846494a39d992b4421d394cd	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Lysobacter	taxa in Da_NC_01_A
ea0aa437baa19997c976eb698cef3a98	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halobacteriaceae		taxa in Da_NC_01_A
ca12f3d204b60340413c5f74b40555d6	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
a2dfacccaf88d61b9e42a5593f9ff9c0	Bacteria	Actinobacteria	Nitriliruptoria	Euzebyales	Euzebyaceae	uncultured	taxa in Da_NC_01_A
8378b20f3e0f79429c184ae2c8cea623	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae		taxa in Da_NC_01_A
a1385696bc8514f1d36c43af0f673d9a	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Halomicrobium	taxa in Da_NC_01_A
e33328af4b87b0209a0ed5a7c93e909b	Bacteria	Actinobacteria	Actinobacteria	Frankiales	uncultured	uncultured bacterium	taxa in Da_NC_01_A
4dcda42bc875db53b5b892340fd776b4	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Kocuria	taxa in Da_NC_01_A
cb20d5169972826afb72496f29291ece	Archaea	Euryarchaeota	Halobacteria	Halobacteriales			taxa in Da_NC_01_A
c7fa3d57dd6cfc8cb7a520ba9ea7d42f	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halococcaceae		taxa in Da_NC_01_A
6b830ab75ebef17fce7cb0a10503171c	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae		taxa in Da_NC_01_A
7814dff06304f1ebbcb0bdbc71718df9	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
08ea6fb8ed6af70196ed3f8efe398156	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae		taxa in Da_NC_01_A
ef35d9a0aac8ad0c45969f72a45c715d	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
a483771c7d6445008d635a3eb9b7873c	Bacteria	Actinobacteria	Nitriliruptoria	Nitriliruptorales	Nitriliruptoraceae	uncultured bacterium	taxa in Da_NC_01_A
de8b32697af29cb80f239ba8aa65192e	Bacteria	Actinobacteria	Thermoleophilia	Solirubrobacterales	Solirubrobacteraceae	Solirubrobacter	taxa in Da_NC_01_A
25c5ae4389e988ace044243e8117fabf	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
154dc9125af543363de9f11dfb3cddba	Bacteria	Actinobacteria	Actinobacteria	Frankiales	uncultured	uncultured bacterium	taxa in Da_NC_01_A
843f5d46dfcb245d0143c40523ceefd0	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halorussus	taxa in Da_NC_01_A
3d39c49eb7fcd60ddb6c296046d31754	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
ae6a3e2525847a5165d5d78f42980f3a	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae		taxa in Da_NC_01_A
46da7b8561867da54e5dd8e6f267c215	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
98eaaf112df3fac36a8142d1ce1a59eb	Bacteria	Actinobacteria	Actinobacteria	Frankiales	uncultured	uncultured bacterium	taxa in Da_NC_01_A
b2c8afceb2f2ff92f759fb517595e639	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Haloterrigena	taxa in Da_NC_01_A
6a1cd575b2e883389e633490125c68ec	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halolamina	taxa in Da_NC_01_A
c9fda3fc58f8cce931681a3e061dbc26	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
b64b463c76b74a24d9b26b7a1e0a0910	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
9cb5afd14415ae40a71b355a1a20c9db	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Cyclobacteriaceae	uncultured	taxa in Da_NC_01_A
adf0920c34952c4350f0504ddf26977e	Bacteria	Proteobacteria	Alphaproteobacteria	uncultured			taxa in Da_NC_01_A
98ede70ebd84dfcd2b67d1cebff2bd7e	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
73d5349e152f7491574b66feeabc527d	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halococcaceae		taxa in Da_NC_01_A
0d695028e1adeffd0b94554abf776bc0	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Gillisia	taxa in Da_NC_01_A
7f82ed2b8224b7fd0616b44cf9e3920e	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in Da_NC_01_A
62ffa7040324da2c4bce082da59068cb	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales			taxa in Da_NC_01_A
006aab8d4ceffcf7752b880f5c53e025	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
23386c5fc9410b34c01e782c16df647c	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in Da_NC_01_A
88f66ade1e5cc4f2a543a8edda9265e6	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Gillisia	taxa in Da_NC_01_A
b514c7885867987a1c59ae3d73e5099f	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Kocuria	taxa in Da_NC_01_A
e3da4135aad0144f5ab28bd4979f450a	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halolamina	taxa in Da_NC_01_A
3f41dfbbde17604942b50bf7f5f9d3c4	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
fd3725d5cc869a6afc0c485fdd985a38	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halobacteriaceae		taxa in Da_NC_01_A
7e17eeeabc78714948aaa02e00aa0168	Bacteria	Proteobacteria	Alphaproteobacteria	uncultured			taxa in Da_NC_01_A
d09e557fd1035de9bf3b03cc1386b26d	Bacteria	Actinobacteria	Nitriliruptoria	Nitriliruptorales	Nitriliruptoraceae		taxa in Da_NC_01_A
083b5a8c4b1377ab348e68dff6e96cd6	Bacteria	Actinobacteria	Acidimicrobiia	Actinomarinales	uncultured		taxa in Da_NC_01_A
6ee464a2f41aca7a778bd439c2be181a	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
7164c63485a36523d4d1f0e58b44cd1a	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A

d4bad5b456c27b6ff44366addb35a5bd	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Arthrobacter	taxa in Da_NC_01_A
00ae16e13ad4b49bdf0c41f2d94ece4c	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
11991f096ac4a796941fd464edec6076	Bacteria	Bacteroidetes	Bacteroidia	Chitinophagales	Chitinophagaceae	Ferruginibacter	taxa in Da_NC_01_A
6e16b6639f29ac86a717fad78c766d78	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
99bc219364189dc129c855920c6a7684	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Roseomonas	taxa in Da_NC_01_A
aaf19d5bdb7491e3cac730f5196f8e5c	Bacteria	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	Roseomonas	taxa in Da_NC_01_A
2ced976965611f689ed096c3d8b321fe	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
366f395ead964db8a392bba429a0b583	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in Da_NC_01_A
68e33e4e685f99fafc191a73317278ba	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
9219ce63d5bd8da7586671684341f78c	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
9eb7402965de0333604a148debbcdae2	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales			taxa in Da_NC_01_A
05ed9ae9de783bc13d3e40025f64ebf6	Bacteria	Actinobacteria	Thermoleophilia	Solirubrobacterales	Solirubrobacteraceae	Solirubrobacter	taxa in Da_NC_01_A
2f1a092bc13d385dadcf1b322ede9b04	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Cyclobacteriaceae	uncultured	taxa in Da_NC_01_A
22f00e11d3887138a48fae792f2ceaee	Archaea	Euryarchaeota	Halobacteria	Halobacteriales			taxa in Da_NC_01_A
67c7684f98fb9dd406ec6446afe1ef8c	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in Da_NC_01_A
9f3c8402964bb3235b33937c7b27bb0b	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	Pontibacter	taxa in Da_NC_01_A
17dc6b2194b628b67552e8572daa28ce	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae		taxa in Da_NC_01_A
9988050d986418da018a0b9a400d3cbc	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	AKYG1722	uncultured bacterium	taxa in Da_NC_01_A
82ea6dbc3b625e01014521bef01bcc81	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bacteriovoracaceae	Peredibacter	taxa in Da_NC_01_A
63ca0d5a43b75ab429c7898ea458877e	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in Da_NC_01_A
d22cdf7e2ee915acb02c103f1dcab9b4	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
370dbaa43d94218a8e97aaf3dd2368bc	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Noviherbaspirillum	taxa in Da_NC_01_A
0b61018b1f7af7955aa3d87a622dc394	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae		taxa in Da_NC_01_A
42cadc47594006096cd334ef6ccdf73c	Bacteria	WS4					taxa in Da_NC_01_A
6408cfae3f02c1b13d5da530d7a3b7fe	Bacteria	Actinobacteria					taxa in Da_NC_01_A
9b0fc0c071765008eef29933476e1461	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Ellin6055	taxa in Da_NC_01_A
bc24d37243dcea8b9f6ace168afe3023	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	AKYG1722	uncultured bacterium	taxa in Da_NC_01_A
cc1ce9bf08e20503cbb019c5582f58db	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
02a3ac6fad81f9a5592bba9d0a8712e4	Bacteria	Gemmatimonadetes					taxa in Da_NC_01_A
76fea45f5ceba571bf72e5412f451e08	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Cellulomonadaceae	Actinotalea	taxa in Da_NC_01_A
96686aae99393e1dbe1d5f90bd03a283	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
1f5d335f15be31accfb5e36015b8c5a0	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
e53f2285f1002a6a8264019e6a343050	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
0359cc5866e9c87b57ef54e52d9d4ae9	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rubellimicrobium	taxa in Da_NC_01_A
9b3d2d77dc0fd7f9b3cf5aab0ec153be	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Gillisia	taxa in Da_NC_01_A
bab132eda429564994f593009e27deb8	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Gillisia	taxa in Da_NC_01_A
b84ce48d735bc5036d49edc7474856b2	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Ellin6055	taxa in Da_NC_01_A
dc8cf969d813326812bfb0966568b811	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Rhizobacter	taxa in Da_NC_01_A
fcf664c29fd9ea040a8b60b6cb70a7ff	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Haloterrigena	taxa in Da_NC_01_A
0a5f06e4e6ecdee23757076d2fbe4dd1	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Devosiaceae	Devosia	taxa in Da_NC_01_A
15a49d28ff81af630d8caea7b015b522	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
52c886f5fcd75fcd598d461fdde12a0f	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
f7929c65a9bba2619b696f40099fb368	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
e4ebfd0579e836394b5e1dd848f8d2bf	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
a3cb960f4d61ef1c0bee98e1e4b404aa	Bacteria	Bacteroidetes	Rhodothermia	Balneolales	Balneolaceae	uncultured	taxa in Da_NC_01_A
ea35646c10019afaac7e60e9ec9d1d32	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	Pontibacter	taxa in Da_NC_01_A
920095bc5cf64c9e9f22835ca0f27ab2	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
af22035a276c3c656c35c8f93c461e50	Bacteria	Actinobacteria	0319-7L14	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
bb4286eb1ed177b74b8b2ac5c3ad7164	Bacteria	Actinobacteria	Thermoleophilia	Gaiellales	uncultured	uncultured bacterium	taxa in Da_NC_01_A
e55523ca352a5500c92e5064502b0d9a	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halolamina	taxa in Da_NC_01_A

4b4084bca6fc16ac0c7b8468861b9c97	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in Da_NC_01_A
535b5f967401c1e3a402064089b4c675	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
b7ffad2c8de917e4920349f3c1951f93	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
ebd79a9fce02412fbbae46e3dddfd44c	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	Pontibacter	taxa in Da_NC_01_A
3ad1b6439c99340b3f71e1581b50fc0c	Bacteria	Actinobacteria	Nitriliruptoria	Nitriliruptorales	Nitriliruptoraceae	uncultured bacterium	taxa in Da_NC_01_A
7b0c7238036948b6aac4618c668c3df4	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
8c07c369f8c4ff49d58825e2ec7687bf	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
c8f8fe5e700007b0c3a60e38613cc9f8	Bacteria	Actinobacteria	Thermoleophilia	Gaiellales	uncultured	uncultured bacterium	taxa in Da_NC_01_A
f5ff4600d1fbdf862c9c7b0ca2834312	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
248f435c99df436486a1a8e33b51459b	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
6d11cca6f784b470ab7f641015755853	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	taxa in Da_NC_01_A
2d299df28a01e6b52bd43767d9cf4b1c	Bacteria	Actinobacteria	Actinobacteria	Frankiales	uncultured	uncultured bacterium	taxa in Da_NC_01_A
59a5dee8011ec1e2d0451d9f748e7cd2	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Haloterrigena	taxa in Da_NC_01_A
77398a3d39b046cd6cb192cf6a8fce18	Bacteria	Firmicutes	Bacilli	Bacillales	Sporolactobacillaceae	Salipaludibacillus	taxa in Da_NC_01_A
02163387f14b805468f98d01d4ff674b	Bacteria	Gemmatimonadetes	AKAU4049				taxa in Da_NC_01_A
32460c7a485e2fa9aa1ceae52c93bb86	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Devosiaceae	Devosia	taxa in Da_NC_01_A
9642287fd37004d45fb384baa1ac9d8d	Bacteria	WS4					taxa in Da_NC_01_A
5a78aa9db081295061e4784ac0d619b3	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
243e3d177aecf970a0276ce514bac781	Bacteria	Actinobacteria	0319-7L14	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
1e17e315455acfaf007fec8e49b64755	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Haloterrigena	taxa in Da_NC_01_A
f28f9fc0bd55c917a63ef7a95e1e934e	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
1c326ee2f89f1b9a68fbd58b2a233c1b	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	uncultured	taxa in Da_NC_01_A
23aeb5bd2bbb3dde7ba0d8004acbea64	Bacteria	Actinobacteria	Thermoleophilia	Gaiellales	uncultured	uncultured bacterium	taxa in Da_NC_01_A
ebc83327e4c64ebc0355c2185b3d5e97	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
2ad6d8e06573257dba48cc9bd8c515ab	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
556940bfd264003060d95b85678a16a8	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
a81c7996272c68078c7d20efe55b2bcc	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Halomicroarcula	taxa in Da_NC_01_A
f88d126e446042f0a1ac74722ef473b0	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
a061cdf34863b56f02377d7149b54fc4	Bacteria	Actinobacteria	Acidimicrobiia	Microtrichales	uncultured		taxa in Da_NC_01_A
d65481be50af410aa7993e067e9882da	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	taxa in Da_NC_01_A
363bbdfccb838670b882c0a4eaceb658	Bacteria	Firmicutes	Bacilli	Bacillales	Planococcaceae	Planococcus	taxa in Da_NC_01_A
75aba4fc6df09200f82da285cf784e55	Archaea	Euryarchaeota	Halobacteria	Halobacteriales			taxa in Da_NC_01_A
9080724a33caf5fa71349d172c97b5a0	Bacteria	Actinobacteria	Acidimicrobiia	Microtrichales	uncultured		taxa in Da_NC_01_A
98d315e5db8cb5a69fa0c2bcca2a4b71	Bacteria	Proteobacteria	Gammaproteobacteria				taxa in Da_NC_01_A
be7f626bb15cb85283a8ae10e4a80f28	Bacteria	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Lysobacter	taxa in Da_NC_01_A
fb36f93bfef5b27f0a9a92a43f5edba6	Bacteria	Actinobacteria	Acidimicrobiia	Microtrichales	uncultured		taxa in Da_NC_01_A
18c67bb3b41fa1088ad5f5d0c3bb7a3d	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bacteriovoracaceae	Peredibacter	taxa in Da_NC_01_A
aa05a4c3795ce4d9dcc1fcfb4332fe46	Bacteria	Actinobacteria	Nitriliruptoria	Euzebyales	Euzebyaceae	uncultured	taxa in Da_NC_01_A
ea4d9aad4de574746622409323fb3b6f	Bacteria	Actinobacteria	Nitriliruptoria	Euzebyales	Euzebyaceae	uncultured	taxa in Da_NC_01_A
3b63707abe67028a97baddeb4714945b	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
515505e93d1dab608b0a72fbd9f91add	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
64d39a04ad97ae264ef63b483d4a3789	Bacteria	Actinobacteria	Nitriliruptoria	Nitriliruptorales	Nitriliruptoraceae		taxa in Da_NC_01_A
97142f6ce9ab4162fd4b9b7765974be4	Bacteria	Bacteroidetes	Rhodothermia	Rhodothermales	Rhodothermaceae	uncultured	taxa in Da_NC_01_A
1d8aa7bfa044bebc592463df936d4157	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Halomicroarcula	taxa in Da_NC_01_A
7bb367bd5d149842e4a37a3e5ade553b	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
e3058a43745ca337793c1328b0e01abe	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	uncultured	taxa in Da_NC_01_A
f2103d699191883b7fe408ce859f7a06	Bacteria	Bacteroidetes	Rhodothermia	Rhodothermales	Rhodothermaceae	uncultured	taxa in Da_NC_01_A
521a1c37381a91e35cd707672cb33f56	Bacteria	Actinobacteria	Acidimicrobiia	uncultured			taxa in Da_NC_01_A
857274b006a81ea77e1eb1c54c6963d8	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Arthrobacter	taxa in Da_NC_01_A
96047c9ccd9197924015d05d9100e753	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Polaromonas	taxa in Da_NC_01_A

dba8ebcb0db51d7ab41539d1554242fd	Bacteria	Chloroflexi	Anaerolineae	Caldilineales	Caldilineaceae	uncultured	taxa in Da_NC_01_A
f1f88b2feae0323423cb9a0767da1ddd	Bacteria	Planctomycetes	Planctomycetacia	Pirellulales	Pirellulaceae	Rhodopirellula	taxa in Da_NC_01_A
0228052edff08fc2451f20da216406c6	Bacteria	Proteobacteria	Alphaproteobacteria	Tistrellales	Geminicoccaceae	uncultured	taxa in Da_NC_01_A
216b6ffbe7f250fa952560cdce2db1af	Bacteria	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	taxa in Da_NC_01_A
83366f07abcd00485c18d8b7dccffe33	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae		taxa in Da_NC_01_A
c398d12775056f96348f6421c39642f6	Archaea	Euryarchaeota	Halobacteria	Halobacteriales			taxa in Da_NC_01_A
e663b1b9075ea02b2e83ee4df2d0665c	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
288223a65e815c3fd95c9351fd92a0a2	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
7c58435b550065cc395b9bee83ba7af1	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	JG30-KF-CM45	uncultured bacterium	taxa in Da_NC_01_A
844610bb1d1bf3af298dfb4e31b19578	Bacteria	Bacteroidetes	Rhodothermia	Balneolales	Balneolaceae	uncultured	taxa in Da_NC_01_A
c9d847a59174589893c319ba737ce680	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	JG30-KF-CM45	uncultured bacterium	taxa in Da_NC_01_A
e4977638da73386b1d84ba26db2cfbe4	Archaea	Euryarchaeota	Thermoplasmata	Marine Group II	uncultured archaeon	uncultured archaeon	taxa in Da_NC_01_A
ab9f479258a28f0fecce8447927cea52	Bacteria	Deinococcus-Thermus	Deinococci	Deinococcales	Deinococcaceae	Deinococcus	taxa in Da_NC_01_A
717633d8c09ef020bee1af3fb27fe113	Bacteria	Deinococcus-Thermus	Deinococci	Deinococcales	Trueperaceae	Truepera	taxa in Da_NC_01_A
7f927a002c6a1f909ca4ac6c0d645358	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
c53e49079c1517053f896a277aea61f9	Archaea	Euryarchaeota	Thermoplasmata	Marine Group II	uncultured archaeon	uncultured archaeon	taxa in Da_NC_01_A
44bacf442b3ed8ab484c7b410e0d3f59	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in Da_NC_01_A
53c3cf050b39ba9d8b1fdba604a3fc67	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae		taxa in Da_NC_01_A
b65d389f2fbb6f2879ba90f7183d28a3	Bacteria	Proteobacteria	Alphaproteobacteria	uncultured	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
8c4f3dcc46199379e64dedf04b535f18	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Anaerobacillus	taxa in Da_NC_01_A
6fb3509007e3a94c1b9a02d99126c4dd	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
c63b35b3c478f47cf311dee1774161f4	Bacteria	Gemmatimonadetes	AKAU4049				taxa in Da_NC_01_A
13ed2de0e2288554a592170a692616c4	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
68a3a0dfe349418f82942d06f08bcf8b	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halococcaceae		taxa in Da_NC_01_A
7a76e3df8dea1e25b3ee84ad600203e1	Bacteria	Chloroflexi	KD4-96	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
86020209fe29e79f6260c99b49a602d2	Bacteria	Actinobacteria	Acidimicrobiia	uncultured	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
d9a27837f383c8a1b7b26b3be0c2faf2	Bacteria	Actinobacteria	Actinobacteria	Streptomycetales	Streptomycetaceae		taxa in Da_NC_01_A
560ba96705030a1eb0f5faa1000445a5	Bacteria	Deinococcus-Thermus	Deinococci	Deinococcales	Deinococcaceae	Deinococcus	taxa in Da_NC_01_A
64c47eb3cd3c8ba77c9a7d788d4212f7	Bacteria	Actinobacteria	Nitriliruptoria	Euzebyales	Euzebyaceae	uncultured	taxa in Da_NC_01_A
901a93d97bd6255e5c80691d09d9f855	Bacteria	Actinobacteria	Acidimicrobiia	Actinomarinales	uncultured		taxa in Da_NC_01_A
b52513bde40785ad8d3a37097af6e5ee	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Halomicrobium	taxa in Da_NC_01_A
8442cbfcd1cab4f2bc305fdecf19c1bd	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halorussus	taxa in Da_NC_01_A
96de32c6f99fdcfc5d8a341c25f49f02	Bacteria	Bacteroidetes	Rhodothermia	Balneolales	Balneolaceae		taxa in Da_NC_01_A
b11248089ffe4e17daf8c867aeb7cd21	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	uncultured	taxa in Da_NC_01_A
586744a416bf534cccda6d2674fe2036	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	MWH-CFBk5		taxa in Da_NC_01_A
8683312215ba5bcd094472c726b3082c	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Pseudarthrobacter	taxa in Da_NC_01_A
f14abebfc04068f33e8fe54540b1fee3	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Haloterrigena	taxa in Da_NC_01_A
f3aad87f8a58a84f8371d754b3ae3a41	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halobaculum	taxa in Da_NC_01_A
fd879e7844b5a7ae019d7476464941e6	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
1bd9d462c1636b60c133116b4268a64f	Bacteria	Actinobacteria	Acidimicrobiia	Microtrichales	Ilumatobacteraceae	uncultured	taxa in Da_NC_01_A
2df0f5365a6f0d0d1a586984b7525473	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae		taxa in Da_NC_01_A
5026d0025db9708ac491ae1dc16779fd	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Halobacillus	taxa in Da_NC_01_A
5641d314839eeb040e22adc92345aa98	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	MWH-CFBk5		taxa in Da_NC_01_A
8e7f6ec2c840ac5f1f528607280ac20d	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bacteriovoracaceae	Peredibacter	taxa in Da_NC_01_A
c5d9c3b4b494e1fd2024d07032a19056	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
dba2648ad8f15ddb2eb5182b2dac229f	Bacteria	Chloroflexi	Anaerolineae	Ardenticatenales	uncultured		taxa in Da_NC_01_A
e94425f8b6c428fb488cb26e3e8ed541	Bacteria	Acidobacteria	Subgroup 6				taxa in Da_NC_01_A

3c5e539b0b82c2bde2511de0c3302d08	Archaea	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	uncultured archaeon	taxa in Da_NC_01_A
4c37c94bf58ddb3d6e7c056332577141	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	JGI 0001001-H03	taxa in Da_NC_01_A
566094ef8f1b644b8f4452327da66205	Bacteria	Chloroflexi	Anaerolineae	Caldilineales	Caldilineaceae	uncultured	taxa in Da_NC_01_A
78ededea2023d6e985b323de588cfe27	Bacteria	Actinobacteria	MB-A2-108	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
c0709248c2a712f312a44fe9788e4549	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
157784f02f80322e35438ebeb706d17a	Archaea	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	uncultured archaeon	taxa in Da_NC_01_A
f3ee8cb25686bbba0cb435de59049993	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halobacteriaceae		taxa in Da_NC_01_A
0ba9eee6ada3e7caf20a427a3ec9bfbd	Bacteria	Halanaerobiaeota	Halanaerobiia	Halanaerobiales	Halanaerobiaceae	Halanaerobium	taxa in Da_NC_01_A
5a422065c56e9ec178eb87bd87a5dc8e	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae		taxa in Da_NC_01_A
962c7cc7f1987a075cf4bf262fe25967	Bacteria	Actinobacteria	Acidimicrobiia	uncultured	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
b80e8bdeac91c84f9f153efef81d8318	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	taxa in Da_NC_01_A
693d34a68ac271d3741c84cdd3a322e6	Bacteria	Planctomycetes	Planctomycetacia	Pirellulales	Pirellulaceae	Rhodopirellula	taxa in Da_NC_01_A
197a9e834a68efbc5363ed6552cd6c3e	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Candidatus Halobonum	taxa in Da_NC_01_A
ea21afbdce4fd4742c3d119a487e47cc	Bacteria	Bacteroidetes	Rhodothermia	Rhodothermales	Rhodothermaceae	uncultured	taxa in Da_NC_01_A
f19af7cbbb068a252b0b2ed3c5dbf069	Bacteria	Actinobacteria	MB-A2-108	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
04e3b75bfd8031147ca8f0c9b8f72952	Bacteria	Proteobacteria	Gammaproteobacteria				taxa in Da_NC_01_A
10e287699e016364b79252707015686b	Bacteria	Proteobacteria	Alphaproteobacteria	Tistrellales	Geminicoccaceae	uncultured	taxa in Da_NC_01_A
3aab967f1a4622c04a490e069e93b385	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Polaromonas	taxa in Da_NC_01_A
6597f34dd1e9eab6a272a34a9922d7b7	Bacteria	Gemmatimonadetes	PAUC43f marine benthic group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
3b4286a8c0c50d87a1e82c88dfae025e	Bacteria	Chloroflexi	Dehalococcoidia	SAR202 clade			taxa in Da_NC_01_A
b8ab1fdf94936a8f60bb11bd0e74e243	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae		taxa in Da_NC_01_A
fc61190baac72958002ee084afbc4fd6	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	uncultured	taxa in Da_NC_01_A
45df7bc4286edc7fa99f34403e7ba51a	Bacteria	Acidobacteria	Holophagae	Subgroup 7			taxa in Da_NC_01_A
87ff98eaf0ed8bdbd1a29b389ea83993	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	taxa in Da_NC_01_A
5f11b1d39b6547b9e2a20c04ffd2baa8	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae		taxa in Da_NC_01_A
78bc068c292f6801448bb3437db07f1f	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	taxa in Da_NC_01_A
95eebff004a59e448b888b0af556d14b	Archaea	Euryarchaeota	Thermoplasmata	Marine Benthic Group D and DHVEG-1	uncultured archaeon	uncultured archaeon	taxa in Da_NC_01_A
99c66e84f7b1ca8f4e5f12fa310f8676	Bacteria	Proteobacteria	Gammaproteobacteria	Ectothiorhodospirales	Ectothiorhodospiraceae	Halofilum	taxa in Da_NC_01_A
e534b2b094aa1636fe26f5e7763ecff0	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Ellin6055	taxa in Da_NC_01_A
f145ede45c488b8487b2284e4c4f5301	Bacteria	Planctomycetes	Planctomycetacia	Isosphaerales	Isosphaeraceae	uncultured	taxa in Da_NC_01_A
7eadd8532b6004358f3a953708f8810a	Bacteria	Bacteroidetes	Rhodothermia	Rhodothermales	Rhodothermaceae	Salinibacter	taxa in Da_NC_01_A
e22f88d70e80d448c874c4ae6277177d	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae		taxa in Da_NC_01_A
f76c4a1a146038d246682a1de8caa2e0	Bacteria	Gemmatimonadetes	PAUC43f marine benthic group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
05a57eace4cba4a0899d6e28eea296a2	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	Adhaeribacter	taxa in Da_NC_01_A
0756724ed4a9ea5635ff4607bbc8dfd3	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
32547cd2518cc7729c5b4d54689885a5	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halorussus	taxa in Da_NC_01_A
6c9ed2fea48f770eb609f90154317af4	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Burkholderia-Caballeronia-Paraburkholderia	taxa in Da_NC_01_A
7ab64b1c3bf942a59adb600e9bd0eb3d	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halorussus	taxa in Da_NC_01_A
e09d6ae28f6a24e56a4c0068a557381c	Bacteria	Proteobacteria	Alphaproteobacteria	uncultured			taxa in Da_NC_01_A
3dfc7a564be9d06e174b55b9db5ca5e3	Bacteria	Proteobacteria	Deltaproteobacteria	Myxococcales			taxa in Da_NC_01_A
4c5022eead9f515224e991bbab4ec2d0	Bacteria	Actinobacteria	Actinobacteria	Frankiales	uncultured	uncultured bacterium	taxa in Da_NC_01_A
87640ca3b5d8ed35ab491390676c18a5	Bacteria	Patescibacteria	Saccharimonadia	Saccharimonadales	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
91c6ca3fe859c54031ae487325e84c7e	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
a2e1e5d26479f40ff131b6175575891c	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
04d731bcee7b7dd7fe435fb2bae68386	Bacteria	Halanaerobiaeota	Halanaerobiia	Halanaerobiales	Halanaerobiaceae	Halanaerobium	taxa in Da_NC_01_A
0500983d603a858f5c24ad17c0bddf7f	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halorussus	taxa in Da_NC_01_A

235edce0db4b197efdea289f9a817b82	Bacteria	Planctomycetes	Planctomycetacia	Isosphaerales	Isosphaeraceae	uncultured	taxa in Da_NC_01_A
4838057848507eca40a59a49203198b7	Bacteria	Firmicutes	Bacilli	Bacillales	Planococcaceae	Planococcus	taxa in Da_NC_01_A
77d2309e6caeb480a14e65e604cfce5d	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in Da_NC_01_A
9002cc5318d269074785d4af8e0a402b	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae		taxa in Da_NC_01_A
971d7e201ebbac84ce12f8d219363555	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halorussus	taxa in Da_NC_01_A
d922d9e0dc73b4d4d1789fc557c8b1a2	Bacteria	Chloroflexi	Dehalococcoidia	SAR202 clade			taxa in Da_NC_01_A
2ae4ef3f33203cf842c7b32a6747681c	Bacteria	Proteobacteria	Deltaproteobacteria	Bdellovibrionales	Bacteriovoracaceae	Peredibacter	taxa in Da_NC_01_A
2db0061bf5e9e8d9ddad9243847f80ec	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	taxa in Da_NC_01_A
bab9577b626f6eb523adc6d15eab2b8a	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
f3c14fcf1a87e34b0a5c365a7c49b0e9	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
f8782909f297c9d702bf577768e93784	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
f885c722b9a0336a09a763bb804881f2	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
1b3a0ad1345475d40f9b3aa940595230	Bacteria	Proteobacteria	Deltaproteobacteria	Bradymonadales	Bradymonadaceae		taxa in Da_NC_01_A
41f37147b4227e55c7c1d09ba0447134	Bacteria	Deinococcus-Thermus	Deinococci	Deinococcales	Trueperaceae	Truepera	taxa in Da_NC_01_A
49aca6b3f64700039c210ce37b7f5eca	Bacteria	Bacteroidetes	Rhodothermia	Rhodothermales	Rhodothermaceae	Salinibacter	taxa in Da_NC_01_A
66924c9233670db606bfbc880cb9efb5	Bacteria	Actinobacteria	Acidimicrobiia	IMCC26256			taxa in Da_NC_01_A
66e188b598e78234edcb778efa0d61b5	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in Da_NC_01_A
80d69091c756c0b12c761cc80539c198	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
8f6e8e04344ce5fa0d0dc81034604e6a	Bacteria	Proteobacteria	Deltaproteobacteria	Bradymonadales	Bradymonadaceae		taxa in Da_NC_01_A
91342de6be41fef798b7a8d5af23d17c	Bacteria	Patescibacteria	Saccharimonadia	Saccharimonadales	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
c46ec12efc42fc7c8867b1621a6f4be8	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
e0ce698d254d46bce5486ce25b80a718	Bacteria	Bacteroidetes	Rhodothermia	Rhodothermales	Rhodothermaceae	Salinibacter	taxa in Da_NC_01_A
e2d1f6770caa20dc2de76fc80fdcb8fb	Bacteria	Proteobacteria	Gammaproteobacteria				taxa in Da_NC_01_A
e8890c8ea3b44f6dfb520f0e9db69cfe	Archaea	Euryarchaeota	Thermoplasmata				taxa in Da_NC_01_A
f13771083c2280f8b5e63339b3906925	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
f32637cd2f84c8f275c285d3813a630e	Bacteria	Acidobacteria	Subgroup 6				taxa in Da_NC_01_A
32334677cffbcfb9705a8bffa32f3a61	Bacteria	Actinobacteria	Nitriliruptoria	Nitriliruptorales	Nitriliruptoraceae		taxa in Da_NC_01_A
5fb1e7e923a9fd15470a1b1e945ba959	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
8c31bc5074cfbfffd9a8243f9246e239	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
28cf018256edaf4f62d83eae6215f414	Bacteria	Proteobacteria	Deltaproteobacteria	Bradymonadales	Bradymonadaceae	uncultured bacterium	taxa in Da_NC_01_A
5b542396ac37f7bba2affc654d88980a	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halorussus	taxa in Da_NC_01_A
8aad50cfbc4bf8440f82f77c925fe18a	Bacteria	Gemmatimonadetes	S0134 terrestrial group				taxa in Da_NC_01_A
ce2caefae50ccbea301c7a08f94abc57	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
5f28f74f18f3d78bba2e88f04a898b9e	Bacteria	Proteobacteria	Alphaproteobacteria	uncultured	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
adbb4ff61a2a098f6f7412abd32bdb1f	Bacteria	Actinobacteria	Acidimicrobiia	uncultured			taxa in Da_NC_01_A
e1debd644c8673dda4b46be939540649	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
2ee27f20e8a02bb565993469d959b146	Bacteria	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Rubritaleaceae	Luteolibacter	taxa in Da_NC_01_A
463db9d0dd3f6f224060bd0f1c87dcbd	Bacteria	Actinobacteria	Acidimicrobiia	Microtrichales	Microtrichaceae	uncultured	taxa in Da_NC_01_A
65614a5ebf7c78d6c84b7bd9b1adf913	Bacteria	Planctomycetes	Planctomycetacia	Isosphaerales	Isosphaeraceae	uncultured	taxa in Da_NC_01_A
8819dd8ee8a1fc8ea880b8da8fccfa73	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Beijerinckiaceae	Methylorosula	taxa in Da_NC_01_A
cf5f3a7a402573f8539880a667a6ae80	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Intrasporangiaceae	Oryzihumus	taxa in Da_NC_01_A
ed957c8014c5ff0ac84c5092e6d7bdc8	Bacteria	Patescibacteria	Saccharimonadia	Saccharimonadales	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
5e1454c2bc922ccb703351ccdebcb615	Bacteria	Planctomycetes	Planctomycetacia	Planctomycetales			taxa in Da_NC_01_A
e01eb7b7e6efddd68e65b377f7644edf	Bacteria	Chloroflexi	Dehalococcoidia	SAR202 clade	uncultured Chloroflexi bacterium	uncultured Chloroflexi bacterium	taxa in Da_NC_01_A
9e755c67d88119d70efefd98b038c56e	Bacteria	Planctomycetes	Phycisphaerae	Phycisphaerales	Phycisphaeraceae	uncultured	taxa in Da_NC_01_A
a2165decc765848262bd3850916b60dc	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
5e578bd71ad372d19fb422cc28b8a249	Bacteria	Patescibacteria	Saccharimonadia	Saccharimonadales	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
7942c68d388826ba33e51d26d2f99e29	Archaea	Euryarchaeota	Thermoplasmata	Marine Group II	uncultured archaeon	uncultured archaeon	taxa in Da_NC_01_A

c56a377df56ca5bcbadd721da5fe040e	Bacteria	Chloroflexi	Chloroflexia	Chloroflexales	Roseiflexaceae	uncultured	taxa in Da_NC_01_A
df596ca813fcf3dc0708f229ec06bbea	Bacteria	Chloroflexi	Anaerolineae	Ardenticatenales	uncultured		taxa in Da_NC_01_A
e16160f5958bdda3e5da5856f424d269	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae 3		taxa in Da_NC_01_A
e3ae1ec998f385550dcdd1eb48f99302	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	AKYG1722	uncultured bacterium	taxa in Da_NC_01_A
2760b737be5b7d47ba374a386b5e2792	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
34cc6fefaefaae389d704c264a7ae301	Bacteria	Actinobacteria	Actinobacteria				taxa in Da_NC_01_A
50f66f9cecf3202f5a7c344e26aa4ce2	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
6500e45a0a974c41b7decf28d1e1f0cf	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
793fcb05e0f02d7843854cde22ca05e1	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
fc74192f73013653cf96c3354080b88a	Bacteria	Actinobacteria	Actinobacteria	Frankiales	Nakamurellaceae	Nakamurella	taxa in Da_NC_01_A
1e3a2a45cd9af71faac9227d71ffd654	Bacteria	Verrucomicrobia	Verrucomicrobiae	Opitutales	Opitutaceae		taxa in Da_NC_01_A
44dd2e23977c27c26fc290da3a1cbf93	Bacteria	Actinobacteria	Actinobacteria	Frankiales	Geodermatophilaceae	Blastococcus	taxa in Da_NC_01_A
45cd8a03769fff4f88c762cd87f564c8	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
566274510ef9fad301d84f9c4e3a7603	Bacteria	Chloroflexi	TK10				taxa in Da_NC_01_A
5714f5524b8d6b10a9b52c0c9f7ec533	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
6ccaf43082a014c111f1314bdf66ac5f	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A
bcdddaccdbe19926cd9883e3acedf59b	Bacteria	Halanaerobiaeota	Halanaerobiia	Halanaerobiales	Halobacteroidaceae	Orenia	taxa in Da_NC_01_A
d1b893eaef0b0750ba2dac385d2cb32f	Bacteria	Actinobacteria	Actinobacteria	Frankiales	Geodermatophilaceae	Blastococcus	taxa in Da_NC_01_A
dd7d8e19f6d4c1cae90efcfeb9c48b27	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in Da_NC_01_A
27eceecef99f07e382f388e610d314fe	Bacteria	Planctomycetes	OM190				taxa in Da_NC_01_A
a892df6c607195dfbe13475813594adf	Bacteria	Bacteroidetes	Rhodothermia	Rhodothermales	Rhodothermaceae	Salinibacter	taxa in Da_NC_01_A
ed9b2c54a1be688d4da8601cfa0b4a0f	Bacteria	Verrucomicrobia	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter	taxa in Da_NC_01_A
58df5170cb1d28df4c90020888eab034	Bacteria	Proteobacteria	Alphaproteobacteria	Tistrellales	Geminicoccaceae	uncultured	taxa in Da_NC_01_A
897e9f37daa8c802f0b0789d2c6328af	Bacteria	Actinobacteria	Acidimicrobiia	Microtrichales	Ilumatobacteraceae	uncultured	taxa in Da_NC_01_A
d7a64cd2022ab639e1718d256cf3bcfa	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halococcaceae		taxa in Da_NC_01_A
e31ff0b4280eb7d6fbfd792c6e9856a4	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halococcaceae		taxa in Da_NC_01_A
336df7abc22a69ecee34028e32791eb4	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	taxa in Da_NC_01_A
3e93d5da49af0d946ca12a1e162ef8d9	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Xanthobacteraceae	uncultured	taxa in Da_NC_01_A
9058165ff7b26dc418014bdbfa611f67	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in Da_NC_01_A
9b1b1391dc6ad6739687ddf36e332f16	Archaea	Euryarchaeota	Halobacteria	Halobacteriales			taxa in Da_NC_01_A
cd68f32d12bbe34c5f34bd75c67af553	Bacteria	Actinobacteria	MB-A2-108				taxa in Da_NC_01_A
1c6b013811a45fdbe79c1220f428998b	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halobacteriaceae	Halobacterium	taxa in Da_NC_01_A
a775c305291fa6e79fb73cec26f6762d	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in Da_NC_01_A
b0b7612f19b1a35bf103260f679d84bf	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Nesterenkonia	taxa in Da_NC_01_A
dc1e46cbe4da5c7ec09f7f82057e90f4	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae		taxa in Da_NC_01_A
5e96389cbac91876a7ebda23f681f4aa	Bacteria	Halanaerobiaeota	Halanaerobiia	Halanaerobiales			taxa in Da_NC_01_A
d227693187a4e21316a582985d3f3c8a	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Halorientalis	taxa in Da_NC_01_A
db089109071fade2a0dd51d310a14322	Bacteria	Planctomycetes	Planctomycetacia	Gemmatales	Gemmataceae	Gemmata	taxa in Da_NC_01_A
0ed2cbcc5ce9a4489a28b724ab6272f2	Bacteria	Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	taxa in Da_NC_01_A
31a9e38cd8dd753023f2e0c4348fd4c4	Bacteria	Actinobacteria	Nitriliruptoria	Nitriliruptorales	Nitriliruptoraceae		taxa in Da_NC_01_A
842965fa44adc9e32a99eb434f1e24d3	Bacteria	Bacteroidetes	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Gillisia	taxa in Da_NC_01_A
b10b17f2b660563f2f71260bebc85900	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	Thermomicrobiaceae		taxa in Da_NC_01_A
ee03047677c3126d150559ce9e8b7cd4	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales Incertae Sedis	uncultured	taxa in Da_NC_01_A
32a48471cb21008b2829e58c2561117c	Bacteria	Chloroflexi	Dehalococcoidia	S085	uncultured bacterium	uncultured bacterium	taxa in Da_NC_01_A
422a584cd2ed91ff1a0eaaf6445dcdac	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
570984ffda613ceb3a8bbf50a3669cff	Archaea	Euryarchaeota	Halobacteria	Halobacteriales			taxa in Da_NC_01_A
897f1433c9095f2214da4334892ea58c	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodovibrionales	Fodinicurvataceae	Rhodovibrio	taxa in Da_NC_01_A
8e24a453e81ca323de076f7f564d8685	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Halobacillus	taxa in Da_NC_01_A
98acabe56b799e85f34cd4d99fd95ddb	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Natronomonas	taxa in Da_NC_01_A
24c25a7e7f9515dd1bdfa2858f1f62c4	Bacteria	Gemmatimonadetes	Longimicrobia	Longimicrobiales	Longimicrobiaceae	uncultured bacterium	taxa in Da_NC_01_A

2df3b2040d88547d28becd5a84e8a91a	Archaea	Euryarchaeota	Halobacteria	Halobacteriales			taxa in Da_NC_01_A
49706c6b2cea07a12d5f6ed669f2f62c	Bacteria	Actinobacteria	Thermoleophilia	Solirubrobacterales			taxa in Da_NC_01_A
712c5dc9626af87416e399cd69900a53	Bacteria	Actinobacteria	Acidimicrobiia	Microtrichales	Ilumatobacteraceae	uncultured	taxa in Da_NC_01_A
ab3afd5a5df206982e57a3963f9c778d	Bacteria	Actinobacteria	Acidimicrobiia	Microtrichales	Ilumatobacteraceae	uncultured	taxa in Da_NC_01_A
c1d7b8df3b6ada04fdce2bd570b0ed62	Bacteria	Firmicutes	Bacilli	Bacillales	Planococcaceae	Jeotgalibacillus	taxa in Da_NC_01_A
d3775c9118720599f0285b54a3460f30	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae		taxa in Da_NC_01_A
0be161f94273c85cf8949ef3adc85017	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	uncultured	taxa in Da_NC_02_B
22bc23d012ffb7384441eba7e7a3cff7	Archaea	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae		taxa in Da_NC_02_B
be6928ce6fd5ae46f0c8ae368a12b1e2	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in Da_NC_02_B
c140d8ffeb3b916f4d6ce9aadf74e407	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in Da_NC_02_B
6f62b739e038e47eb2bd1063523f57f7	Bacteria	Proteobacteria	Deltaproteobacteria	Bradymonadales	Bradymonadaceae		taxa in Da_NC_02_B
6602c611485023ec6a11368a4556f2b0	Bacteria	Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	taxa in Da_NC_02_B
81bc2cb2301086153dbeb705b71dc630	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	uncultured	taxa in Da_NC_02_B
2ec6eb3de57fc12eaef4cbfc6c3314a7	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Halobacillus	taxa in PCR_NC_B
b92c12801a462abb993cf52f6c3d9406	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
e97198f6f59eab9c5bf70f3c64be5712	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
813229d7f5d1d1fea8d9b95d42bda0e7	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
97af0565c040d8c28eb495a49fa73dad	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in PCR_NC_B
2c2bea982ae5be8246700e96458987bb	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Haloterrigena	taxa in PCR_NC_B
8cd8b6bfed8d925d71a7510a58ddb32c	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Halobacillus	taxa in PCR_NC_B
bef86d3a715f57669b7acf09cc874847	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae		taxa in PCR_NC_B
272b57ad42945ade456620056fc3ce80	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
c338cee355e9b1189c86476c426f8eb9	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
cdebed842c7ea83145a4f2258d959a50	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Pontibacillus	taxa in PCR_NC_B
daad403a3c8593cb7e0fc9ae64507503	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	Pontibacillus	taxa in PCR_NC_B
5164626fccbddfab353196d453bb5e3e	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Haloarcula	taxa in PCR_NC_B
7a8f286934bf73d26f00d06fad86acc8	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
5d1e8c6d35bb855ed4f9c4e9b6c832ed	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Haloterrigena	taxa in PCR_NC_B
5a77d84c7594224338252b20d0e5006e	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
403abb56bd1f13c8b26a7a7f7b52770d	Archaea	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	uncultured archaeon	taxa in PCR_NC_B
94bafe6e7b0fd5ece4aa353081b6d798	Bacteria	Proteobacteria	Gammaproteobacteria	Alteromonadales	Marinobacteraceae	Marinobacter	taxa in PCR_NC_B
cd32bf961dd25e8f25abd7037f744d9d	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
2a5719b4cb3d67f798b055bb0bfebd18	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Nesterenkonia	taxa in PCR_NC_B
98d9ee0239c0c750dfd4d6879a6a8b53	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
853219b52702aa68d9c6f5c8160380ea	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
032e7ad856fefbc54db50dab1b785863	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Haloarcula	taxa in PCR_NC_B
58167eb7b926e3a98c25ebb2ff763bf1	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
21a3a87a0bf673cfeb467e9343c01363	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Haloarcula	taxa in PCR_NC_B
99d6c97560bb5fac4debaa412aea0621	Bacteria	Firmicutes	Bacilli	Bacillales	Planococcaceae	Planococcus	taxa in PCR_NC_B
8124867736ddf5b9c60e14e300b0505f	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	Haloarcula	taxa in PCR_NC_B
25007237c14777654a01b65df4f24696	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	JGI 0001001-H03	taxa in PCR_NC_B
02cb0b0f8685661d3f3788663ac43597	Archaea	Thaumarchaeota	Nitrososphaeria	Nitrososphaerales	Nitrososphaeraceae	uncultured archaeon	taxa in PCR_NC_B
4f238b241370a477cd4a5f2cf5324059	Bacteria	Acidobacteria	Subgroup 6				taxa in PCR_NC_B
055bf88af1e8cc581183ccef5e690640	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Nesterenkonia	taxa in PCR_NC_B
029d7672a259e3d1c6f182afb9d4130f	Bacteria	Bacteroidetes	Rhodothermia	Balneolales	Balneolaceae	uncultured	taxa in PCR_NC_B
12a727cf3709b3d6724099bf6999489e	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Rubellimicrobium	taxa in PCR_NC_B
db01b0746f06c12a4bb88e819fc4d638	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Haloferacaceae	Halolamina	taxa in PCR_NC_B
12df2e0dd541423665509febfdcf4567	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	uncultured	taxa in PCR_NC_B
5fe39e8826654b0887c4a2e18f3bbc69	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
16a9139c090a446e7ece560eff54335c	Bacteria	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	taxa in PCR_NC_B

e9484318818cee93323164de8a2fe4a5	Bacteria	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Halomonadaceae	Halomonas	taxa in PCR_NC_B
ec3ff0b35da56b8930611cb70223dc66	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in PCR_NC_B
d5d62cf8a16e1144e82e3614e6c2395b	Bacteria	Acidobacteria	Blastocatellia (Subgroup 4)	Blastocatellales	Blastocatellaceae	JGI 0001001-H03	taxa in PCR_NC_B
3fc0ee6ef16d4977e3811396cb14dbea	Bacteria	Chloroflexi	KD4-96	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC_B
46c4acd1b5b9c278af45975425ca19c1	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC_B
9d41cbca8b37373a79f2cb16d7b60458	Bacteria	Chloroflexi	KD4-96	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC_B
192f2a3df9dc8a830720b0892bffeca2	Bacteria	Bacteroidetes	Bacteroidia	Cytophagales	Hymenobacteraceae	Pontibacter	taxa in PCR_NC_B
569b7dc34beb98ab4d511f126dfcaee5	Bacteria	Bacteroidetes	Rhodothermia	Balneolales	Balneolaceae	uncultured	taxa in PCR_NC_B
59c267e2d6fd1ca6477bed8aedcf016c	Bacteria	Verrucomicrobia	Verrucomicrobiae	Verrucomicrobiales	Rubritaleaceae	Luteolibacter	taxa in PCR_NC_B
eb7e701b11ff0ad37c95d0a32d70cad3	Bacteria	Chloroflexi	Anaerolineae	Caldilineales	Caldilineaceae	uncultured	taxa in PCR_NC_B
80c9b35bcf4ac13dc1101386a779067f	Archaea	Euryarchaeota	Halobacteria	Halobacteriales	Halomicrobiaceae	uncultured	taxa in PCR_NC_B
13ea357817dd4276afdc01a860d20945	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	Kocuria	taxa in PCR_NC_B
6168edba7206a2b4220a7d6d8648330a	Bacteria	Proteobacteria	Deltaproteobacteria	Desulfuromonadales	Geobacteraceae	Geobacter	taxa in PCR_NC_B
fa1620623cfa2a2030f277b1695e01bf	Bacteria	Gemmatimonadetes	BD2-11 terrestrial group	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC_B
0c261cb300d4e1cfc5c59ae3099692cc	Bacteria	Chloroflexi	OLB14	uncultured bacterium	uncultured bacterium	uncultured bacterium	taxa in PCR_NC_B
b2a8d4f63fddc9918a507789c9119502	Bacteria	Chloroflexi	Chloroflexia	Thermomicrobiales	JG30-KF-CM45	uncultured bacterium	taxa in PCR_NC_B
dc9b0496419e8f4ef082527e462207eb	Bacteria	Nitrospirae	Nitrospira	Nitrospirales	Nitrospiraceae	Nitrospira	taxa in PCR_NC_B
c5227d674bb2db2fab607c7861a2ef02	Bacteria	Chloroflexi	TK10				taxa in PCR_NC_B
bc9fc445f23e0b544dc9f98339927cd4	Bacteria	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Paenalcaligenes	taxa in PCR_NC_B
d36b9db25cf20829de50a8c61cfbd3ed	Bacteria	Verrucomicrobia	Verrucomicrobiae	Chthoniobacterales	Chthoniobacteraceae	Candidatus Udaeobacter	taxa in PCR_NC_B

3. Bowtie2 , Anvi'o + RGI protocol

#megahit

/home/berechnung/miniconda3/pkgs/megahit-1.2.9-h8b12597_0/bin/megahit --presets meta-sensitive -1 /home/berechnung/Documents/Max/All_Dachstein_Aral_Sea_trimmed/A51_fwd_paired.fastq.gz -2 /home/berechnung/Documents/Max/All_Dachstein_Aral_Sea_trimmed/A51_rev_paired.fastq.gz -m 0.9 --mem-flag 2 -o /home/berechnung/Documents/Max/All_Dachstein_Aral_Sea_assembled_megahit/A51 -t 18

#anvi'o

#fasta fix

anvi-script-reformat-fasta '/home/berechnung/Documents/max+julia/A52.contogs.500nt.fa' -o A52.contogs.500nt_fixed.fa -l 0 --simplify-names

#bowtie2

cutadapt contigs.fa -m 500 -o contigs.500nt.fa

bowtie2-build --seed 1605 '/home/berechnung/Documents/max+julia/contigs_fixed/A52.contogs.500nt_fixed.fa' A52.contogs.500nt_fixed_index

bowtie2 --sensitive-local -p 20 --seed 1605 -x '/home/berechnung/Documents/max+julia/A52.contogs.500nt__fixed.index' -1 '/home/berechnung/Documents/max+julia/trimmed/A52_fwd_paired.fastq' -2 '/home/berechnung/Documents/max+julia/trimmed/A52_rev_paired.fastq' -S A52.contigs.500nt.sam -q

samtools faidx A52.contogs.500nt_fixed.fa

samtools import A52.contigs.500nt_fixed.fa.fai '/home/berechnung/Documents/max+julia/contigs_fixed/A52.contigs.500nt_fixed.sam' A52.contigs.500nt_fixed.bam

samtools sort -@ 4 '/home/berechnung/Documents/max+julia/contigs_fixed/A52.contigs.500nt_fixed.bam' -o A52.contigs.500nt.sorted.bam

#anvi'o

#database creation

anvi-gen-contigs-database -f '/home/berechnung/Documents/max+julia/contigs_fixed/A52.contogs.500nt_fixed.fa' -o A52ontigs_2.db -n 'A52 contigs datbase'

#run NCBI

anvi-run-ncbi-cogs -c A52ontigs_2.db --num-threads 25 --sensitive

#taxonomy

anvi-get-sequences-for-gene-calls -c A52ontigs.db -o A52.contogs.500nt.sorted.init_gene_calls.fa

#kaiju

#makedb

'/home/berechnung/miniconda3/pkgs/kaiju/bin/kaiju-makedb' -s nr -t 30

#classification

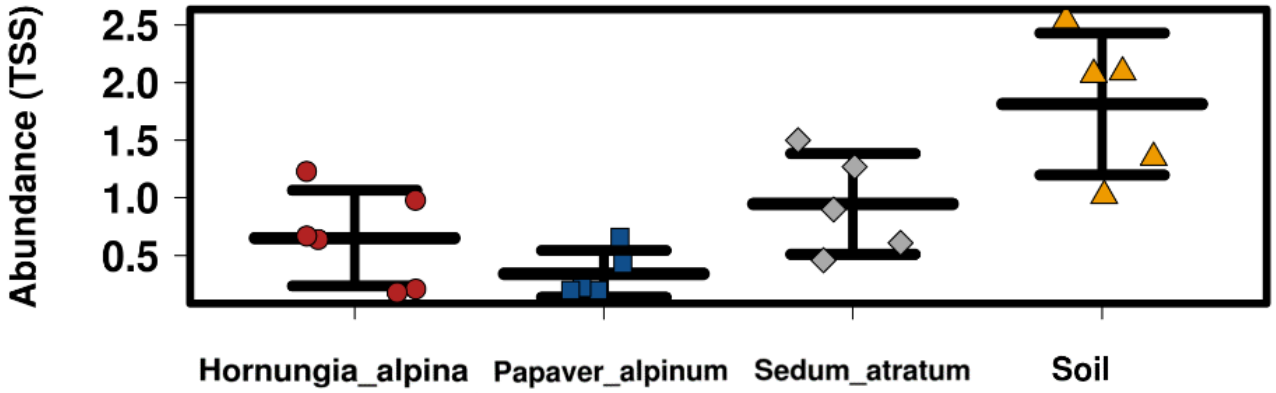
/home/berechnung/miniconda3/pkgs/kaiju/bin/kaiju -t '/home/berechnung/miniconda3/pkgs/kaiju/bin/nodes.dmp' -f '/home/berechnung/miniconda3/pkgs/kaiju/bin/kaiju_db_nr.fmi' -i '/home/berechnung/Documents/max+julia/anvio/A52.contogs.500nt.sorted.init_gene_calls.fa' -o '/home/berechnung/Documents/max+julia/taxnomy/A52.contogs.500nt.sorted.init_fixed_gene_calls_nr.out' -z 30 -v

#add taxon names
/home/berechnung/miniconda3/pkgs/kaiju/bin/kaiju-addTaxonNames -t '/home/berechnung/miniconda3/pkgs/kaiju/bin/nodes.dmp' -n '/home/berechnung/miniconda3/pkgs/kaiju/bin/names.dmp' -i '/home/berechnung/Documents/max+julia/taxnomy/A52.contogs.500nt.sorted.init_fixed_gene_calls_nr.out' -o '/home/berechnung/Documents/max+julia/taxnomy/A52.contogs.500nt.sorted.init_fixed_gene_calls_nr.out.names' -r superkingdom,phylum,order,class,family,genus,species
#import-taxonomy
anvi-import-taxonomy-for-genes -i '/home/berechnung/Documents/max+julia/taxnomy/A52.contogs.500nt.sorted.init_fixed_gene_calls_nr.out.names' -c '/home/berechnung/Documents/max+julia/anvio/A52ontigs_2.db' -p kaiju --just-do-it
profiling mit contig clustering
anvi-init-bam '/home/berechnung/Documents/Max/anvio/bam/A52.contigs.500nt.sorted.bam' -o A52.contigs.500nt.sorted.init.bam
anvi-profile --input-file '/home/berechnung/Documents/Max/anvio/bam/index/A52.contigs.500nt.sorted.init.bam' --contigs-db /home/berechnung/Documents/Max/anvio/A52ontigs_2.db --sample-name A5_2 -o '/home/berechnung/Documents/Max/anvio/profiles' --cluster-contigs --min-contig-length 1000 --num-threads 44

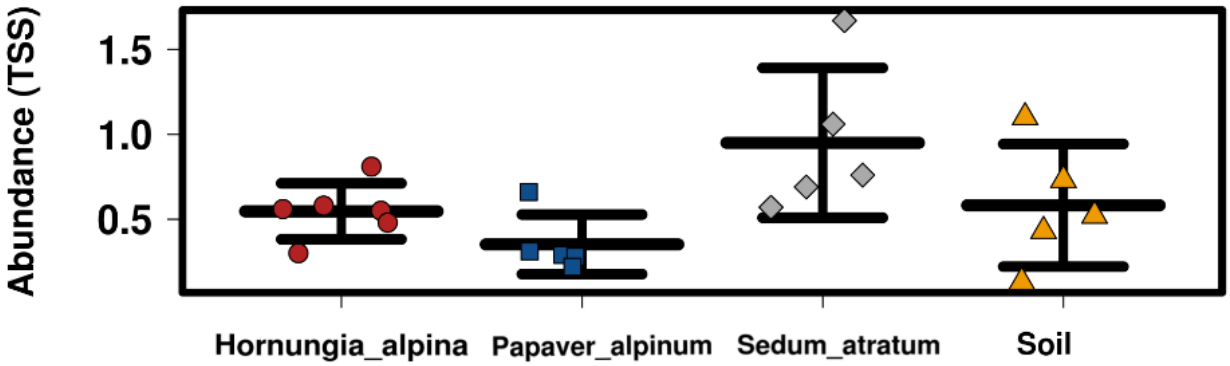
#kaiju
#kaiju2krona
/home/berechnung/miniconda3/pkgs/kaiju/bin/kaiju2krona -t '/home/berechnung/miniconda3/pkgs/kaiju/bin/nodes.dmp' -n '/home/berechnung/miniconda3/pkgs/kaiju/bin/names.dmp' -i '/home/berechnung/Documents/max+julia/taxnomy/A52.contogs.500nt.sorted.init_fixed_gene_calls_nr.out' -o A52_kaiju.out.krona
#export krona as html
ktImportText -o A52_kaiju.out.html A52_kaiju.out.krona
#anvi'o
#interactive
anvi-interactive -p '/home/berechnung/Documents/max+julia/anvio/A52.contigs.500nt.sorted.init.bam-ANVIO_PROFILE' -c '/home/berechnung/Documents/max+julia/anvio/A52ontigs_2.db'
#RGI bwt
rgi bwt --read_one /home/berechnung/Documents/Max/Aral_Sea_rgi/rgi_bwt_approach/A402_fwd_paired.fastq.gz --read_two /home/berechnung/Documents/Max/Aral_Sea_rgi/rgi_bwt_approach/A402_rev_paired.fastq.gz --aligner bowtie2 --output_file /home/berechnung/Documents/Max/Aral_Sea_rgi/rgi_bwt_approach/A402raw_out --threads 32 --include_wildcard

4. P-values (ANOVA) of LEfSe plot taxa

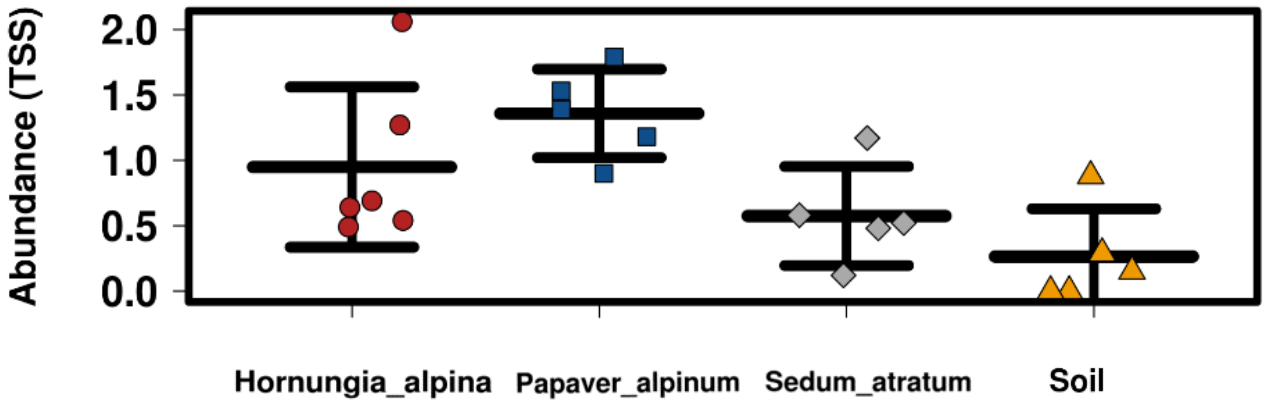
4.1 Dachstein:



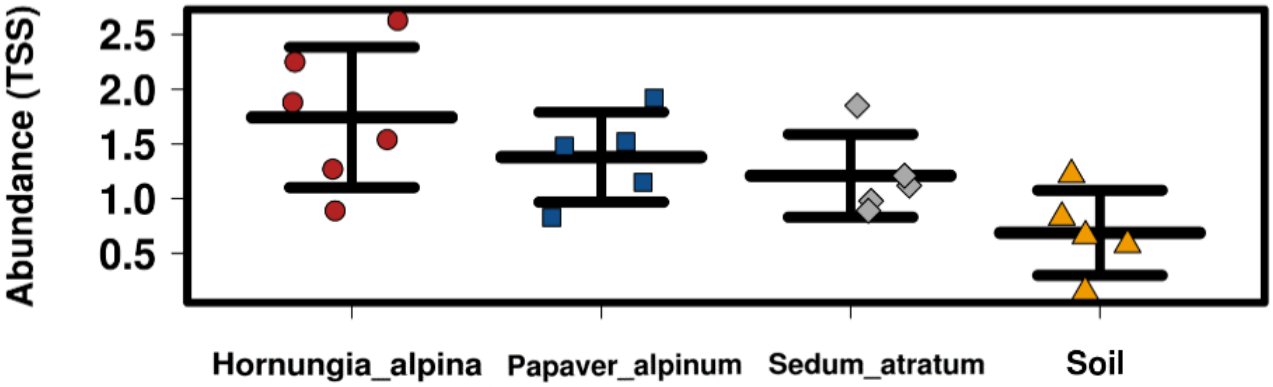
Dachstein alpha diversity Shannon index distribution of *Blastocatella* on ASV level. Samples clustered in plant species *Hornungia alpina, Papaver alpinum, Sedum atratum* and soil; each symbol represents one sample; p-value (ANOVA) = 0.00039.



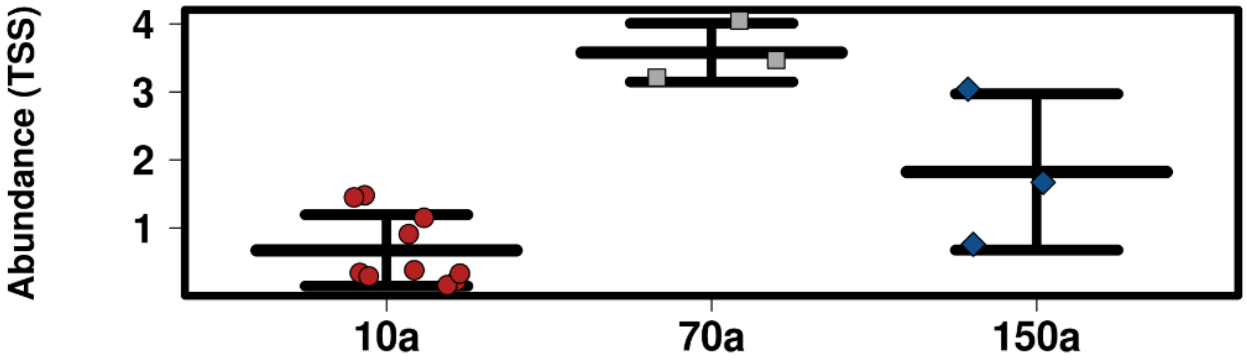
Dachstein alpha diversity Shannon index distribution of uncultured *Rhodanobacteraceae* on ASV level. Samples clustered in plant species *Hornungia alpina, Papaver alpinum, Sedum atratum* and soil; each symbol represents one sample; p-value (ANOVA) = 0.041.



Dachstein alpha diversity Shannon index distribution of *Pseudomonas* on ASV level. Samples clustered in plant species *Hornungia alpina, Papaver alpinum, Sedum atratum* and soil; each symbol represents one sample; p-value (ANOVA) = 0.0076.

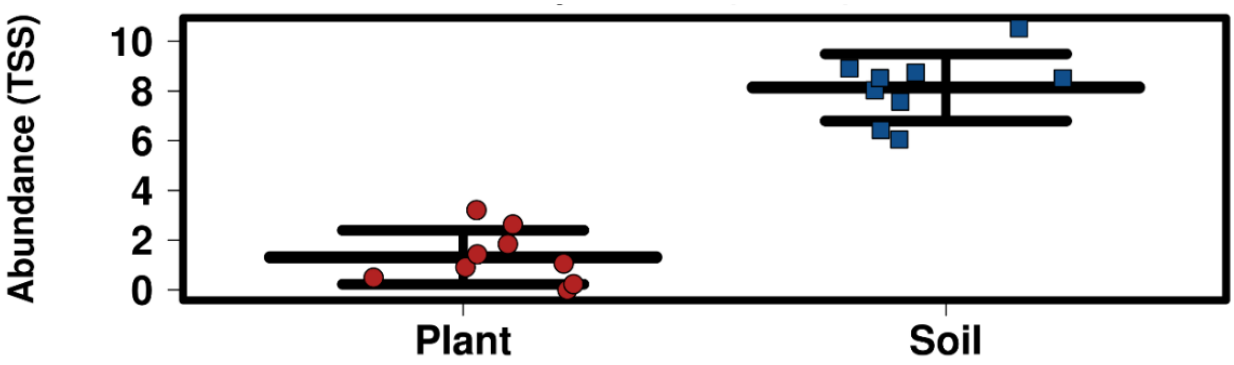


Dachstein alpha diversity Shannon index distribution of *Nitrosomonadaceae Ellin6067* on ASV level. Samples clustered in plant species *Hornungia alpina, Papaver alpinum, Sedum atratum* and soil; each symbol represents one sample; p-value (ANOVA) = 0.017.

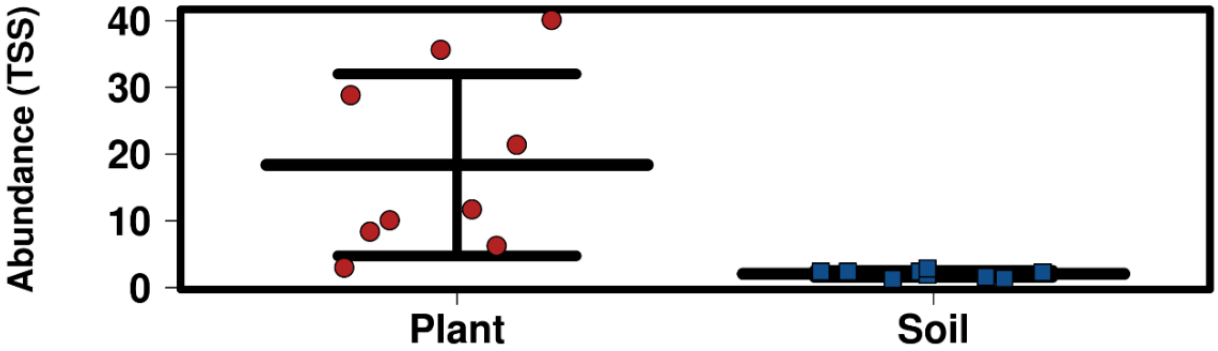


Dachstein alpha diversity Shannon index distribution of uncultured *Nitrososphaeraceae* on ASV level. Samples clustered in years without ice; 10a: 10 years since ice receded; 70a: 70 years since ice receded; 150a: 150 years since ice receded; each symbol represents one sample; p-value (ANOVA) = 4.5e-05.

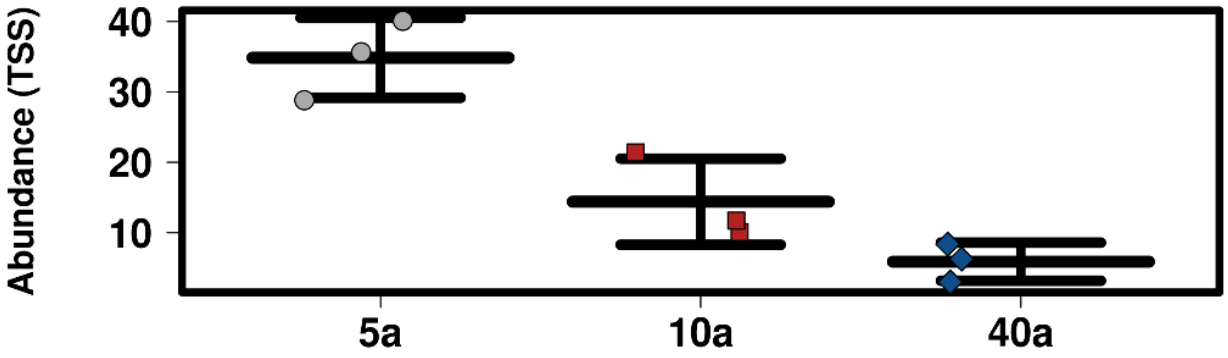
4.2 Aral Sea:



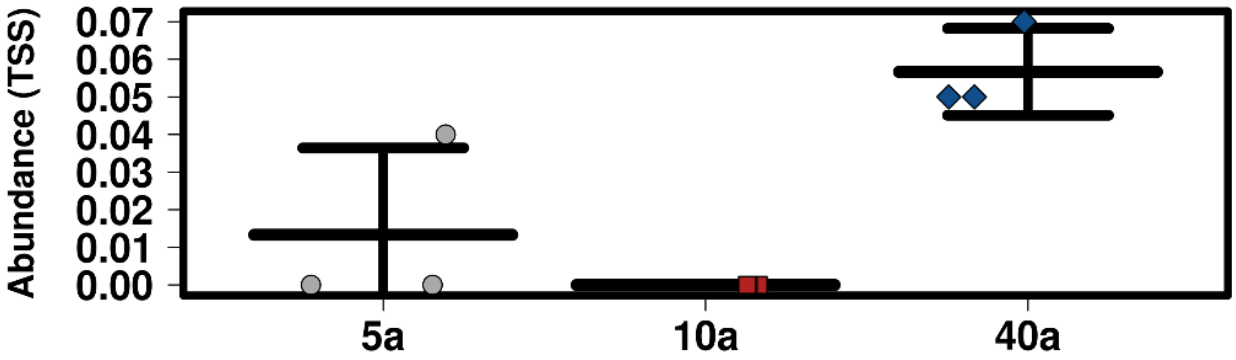
Aral Sea alpha diversity Shannon index distribution of uncultured *Halomicrobiaceae* on ASV level. Samples clustered in plant associated and soil; each symbol represents one sample; p-value (ANOVA) = 4.1e-09.



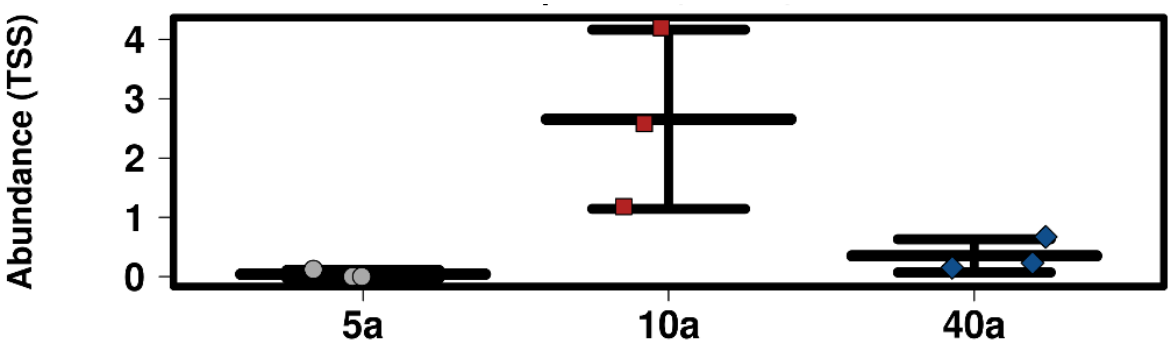
Aral Sea alpha diversity Shannon index distribution of *Halomonas* on ASV level. Samples clustered in plant associated and soil; each symbol represents one sample; p-value (ANOVA) = 0.007.



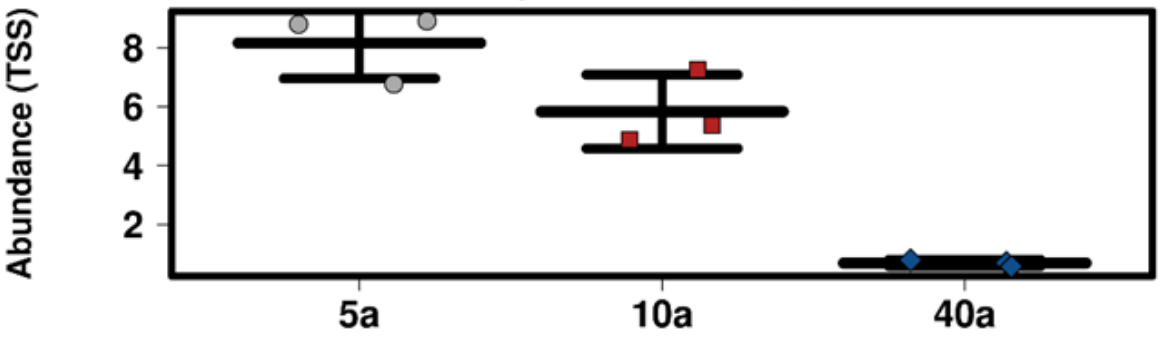
Aral Sea alpha diversity Shannon index distribution of *Halomonas* on ASV level. Samples clustered in years without water; 5a: 5 years since water receded; 10a: 10 years since water receded; 40a: 40 years since water receded; each symbol represents one sample; p-value (ANOVA) = 0.007.



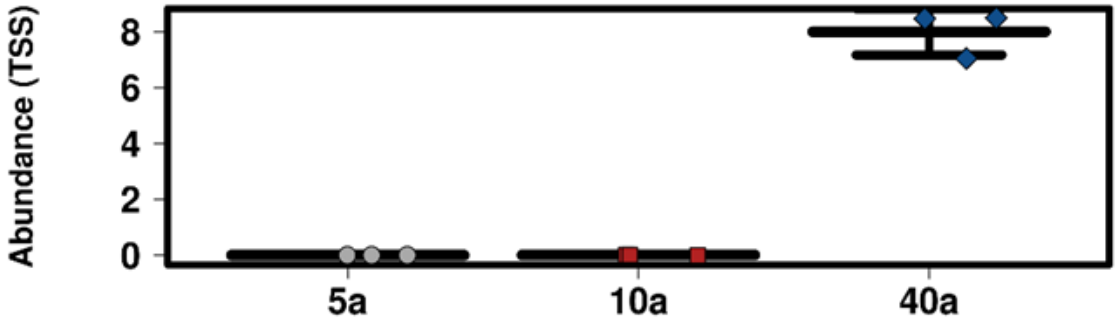
Aral Sea alpha diversity Shannon index distribution of *Solirubrobacter* on ASV level. Samples clustered in years without water; 5a: 5 years since water receded; 10a: 10 years since water receded; 40a: 40 years since water receded; each symbol represents one sample; p-value (ANOVA) = 0.0082.



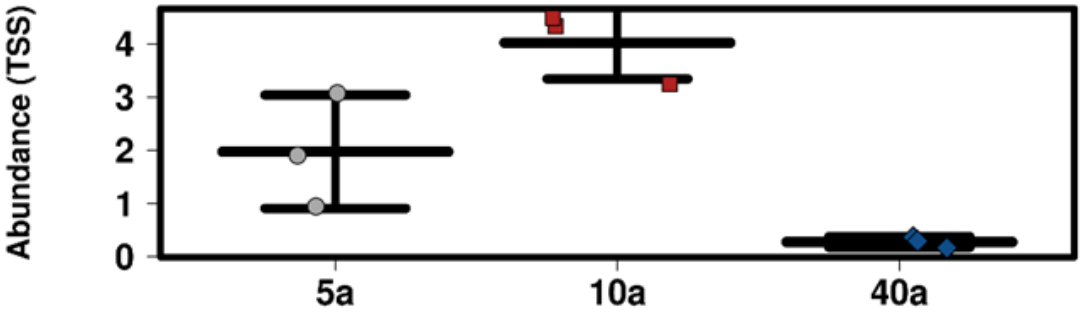
Aral Sea alpha diversity Shannon index distribution of *Myceligenerans* on ASV level. Samples clustered in years without water; 5a: 5 years since water receded; 10a: 10 years since water receded; 40a: 40 years since water receded; each symbol represents one sample; p-value (ANOVA) = 0.022.



Aral Sea alpha diversity Shannon index distribution of uncultured *Gammaproteobacteria* on ASV level. Samples clustered in years without water; 5a: 5 years since water receded; 10a: 10 years since water receded; 40a: 40 years since water receded; each symbol represents one sample; p-value (ANOVA) = 0.00027.



Aral Sea alpha diversity Shannon index distribution of uncultured *Longimicrobiaceae* on ASV level. Samples clustered in years without water; 5a: 5 years since water receded; 10a: 10 years since water receded; 40a: 40 years since water receded; each symbol represents one sample; p-value (ANOVA) = 1.2e-06.



Aral Sea alpha diversity Shannon index distribution of uncultured *PAUC43f marine benthic group* on ASV level. Samples clustered in years without water; 5a: 5 years since water receded; 10a: 10 years since water receded; 40a: 40 years since water receded; each symbol represents one sample; p-value (ANOVA) = 0.0023.

5. CARD and wildCARD reads

	A52	A420	A-Soil5	A-Soil40	Ho1
CARD	11,953	44,902	6,946	13,471	12,687
in silico perfect	857	432	388	233	2,183
in silico strict 100%	9,247	8,614	4,760	794	1,990
in silico strict 98 -100%	8,938	2,495	2,806	1,376	7,402
in silico strict 95 -98%	4,344	3,097	1,090	381	2,270
in silico strict 90 -95%	7,386	2,612	8,190	7,100	5,251
in silico strict 80 -90%	569	725	318	106	1,306
in silico strict 50 -80%	37,638	6,641	4,656	1,736	13,839
in silico strict <50%	5,070	2,258	1,796	877	7,094
total	86,002	71,776	30,950	26,074	54,022

VIII. **Abbreviations**

AMR	antimicrobial resistance
ANOVA	analysis of variance

AR	antibiotic resistance
ARG	antibiotic resistance gene
ASV	amplicon sequence variant
C.A.Mey.	Carl Anton von Meyer
L.	Carl Linnaeus
LEfSe	linear discriminant analysis effect size
Moq.	Alfred Moquin-Tandon
MRSA	methicillin resistant Staphylococcus aureus
O. Appel	Oliver Appel
PCoA	principal coordinates analysis
PCR	polymerase chain reaction
PERMANOVA	permutational multivariate analysis of variance
PPS	protein precipitation solution
qPCR	quantitative polymerase chain reaction