2022

# Investigating the microbiota associated with maerl beds located in the Fal estuary, Cornwall

Scales, Anthony

UNIVERSITY OF
PLYMOUTH

# Investigating the microbiota associated with maerl beds located in the Fal estuary, Cornwall

By ANTHONY SCALES

A thesis submitted to the University of Plymouth in partial fulfilment for the degree of

RESEARCH MASTERS

School of Biological and Marine Sciences

July 2022

## Acknowledgements

With special acknowledgments to: Dr Kelly Haynes, Prof Jason Hall-Spencer, DrKaren Moore And Abi Outred. Also, special acknowledgments to Exeter university and Seale-Hayne Education Trust for proving funding.

Author's Declaration

At no time during the registration for the degree of Research Masters has the author been registered for any other University award without prior agreement ofthe Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has notformed part of any other degree either at the University of Plymouth or at another establishment.

This research has been conducted under a formal agreement with Plymouthuniversity and Cornwall College for which a joint award will be awarded

This study was self-financed and carried out in collaboration with ExeterUniversity and the Seal-Hayne Educational Trust.

A programme of advanced study was undertaken, which included taught modulestaken, -A programme of advanced study was undertaken, which included taught modules taken – BIO5131 post graduate research skills and methods, BIOL5001 advanced postgraduate skills, GSRBIOL1 research biological sciences, GSRBIOL2 research biological sciences, GSRBIOL3 research biological sciences and ResM biological sciences (4066).

The following external institutions were visited for consultation purposes: ExeterUniversity

Presentations at conferences:Cornwall College

Falmouth Marine School

Word count of main body of thesis: 11051

Signed:  Dated: 09/07/2022

TITLE:

# Investigating the microbiota associated with maerl beds located in the Fal estuary, Cornwall

By ANTHONY SCALES

## Abstract

The microbiota of maerl has the potential to contain diverse communities of bacteria which are thought to aid in maintaining maerl health and potentially be a driving factor in the settlement and development of ecological and economical important bi-valve species. However, little is currently known regarding the composition and structure of these communities, leading to uncertainty of their importance and their direct association to maerl beds. Using 16s amplicon sequencing analysis, the bacterial community composition associated with live and dead maerl (*Phymatolithon calcareum*) has been examined and compared to the communities associated with those of the surrounding environment - adjacent sediments and overlaying seawater. Four unique communities were present on the following samples: two associated with each of the surrounding environments (overlaying seawater and adjacent sediments), one associated with live and dead whole maerl nodules and the remaining observed on the surface layers of living maerl nodules. Sediment communities were seen to contain the most diversity, with enriched members of Moraxellaceae and Desulfobulbaceae present. Seawater communities in comparison harboured the lowest diversity and was dominated with members from Clade_I, Flavobacteriaceae, Rhodobacteraceae. Dead and living maerl nodules contained similar communities to one another with increased members of Flavobacteriaceae, Pirellulaceae, Nitrosopumilaceae witnessed across both types of nodules. Surface

layers of the living maerl nodule contained a unique community compared to the that of the living whole nodule, enriched with members of Flavobacteriaceae, Rhodobacteraceae, Vibrionaceae.

This thesis concludes that maerl beds studied in the Fal estuary contain two unique bacterial communities. Whilst the analysis of bacterial functions is beyond the scope of this study, it is highly speculated that members of either community are providing vital ecological roles in supporting the maerl bed and other associated species observed within maerl habitats. It is therefore the authors belief that greater protection should be afforded to maerl beds to protect ecological and economic interests.

*List of contents:*

## List of Illustrations

*1 - Introduction*

'Maerl' is the term use throughout this thesis for unattached non- geniculate coralline red algae thalli. Accumulations of these slow growing thalli entwine to form a lattice of living and dead maerl on the sea floor forming habitats called 'maerl beds'. These beds are typically composed of a dead base layer with a living maerl layer growing on top. When well established, maerl beds are complex three-dimensional habitats for many species and are often considered to be biodiversity hotspots (Hall-Spencer 2010). Some maerl beds are thousands of years old, with individual thalli growing just 0.5–1.5 mm per tip per year, usually breaking and growing in an asexual fragmentation cycle (Blake and Maggs, 2003). Fragmentation is the main reproduction method with sexual thalli rarely seen, for example over a two-year monthly observation of maerl beds in Galway Bay (Maggs, 1983). There are several accounts of multiporate sporangial conceptacles in *Phymatolithon calcareum* and *Lithothamnion corallioides*, but sexual reproduction is rare (Peña *et al.*, 2014a). The complex three-dimensional habitats formed by the entwining maerl provide a habitats for an exceptionally wide range of other benthic species (Hall-Spencer., 2010). These include epiphytes attached to the maerl (Qui-Minet *et al.*, 2018), arthropods and annelids living within the 3D structure of this biogenic sediment and vertebrate species swimming over the bed (Foster *et al.*, 2013). Maerl habitats also provide economic benefits, providing suitable settlement and growing areas for commercially important bivalves. Some bivalve molluscs, such as the Pacific calico scallop (*Argopecten ventricosus*)preferentially settle on maerl and demonstrated faster growth (Steller, D. L. and Cáceres-Martínez, 2009). Maerl can induce settlement of bivalves through production of chemical cues such as

GABA (γ-aminobutyric acid) detected on the surface layers of maerl (Morseand Morse, 1984). It is not known whether these cues are produced directly by the maerl species or by the microbial community that lives on their surfaces. Microbes on corals and macroalgae can produce chemical cues that induce bivalve settlement (Thompson *et al.*, 2015) and macroalgal mucus exudates support microbial communities (Armstrong *etal.*, 2001).

Many studies have investigated the microbiota of corals and other major ecosystem engineers, but there has been a distinct lack of investigation of maerl beds and other calcified macroalgae (Hollants *et al.*, 2013b). Overall,the microbiota found on coral and macroalgae species appear to be either mutualistic or communalistic by nature, providing varied and numerous benefits to the host species. Such benefits include the recycling of nutrientsinto new bioavailable forms enabling enhanced growth as an indirect consequence, or producing chemicals which stimulates growth (Goecke *et al.*, 2010). Whilst also encouraging growth, certain chemicals are vital in the formation of the standard structure of the host species and without these chemical the host grows in an irregular manner (Matsuo *et al.*, 2003).

Other microbes within similar communities are known to produce cues which deter the settlement of pathogenic species to the host and in doing so providing a first line of defence against host associated diseases (Lam, Stang and Harder, 2008). Other microbial community members are known to produce chemical cues which attract beneficial macro species such as bi-valves and other macro species, potentially increase the surrounding waterquality and/or reducing other algal competitors and thus enhancing the living conditions of the maerl ( Steller, D. L. and Cáceres-Martínez, 2009). However, whilst there is evidence supporting beneficiary microbial roles being

performed, there has been observations of parasitic activity, particularly targeting species already in a vulnerable state (Busetti, Maggs and Gilmore, 2017) Maerl bed habitats are found in tropical, temperate and polar regions at depths ranging from the low intertidal zone to over 150 m (Rendina *et al.*, 2020). Maerl habitats in the lower Fal, St Mawes Bank and the Helford river, Cornwall UK, are around 4000 years old (Bosence and Wilson, 2003) and were a key feature in assigning the area with SAC status in 2004 (Allen *et al.*, 2013). The Fal estuary receives nutrient inputs from various sources, some of these are natural - coastal erosion, precipitation run-off, tidal flow and wild animal defaecation (Jickells *et al.*, 2014) - whilst others are anthropogenic – derived from boats or sewage discharge (Paerl *et al.*, 2006) . Nutrient overloading causes eutrophication induced algal blooms (Kennish, 2002) and can stimulate the growth of an array of species from microbes to macrofauna and flora (Statham, 2012). Nitrogen, phosphate and carbon that is locked within complex organic matter is utilised by organisms, resulting in recycled bioavailable nutrients (Nielsen, Banta and Pedersen, 2007). Additionally, these nutrient inputs can provide methods of transport for microbial life which may aid or hinder native microbiota and host species (Litchman, 2010; Abu-Bakar, Ahmadian and Falconer, 2017). On the other hand, certain inputs contain concoctions of toxic elements and compounds which can cause profound effects on the ecosystem. One such notable toxic pollutant is the high concentrations of heavy metals that are locked with the sediments. These heavy metals include iron, lead, manganese, zinc, arsenic, , copper and cadmium have been significantly linked to the rich history of mining in the surrounding area (Whitehead and Prior, 2005).Whilst the surrounding sediments already contain heighted concentrations of heavy metals, the regular pumping of excess mine water further exacerbated these concentrations. The largest influxes of pollutants occurred

in an accidental acid mine water release such as the Wheal Jane incident during 1992 (Pirrie *et al.*, 2003; Whitehead and Prior, 2005). As a result from this accident, an approximate 50 million litres of metal laden acid mine water – mainly iron oxides though small amounts of other metals were detected - was released directly into the surrounding water ways. However, species such as *Nereis diversicolor* found within the esturary have demonstrated a higher tolerance to Cu and Zn compared to individuals from non-pollinated areas (Bryan *et al.*, 1987). Similar tolerances are also observed within *Ostrea edulis* populations, indicating species can become tolerate to certain heavy metals and inhibit areas polluted with them (Bryan *et al.*, 1987). Whilst some species haveadapted, others such as the *Cerastoderma edule* are now absent. Moreover, *Nucella lapillus* are still present within the area, however, present varying degrees of imposex - the appearance of male anatomy in females – caused by not only mining pollutants but tributyltins derived frommarine craft antifouling paints (Bryan et al., 1987; Nicolaus and Barry, 2015). Although the microbial communities of the Fal environments were never documented before pollution occurred, it is assumed that their presence has and can still influence microbial communities. Inputs within the estuary can carry microbial loads which in addition to adding to nutrient or pollute loads can also introduce new invasive microbe species,potentially causing temporary shifts within native microbial communities. Known sources originate from sewage discharge, overflow pipes and smaller boats, which are all known to operate within the fal estuary area (Kershaw and Campos, 2010). The largest contributor to this input is the Falmouth Sewage treatment works (STW) which discharges its consented 9,500 $m^3$/day tertiary treated effluent at St Mawes. Although this is the largest STW in the area, there are 3 other STW`s also operating in the area contributing to the enrichment of microbial loading

during periods where sewage is produced above the threshold level (Kershaw and Campos, 2010). These impacts of enriched microbial loading can be visually witnessed within the shellfish industry that operates within the estuary with numerous accounts of temporary/permanent closures due to microbial loads above safe consumer thresholds. While the shellfish industry and the surrounding water body are monitored routinely – due to the EC Water framework directive and EC habitats directive -, numerous habitats such as the maerl beds situated within the area are not monitoredand may experience microbial community shifts during or shortly after periods of input. However, the maerl beds are generally located in areas ofhigh-water movements and so could mitigate or limit the effect of invasivemicrobes.

This thesis`s aim was to provide an insight into the bacterial community composition situatedon living and dead maerl thalli via the use of 16S rDNA. This insight was gained and tested via three hypotheses; 1) there is a difference in the bacterial community composition present on the maerl compared to that which is associated with the surrounding sediment and seawater, 2) there are community differences between the surface layers and their corresponding total maerl nodules, 3) there is a difference in community profiles between dead maerl nodules and living maerl nodule.

## 2 - *Methods*

### 2.1 - *Sample collection*

Sampling took place on the 3rd of June 2018 at the maerl beds located in St Mawes Bank (Falmouth, Cornwall). This site was chosen due to being established and very well -studied (Howson *et al.* 2004, Allen *et al.*, 2014). As St Mawes Bank is within a Special Area of Conservation and the maerl is protected by The Conservation of Habitats and Species Regulations 2010 (EC habitats directive, 1992), sampling permission was obtained from NaturalEngland – reference number 243974. Sediment collection was taken outside the maerl bed to minimise disruption, at the request of Habitats Regulations Assessment.

Highly trained SCUBA divers collected the samples, minimising potential damage to the maerl bed. Live maerl sample collection occurred at 50° 09.992'N, 05° 01.957'W (Dive site 04, Fig. 1). Three live maerl thalli, each estimated to be 5 cm long, were collected using an inverted food-grade ziplock bag to collect and isolate individual nodules. Each thallus was bright pink/red in colour and devoid of macroscopic epiphytes. A seawater sample was taken from one metre above the location of the sampled thalliin a sterile 15 ml falcon tube. Three dead maerl thalli and their respective seawater samples were collected as above at site 8 (Fig. 1).. Six sediment core samples spaced 3 m apart were collected using 15 ml sterile falcon tubes pushed into the sediment to a depth of 44mm. The last sediment sample location was recorded, 50° 09.819N, 05° 02.085W, and all samples were transported on ice inside a thermal box to Cornwall College Newquay and stored at -20$^{o}$C until DNA extraction was carried out.

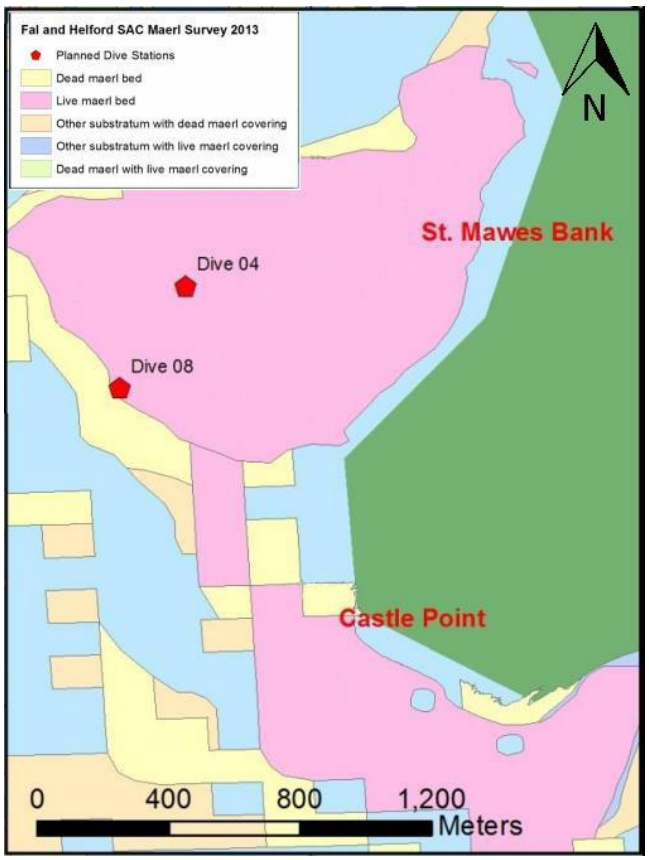


*Figure 1: A map of the maerl bed location at St. Mawes Bank. The collection of samples started fromthe entrance point – dive 04 – and ended at the exit point - dive 08. This map was modified from Allen et al., 2014.*

## 2.2 - *Sample and equipment preparation for genetic identification and Microbial DNA*

All stored samples (-20°C) were defrosted at 4°C for 15.5hrs to allow to defrost throughout. All equipment was autoclaved at 121°C for 80 minutes (Suyama and Kawaharasaki, 2013), unless stated differently. During sample preparation, raw samples were stored on ice whilst subsamples were takenand immediately returned to -20°C storage after use. Sub-samples were stored at 4°C whilst other sub-samples were created, allowing DNA extraction to be carried out in unison.

## 2.3 - *Genetic identification of maerl*

### 2.3.1 - *DNA extraction and amplification*

DNA was extracted from both the living and dead maerl nodules as to perform genetic identification. Maerl samples were ground and prepared following the same protocols as discussed within the *Total community subsamples section.* DNA was extracted using the DNeasy mini plant kit by Qiagen and followed manufacturers recommended protocols (appendix 1)The two independent gene regions< Mitochondrially encoded Cytochrome C Oxidase I (*cox1)* and Photosystem II protein D1 (*psbA)*, were used for identification. Primer sets for *cox1*: GazF1 5' TCAACAAATCATAAAGATATTGG 3' and GazR1 5' ACTTCTGGATGTCCAAAAAAYCA 3 (Saunders, 2005) and *psbA*-F 5'ATGACTGCTACTTTAGAAAGACG 3' and *psbA*-R2 5' TCATGCATWACTTCCATACCTA 3' (Yoon, Hackett and Bhattacharya, 2002) were used. PCR was carried out independently for each gene but followedthe same protocol; 25 µl TopTaq Master Mix (Qiagen), 2.5 µl forward primer (0.5 µmol), 2.5 µl reverse primer (0.5 µmol), 15 µl nuclease free water (Qiagen) and 5 µl of template DNA. *cox1*followed a modified PCR protocol of an initial denaturation 94°C for 2 min; then 5 cycles of denaturation at 94°C for 30 s; annealing at 45°C for 30 s; extension at 72°C

for 1 min; a further 35 cycles of denaturation at 94°C for 30 s; annealing at 46.5°C for 30 s; extension at 72°C for 1 min; final extension at 72°C for 7 min and then a final hold at 10°C) (Saunders and Moore, 2013). *psbA* followed a standard PCR protocol of an initial denaturation 94°C for 5 min, then 35 cycles of denaturation at 94°C for 30 s; annealing at 50°C for 30 s; extension at 72°C for 1 min; final extension at 72°C for 5 min and final holdat 10°C) was used for *psbA*(Melbourne et al., 2017). Amplifications

were visualised using 2% agarose gel electrophoresis. Successful live maerl amplifications were purified using purelink PCR purification kit (Thermo Fisher Scientific) following manufacturers recommended protocol. Aliquotsof purified live maerl samples were sent for sanger sequencing at Macrogen (Amsterdam).

### 2.3.2 - *Maerl identification using sequence data*

Consensus *psbA* sequences were trimmed and formed using Bioedit version 7.0. Identification of closest matches was gained via inputting sequences into the Basic Local Alignment Search Tool (BLAST). All live maerlsample sequences returned with a 97-100% match for *Phymatolithon calcareum psbA* gene, accession number JQ896231.1. The next closest matched species was *Phymatolithon laevigatum* with a 94% match

## 2.4 – *Environmental sample preparation*

### 2.4.1 – *Sediment microbiota subsample*

The entire sediment samples were vortexed at 24,000rpm, creating a homologous suspension. This suspension was then centrifuged at 6000rpmfor 15 secs producing a homologous sediment pellet. 50 mg of sediment was removed from the pellet, forming a single subsample. Each subsamplewas weighed and kept in a 1.5 ml tube ready for DNA extraction. This was repeated so that each sediment sample had one 50 mg subsample.

2.4.2 – *Seawater microbiota samples*

Seawater samples were vortexed at 29 RCF (2,400 rpm Fisherbrand™ Classic Vortex Mixer) and concentrated onto 0.22 µm filter paper (mixed cellulose esters membrane, Millipore) using a sterile Büchner filter and vacuum pump. Filterpapers were cut into smaller fragments and stored within a 15 ml tube.

Sterilisation of the Büchner filter flask was achieved by washing with 70% ethanol and allowing to saturate for 10 seconds before drawing the solution through the filter operating the pump. Ethanol residue was removed from the Büchner filter flask by washing with autoclaved distilled water ($ADH_2O$) and then drawing the solution through the filter. Two bunsen burners were lit either side of the vacuum filter to mitigate air- borne contamination and maintain aseptic conditions.

*2.5 – Microbiota maerl subsample preparation and observations*

Individual maerl nodules (live and dead) were used to create two subsamples, one representing the surface microbial communities and theother a looking at the whole community found on and within the nodule.

Each nodule was visually inspected for epiphytic organisms or signs of abnormalities and only visually acceptable sections were utilised.

*2.5.1 – Surface maerl microbiota subsamples*

Short surface layer sections (<5 mm) of maerl were carefully incised, compiling a combined weight of 50 mg. Sterilisation of the scalpel blade was achieved via flaming and was incorporated between removing sectionsof maerl from the same sample. A new sterile scalpel blade was used for each new sample. Surface layer

sections were submerged in 150 ml $ADH_2O$ and vortexed at 45 RCF (3,000rpm Fisherbrand™ Classic Vortex Mixer) for 30 secs. The supernatant was filtered through a 0.22 μm filter paper and prepared in identical manner as the seawater samples previously described.

*2.5.2 - Total maerl microbiota subsamples*

Whole maerl branches were transferred into a new sterile zip lock bag and pulverised into a fine powder using a pestle and mortar, 50 mg of maerl powder was weighed and stored in a 1.5 ml tube. Containment within the sterile zip lock bag whilst being pulverised prevented cross contaminationbetween samples. Blue paper towel was used to line the inside of the mortar and to cover the ziplock bag ensuring the integrity of the bag.

2.6 - *DNA Extraction of microbial communities*

All samples were extracted and purified using the Qiagen DNeasy blood and tissue kit in accordance with manufacturer's instructions. Steps within the Pre-treatment for Gram-Positive Bacteria Protocol (pg52) and purification protocol (pg 32) were followed, however, modifications of the volumes and/or incubation times were used.Samples which contained filter paper had volumes of lysis buffer, ATL buffer, proteinase K, Al buffer and ethanol multiplied by a factor of 3.15 (Renshaw *et al.*, 2014), ensuring the complete emersion of the filter paper. All samples were subjected to a modified Buffer ATL and proteinase K incubation stage, with increases of temperature of 65°C (from 56°C) and extended incubation time of 2 hours

(increased from 30mins) (Boström *et al.*, 2004; Renshaw *et al.*, 2014). This extended incubation was undertaken due increased DNA yields with an increased incubation compared to the manufactures specified time, as demonstrated by to Boström *et al.*, (2004).Samples which contained higher volumes of extraction supernatant – samples containing filter papers – required 3 additional iterations of step 4 of the purification protocol (pg 34). This was due to the spin column havinga maximum volume capacity less than the total volume of extraction supernatant and so required step 4 to be repeated multiple times, ensuringthe total volume of extraction supernatant was concentrated onto the spin column. After which point, the standard protocol was followed. Elution wasperformed with 100 µl of buffer AE to increase DNA concentration. This step was repeated with another sterile microcentrifuge tube to create a second elution, ensuring maximum yield of DNA whilst each elution contained the highest concentration possible.

2.7 - *DNA amplification and sequencing*

Aliquots of extracted purified DNA were sent to Exeter Sequencing Servicefor DNA quantification, PCR, purification, and massively parallel sequencing. Details of all used protocols, reagents, and the construction ofthe PV4 primers can be found within appendix 1. These protocols and regents are adapted from the Exeter Sequencing Service's 'Standard protocol for Illumina MiSeq-based 16S rRNA gene studies'. DNA concentrations (appendix 2) were quantified using the Agilent High Sensitivity D1000 ScreenTape system. PV4 Primers were used in PCR and PV4 protocols were followed to amplify the v4 region of 16S rDNA. Sequencing was carried out using the Illumina MiSeq platform with v2 chemistry (PE250).

2.8 - *Bioinformatics*

Sequencing sample files were demultiplexed and cleaned of non-biological nucleotides by Exeter sequencing service and then returned for further bioinformatic analysis. Samples were organised into an ASV table using DADA2 in R, following the DADA2 pipeline Tutorial 1.12 created by Benjamin J Callahan (Callahan BJ *et al.*, (2016).

A phyloseq variable (McMurdie and Holmes (2013) was created using the Bonus: Handoff to phyloseqtutorial listed within the DADA2 tutorial with the addition of adding a phylogenetic tree, enabling the use of the unifrac distance measurements for further analysis. The R script appendix 3 was used for data analysis andvisual aid creation Alpha diversity was assessed using a box and whiskers plot which incorporated the Shannon-weiner measurement (H values). The dataset used was untrimmed and unfiltered as not to distort the alpha values by removing singletons. An ANOVA was used on generated Shannon H values to assess if significant differences lay between samples. Beta diversity was investigated using a Hierarchical cluster analysis, which utilised non- weighted unifrac distance measurements and used the Lance–Williams dissimilarity updated formula to create clusters from the complete dataset converted into relative abundance. The Hierarchical cluster analysis was used to assess the dissimilarities between the samples to a greater accuracy compared to the NMDS, which can suffer from distortion of plotted points. The dataset was first transformed into relative abundances as to mitigate biases caused by differences sequence counts between sample types. Microbial compositions were first examined by using a horizontal stacked bar plot and

rank table which used the ten most abundant families from each sample types. The taxonomic level of family was chosen due to retaining high-resolution differences between sample types whilst retaining classifications in as many of the taxa observed as possible. Differences between these families were then assessed using a heat map which investigated the 50 most abundant ASV`sfrom the entire dataset arranged to family taxonomic levelling. Whilst this does not tell us exactly what genus or species the ASV represents, it does highlight differences within the families whilst remaining identifiable.

## 3 – *Results*

### *3.1 Overview*

Microbial community Alpha diversity was analysed using the Shannon Wiener index (H values) and highlighted that sediment samples contained the richness diversity 6.87 (H), whilst seawater samples collected over dead maerl nodules contained the lowest richness diversity (4.38 H). All maerl samples, except for live maerl exterior samples, contained a similar H value. The live maerl exterior contained a noticeable lower diversity value to that of other maerl sample types. Beta diversity was analysed using hierarchical cluster analysis by means of the non-weighted unifrac distance metric. From this analysis, four main different microbial communities were revealed indicted by the clustering of sample types. The four clustered groups were composed of the following sample types: group 1 – sediment only, group 2 – live maerl exterior only, group 3 – dead maerl exterior, dead and live maerl total communities, and group 4 – seawater collected over both dead and live maerl nodules. Lastly, investigating the similaritiesand differences of the four main distinctive microbial

communities was carried out in two ways. The first was by ranking the most dominant taxonomic families by relative abundance in each sample type, highlighting either a highly dominant community or an evenly distributed community. The ranking of the families were then compared against other sampletypes to see similarities or difference in community composition. The second method used a heat map and investigated the 50 most abundant ASVs from all samples. These ASV`s were classified to family taxonomic level and the abundances were compared between each sample. Whilst similar to the ranking method, the heat map allowed finer resolution as it made use of ASVs and as such highlighted differences between sample types which contained similar families in the ranking table.

*3.2 - Alpha diversity*

Seawater samples collected over dead maerl nodules - Dead seawater – contained the lowest median microbial alpha diversity (H values) score (4.38 Fig 2). This was similar to the diversity observed within the seawatercollected over live maerl nodules – Live seawater (4.48). Live maerl exterior -LMEX - contained the lowest median diversity (5.33) of the maerlsample types, however, a single replicate from both Live maerl total communities (LMT) and dead maerl total communities (DMT) contained values lower than LMEX. However, both DMT and LMT contained much higher median values (6.46 and 6.60, respectively) and their values were more comparable to that of the dead maerl exteriorcommunities - DMEX (6.77). Although,DMEX contained a much higher diversity compared to LMEX (1.54 median difference). Sediment contained the highest diversity (6.87), though this was only marginally higher than that of DMEX and was the only sample type which contained

values which overlapped.Whilst the differences of alpha diversity do represent differences in richness of the sample type, it does not directly support or reject the studies hypothesis. However, it aids supporting differences perceived during beta diversity analysis.

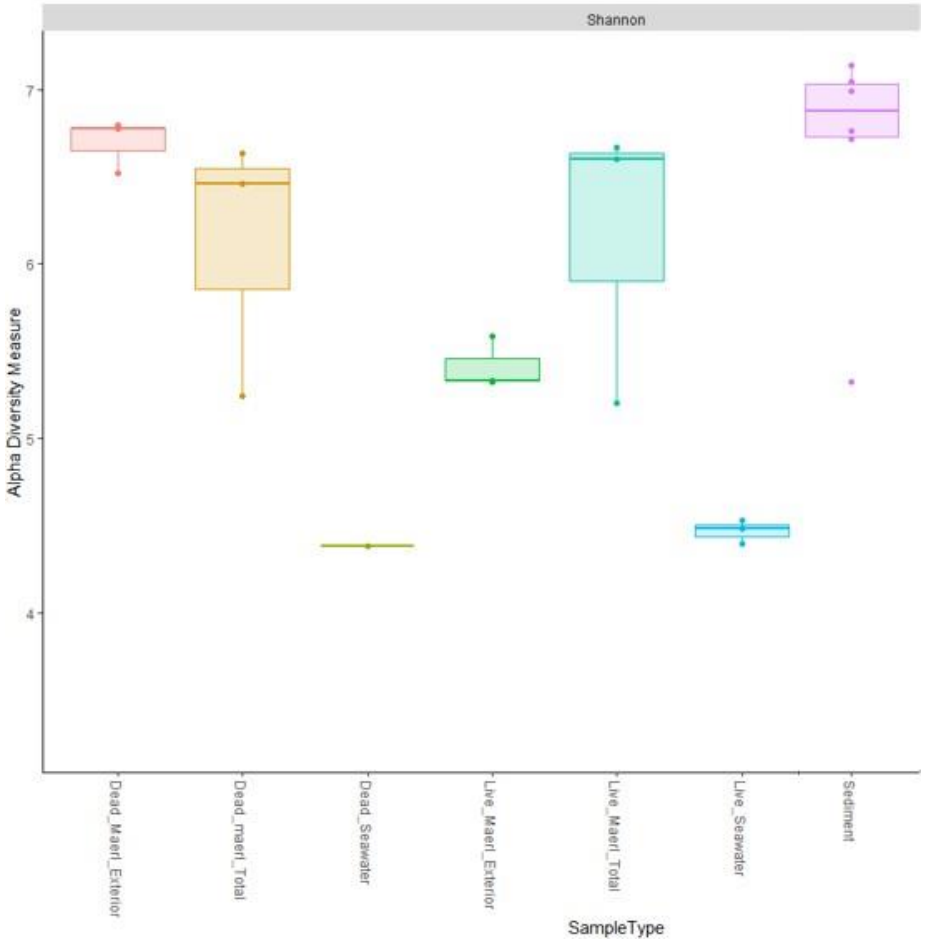


*Figure 2: Alpha diversity was measured using the Shannon wiener index (H values). The boxplot compactly displays the distribution of a continuous variable. It visualises five summary statistics (the median, two hinges and two whiskers), and all "outlying" points individually. The lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles). The upper whisker extends from the hinge to the largest value no further than 1.5 * IQR from the hinge (where IQR is the inter-quartile range, or distance between the first and third quartiles). The lower whisker extendsfrom the hinge to the smallest value at most 1.5 * IQR of the hinge. Data beyond the end of the whiskers are called "outlying" points and are plotted individually. Significance based on alphadiversity measurements between sample types ( P-value = 0.0039, DF Sampletype = 8, Residuals DF = 16)*

*3.3 - Beta Diversity*

A cluster dendrogram (Fig 3) using a non-weighted Unifrac distance metric and clustered by the Lance–Williams dissimilarity update formula was used to determine similarities and differences in microbial communities within sample types. Four distinctive groupings were formed from this, representing four unique bacterial communities . Sediment and LMEX samples formed two independent groupings, which will be known as groups 1 and 2 respectively. These two groups were not attached directly together but instead were joined via the second adjacent branch. However, Group 3 - composed of the remaining maerl sample types – DMEX, DMT and LMT – was joined directly to an adjacent branch to that of group 1 . The last isolated cluster was composed of both dead and live seawater samples and was not conjoined to any other groups via adjacent branches. Interesting, DMT sample replicates were not directly branched to one another as observed with other sample type replicates, indicating differences in communities situated within these replicates.

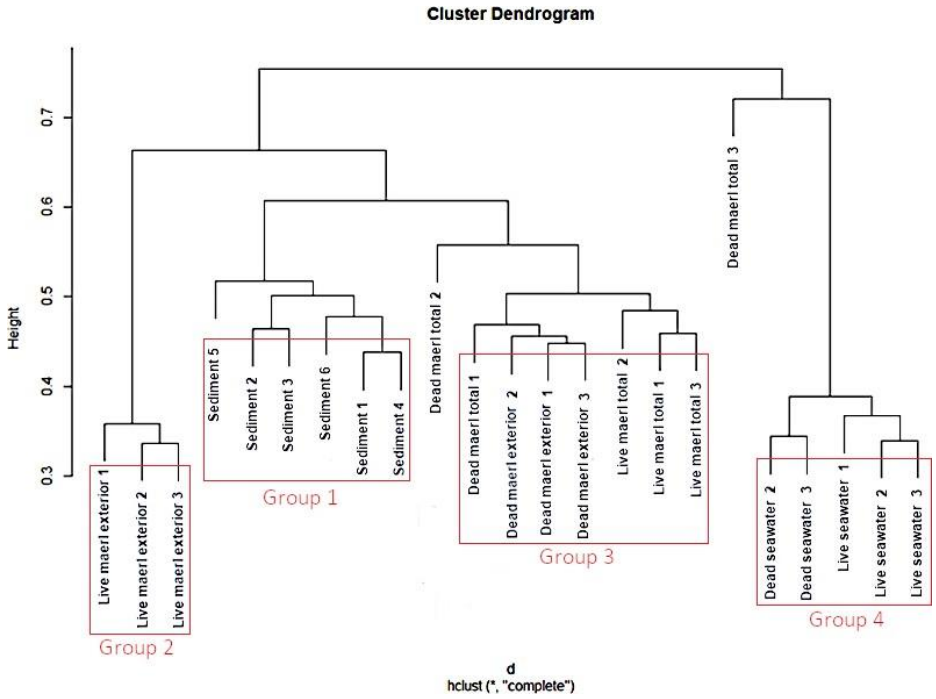


*Figure 3: Hierarchical cluster analysis using non-weighted Unifrac distance metric to determine sample relatedness Clusters are computed by the Lance-Williams dissimilarity update formula. Four distinct groupings can be witnessed, group one composed of the sediment. Group 2 of the Live maerl exteriors. Group 3 formedby a collection of Live maerl total, Dead maerl exterior and two replicates of Dead maerl total. Group4 consisted of both seawater types, seawater taken from above live or dead maer*

### *3.4 - Microbial compositions*

The ten most relative abundant taxa – classified to family level - were compared between sample groups, with similarities between live samples of all groups and dead samples of all groups (Table 1 and Fig 3). Vibrionaceae was the only family seen in all samples with relatively high abundance in all but more abundant in the living samples (highest abundances within LMEX, 9.41%, ranked 3$^{rd}$). Thiohalorhabdaceae was present in all Maerl samples, with relative abundance highest in LMEX (5.54%, ranked 6$^{th}$), but absent inseawater. Flavobacteriaceae was present in high abundances in Live seawater, LMT and LMEX and was witnessed lower abundances in all othersamples. Whilst the highest abundances were recorded within Live Seawater (19.58%, ranked 2$^{nd}$), it was the dominantfamilywithin LMEX (17.57%, ranked 1$^{st}$). Moracellaceae was the

Dominant families within sediment (10.1%) and DMT (6.21%) with verylittle abundance being observed within all other sample types. Sediments also contained Pasteurellacaea (3.62%) and Desulobacteraxeae (4.03%) which were predominantly only found within the sediments. Whilst DMT was similar to other sample types, it did contain Nostocaceae, which is only found within this sample type. The dominant families within both dead and live seawater were Clade_1, Flavobacteriacea and Rhodobacteraceae. Whilst Clade_1 was predominantly found only in the seawater samples, Flavobacteriacea and Rhodobacteraceae were observedin the other sampletype with higher abundances seen in LMEX. Interestingly. Whilst LMEX contained the highest abundances of Sphingomonadaceae (3.27%), Rhizobiaceae (3.77%), Microtrichaceae (3.36%), Arenicellaceae (2.96%), similar relative abundances were observed within dead and live maerl total and dead maerl exterior.

| Ranking | Sediment | Seawater over dead maerl | Dead maerl Total commu |
|---|---|---|---|
| 1 | Moraxellaceae 10.10% | Clade_I 19.14% | Moraxellaceae 6.21% |
| 2 | Desulfobulbaceae 8.62% | Flavobacteriaceae 17.30% | Pirellulaceae 5.84% |
| 3 | Pirellulaceae 4.89% | Rhodobacteraceae 14.77% | Flavobacteriaceae 5.34% |
| 4 | Flavobacteriaceae 4.30% | Vibrionaceae 7.22% | Nostocaceae 4.43% |
| 5 | Halieaceae 4.24% | SAR116_clade 5.59% | Halieaceae 4.05% |
| 6 | Desulfobacteraceae 4.03% | Halieaceae 4.53% | Vibrionaceae 3.19% |
| 7 | Pasteurellaceae 3.62% | Clade_II 2.87% | Nitrosopumilaceae 3.08% |
| 8 | Unknown_Family 2.90% | Ectothiorhodospiraceae 2.70% | Sandaracinaceae 2.88% |
| 9 | Anaerolineaceae 2.07% | Cyanobiaceae 2.32% | Woeseiaceae 2.83% |
| 10 | Woeseiaceae 1.69% | Methylophilaceae 2.27% | Enterobacteriaceae 2.35% |

| Ranking | Dead maerl surface community | Seawater over Live Maerl | Live maerl surface community | Live maerl Total Communit |
|---|---|---|---|---|
| 1 | Flavobacteriaceae 8.56% | Clade_I 21.14% | Flavobacteriaceae 17.57% | SAR116_clade 16.20% |
| 2 | Pirellulaceae 7.51% | Flavobacteriaceae 19.58% | Rhodobacteraceae 9.59% | Flavobacteriaceae 10.81% |
| 3 | Woeseiaceae 4.99% | Rhodobacteraceae 13.68% | Vibrionaceae 9.41% | Pirellulaceae 5.36% |
| 4 | Halieaceae 4.81% | SAR116_clade 4.88% | Pirellulaceae 6.73% | Nitrosopumilaceae 3.96% |
| 5 | Nitrosopumilaceae 4.45% | Vibrionaceae 4.56% | Thiohalorhabdaceae 5.54% | Rhodobacteraceae 3.67% |
| 6 | Rhodobacteraceae 2.70% | Halieaceae 3.87% | Rhizobiaceae 3.77% | Woeseiaceae 3.53% |
| 7 | Thiohalorhabdaceae 2.60% | Clade_II 3.01% | Microtrichaceae 3.36% | Halieaceae 3.12% |
| 8 | Saprospiraceae 2.33% | Cyanobiaceae 2.84% | Sphingomonadaceae 3.27% | Vibrionaceae 3.11% |
| 9 | Sandaracinaceae 2.31% | AEGEAN-169_marine_group 2.02% | Arenicellaceae 2.96% | Thiohalorhabdaceae 2.18% |
| 10 | Rhizobiaceae 2.18% | Ectothiorhodospiraceae 2.00% | Halieaceae 2.81% | Rhizobiaceae 2.00% |

*Table 1: Table representing the 10 highest relative abundant families foundwithin each sample type.*

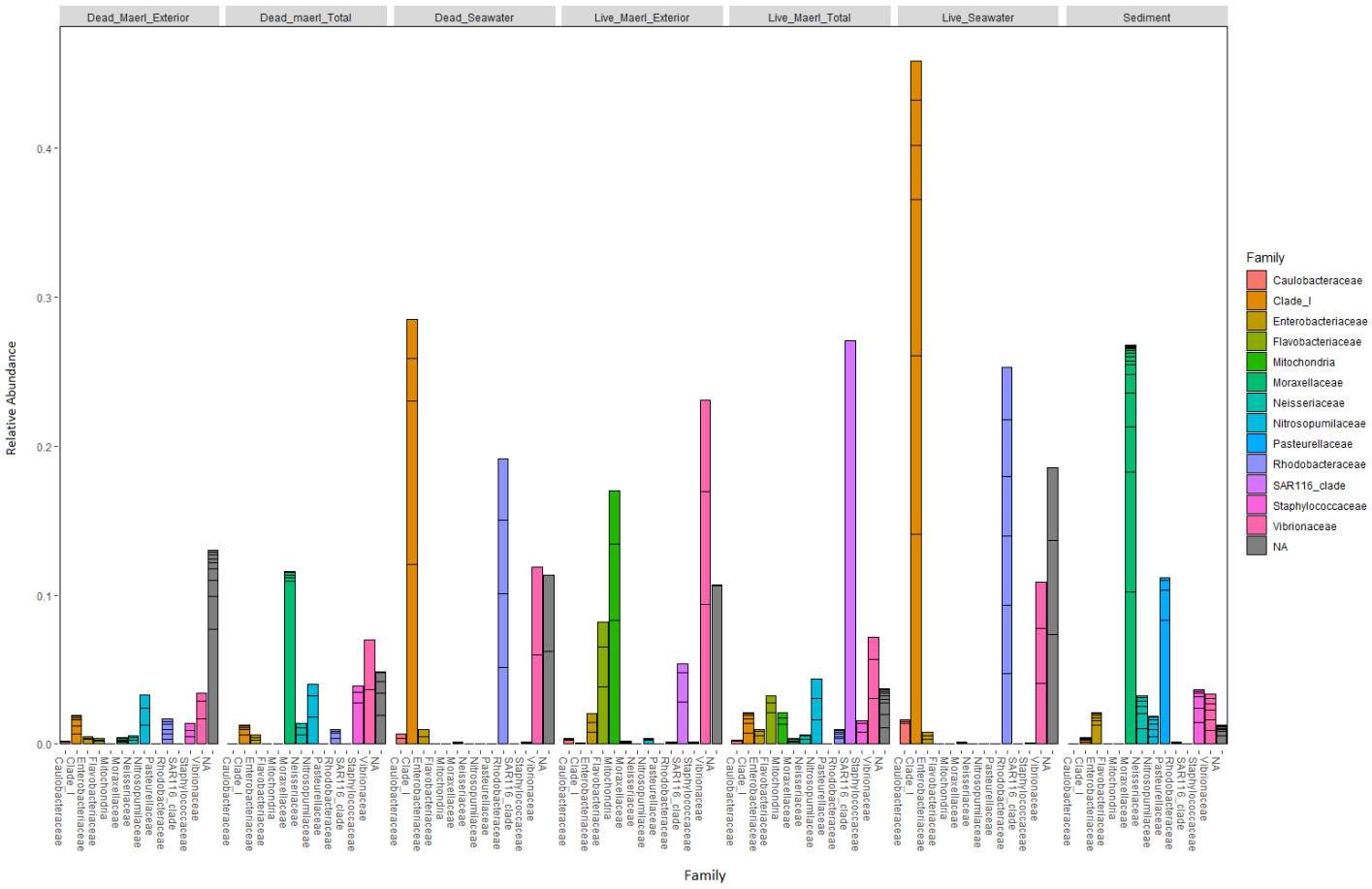


*Figure 5: Relative abundance of the 20 most abundant families found within each sampletype .*

A heat map examining the 50 most abundant ASV`s (log 4 scale) arranged to a family level taxonomy was used to investigate the microbial composition of the sample-types with greater resolution (fig 6). Examining the differences of taxa between each sample type further highlights the four distinct groups previously mentioned. Group 1 contained The following enriched Taxa:Desulfobulbaceae , Sulfurovaceae , Desulfobulbaceae, Spongiibacteraceae,3 unknown family classisfed ASV`s and Carnobacteriaceae.. This group shared a similar abundance of taxa listed in the Y axis starting from Thiohalorhabdaceae till Pirellulaceae between LMT, DMT and DMEX. Thesetaxa as well as those that were enriched were absent or present in low abundances in the seawater and LMEX samples.

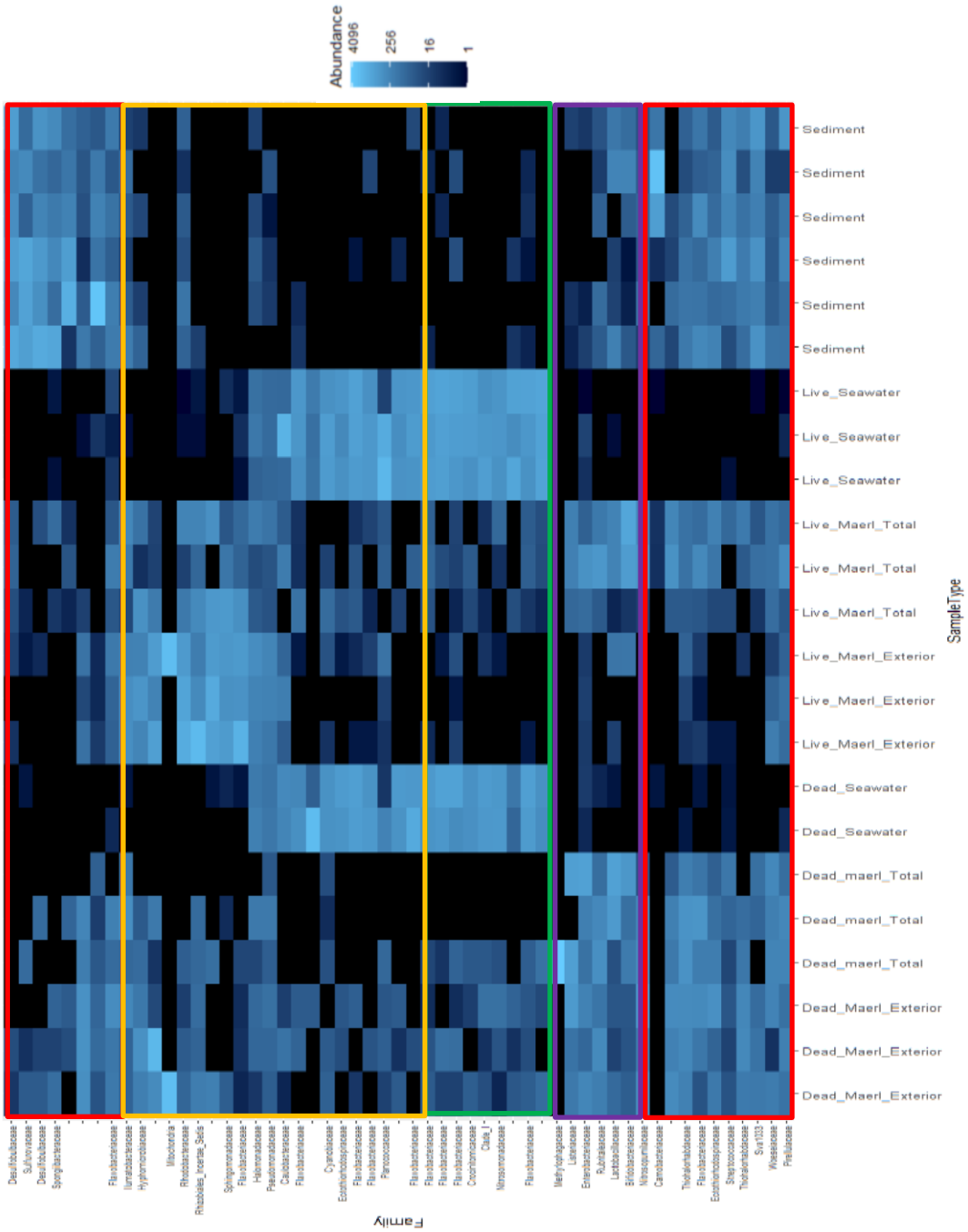


*Figure 6: The 50 highest abundant ASVs (Log 4 scale) ,classified to family taxonomic level. The red box highlights ASVs predominantly found in group 1 -Sediment  The green box highlights ASVs* predominantly found within group 2 – live maerl surface layers. The purple box highlights ASVs predominantly found with group 3 – live and dead maerl whole nodules and dead maerl surface layers. The purple box *highlights ASVs predominantly found in group 4 – Seawater found over dead and live maerl nodules*

Group 2 contained the highest log 4 abundances of Ilumatobacteraceae, Hyphomicrobiaceae, Rhodobacteraceae, Rhizobiales_Incertae_Sedis, Sphingomonadaceae and Flavobacteriaceae. These are only marginally higher in abundance (500-20 difference) compared to live maerl total samples however are notably different to both the dead maerl exterior and total samples. Whilst seawater samples appear to contain very few of these taxonomic groups, sediment samples contain Ilumatobacteraceae, Hyphomicrobiaceae, Rhodobacteraceae in similar concentrations as both
the dead maerl samples. Rhodobacteraceae, an unknown ASV, Rhizobiales_Incertae_Sedis, Sphingomonadaceae and Flavobacteriaceae taxa present on the live maerl exterior appear to be particularly enrichedcompared to other samples. (Fig 6)

Group 3 contained an even abundance across all taxa, however, Listeriaceae, Enterobacteriaceae, Rubritaleaceae, Lactobacillaceae, Bifidobacteriaceae and Nitrosopumilaceae taxa recorded the highest abundances in all samples (Fig 6). Group 1 contained approx. 1000-100 lower log 4 abundances of these taxa, with the exception of Nitrosopumilaceae whose abundance is only slightly lessened, approx. 150 difference. Groups 2 and 4 contain noticeable difference in abundance of these taxa, with many of them being absent. Interestingly, Enterobacteriaceae is present in the seawater samples, live maerl exterior and some of the sediment samples in very similar abundance compare to one another, however, is enriched on the live maerltotal and both the dead maerl samples.

Group 4 contained highly enriched taxa listed in fig 6, with the exception of: Halomonadaceae, Pseudomonadaceae and Caulobacteraceae whose abundances were similar to the maerl species butnot group 1. Whilst Group 4 was observed to contain these highly dominant taxa, it however harboured very lowabundance of taxas present in other groups.

*3.5- Conclusion*

The four distinctive groups observed fig 3 and examined in fig 6 provides evidence to support two hypothesis and one null hypothesis. The first hypothesis it supports is: There is a difference in the bacterial composition of maerl compared to the surrounding environment. This is due to all the maerl samples, groups 2 and 3, not being clustered near to either sedimentor the seawater samples; as well as clear differences in taxa witnessed.

Additionally, live maerl exterior displayed the highest degree of difference from the seawater and sediment samples with live maerl total and both dead maerl samples being slightly similarity to the sediment bacterial composition. The second hypothesis it supports is: there is a difference between the surface and the total maerl microbial composition. Whilst thisdoes highlight only very slight differences between the exterior of dead maerl and the total of the dead maerl nodule, it does provide evidence of the differences between the exterior of the live maerl and the entire live maerl nodule. The null hypothesis that it supports is: there is no difference between dead maerl nodules and living maerl nodules. This conclusion wasmade since dead maerl and live maerl totals were clustered

very closely together in fig 3, contained similar diversity levels fig 2 and contained very similar microbial composition fig 6. However, it should be noted that the surface layers of live maerl contain substantial differences to the other maerl samples and if you solely compare these surface layers then it would provide evidence against this null hypothesis.

# 4 - *Discussion*

## *4.1 - Summary*

There are noticeable differences in microbiota found on maerl nodules - *Phymatolithon calcareum,* both live and dead nodules - compared to the communities found within the surrounding seawater and sediment. These differences are observed within the alpha and beta diversity as well as the composition of dominant families. However, little variation is witnessed when comparing the communities found on the surface of dead maerl nodules and the total communities situated in complete live and dead nodules. Furthermore, a unique community was documented on the surface layers of the living maerl nodule, different from other communities observed on live and deadmaerl nodules and the surrounding environment.

## *4.2 – Group 1: Sediment communities*

The sediment samples recorded the richest microbiota diversity of all tested samples, which is in line with other estuarian sediments (Rublee, 1982; Chen *et al.*, 2019). The top abundant families witnessed included Moraxellaceae, Desulfobulbaceae and Flavobacteriaceae, which are associated with

decomposer roles, such as the recycling of various nutrients, including carbon, nitrogen, sulphur and iron reduction (Baker *etal.*, 2015; Reyes *et al.*, 2016; Hinger, Pelikan and Mußmann, 2019). This recycling of nutrients could potentially aid in supporting a complex community highlighting why a broad community was observed. Which leads to the hypothesis that no single factor is primarily responsible for enhancing the richness and abundance of these families. Principally, it is understood that copious assorted amounts of nutrients are readily available, derived either from rich coastal waters or inputs from the land (Kemp and Boynton, 1984). In addition, anthropogenic inputs such as sewage discharge and agricultural runoff further increases available nutrients (Howarth, 2008; Rabalais *et al.*, 2009), both of which are known to impact areas adjacent to this study area- Fal river and Tolverne regions - (Walker, 2017). Instead of nutrient limitations, it is possible that microbe presence is due to certain abiotic and biotic stress factors and/or being outcompeted by better suited species. . Microbes are known to access required nutrients via several methods, these adaptions to access nutrients may help explain the presences of certain families. For example, Species members of the Desulfobulbaceae family are known to grow long filaments, known as electron transport cables, which carry out sulphur oxidation (Pfeffer *et al.*, 2012; Malkin *et al.*, 2014). These cables are documented to reach lengths of up to 1.5 cm long, demonstrating the extreme ranges at which sulphur cycling can occur and potentially how thisfamily can out compete other individuals situated with the sediments. On the other hand, members of Flavobacteriaceae produce extracellular enzymes and toxic compounds to break down complex exopolysaccharide (Bowman, 2006). Broken down

substrates enables access to nutrients which were locked in complex structures, such as detritus, from various origins. The contradictory methods of accessing nutrients may offer suggestions as why both families were witnessed within the sediment andable to thrive in relative high abundances. However, whilst these familiesdo appear to coexist, production of metabolites are known to be deleterious and enable one family to dominate over others. (Egan *et al.*, 2000).

Stress factors caused by the presence of several types of heavy metals in high concentrations found within the sediments may further influence the witnessed community composition (Sheeba *et al.*, 2017). Heavy metal presence has been demonstrated to cause harmful effects on microbes (Prabhakaran, Ashraf and Aqma, 2016). For example, the presence of Zinc (II) or copper are known to either induce cytostasis or cell death (Gikas *etal.*, 2009; Utgikar *et al.*, 2003). Whilst cadmium exhibits similar effects on microbes, it can potentially incur mutagenic impacts (Bischoff, 1982; Wang,Lu and Shen, 2007), rendering microbial communities vulnerable to other environmental stressors and/or bacteriophages. Furthermore, microbial communities within Jiaozhou Bay were seen to shift in composition depending on heavy metal concentrations within the sediments (Yao *et al.*, 2017). This bay harbours various heavy metals including zinc, arsenic, lead, copper and cadmium and has demonstrated to strongly influence the microbiota present within specific localised regions. Considering the Fal estuary contains a similar composition of heavy metals with the addition ofiron, lead, manganese and tin (Bryan, 1985; Pirrie *et al.*, 2003), it can be assumed that the presence of these

heavy metals are highly likely to influence the microbial communities recorded within this study. In addition, there is evidence that communities can shaped dependant on community member tolerances, advantageous members with higher tolerances outcompeting (Nies, 1999; Gubelit *et al.*, 2016).

*4.3 - Group 2: Live maerl surface communities (LMEX)*

Whilst Flavobacteriaceae and Rhodobacteraceae have the highest relative abundance within LMEX, there are high counts of these families within the seawater samples. However, Fig6 highlights several unique ASVs belonging to Flavobacteriaceae and Rhodobacteraceae witnessed only within LMEX and LMT at relatively high abundances. This may indicate an association between live maerl and these ASVs. The families which appear in the ranking table are believed to be the dominant familes, representing 70.36% of the microbiota situated on the surface of live maerl, suggesting an associated community. Additionally, LMEX recorded the lowest alpha diversity score of all the maerl samples (LMEX 5.33) suggesting mechanisms influencing community richness and composition. To further demonstrate this point, the microbiota of live maerl - *Sporolithon austral*- are documented to remain stable and unchanged even when exposed to simulated stressful conditions caused by increased $pCO_2$ levels (Cavalcanti *et al.*, 2018). Interestingly, different abundances of microbial phylawere observed on *Sporolithon austral (*Cavalcanti*et al.*, 2014), which may indicate microbial communities association to specific maerl species or community variation caused by localised environmental conditions. Whilst Proteobacteria remained the dominantphyla, all other groups either changed rank order or was not present in dominant abundances, Planctomycetes and

Verrucomicrobia were present within this study whilst Firmicutes were not. Shifts in dominant groups could be due to numerous factors, such as: Differences inmaerl species, variations in regional conditions, seasonality differences, presence/absence of key flora/fauna and anthropogenic activities. These differences could have also been influenced by the inherent bias of the executed molecular techniques chosen in each study (Edet *et al.*, 2017).

The dominant microbial families witnessed within LMEX are believed to livein either a mutualistic or commensal symbiotic relationship due the perceived overall good health of bed (Allen *et al.*, 2014). However, various individuals from the familiesof Rhodobacteraceae, Flavobacteriaceae and Vibrionaceae – all of which where found on LMEX - have been identified as opportunistic pathogens with the capabilities to cause various diseases to red algal and coral species (Egan *et al.*, 2013; Krediet *et al.*, 2013; Feng *et al.*, 2020). However, further Investigations into microbial interactions within maerl beds are required as to ascertain microbe-host relationships. There is further evidence of host calcifying and non-calcifying macroalgae associated microbes ability to recycle carbon, nitrogen, sulphate and phosphate, as well as the degrading of (poly)aromatic compounds into bioavailable metabolites (Goecke *et al.*, 2010; Aires *et al.*, 2019; Valdespino-Castillo *et al.*, 2021). These metabolities can then be directly utilised by the macroalgae host for its specific needs – growing and reproduction - or made bioavailable via another microbial loop (Hollants *et al.*, 2013a) . Whilst very little research on the functionality of microbial activities associated with maerl has been carried out, using studies from other

macroalgae and other host organismscan indicate what recorded dominantfamilies may be doing. It is assumed that the dominant families recorded within live maerl exterior are performing roles in the recycling of essential compounds. For example, pirellulaceae family members discovered on varying species of corals and sponges have been revealed to convert fixed nitrogen in the form of ammonia into nitrates (Mohamed *et al.*, 2010; Kellogg, 2019).

Interestingly, nitrate availability is one of the main drivers for enhancing rhodolith growth and abundance (Carvalho *et al.*, 2020). On the other hand, ammonium in concentrations higher than 5 µM restricts the uptake of nitrates by up to 50% of the seaweeds *Enteromorpha intestinalis* and *Gracilaria pacifica* (Thomas and Harrison, 1987). This potentially suggests that the abundance of Pirellulaceae species are playing a vital double role by removing the disadvantages ammonia and replacing it with beneficial nitrates. Removing the negative effects of ammonia and increasing the concentrations of nitrate may facilitate accelerated growth rates in maerl. Considering maerl`s skeletons are composed primarily of calcium carbonate, it can be assumed that the associated microbiota may provide a role in supporting the hosts requirements in suppling either precipitated calcium carbonate or the necessary required nutrients. It has been shown that several individuals within cyanobacteria and other specialised species outside of this phylum such as *Pseudomonas aeruginosa* and *Diaphorobacter nitroreducens* have the ability toconduct Precipitation Calcium Carbonate (PCC) (Erşan, de Belie and Boon,2015; Castro-Alonso *et al.*, 2019). Whilst these species that are known to PCC are not witnessed, by using “the coral probiotic hypothesis” (Reshef *etal.*, 2006) it can be suggested which

microbial groups are most important. This hypothesis suggests host species select the most advantageous holobiont depending on environmental stress factors (Peixoto *et al.*, 2017; Rosado *et al.*, 2019). Under increased $CO_2$ concentrations different microbial groups associated with maerl increased in abundances whilst others declined (Cavalcanti *et al.*, 2018). Familiesthat showed an increase and were present within this study are: Rhodobacteraceae (+ 117%), Rhizobiaceae (+ 311%) and Flavobacteriaceae (+ 334%). As there was no notable calcium carbonate biomass loss but there was a shift in the microbiota strongly indicts those microbes which saw a notable increase are either directly or indirectly involved with the biomineralisationof calcium carbonate. Considering these familieswere also witnessed within LMEX of this study, strongly suggests the importance these microbialfamilies are to maerl species.

In addition to the cycling of nutrients to facilitate growth, the dominant families could also be providing direct cross host-microbial communication which has been observed in other algal species, particularly within the Ulvagenus. These communications between host and microbiota appear to be vital for the functionality of individuals, *Ulva* species have been shown to have limited growth in comparison to their typical morphology, instead developing into a mass of undifferentiated callus cells whilst grown in axenic conditions (Fries, 1975; Spoerner *et al.*, 2012). Whilst growing *U. linza* 13 different bacterial isolates returned atypical morphology to normalwith five significantly enhancing growth rates (Marshall *et al.*, 2006), demonstrating multiple taxa have the potential for host-microbe communication. Furthermore, species

belonging to Rhodobacteraceae and Flavobacteriaceae familieshave been revealed to completely recover the growth and morphology of *U. mutabilis*, signifying the importance of these family to the host species. The high abundance of both families within the live maerl exterior samples of this study could indict a similar level of importance between maerl species and microbes. Microbe to algaesignalling could also extend from playing advantageous roles of the parent plant to also effecting zoospores settlement sites. Zoospores originating from *U. linza* have been revealed to accumulate in areas of bacterial derived N-(3-oxododecanoyl)-homoserine lactone (AHL) by a chemokinesis response (Wichard, 2015). Instead of being directly attracted to AHL (chemotactic response) it has been revealed that *U. linza* zoospores dramatically decrease swimming speeds when AHL is detected, resulting inthe tumbling and the accumulation of abundant zoospores (Tait, Karen ref 2004). Tumbling and settlement in this manner suggests quick colonisationof advantageous microbiota on the new individual, providing similar benefit as mentioned for the parent individual. Whilst the sexual reproduction cycles of maerl are currently not well understood, conceptacles have been recorded indicting sexual reproduction occurrence(Peña *et al.*, 2014b). Whilst red algae are known to produce carpospores and tertraspores and not zoospores, the microbial groups present may stillcommunicate via similar chemical cues which may induce production of gametophytes. Furthermore, spores are released in a broadcast manner into the environment and are non-motile like that of certain corals. Microbial loads containing high abundances of core microbiota have been recorded to be released alongside broadcast gametes (Leite *et al.*, 2017), which is believed to aid in the transmission of advantageous microbes from

parents to offspring. A similar approach of releasing microbial loads may also be carried out by the maerl during sexual reproduction. Although there is evidence of sexual reproduction, it is widely believed the primary method of reproduction is asexually via fragmentation, where clone maerl individuals develop from broken fragments of existing individuals (Martin and Hall-Spencer, 2017). With this approach in mind, core microbial community members from the original individual is assumed to still be attached to the fragment, thus the microbiota could develop alongside the growing clone individual. The carrying over of the microbiota in this fashionhas been document in corals species (Thompson *et al.*, 2015) and so is highly likely to also occur with maerl fragments. Regardless of reproductiontype, the transmission of core advantageous microbial groups from parentsis high suggestive of improving new plant success. The transmission of the abundant bacterial family Sphingomonadaceae may aid in the recruitment of other probiotic microbes by providing the initial colonisation and construction of a biofilm which allows the harbouring of a diverse range of species ( de Vries *et al.*, 2019). Additionally, several species have been found to degrade a range of xenobiotic compounds derived from anthropogenic and natural origins, producing bioavailable metabolities which could be utilised as nutrients for either the growth of other microbesor promoting plant growth (Glaeser and Kämpfer, 2014).

Alongside varying abilities to promote growth, certain members of the microbiota are believed to be involved in the inhibition of pathogenic species (Pringgenies *et al.*, 2020) . Inhibition of certain microbes may be due to a host to bacterial relationship or may be the result of microbial competition derived

from securing an area on the nutrient rich environment that the host provides. This inhibition of pathogenic growth isbelieved to occur in three main ways. The first method is by controlling anddisrupting Quorum signalling (QS) signals between epibacterial individuals, the host and microbiota by inducing signalling compounds, such as N-acyl homoserine lactones (AHL). Rhodobacteraceae strains originating from

*Fucus spiralis* were found to emit high quantities of indole which is known to infer with QS regulator folding (Kim and Park, 2013). Furthermore, their own AHLs were composed of 10-18 carbon atoms which has a greater resistance to hydrolysis whilst in an alkaline pH, which correlates with the pH found on the thallus of *F. spiralis* (Dogs *et al.*, 2017). The production of compounds which inhibit QS whilst having QS compounds adapted to match the hosts conditions strongly suggests the controlling of epibacterialactivities. A similar situation maybe occurring within the surface layers of maerl due to the high abundances of Rhodobacteraceae species witnessed.The second method of controlling epibacterial communities is via the production of bioactive compounds which are detrimental for microbial integrity (Singh, Kumari and Reddy, 2015).The production of antimicrobial compounds have been documented to be produced by a range of bacterial clades and is not restricted to “elite” species, the range of classes include: Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Flavobacteria, Actinobacteria and Bacill (Hollants *et al.*, 2013). Bacterial antimicrobial metabolites produced may include molecules of Pyrrol whichare known to be produced by *Bacillus* spp. originating from *Dysidea fragilis* (Mohan, Thangappanpillai and Ramasamy, 2016). The last method of

prevention of pathogenic growth is by occupying all suitable surfaces and producing a non-attachable biofilm. This incompatibility of attachment to the biofilm prevents settlement and growth of other microbial groups before a foothold can be gained (Gu, 2018). *Sphingomonadaceae* spp. have previously been highlighted to exhibit the ability to be the initial colonisers of a surface and to construct a biofilm. However, this early and rapid biofilm development has the potential to prevent other incompatible microbial groups from settling. Whilst it is currently unclear whether maerl individuals produce chemical cues to mitigate specific unwanted microbial groups, it is highly suggestable that the associated microbiota witnessed with this study can play a role in composition of the holobiont and as a bi product prevent host infections and diseases. While the microbiota has been demonstrated to potentially provide a range of benefits to the host, it may further be responsible for enhancing the settlement of other macro-species associated with maerl beds. For example, larval Pacific calico scallops, *Argopecten ventricosus,* were seen tobe 30 to 35 times more likely to settle on living maerl vs non-coralline sedimentary substrates. Additionally, lavae were also 3.3 times more likely to settle on live vs dead maerl nodules (Steller, D. L. and Cáceres-Martínez, 2009). One of the suggestions for this enhanced settlement was strong chemical cues produced by the holobiont attracting and inducing the observed settlement. In addition, two subspecies of Abalone - *Haliotis discus discus* and *Haliotis discus hannai* - larvae have also seen similar induced metamorphosis when in the presence of maerl (Suenaga *et al.*, 2004). G-aminobutyric acid (GABA) 1 and thyroid hormones two were found to be responsible for inducing the metamorphosis, however, these chemical compounds are not associated with maerl. GABA concentrations ($10^{-4}$

M) have also been documented to induce the settlement and metamorphosis (S&M) of four other bi-value species *(Mytilus galloprovincialis, Venerupis pullastra , Ruditapes philippinarum* and *Ostrea edulis),* with lower concentrations of GABA also causing S&M but to a lower effectiveness (García-Lavandeira *et al.*, 2005). Whilst GABA and similar chemicals may have been produced by other macro-species in the surrounding environment, natural biofilms on glass slips have been documented to induce S&M in *Mytilus galloprovincialis* (Bao *et al.*, 2007). Furthermore, significant increases in S&M were noted with biofilm age andhigher dry weight, heavily suggesting microbial activities were responsible. Whilst GABA and other similar chemical production has been observed within bacterial and fungal species (MASUDA *et al.*, 2008; Dhakal, Bajpai and Baek, 2012), it can only be suggested that maerl microbiota members can also perform similar roles. Further investigations into production of chemical cues and S&M from microbial species revealed within the maerl samples of this study is required. These additional studies, alongside the microbial community evidence provided in this study, would highly likely signify the underrepresented importance which maerl microbes play in supporting the wider ecosystem and shellfish fisheries and as such highlights the vital importance to conserve maerl beds for ecological andeconomic reasons.

It is not unsurprising to witness such a diverse and rich bacterial relative abundance associated with the surface of maerl. As seaweed surfaces are known to provide a sheltered nutritional 'hot spot' for microbial opportunist that thrive wherever organic material is available (Armstrong *et al.*, 2001). With

algal exudates found to contain enriched concentrations of dissolved combined neutral sugars (DCNS), such as fucose discovered in Ochrophyta exudates and galactose revealed in Rhodophyta exudates (Nelson *et al.*, 2013). These exudates were witnessed to enhance bacterial growth with results showing significant higher growth compared to coral exudates and seawater treatments. Indicting the nutrient loads and growing potential algal exudates have on microbial groups. However, it appears that there are underlaying mechanisms at work which dictate which microbial familiescan thrive in these rich havens and which groups are removed. What appears to be a bias towards certain microbial familieslocated on thesurface of maerl, may in fact be a bias towards functional gene groups rather than species or taxonomic groups. *Ulva australis* were demonstratedto individually harbour a unique associated microbiota (Burke *et al.*, 2011), whilst these associated microbes were different in taxonomic groups there were similarities in the functional gene guilds, suggesting host to functionalgene bias. Although this does not appear to be the case with the surface of maerl, there may still be difference between genus and/or species as well as local regional differences. This was outside the scope of this study and requires further investigations into the functional gene guilds of the microbiota of multiple different maerl beds. QS sensing chemical such AHL have previously been discussed, although it should be stressed that different microbial groups possess the ability to create sensing chemical which are better suited to the hosts conditions such as pH. These adaptionsand conditions will favour those microbial groups which are better adapted to the hosts environment compared to those that are not, potentially causing the bais in the microbial groups that are present. Lastly, macro- algae are known to

possess concoctions of biochemical compounds which will severely impact the epimicrobial composition (Lachnit, Wahl and Harder, 2010; Hollants *et al.*, 2013a). These bioactive compounds are designed to help alleviate unwanted pathogens, grazers, and biofouling which cause deleterious impacts on the macroalgae health and growth rates. One such release trigger of these bioactive compounds is believed tobe between algal cell wall components and secondary metabolites sensing,instead of the sensing of the microbe presence (Engel, Jensen and Fenical, 2002)

### *4.4 - Group 3: Live and dead maerl total communities (LMT and DMT) with dead maerl surface communities (DMEX)*

Group 3 is assembled from a combination of LMT, DMT and DMEX sample types, indicating very few differences in the microbial communities associated between them. However, there are noticeable differences between the other sample types which highlights the unique but shared microbiota that is associated within this group. This shared microbiota between living maerl total and the dead maerl samples demonstrates that the microbiota found within living maerl nodules remains after host death,with the only noteworthy microbial shift witnessed in the living surface layers. However, the duration that the microbiota remains after maerl death is not investigated within this study as the age of the dead maerl nodules are unknown. The microbial shift from LMEX to group 3 could be an interesting further investigation where the results of this research may be used to indicate the health and condition of maerl, with potential results highlighting when maerl beds are stressed or

declining in health. Returning to this study, the similarities shared within group 3 highlights theimportance the habitats created by maerl skeletons even after death are for microbial ecology, which is highly likely playing a wider role in the surrounding ecosystem. The main similarities of highest abundant families within this sample group are: Nitrosopumilaceae, Sandaracinaceae, Woeseiaceae, Flavobacteriaceae and Pirellulaceae. However, whilst it does appear that the Pirellulaceae species are shared amongst all four maerl samples (LMT, DMT, LMEX and DMEX), several Flavobacteriaceae species are not. Suggesting that Flavobacteriaceae plays a vital role in the maerl holobioant with certain species being better suited to the different regions of maerl. This difference in species location could be due to differences in conditions and/or nutrients these regions receive. Whilst the potential differences of pH, oxygen concentrations and light intensities may be having an impact, the primary cause of species difference is considered to be due to the absence of algal exudates being produced. LMT and DMEX represent a greater similarity of microbiota compared to DMT, highlighting that it is not the interior or exterior conditions of the maerl skeleton causing this difference but the fact that the maerl is no longer alive creating extudates. After maerl death the microbiota of the living interior are perceived to take over the surface layers of the dead nodule as nutrientloads are no longer suitable to support the microbial species that once thrived on the living surface. The dead maerl interior microbial species thenmay start to decline as nutrient loads start to wane and opportunistic microbes move in, this is represented by the varied sample replications witnessed within the beta diversity analysis. Whilst moraxellaceae the mostabundance family found in DMT is also abundant in the surrounding

sediment, the other dominant families do not match the sediments and so it is believed that the dead interior does not shift towards a microbiota reflecting that of the surrounding sediments. Instead, the appearance of a new family– Nostocaceae – and the reorganisation of existing families potentially indicates the creation of a new microbiota. The creation of this new microbiota could provide vital ecosystem roles and as such dead maerl beds need to be considered as important as living beds for microbial ecology. Whilst this needs further investigations, the preliminary results of this study of maerl microbiota suggests the potential emergence of a new microbiota dissimilar to that ofthe living maerl beds and the surrounding environment. In addition to the roles that have been previously discussed and suggestedduring section 4.3, it is believed many of the microbes present in group 3 are performing nutrient cycling. Either by providing for the maerl host - whilst alive - or as part of a microbial loop returning nutrients back to the environment in the form of metabolites. Sandaracinaceae play vital roles in the degrading and recycling of carbon, nitrogen and phosphates by utilising low-molecular-weight organic compounds such as ethanol, hydrogen, butyrate, and acetate (Probandt *et al.*, 2017). Considering both dead and live maerl have been documented to contain high abundances oflow molecular weight carbohydrates mostly derivatives of trimethylsilyl- ether (O'Reilly *et al.*, 2012), potentially explains the high dominant abundance of Sandaracinaceae. On the other hand, Nitrosopumilaceae are known oxidisers of ammonia into nitrite (Könneke *et al.*, 2005) and so could be playing a similar double role as Pirellulaceae performed in the live surface layers of maerl. The last shared high abundant families, Woeseiaceae, are known sulphur, nitrate and nitrite reducers with the ability

to produce high concentrations of hydrogen creating anoxic conditions (Kessler *et al.*, 2019). The creation of new conditions could allow the flourishing of anaerobic microbes in a symbiotic consortium. Whilst Woeseiaceae have been documented to create anoxic conditions in other environments, it has not been found to produce similar conditions within maerl beds and as such will need to be further investigated.

Sar116_clade was highly dominant within LMT, however, did not appear abundant in either DMT or DMEX. Members of this family are known to reduce nitrate and sulfate and uptake inorganic phosphates (Oh *et al.*, 2010), potentially recycling vital nutrients for its living host. The sudden disappearance of sar116 once the maerl dies strongly suggests their presence is linked to the life status of the maerl and may indicate the potential close association between species of sar116 the maerl interior microbiota. On the other hand, once the maerl deceases the family Nostocaceae appears. Thisappearance of this group could be due to a "gap" left behind in the microbiota by key microbial groups, such as nitrogen fixing as Nostocaceae are known nitrogen fixers (Rush and Sinninghe Damsté, 2017). Whilst this could be true, alternatively Nostocaceae could opportunistically settle on the skeletons due to the absence of bioactive compounds produced by the maerl and its associated microbiota. However, the addition of this family has the potential to inflict wide ramifications on the micro and macro inhabitants of the surrounding area due to the assorted toxins this family are known to produce (Řezanka and Dembitsky, 2006).

*4.5 – Group 4: Seawater communities*

The separation of Group 4 from the other groups highlights another distinctive microbial community from the other communities observed in the other sample types. This group was solely composed of both dead and live seawater samples, demonstrating the surrounding water column comprised similar microbiota regardless of the condition of the maerl bed below. These differences were notable in the beta diversity as well as the alpha diversity, with both dead and live seawater samples supporting the lowest H values of all samples (4.3 and 4.4). This observation was unsurprising considering other studies have also witnessed similar alpha diversity levels (Barott *et al.*, 2011)**,** with results believed to be due to pooravailability of vital resources required in supporting a high diversity (Traving *et al.*, 2017) **.** The families: Methylophilaceae, cyanbiaceae, ectothiorhodos piraceae, clade II and Clade I were found primarily within group 4 and were not witnessed in similar abundances in the other groups. These families disappearance within other sample types suggests thatthey are not suitable within these environments or were being repelled by biologicals mechanisms. Further research is required investigating the factors causing these families to disappear whilst other families appear to be enriched. This could be carried out within microcosms examining which factors increases abundances of these family members and which factors hamper their abundance. However, there were families that were dominant within the seawater which also appeared to be enriched within all maerl samples, these are: Sar116_clade, rhodobacteraceae and flavobacteriaceae. The presence of these families within all maerl samples but

not sediment could potentially suggest a transportation link and indicate how some of these enriched families within the maerl were introduced and were suitable for living within the habitat created by the maerl. However, this is outside the constraints of this study and further investigations will need to be carried out to determine the source of the microbial communities witnessed living on the maerl. Interesting Halieaceae was found in similar abundances (3-4%) in all sample types except from live maerl exterior, highlighting the potentially widespread andadaptability of this family whilst illuminating the carefully associated community found on the live maerl exterior.

*4.6 - Further research*

Throughout this discussion there have been indications of potential areas for further research, this section will detail further vital investigations and why I believe in their importance. Whilst this study highlighted the microbiota differences between the live maerl exterior and dead maerl exterior, further investigations examining transition time, levels of stabilityand drivers enforcing the community shift are required. This exploration

could also be paired with recording different stressors and their effects on the microbiota, which will develop our understanding of the effects varyingstressors may of on the holobiont leading deleterious impacts on the maerlhost. Furthermore, exploring these shifts could aid in monitoring maerl bedcommunity health during emergency actions such as oil spills or aid in further accessing ecosystem health during assessments. Another possible area which could be further investigated is the microbiota across different species of maerl in the uk or the same species of maerl in different parts of the world (local and global). This will broaden our knowledge of the microbiota associated with a key ecosystem engineer. Furthering this knowledge could aid in conserving this biodiversity hotspot which in turn could improve local economies via increasing productivity in the fishing industry and/or increase tourism within the region. Lastly, many sequenceswere unable to be matched with known genomes from a reference database, representing potential undiscovered microbes. The highest

counts were witnessed within the sediment samples, however there were

key abundant species within samples such as live maerl exterior that couldbe key holobiant species which are currently unknown. Exploring these unknown species could lead to the discovery of new microbial species which in turn has the potential for further studies to carried investigating functions and specific attributes. These investigates have the potential to discover new technologies or products which can beneficial in numerous industries such medical or food.

*4.7 – Conclusion*

The surrounding environment samples – sediments and seawater - were found to contain different microbial communities from one another. These environment communities were different from the communities observed on the whole live and dead maerl nodules as well as the surface layers of the dead maerl nodule. However, the surface layers of the living maerl contain a unique communities composition that was different compared all

other samples. These two different communities found within the maerl bed suggests a strong maerl-microbe association. Whilst factors of determining why this was observed is outside of the study, it is highly probable that maerl and/or its derivatives play an important role in supporting the witnessed microbial communities. The presences of these communities further suggests the important roles which microbes may perform, not just for the maerl host but also providing ecological and economic benefits to the region. For these reasons, it is the belief of the author that maerl beds are critically important and should be conserved with the upmost management.

Bibliography

Abu-Bakar, A., Ahmadian, R. and Falconer, R. A. (2017) 'Modelling the transport and decay processes of microbial tracers in a macro-tidal estuary', *Water Research*. Elsevier Ltd, 123, pp. 802–824. doi: 10.1016/j.watres.2017.07.007.

Aires, T. *et al.* (2019) 'Seaweed loads cause stronger bacterial community shifts in coastal lagoon sediments than nutrient loads', *Frontiers in Microbiology*. Frontiers Media S.A., 10(JAN), p. 3283. doi: 10.3389/fmicb.2018.03283.

Allen, C. *et al.* (2014) *Fal and Helford SAC Maerl Drop-down Video and Dive Survey 2013*.

Armstrong, E. *et al.* (2001) 'The symbiotic role of marine microbes on living surfaces', in *Hydrobiologia*. Springer, pp. 37–40. doi: 10.1023/A:1012756913566.

Baker, B. J. *et al.* (2015) 'Genomic resolution of linkages in carbon, nitrogen, and sulfur cycling among widespread estuary sediment bacteria', *Microbiome*. Springer Nature, 3(1). doi: 10.1186/s40168-015-0077-6.

Bao, W. Y. *et al.* (2007) 'Larval settlement and metamorphosis of the mussel *Mytilus galloprovincialis* in response to biofilms', *Marine Biology*. Springer, 150(4), pp. 565–574. doi:10.1007/s00227-006-0383-4.

Barott, K. L. *et al.* (2011) 'Microbial diversity associated with four functional groups of benthic reef algae and the reef-building coral *Montastraea annularis'*, *Environmental Microbiology*.Blackwell Publishing Ltd, 13(5), pp. 1192–1204. doi: 10.1111/j.1462-2920.2010.02419.x.

Bischoff, B. (1982) 'Effects of cadmium on microorganisms', *Ecotoxicology and Environmental Safety*. Academic Press, 6(2), pp. 157–165. doi: 10.1016/0147-6513(82)90003-3.

Blake, C. and Maggs, C. A. (2003) 'Comparative growth rates and internal banding periodicity of maerl species (*Corallinales, Rhodophyta*) from northern Europe', *Phycologia*. International Phycological Society, 42(6), pp. 606–612. doi: 10.2216/i0031-8884-42-6-606.1.

Bosence, D. and Wilson, J. (2003) 'Maerl growth, carbonate production rates and accumulation rates in the ne atlantic', *Aquatic Conservation: Marine and Freshwater Ecosystems*. John Wiley & Sons, Ltd, 13(S1), pp. S21–S31. doi: 10.1002/aqc.565.

Boström, K. H. *et al.* (2004) 'Optimization of DNA extraction for quantitative marine bacterioplankton community analysis', *Limnology and Oceanography: Methods*, 2(11), pp. 365–373. doi: 10.4319/lom.2004.2.365.

Bowman, J. P. (2006) 'The Marine Clade of the Family Flavobacteriaceae: The Genera *Aequorivita, Arenibacter, Cellulophaga, Croceibacter, Formosa, Gelidibacter, Gillisia, Maribacter, Mesonia, Muricauda, Polaribacter, Psychroflexus, Psychroserpens, Robiginitalea, Salegentibacter, Tenacibaculum, Ulvibacter, Vitellibacter and Zobellia*', in *The Prokaryotes*. New York, NY: Springer New York, pp. 677–694. doi: 10.1007/0-387-30747-8_26.

Bryan, G. W. (1985) '*A Guide to the assessment of heavy-metal contamination in estuaries usingbiological indicators'*. Marine Biological Association of the U.K.

Bryan, G. W. *et al.* (1987) 'Copper, zinc, and organotin as long-term factors governing the distribution of organisms in the fal estuary in Southwest England', *Estuaries*. Springer, 10(3), pp. 208–219. doi: 10.2307/1351849.

Burke, C. *et al.* (2011) 'Bacterial community assembly based on functional genes rather than species', *Proceedings of the National Academy of Sciences of the United States of America*.

National Academy of Sciences, 108(34), pp. 14288–14293. doi: 10.1073/pnas.1101591108.

Busetti, A., Maggs, C. A. and Gilmore, B. F. (2017) 'Marine macroalgae and their associated microbiomes as a source of antimicrobial chemical diversity', *European Journal of Phycology*. Taylor and Francis Ltd., 52(4), pp. 452–465. doi: 10.1080/09670262.2017.1376709.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016). "DADA2: High-resolution sample inference from Illumina amplicon data." Nature Methods, 13, 581-583. doi: 10.1038/nmeth

Carvalho, V. F. *et al.* (2020) 'Environmental drivers of rhodolith beds and epiphytes community along the South Western Atlantic coast', *Marine Environmental Research*. Elsevier Ltd, 154, p. 104827. doi: 10.1016/j.marenvres.2019.104827.

Castro-Alonso, M. J. *et al.* (2019) 'Microbially induced calcium carbonate precipitation (MICP) and its potential in bioconcrete: Microbiological and molecular concepts', *Frontiers in Materials*. Frontiers Media S.A., p. 126. doi: 10.3389/fmats.2019.00126.

Cavalcanti, G. S. *et al.* (2014) 'Physiologic and metagenomic attributes of the rhodoliths forming the largest CaCO3 bed in the South Atlantic Ocean', *The ISME Journal*. Nature Publishing Group, 8(1), pp. 52–62. doi: 10.1038/ismej.2013.133.

Cavalcanti, G. S. *et al.* (2018) 'Rhodoliths holobionts in a changing ocean: Host-microbes interactions mediate coralline algae resilience under ocean acidification', *BMC Genomics*. BioMed Central Ltd., 19(1). doi: 10.1186/s12864-018-5064-4.

Ceh, J. *et al.* (2012) 'Coral-bacterial communities before and after a coral mass spawning event on Ningaloo Reef', *PLoS ONE*. Public Library of Science, 7(5). doi: 10.1371/journal.pone.0036920.

Chen, J. *et al.* (2019) 'A pollution gradient contributes to the taxonomic, functional, and resistome diversity of microbial communities in marine sediments', *Microbiome*. BioMed Central Ltd., 7(1), p. 104. doi: 10.1186/s40168-019-0714-6.

*Conservation of Habitats and Species Regulations 2010 - Summary* (2010). Available at: http://adlib.everysite.co.uk/adlib/defra/content.aspx?id=2RRVTHNXTS.8CYX96C65K5BK (Accessed: 18 March 2021).

Dhakal, R., Bajpai, V. K. and Baek, K. H. (2012) 'Production of GABA (γ-aminobutyric acid) by microorganisms: A review', *Brazilian Journal of Microbiology*. SBM, 43(4), pp. 1230–1241. doi: 10.1590/S1517-83822012000400001.

Dogs, M. *et al.* (2017) 'Rhodobacteraceae on the marine brown alga *Fucus spiralis* are abundant and show physiological adaptation to an epiphytic lifestyle', *Systematic and Applied Microbiology*. Elsevier GmbH, 40(6), pp. 370–382. doi: 10.1016/j.syapm.2017.05.006.

Edet, U. O. *et al.* (2017) 'An Overview of Cultural, Molecular and Metagenomic Techniques in Description of Microbial Diversity', *Journal of Advances in Microbiology*, pp. 1–19.

Egan, S. *et al.* (2000) 'Phylogenetic relationship and antifouling activity of bacterial epiphytes from the marine alga *Ulva lactuca'*, *Environmental Microbiology*. Environ Microbiol, 2(3), pp. 343–347. doi: 10.1046/j.1462-2920.2000.00107.x.

Egan, S. *et al.* (2013) 'The seaweed holobiont: understanding seaweed–bacteria interactions', *FEMS Microbiology Reviews*, 37(3), pp. 462–476. doi: 10.1111/1574-6976.12011.

Engel, S., Jensen, P. R. and Fenical, W. (2002) 'Chemical ecology of marine microbial defense', *Journal of Chemical Ecology*. Springer, pp. 1971–1985. doi: 10.1023/A:1020793726898.

Erşan, Y. Ç., de Belie, N. and Boon, N. (2015) 'Microbially induced CaCO3 precipitation through

denitrification: An optimization study in minimal nutrient environment', *Biochemical Engineering Journal*. Elsevier B.V., 101, pp. 108–118. doi: 10.1016/j.bej.2015.05.006.

Feng, L. *et al.* (2020) 'Investigating the composition and distribution of microplastics surface biofilms in coral areas', *Chemosphere*. Elsevier Ltd, 252, p. 126565. doi: 10.1016/j.chemosphere.2020.126565.

Foote, A. D. *et al.* (2012) 'Investigating the Potential Use of Environmental DNA (eDNA) for Genetic Monitoring of Marine Mammals', *PLoS ONE*. Edited by S. Lin. Public Library of Science, 7(8), p. e41781. doi: 10.1371/journal.pone.0041781.

Foster, M. S. *et al.* (2013) 'Rhodoliths and Rhodolith Beds', *Smithsonian contributions to the marine sciences*, 39, pp. 143–155. Available at: https://core.ac.uk/download/pdf/16462564.pdf (Accessed: 13 January 2018).

Fries, L. (1975) 'Short Communication Some Observations on the Morphology of *Enteromorpha linza* (L.) J. Ag. and Enteromorpha compressa (L.) Grev. in Axenic Culture', *Botanica Marina*, 18(4), pp. 251–253. doi: 10.1515/botm.1975.18.4.251.

García-Lavandeira, M. *et al.* (2005) 'Effects of GABA and epinephrine on the settlement and metamorphosis of the larvae of four species of bivalve molluscs', *Journal of Experimental Marine Biology and Ecology*. Elsevier, 316(2), pp. 149–156. doi: 10.1016/j.jembe.2004.10.011.

Gikas, P. *et al.* (2009) '*The effects of heavy metals and temperature on microbial growth and lag*' Environ Toxicol Chem. doi: 10.1897/08-273.1

Glaeser, S. P. and Kämpfer, P. (2014) 'The family Sphingomonadaceae', in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*. Springer-Verlag Berlin Heidelberg, pp. 641–707. doi: 10.1007/978-3-642-30197-1_302.

Goecke, F. *et al.* (2010) 'Review chemical interactions between Marine macroalgae and bacteria', *Marine Ecology Progress Series*, pp. 267–300. doi: 10.3354/meps08607.

Gray, M. W., Lang, B. F. and Burger, G. (2004) 'Mitochondria of Protists', *Annual Review of Genetics*. Annual Reviews, 38(1), pp. 477–524. doi: 10.1146/annurev.genet.37.110801.142526.

Gu, J. D. (2018) 'Microbial biofilms, fouling, corrosion, and biodeterioration of materials', in *Handbook of Environmental Degradation Of Materials: Third Edition*. Elsevier Inc., pp. 273–298. doi: 10.1016/B978-0-323-52472-8.00014-9.

Gubelit, Y. *et al.* (2016) 'Nutrient and metal pollution of the eastern Gulf of Finland coastline: Sediments, macroalgae, microbiota', *Science of the Total Environment*. Elsevier, 550, pp. 806–819. doi: 10.1016/j.scitotenv.2016.01.122.

Hinger, I., Pelikan, C. and Mußmann, M. (2019) 'Role of the ubiquitous bacterial family Woeseiaceae for N2O production in marine sediments'. Geophysical Research Abstracts

Hollants, J. *et al.* (2013a) 'What we can learn from sushi: a review on seaweed-bacterial associations', *FEMS Microbiology Ecology*, 83(1), pp. 1–16. doi: 10.1111/j.1574-6941.2012.01446.x.

Howarth, R. W. (2008) 'Coastal nitrogen pollution: A review of sources and trends globally and regionally', *Harmful Algae*. Elsevier, pp. 14–20. doi: 10.1016/j.hal.2008.08.015.

Jickells, T. D. *et al.* (2014) 'Nutrient transport through estuaries: The importance of the estuarine geography', *Estuarine, Coastal and Shelf Science*. Academic Press, 150(PB), pp. 215–229. doi: 10.1016/j.ecss.2014.03.014.

Kellogg, C. A. (2019) 'Microbiomes of stony and soft deep-sea corals share rare core bacteria',

*Microbiome*. BioMed Central Ltd., 7(1), pp. 1–13. doi: 10.1186/s40168-019-06973

Kemp, M. W. and Boynton, W. R. (1984) 'Spatial and Temporal Coupling of Nutrient Inputs to Estuarine Pri: Ingenta Connect', *Marine Science*, 35(3), pp. 522–535. Available at: https://www.ingentaconnect.com/content/umrsmas/bullmar/1984/00000035/00000003/art00023 (Accessed: 21 May 2020).

Kennish, M. J. (2002) 'Environmental threats and environmental future of estuaries', *Environmental Conservation*, 29(1), pp. 78–107. doi: 10.1017/S0376892902000061.

Kershaw, S. and Campos, C. (2010) '*S*anitary survey report lower Fal estuary Queen Scallops (*A. opercularis*) at East Bank and overall review of production areas in the lower Fal Estuary and Percuil River, lower Fal estuary Queen scallops *(A. opercularis)*at east bank and overall review of production areas in the lower fal estuary and percuil river'. CEFAS

kessler, a. j. *et al.* (2019) 'bacterial fermentation and respiration processes are uncoupled in anoxic permeable sediments', *Nature Microbiology*. Nature Publishing Group, 4(6), pp. 1014–1023. doi: 10.1038/s41564-019-0391-z.

Kim, J. and Park, W. (2013) 'Indole inhibits bacterial quorum sensing signal transmission by interfering with quorum sensing regulator folding', *Microbiology (United Kingdom)*. Microbiology Society, 159(part 12), pp. 2616–2625. doi: 10.1099/mic.0.070615-0.

Könneke, M. *et al.* (2005) 'Isolation of an autotrophic ammonia-oxidizing marine archaeon', *Nature*. Nature Publishing Group, 437(7058), pp. 543–546. doi: 10.1038/nature03911.

Kramarsky-Winter, E. *et al.* (2006) 'Identification of a protist-coral association and its possible ecological role', *Marine Ecology Progress Series*, 317, pp. 67–73. doi: 10.3354/meps317067.

Krediet, C. J. *et al.* (2013) 'Coral-associated micro-organisms and their roles in promoting coral health and thwarting diseases', *Proceedings of the Royal Society B: Biological Sciences*. Royal Society. doi: 10.1098/rspb.2012.2328.

Lachnit, T., Wahl, M. and Harder, T. (2010) 'Isolated thallus-associated compounds from the macroalga *Fucus vesiculosus* mediate bacterial surface colonization in the field similar to that on the natural alga.', *Biofouling*. Biofouling, 26(3), pp. 247–255. doi: 10.1080/08927010903474189.

Lam, C., Stang, A. and Harder, T. (2008) 'Planktonic bacteria and fungi are selectively eliminated by exposure to marine macroalgae in close proximity', *fems microbiology ecology*. femS Microbiol Ecol, 63(3), pp. 283–291. doi: 10.1111/j.1574-6941.2007.00426.x.

Leite, D. C. A. *et al.* (2017) 'Broadcast Spawning Coral *Mussismilia hispida*: Can Vertically Transfer its Associated Bacterial Core', *Frontiers in Microbiology*. Frontiers, 8, p. 176. doi: 10.3389/fmicb.2017.00176.

Litchman, E. (2010) 'Invisible invaders: non-pathogenic invasive microbes in aquatic and terrestrial ecosystems', *Ecology Letters*. Blackwell Publishing Ltd, 13(12), pp. 1560–1572. doi: 10.1111/j.1461-0248.2010.01544.x.

Malkin, S. Y. *et al.* (2014) 'Natural occurrence of microbial sulphur oxidation by long-range electron transport in the seafloor', *The ISME Journal*, 8, pp. 1843–1854. doi: 10.1038/ismej.2014.41.

Marshall, K. *et al.* (2006) 'Effect of marine bacterial isolates on the growth and morphology of axenic plantlets of the green alga *Ulva linza*', *Microbial Ecology*. Springer, 52(2), pp. 302–310. doi: 10.1007/s00248-006-9060-x.

Martin, S. and Hall-Spencer, J. M. (2017) 'Effects of ocean warming and acidification on

rhodolith/maërl beds', in *Coastal Research Library*. Springer, pp. 55–85. doi: 10.1007/978-3-319-29315-8_3.

MASUDA, K. *et al.* (2008) 'Isolation of Marine Yeasts Collected from the Pacific Ocean Showing a High Production of γ-Aminobutyric Acid', *Bioscience, Biotechnology, and Biochemistry*. Japan Society for Bioscience, Biotechnology, and Agrochemistry, 72(12), pp. 3265–3272. doi: 10.1271/bbb.80544.

Matsuo, Y. *et al.* (2003) 'Isolation and phylogenetic characterization of bacteria capable of inducing differentiation in the green alga *Monostroma oxyspermum'*, *Environmental Microbiology*. Blackwell Science Ltd, 5(1), pp. 25–35. doi: 10.1046/j.1462-2920.2003.00382.x.

McMurdie and Holmes (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE. 8(4):e61217

Melbourne, L. A. *et al.* (2017) 'There is more to maerl than meets the eye: DNA barcoding reveals a new species in Britain, *Lithothamnion erinaceum* sp. nov. (*Hapalidiales, Rhodophyta*)', *European Journal of Phycology*. Taylor & Francis, 52(2), pp. 166–178. doi: 10.1080/09670262.2016.1269953.

Mohamed, N. M. *et al.* (2010) 'Diversity of aerobic and anaerobic ammonia-oxidizing bacteria in marine sponges', *ISME Journal*. Nature Publishing Group, 4(1), pp. 38–48. doi: 10.1038/ismej.2009.84.

Mohan, G., Thangappanpillai, A. K. and Ramasamy, B. (2016) 'Antimicrobial activities of secondary metabolites and phylogenetic study of sponge endosymbiotic bacteria, Bacillus sp. at Agatti Island, Lakshadweep Archipelago', *Biotechnology Reports*. Elsevier B.V., 11, pp. 44–52. doi: 10.1016/j.btre.2016.06.001.

Morse, A. n. c. and Morse, D. E. (1984) 'Recruitment and metamorphosis of *Haliotis* larvae induced by molecules uniquely available at the surfaces of crustose red algae', *Journal of Experimental Marine Biology and Ecology*. Elsevier, 75(3), pp. 191–215. doi: 10.1016/0022-0981(84)90166-7.

Nelson, C. E. *et al.* (2013) 'Coral and macroalgal exudates vary in neutral sugar composition and differentially enrich reef bacterioplankton lineages', *ISME Journal*, 7(5), pp. 962–979. doi: 10.1038/ismej.2012.161.

Nicolaus, E. E. M. and Barry, J. (2015) 'Imposex in the dogwhelk (Nucella lapillus): 22-year monitoring around England and Wales', *Environmental Monitoring and Assessment*. Springer International Publishing, 187(12), pp. 1–14. doi: 10.1007/s10661-015-4961-0.

Nielsen, S. L., Banta, G. T. and Pedersen, M. F. (2007) 'Decomposition Of Marine Primary Producers: Consequences For Nutrient Recycling And Retention In Coastal Ecosystems', in *Estuarine Nutrient Cycling: The Influence of Primary Producers*. Springer Netherlands, pp. 187–216. doi: 10.1007/978-1-4020-3021-5_7.

Nies, D. H. (1999) 'Microbial heavy-metal resistance', *Applied Microbiology and Biotechnology*. Springer, pp. 730–750. doi: 10.1007/s002530051457.

O'Reilly, S. *et al.* (2012) 'Chemical and physical features of living and non-living maerl rhodoliths', *Aquatic Biology*, 15(3), pp. 215–224. doi: 10.3354/ab00431.

Oh, H.-M. *et al.* (2010) 'genome announcement Complete Genome Sequence of "*Candidatus Puniceispirillum marinum*"IMCC1322, a Representative of the SAR116 Clade in the Alphaproteobacteria', *journal of bacteriology*, 192(12), pp. 3240–3241. doi: 10.1128/JB.00347-10.

Paerl, H. W. *et al.* (2006) 'Anthropogenic and climatic influences on the eutrophication of large estuarine ecosystems'. *Limnol. Oceanogr., 51(1, part 2),448–462*, *Limnol. Oceanogr*.

Peixoto, R. S. *et al.* (2017) 'Beneficial microorganisms for corals (BMC): Proposed mechanisms

for coral health and resilience', *Frontiers in Microbiology*. Frontiers Media S.A., p. 341. doi: 10.3389/fmicb.2017.00341.

Peña, V. *et al.* (2014a) ' Detection of Gametophytes in the Maerl-Forming Species *Phymatolithon calcareum* (*Melobesioideae, Corallinales*) Assessed by DNA Barcoding ', *Cryptogamie, Algologie*. Museum National d'Histoire Naturelle, Paris, France, 35(1), pp. 15–25. doi: 10.7872/crya.v35.iss1.2014.15.

Pfeffer, C. *et al.* (2012) 'Filamentous bacteria transport electrons over centimetre distances', *Nature*, 491(7423), pp. 218–221. doi: 10.1038/nature11586.

Pirrie, D. *et al.* (2003) 'The spatial distribution and source of arsenic, copper, tin and zinc within the surface sediments of the Fal Estuary, Cornwall, UK', *Sedimentology*, 50, pp. 579–595.

Prabhakaran, P., Ashraf, M. A. and Aqma, W. S. (2016) 'Microbial stress response to heavy metals in the environment', *RSC Advances*. Royal Society of Chemistry, pp. 109862–109877. doi: 10.1039/c6ra10966g.

Pringgenies, D. *et al.* (2020) 'Symbiotic microbes from various seaweeds with antimicrobial andfermentative properties'. Available at: http://www.bioflux.com.ro/aacl (Accessed: 7 March 2021).

Probandt, D. *et al.* (2017) 'Permeability shapes bacterial communities in sublittoral surface sediments', *Environmental Microbiology*. Blackwell Publishing Ltd, 19(4), pp. 1584–1599. doi: 10.1111/1462-2920.13676.

Qui-Minet, Z. N. *et al.* (2018) 'The role of local environmental changes on maerl and its associated non-calcareous epiphytic flora in the Bay of Brest', *Estuarine, Coastal and Shelf Science*. Academic Press, 208, pp. 140–152. doi: 10.1016/j.ecss.2018.04.032.

Rabalais, N. N. *et al.* (2009) 'Global change and eutrophication of coastal waters', *ICES Journal of Marine Science*. Oxford Academic, 66(7), pp. 1528–1537. doi: 10.1093/icesjms/fsp047.

Rendina, F. *et al.* (2020) 'Distribution and Characterization of Deep Rhodolith Beds off the Campania coast (SW Italy, Mediterranean Sea)', *Plants*. MDPI AG, 9(8), p. 985. doi: 10.3390/plants9080985.

Renshaw, M. A. *et al.* (2014) 'The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol-chloroform-isoamyl alcohol DNA extraction'. Mol Ecol Resour. doi:10.1111/1755-0998.12281.

Reshef, L. *et al.* (2006) 'The Coral Probiotic Hypothesis', *Environmental Microbiology*. John Wiley & Sons, Ltd, 8(12), pp. 2068–2073. doi: 10.1111/j.1462-2920.2006.01148.x.

Reyes, C. *et al.* (2016) 'Bacterial communities potentially involved in iron-cycling in Baltic Sea and North Sea sediments revealed by pyrosequencing', *FEMS Microbiology Ecology*, 92(4). doi: 10.1093/femsec/fiw054.

Řezanka, T. and Dembitsky, V. M. (2006) 'Metabolites produced by cyanobacteria belonging to several species of the family Nostocaceae', *Folia Microbiologica*. *Folia Microbiol* (Praha), pp. 159–182. doi: 10.1007/BF02932119.

Rosado, P. M. *et al.* (2019) 'Marine probiotics: increasing coral resistance to bleaching through microbiome manipulation', *ISME Journal*. Nature Publishing Group, 13(4), pp. 921–936. doi:

10.1038/s41396-018-0323-6.

Rublee, P. A. (1982) 'bacteria and microbial distribution in estuarine sediments', in *Estuarine Comparisons*. Elsevier, pp. 159–182. doi: 10.1016/b978-0-12-404070-0.50016-8.

Rush, D. and Sinninghe Damsté, J. S. (2017) 'Lipids as paleomarkers to constrain the marine nitrogen cycle', *Environmental Microbiology*. Blackwell Publishing Ltd, 19(6), pp. 2119–2132.doi: 10.1111/1462-2920.13682.

Saunders, G. W. (2005) 'Applying DNA barcoding to red macroalgae: A preliminary appraisalholds promise for future applications', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society, 360(1462), pp. 1879–1888. doi: 10.1098/rstb.2005.1719.

Saunders, G. W. and Moore, T. E. (2013) 'Refinements for the amplification and sequencing of red algal DNA barcode and RedToL phylogenetic markers: a summary of current primers, profiles and strategies', *ALGAE*, 28(1), pp. 31–43. doi: 10.4490/algae.2013.28.1.031.

Sheeba, V. A. *et al.* (2017) 'Role of heavy metals in structuring the microbial community associated with particulate matter in a tropical estuary', *Environmental Pollution*. Elsevier Ltd,231, pp. 589–600. doi: 10.1016/j.envpol.2017.08.053.

Singh, R. P., Kumari, P. and Reddy, C. R. K. (2015) 'Antimicrobial compounds from seaweeds-associated bacteria and fungi', *Applied Microbiology and Biotechnology*. Springer Verlag, pp.1571–1586. doi: 10.1007/s00253-014-6334-y.

Spoerner, M. *et al.* (2012) 'Growth and Thallus Morphogenesis of *Ulva mutabilis* (Chlorophyta)Depends on A Combination of Two Bacterial Species Excreting Regulatory Factors', *Journal of Phycology*. John Wiley & Sons, Ltd, 48(6), pp. 1433–1447. doi: 10.1111/j.1529- 8817.2012.01231.x.

Statham, P. J. (2012) 'Nutrients in estuaries - An overview and the potential impacts of climatechange', *Science of the Total Environment*. Elsevier, pp. 213–227. doi: 10.1016/j.scitotenv.2011.09.088.

Steller, D.L. and Cáceres-Martínez, C. (2009) 'Coralline algal rhodoliths enhance larval settlementand early growth of the Pacific calico scallop *Argopecten ventricosus*', *Marine Ecology ProgressSeries*, 396, pp. 49–60. doi: 10.3354/meps08261.

Steller, D. L. and Cáceres-Martínez, C. (2009) 'Coralline algal rhodoliths enhance larval settlement and early growth of the Pacific calico scallop *Argopecten ventricosus'*, 396, pp. 49–60. doi: 10.3354/meps08261.

Suenaga, K. *et al.* (2004) 'Inducing substance for abalone larval metamorphosis from the crustose coralline alga *Hydrolithon samoense'*, *Fisheries Science*, 70(2), pp. 342–344. doi:10.1111/j.1444-2906.2003.00811.x.

Suyama, T. and Kawaharasaki, M. (2013) 'Decomposition of waste DNA with extendedautoclaving under unsaturated steam', *BioTechniques*, 55(6), pp. 296–9. doi: 10.2144/000114113.

Thomas, T. E. and Harrison, P. J. (1987) 'Rapid ammonium uptake and nitrogen interactions in five intertidal seaweeds grown under field conditions', *Journal of Experimental Marine Biologyand Ecology*. Elsevier, 107(1), pp. 1–8. doi:

10.1016/0022-0981(87)90118-3.

Thompson, J. R. *et al.* (2015) 'Microbes in the coral holobiont: partners through evolution, development, and ecological interactions', *Frontiers in Cellular and Infection Microbiology*.Frontiers, 4, p. 176. doi: 10.3389/fcimb.2014.00176.

Traving, S. J. *et al.* (2017) 'The Effect of Increased Loads of Dissolved Organic Matter on Estuarine Microbial Community Composition and Function', *Frontiers in Microbiology*.Frontiers Research Foundation, 8(MAR), p. 351. doi: 10.3389/fmicb.2017.00351.

Utgikar, V. P. *et al.* (2003) 'Quantification of toxic and inhibitory impact of copper and zinc onmixed cultures of sulfate-reducing bacteria', *Biotechnology and Bioengineering*. Biotechnol Bioeng, 82(3), pp. 306–312. doi: 10.1002/bit.10575.

Valdespino-Castillo, P. M. *et al.* (2021) 'Interplay of microbial communities with mineralenvironments in coralline algae', *Science of the Total Environment*. Elsevier B.V., 757, p.143877. doi: 10.1016/j.scitotenv.2020.143877.

de Vries, H. J. *et al.* (2019) 'Isolation and characterization of Sphingomonadaceae from fouledmembranes', *npj Biofilms and Microbiomes*. Nature Publishing Group, 5(1), pp. 1–9. doi: 10.1038/s41522-018-0074-1.

Walker, D. (2017) *Fal Estuary, Tolverne Provisional RMP Assessment*. Available at:www.cefas.defra.gov.uk (Accessed: 6 March 2021).

Wang, J., Lu, Y. and Shen, G. (2007) 'Combined effects of cadmium and butachlor on soil enzyme activities and microbial community structure', *Environmental Geology*. Springer, 51(7),pp. 1221–1228. doi: 10.1007/s00254-006-0414-y.

Whitehead, P. G. and Prior, H. (2005) 'Bioremediation of acid mine drainage: An introduction to the Wheal Jane wetlands project', *Science of the Total Environment*. Elsevier, 338(1-2 SPEC.ISS.), pp. 15–21. doi: 10.1016/j.scitotenv.2004.09.016.

Wichard, T. (2015) 'Exploring bacteria-induced growth and morphogenesis in the greenmacroalga order Ulvales (Chlorophyta)', *Frontiers in Plant Science*. Frontiers Research Foundation, 6(MAR), pp. 1–19. doi: 10.3389/fpls.2015.00086.

Yao, X. feng *et al.* (2017) 'The effect of heavy metal contamination on the bacterial community structure at Jiaozhou Bay, China', *Brazilian Journal of Microbiology*. Elsevier Editora Ltda, 48(1),pp. 71–78. doi: 10.1016/j.bjm.2016.09.007.

Yoon, H. S., Hackett, J. D. and Bhattacharya, D. (2002) 'A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis', *Proceedingsof the National Academy of Sciences of the United States of America*. National Academy of Sciences, 99(18), pp. 11724–11729. doi: 10.1073/pnas.172234799.

*AppendicesAppendix 1:*

Quick-Start Protocol March 2016

# DNeasy® Plant Mini Kit

The DNeasy Plant Mini Kit (cat. nos. 69104 and 69106) can be stored at room temperature (15–25°C) for up to 1 year if not otherwise stated on label.

## Further information

- *DNeasy Plant Handbook*: **www.qiagen.com/HB-1166**
- Safety Data Sheets: **www.qiagen.com/safety**
- Technical assistance: **support.qiagen.com**

## Notes before starting

- Perform all centrifugation steps at room temperature (15–25°C).
- If necessary, redissolve any precipitates in Buffer AP1 and Buffer AW1 concentrates.
- Add ethanol to Buffer AW1 and Buffer AW2 concentrates.
- Preheat a water bath or heating block to 65°C.

1. Disrupt samples (≤100 mg wet weight or ≤20 mg lyophilized tissue) using the TissueRuptor®, the TissueLyser II or a mortar and pestle.
2. Add 400 µl Buffer AP1 and 4 µl RNase A. Vortex and incubate for 10 min at 65°C. Invert the tube 2–3 times during incubation.

   **Note**: Do not mix Buffer AP1 and RNase A before use.
3. Add 130 µl Buffer P3. Mix and incubate for 5 min on ice.
4. **Recommended**: Centrifuge the lysate for 5 min at 20,000 x *g* (14,000 rpm).
5. Pipet the lysate into a QIAshredder spin column placed in a 2 ml collection tube. Centrifuge for 2 min at 20,000 x *g*.

Sample to Insight QIAGEN

6. Transfer the How-through into a new tube without disturbing the pellet ii present. Add 1.5 volumes of Buller AW l , and mix by pipetting.
7. . Transfer 650 pl of the mixture into a DNeosy Mini spin column placed in a 2 ml collection tube. Centrifuge for 1 min at ;,6000 x *g* (;,8000 rpm). Discord the flow• through. Repea t this step with the remaining sample.
8. Place the spin column into a new 2 ml collection tube. Add 500 pl Buller AW 2, and centrifuge for 1 min at ;,6000 x *g*. Discord the flow through.
9. Add another *500* pl Buller AW2. Centrifugefor 2 min at 20,000 x *g*.

   **Note:** Remove the spin column from the collection tube carefully so that the column doesnot come into contact with the flow-through.
10. Tronsfer the spin column to a new 1. 5 ml or 2 ml microcentriluge tube.
11. Add 100 pl Buller AE for elution. Incuba te for *5* min at room temperature (15- 25°C).Centri fuge for 1 min at ;,6000 x *g*.
12. Repeot step 11.

Scan **QR** code for handbook.

**For up-to-dote licensing information and product-specific disclaimers, see the respective** QIAGEN kit handbook or user manual.

Appendix 3:

# Standard protocol for 16S rRNA (V4) and 18S rRNA (V9) ampliconsequencing on the MiSeq

## *One-step indexing with custom primers*

This protocol is adapted from the Exeter Sequencing Service's 'Standard protocol for Illumina MiSeq-based 16S rRNA gene studies', using the custom dual-index 1-step PCR system from theSchloss lab at the University of Michigan. Sequencing is carried out on the Illumina MiSeq platform with v2 chemistry, paired-end 250 bp, with 384 samples (4 x 96-well plates) multiplexed per run.

*Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data onthe miseq illumina sequencing platform. Appl. Environ. Microbiol. 79, 5112–5120 (2013)*

*https://github.com/SchlossLab/MiSeq_WetLab_SOP/blob/master/MiSeq_WetLab_SOP_v4.md*

Steps in the protocol:

1. Quantify sample DNA and dilute to 2 ng/µL. Arrange in 96-well plates.
2. PCR with custom indexed target-specific primers (16S rRNA V4 or 18S rRNA V9) thatinclude Illumina adapters on their 5' end.
3. Check PCRs on gel(s). Repeat those that failed.
4. PCR clean-up with DIY version of AMPure magnetic beads, to remove free nucleotides,free primers, polymerase, etc.
5. Quantify amplicons and check size on TapeStation. Dilute amplicons to 4 nM.
6. Pool all amplicon libraries and submit for sequencing.

Kits required:

1. Promega QuantiFluor ONE dsDNA quantification kit (E4870) with appropriate black 96-well flat-bottom plates that fit in the Promega GloMax instrument (or an equivalent fluorometry-based DNA quantification kit, e.g. Qubit)

2. NEBNext high-fidelity PCR master mix (New England Biolabs M0541)
3. Agencourt AMPure XP beads (Beckman Coulter A63881) or the DIY version from Jolivetand Foley (Appendix B)
4. High Sensitivity D1000 ScreenTape and reagents (Agilent 5067-5584, 5067-5585, 5067-5587)
5. Optional: gel clean-up kit (e.g. Qiagen MinElute, Promega Wizard SV kit)

# Protocol

## *1. Quantify and dilute sample DNA*

Use the Promega QuantiFluor ONE dsDNA kit (E4870) to quantify sample DNA. This requires a GloMax fluorescence plate reader and suitable 96-well black flat-bottom plates. Samples can also be quantified individually with an alternative fluorometry kit (e.g. Qubit).

Follow the kit protocol to prepare DNA standards and quantify samples. The best standard curves are obtained if all reagents are at room temperature, and each standard is vortexed for 30s and left to equilibrate for 5 minutes before preparing the next standard.

If you expect samples to have 200 ng/µL DNA or less, you can use 100 µL QuantiFluor reagent instead of 200 µL per well (a considerable saving).

Dilute DNA to 2 ng $\mu L^{-1}$ with 10mM Tris-HCl pH 8.5 (e.g. Qiagen buffer EB). Arrange diluted samples in 96-well plates and store at -20°C. Each plate should have a minimum of two empty wells for a positive and negative control (additional empty wells can be left for other controls where appropriate - extraction blanks, mock community extractions, etc.).

## *2. PCR with target-specific primers*

PCR Reagents: NEBNext high-fidelity PCR master mix (New England Biolabs – M0541S/L)

Each 50 uL PCR reaction contains the following components:

| Component | Amount per 50 uL rxn | Final concentration |
|---|---|---|
| NEBNext PCR mix | 25.0 µL | x1 |
| Water | 17.5 µL | |
| 10 µM F primer | 2.5 µL | 0.5 µM |
| 10 µM R primer | 2.5 µL | 0.5 µM |
| Template DNA (2 ng/µL) | 2.5 µL | 5 ng / rxn |

Since a different pair of indexed primers is used for each sample, the primers are not included inthe PCR master mix but are added separately. Primer pairs should already be arranged in 96- well index primer plates (see Appendix A for details).

There are four index primer plates for each marker gene:

PV4 plates A, B, C, D (PV4 means prokaryotes 16S

rRNA V4)EV9 plates A, B, C, D (EV9 means

eukaryotes 18S rRNA V9)

Steps to set up PCR:

1. In a sterile 15 mL Falcon tube, prepare master mix of NEBNext PCR mix and water only,allowing extra for loss due to using reagent troughs:

One plate: For 100 reactions, mix 2500 µL NEBNext with 1750

µL water. Two plates: For 196 reactions, mix 4900 µL NEBNext

with 3430 µL water.*Chill the NEBNext/water mix on ice.*

2. Transfer 5 µL from index primer plate into new plate. Keep the new plate on a cold block oron ice throughout.

3. Add 2.5 µL from diluted DNA plate to the new plate containing 5 µL primer mix (previousstep).

4. To each well, add chilled 42.5 µL NEBNext/water mix. Seal plate and spin briefly to remove airbubbles. Keep plate cold until PCR machine is ready and pre-heated.

PCR conditions:

98°C for 30s

30 cycles of:

- 98°C for 10s
- annealing temp for 30s (55°C for PV4, 60°C for EV9)
- 72°C for 30s

72°C for 2 min

hold at 10°C

Pre-heat the PCR block to 98°C before inserting plate. This can be done by starting the programme and pausing the machine once the temperature reaches 98°C.

## *3. Check amplification on gel, repeat failed PCRs*

If the samples generally amplify well, a subset of each plate can be run on an agarose gel to verify that the PCR worked. Positive and negative controls should always be checked on a gel.

Environmental or tissue samples often contain varying amounts of PCR inhibitors, making amplification unreliable. In these cases, the whole plate should be run on a gel.

Loading 4 µL of PCR product onto a 1% or 1.5% agarose gel is generally sufficient to assess whether amplification occurred. Faint bands are acceptable, as very little PCR product is needed per sample in the final library pool. Expected sizes are approx. 400 bp for PV4 and approx. 270 bp for EV9.

When no amplification occurred, this is usually due to PCR inhibitors in the DNA sample. PCR should be repeated with less template DNA (2 ng per reaction):

| Component | Amount per 50 uL rxn | Final concentration |
|---|---|---|
| NEBNext PCR mix | 25.0 µL | x1 |
| Water | 19.0 µL | |
| 10 µM F primer | 2.5 µL | 0.5 µM |
| 10 µM R primer | 2.5 µL | 0.5 µM |
| Template DNA (2 ng/uL) | 1.0 µL | 2 ng / rxn |

This usually solves the problem, but particularly difficult sets of samples may need to go through an additional DNA clean-up step prior to dilution and amplification. The Zymo Genomic DNA Clean & Concentrator kit is good for cleaning particularly difficult samples.

## *4. Clean up PCRs*

This step uses Agencourt AMPure XP magnetic beads (or the DIY version from Jolivet and Foley, see Appendix B) to purify the amplicons away from free primers and primer dimer species.

Bring the magnetic beads to room temperature for 30 minutes before use.

a) Vortex the magnetic beads for 30 seconds to make sure that the beads are evenly dispersed. Add an appropriate volume of beads to a trough depending on the number of samples processing.
b) Using a multichannel pipette, add 0.8 volumes (36.8 µl) of magnetic beads to each 46 µl PCR reaction (50 µL PCR reaction minus 4 µL used for gel). Change tips between tubes/columns.
c) Gently pipette entire volume up and down 10 times.
d) Incubate at room temperature without shaking for 5 minutes.
e) Place the plate on a magnetic stand for 2 minutes or until the supernatant has cleared.
f) With the PCR plate on the magnetic stand, use a multichannel pipette to remove and discard the supernatant. Change tips between samples.
g) With the PCR plate on the magnetic stand, add 200 µl of freshly prepared 80% ethanol to each sample well.
h) Incubate the plate on the magnetic stand for 2 minutes.
i) Carefully remove and discard the supernatant.
j) With the PCR plate on the magnetic stand, perform a second ethanol wash: add 200 µl of freshly prepared 80% ethanol to each sample well.
k) Incubate the plate on the magnetic stand for 2 minutes.
l) Carefully remove and discard the supernatant.
m) Use a P10 multichannel pipette with fine pipette tips to remove excess ethanol.
n) With the PCR plate still on the magnetic stand, allow the beads to air-dry for 10 minutes.
o) Remove the PCR plate from the magnetic stand. Using a multichannel pipette, add 22.5 µl of TE+Tween buffer (see Appendix B) to each well of the PCR plate. Gently pipette mix up and down at least 10 times, or until beads are fully resuspended.
p) Incubate at room temperature for 2 minutes.
q) If droplets are present on the sides of the wells, spin briefly to collect liquid in the bottom of each well. Place the plate on the magnetic stand for 2 minutes or until the supernatant has cleared.
r) Using a multichannel pipette, carefully transfer 20 µl of the supernatant from the PCRplate to a new 96-well PCR plate. Change tips between samples to avoid cross-contamination.

*Store at -20°C until quantification.*

## *5. Quantify PCRs, check size on TapeStation and dilute to 4 nM*

a) Use the Promega QuantiFluor ONE dsDNA kit to quantify each cleaned PCR product (asdescribed above in step 1, 'Quantify sample DNA').

b) For a small subset of samples, check the size of the PCR product on a TapeStation instrument using High Sensitivity D1000 ScreenTape and reagents. This could also be done on a Bioanalyzer DNA 1000 chip, if that is what is available.

c) Calculate PCR product concentration of each sample in nM, based on the size of theamplicon:

*(conc. in ng/µl) / (660 g/mol × PCR product size)× $10^6$ = concentration in nM*

*For example:*
15 *ng/µl / (660 g/mol × 500 bp) × $10^6$ = 45 nM*

d) In a new plate, dilute each final PCR product to 4 nM with 10 mM Tris pH 8.5 or TE+Tween.

## *6. Final pooling and submission for sequencing*

Combine 5 µL from each diluted 4 nM PCR product into a single tube. This can be done in stages:

a) Prepare a pool from each 96-well plate, into a 0.5 mL lo-bind Eppendorf tube.
b) Mix well by vortexing briefly and allowing to equilibrate at room temperature for a minimum of five minutes.
c) Prepare the final pool by taking 100 µL from each of the four plate pools and combining them in a new 0.5 mL lo-bind Eppendorf tube. Mix well.

Check the final pool on the Bioanalyzer or TapeStation. If primer dimers remain, gel purify the final pool.

Obtain a sequencing quote from the Exeter Sequencing Service and enter the sample details on the LIMS. When filling out the LIMS spreadsheet, the first Adaptor Name column should contain the Sx7 index names, and the first Barcode Sequence column should contain the *reverse complement* of the Sx7 indices. Adaptor2 Name should contain the Sx5 index names, and Barcode2 Sequence should contain the Sx5 indices (not their reverse complement).

Submit 50-100 µL of the final pool to the sequencing service for sequencing on the Illumina MiSeq platform with v2 chemistry (PE250), specifying that the amplicons were prepared with custom primers.

# Appendix A: Custom indexed primers

Custom primer design is based on the system developed by the Schloss lab. Full details can be found here:

*Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. Appl. Environ. Microbiol. 79, 5112–5120 (2013)*

*https://github.com/SchlossLab/MiSeq_WetLab_SOP/blob/master/MiSeq_WetLab_SOP_v4.md*

Briefly, the target (16S rRNA V4 or 18S rRNA V9) is amplified by primers that include the following components, listed in 5' to 3' order:

*Forward amplification primer:*
<Illumina p5 adapter><S5 8 bp index><10 bp pad><2 bp link><target-specific F primer>

*Reverse amplification primer:*

<Illumina p7 adapter><S7 8 bp index><10 bp pad><2 bp link><target-specific R primer>

The Illumina p5/p7 adapters allow binding of the amplicon to the flow cell. The 8-bp index sequences allow multiplexing of many samples per run. The 10-bp pad sequence increases the annealing temperature of the amplicon to approx. 65°C, as required for the MiSeq. The 2-bp linker sequence is chosen so as not to match any sequences from reference databases at the positions immediately upstream of the target-specific primers. Finally, the target-specific primers are the only parts that anneal to the sample DNA: 515fB/806rB for 16S V4, 1391f/EukB for 18S V9.

In addition to the amplification primers, which are used with the sample DNA, there are three custom primers that are used only during the MiSeq run. These primers are held by the Exeter Sequencing Service so there is no need to submit them every time.

*Read 1 sequencing primer:*

<10 bp pad><2 bp link><target-specific F primer>

*Read 2 sequencing primer:*

<10 bp pad><2 bp link><target-specific R primer>

*Index read primer:*

reverse complement of <10 bp pad><2 bp link><target-specific R primer>

Primer elements used for our custom primers sets (PV4 and EV9):

PV4 - 16S rRNA V4

Forward (515fB): GTGYCAGCMGCCGCGGTAA

link: GT

pad: TATGGTAATT

p5: AATGATACGGCGACCACCGAGATCTACAC

seq.primer: TATGGTAATTGTGTGYCAGCMGCCGCGGTAA

Reverse (806rB): GGACTACNVGGGTWTCTAAT

link: CC

pad: AGTCAGTCAG

p7: CAAGCAGAAGACGGCATACGAGAT

seq.primer: AGTCAGTCAGCCGGACTACNVGGGTWTCTAAT

EV9 - 18S rRNA V9

Forward (1391f): GTACACACCGCCCGTC

link: CG

pad: TATCGCCGTT

p5: AATGATACGGCGACCACCGAGATCTACAC

seq.primer: TATCGCCGTTCGGTACACACCGCCCGTC

Reverse (EukB): TGATCCTTCTGCAGGTTCACCTAC

link: CA

pad: AGTCAGCCAG

p7: CAAGCAGAAGACGGCATACGAGAT

seq.primer: AGTCAGCCAGCATGATCCTTCTGCAGGTTCACCTAC

Index sequences:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SA501 | ATCGTACG | SB501 | CTACTATA | SA701 | AACTCTCG | SB701 | AAGTCGAG |
| SA502 | ACTATCTG | SB502 | CGTTACTA | SA702 | ACTATGTC | SB702 | ATACTTCG |
| SA503 | TAGCGAGT | SB503 | AGAGTCAC | SA703 | AGTAGCGT | SB703 | AGCTGCTA |
| SA504 | CTGCGTGT | SB504 | TACGAGAC | SA704 | CAGTGAGT | SB704 | CATAGAGA |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SA505 | TCATCGAG | SB505 | ACGTCTCG | SA705 | CGTACTCA | SB705 | CGTAGATC |
| SA506 | CGTGAGTG | SB506 | TCGACGAG | SA706 | CTACGCAG | SB706 | CTCGTTAC |
| SA507 | GGATATCT | SB507 | GATCGTGT | SA707 | GGAGACTA | SB707 | GCGCACGT |
| SA508 | GACACCGT | SB508 | GTCAGATA | SA708 | GTCGCTCG | SB708 | GGTACTAT |
| | | | | SA709 | GTCGTAGT | SB709 | GTATACGC |
| | | | | SA710 | TAGCAGAC | SB710 | TACGAGCA |
| | | | | SA711 | TCATAGAC | SB711 | TCAGCGTT |
| | | | | SA712 | TCGCTATA | SB712 | TCGCTACG |

*Indexed primers are arranged in 4x96-well plates. Each well has a unique Sx5/Sx7 pair:*

Plate A: SA5 + SA7

| | | SA 701 | SA 702 | SA 703 | SA 704 | SA 705 | SA 706 | SA 707 | SA 708 | SA 709 | SA 710 | SA 711 | SA 712 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| SA501 | A | | | | | | | | | | | | |
| SA502 | B | | | | | | | | | | | | |
| SA503 | C | | | | | | | | | | | | |
| SA504 | D | | | | | | | | | | | | |
| SA505 | E | | | | | | | | | | | | |
| SA506 | F | | | | | | | | | | | | |
| SA507 | G | | | | | | | | | | | | |
| SA508 | H | | | | | | | | | | | | |

Plate B: SB5 + SA7

| | | SA 701 | SA 702 | SA 703 | SA 704 | SA 705 | SA 706 | SA 707 | SA 708 | SA 709 | SA 710 | SA 711 | SA 712 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| SB501 | A | | | | | | | | | | | | |
| SB502 | B | | | | | | | | | | | | |
| SB503 | C | | | | | | | | | | | | |
| SB504 | D | | | | | | | | | | | | |
| SB505 | E | | | | | | | | | | | | |
| SB506 | F | | | | | | | | | | | | |
| SB507 | G | | | | | | | | | | | | |
| SB508 | H | | | | | | | | | | | | |

Plate C: SA5 + SB7

| | | SB 701 | SB 702 | SB 703 | SB 704 | SB 705 | SB 706 | SB 707 | SB 708 | SB 709 | SB 710 | SB 711 | SB 712 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| SA501 | A | | | | | | | | | | | | |
| SA502 | B | | | | | | | | | | | | |
| SA503 | C | | | | | | | | | | | | |
| SA504 | D | | | | | | | | | | | | |
| SA505 | E | | | | | | | | | | | | |
| SA506 | F | | | | | | | | | | | | |
| SA507 | G | | | | | | | | | | | | |
| SA508 | H | | | | | | | | | | | | |

Plate D: SB5 + SB7

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | SB | SB | SB | SB | SB | SB | SB | SB | SB | SB | SB | SB |

| | | 701 | 702 | 703 | 704 | 705 | 706 | 707 | 708 | 709 | 710 | 711 | 712 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| SB501 | A | | | | | | | | | | | | |
| SB502 | B | | | | | | | | | | | | |
| SB503 | C | | | | | | | | | | | | |
| SB504 | D | | | | | | | | | | | | |
| SB505 | E | | | | | | | | | | | | |
| SB506 | F | | | | | | | | | | | | |
| SB507 | G | | | | | | | | | | | | |
| SB508 | H | | | | | | | | | | | | |

Primers are ordered from Eurofins in individual tubes, 0.05 µmol synthesis scale, PTO modification near the 3' end, and salt-free purification. They are first resuspended to 100 µM,and then dilutions to 10 µM are prepared (all in 10 mM Tris pH 8.5).

The four index primer plates are then prepared by transferring 20 µL of each 10 µM diluted primer into the appropriate well, according to the plate maps on the previous page. These indexprimer plates, in which each well contains a unique F and R primer pair, are used during PCR
set-up. They can be stored at -20°C.

# Appendix B: DIY magnetic beads


Solutions for purifying nucleic acids by solid-phase reversible immobilization (SPRI)

Philippe Jolivet and Joseph W. Foley

Ludmer Centre for Neuroinformatics and Mental HealthOctober 21, 2015

Based on DeAngelis MM, Wang DG, Hawkins TL, "Solid-phase reversible immobilization for the isolation of PCR products." *Nucleic Acids Res* 1995, 23:4742–4743.

Abstract

This protocol describes the preparation of stocks and buffers for inexpensive, convenient, and scalable DNA and RNA purification from aqueous solutions by solid-phase reversible immobilization (SPRI) on carboxylated paramagnetic beads. It also describes how to validate the effectiveness of the mixes before use.

The bead mixes described in this protocol are drop-in substitutes for AMPure XP and RNAClean XP beads (Beckman Coulter), but at about 1/100 of the cost (~$0.55/mL vs. $15–$70/mL at current Canadian prices).

Materials

*Beads*

The beads we use are 0.1% carboxyl-modified ThermoScientific Particle TechnologySeraMagnetic Speed Beads obtained from Fisher using the code 12326433 .

*Chemicals (molecular biology grade)*

- Sodium chloride (NaCl)
- Poly(ethylene glycol), avg. mol. wt. 8000 (PEG 8000)
- Polysorbate 20 (Tween 20)
- Hydrochloric acid (HCl) concentrate
- Nuclease-

free waterFor

DNA mix

- Tris(hydroxymethyl)aminomethane (Tris base)

- Disodium

ethylenediaminetetraacetate (EDTA)For

RNA mix

- Trisodium citrate dihydrate

*Consumables*

- 50 mL conical tubes
- 1.5 mL microcentrifuge tubes
- Disposable weighing vessels
- Disposable Pasteur pipettes
- Parafilm
- 0.22 µm syringe filters
- 10 mL disposable syringes
- 25 mL, 10 mL, 5 mL serological pipettes
- 1000 µL, 200 µL micropipette tips

*Equipment*

- Milligram-range balance
- Funnels
- Spatulas
- Heating plate
- Rotary mixer
- Microcentrifuge
- 25 mL graduated cylinder
- 50 mL volumetric flasks and stoppers
- 1000 µL, 200 µL adjustable-volume micropipettes
- Squirt bottle
- Magnetic separation block for 1.5 mL microcentrifuge tubes

Stock solutions

Common solutions

- 1 N HCl
- 5 M NaCl
- 10% (v/v) Tween 20
- 50% (w/v) PEG 8000

DNA solutions

- 1 M Tris base
- 0.1 M EDTA

RNA solution

- 1 M trisodium citrate

*Making 1 N HCl*

Prepare at least 10 mL 1 N HCl in a glass bottle from available concentrated stock.

*Making 50 mL of 1 M Tris base, 0.1 M disodium EDTA, 1 M trisodium citrate and 5 M NaCl stocks*

In 50 mL volumetric flasks, prepare a separate stock solution for each of the following components with the specified weights of solids.

| *Common solution* | | *DNA solutions* | | *RNA solution* | |
|---|---|---|---|---|---|
| 5 M NaCl | 14.610 g | 1 M Tris base | 6.057 g | 1 M $Na_3$-citrate·$2H_2O$ | 14.705 g |
| | | 0.1 M $Na_2$-EDTA | 1.861 g | | |

Some gentle heating may be necessary. Ensure the solution comes back to room temperature before completing the volume to the mark on the flask. Store in 50 mL conical tubes. Optional: filter the stocks with the syringes and filters to remove undissolved solids. It is strongly recommended to filter the solutions used for making RNA mix for sterilization.

*Making 50 mL of 10% (v/v) Tween 20 stock*

1. Place a labeled 50 mL conical tube on the balance and tare it.
2. With a new disposable Pasteur pipette, aspirate approximately 0.5 mL of Tween 20.
3. Slowly dispense the Tween 20 into the 50 mL conical tube to reach 5.475 g.
4. Remove the tube from the balance and add 45.0 mL of water with a 25 mL serologicalpipette.
5. Cap the tube and mix on a rotary mixer for one hour to dissolve the viscous liquid.

*Making 25 mL of 50% (w/v) PEG 8000 stock*

1. Place the 25 mL graduated cylinder on the balance and tare it.
2. Weigh 12.5 g of PEG 8000 powder directly into the cylinder. It is recommended to use latex gloves instead of nitrile to reduce static charges that make the powder fly off the spatula.
3. Add no more than 14 mL of nuclease-free water with a serological pipette on top of the PEG powder in the cylinder. The water level will reach over the 25 mL mark as the cylinder already contains about 20 mL of dry powder. Be sure not to fill the cylinder completely, as some air is required to make mixing possible.
4. Seal the cylinder with a double layer of Parafilm.
5. Shake vigorously to suspend the powder in the water until there are no more lumps of dry solid sticking to the cylinder wall. It will be very viscous and clumpy.
6. Let the suspension stand at room temperature for at least an hour to allow the solids to dissolve and the air bubbles to rise.
7. Remove the Parafilm and complete the volume with nuclease-free water to the 25 mL mark.
8. Seal the cylinder again and mix well by inverting. The solution is very viscous and homogenizing it can take a while.
9. Transfer the solution to a 50 mL conical tube for storage.

This recipe can be scaled up with larger cylinders to make mixing easier, for example making 50 mL of solution in a 100 mL cylinder.

Buffer recipes

*Nucleic acid elution and storage buffers*

*DNA buffer:* TE+Tween (10 mM Tris base, 1 mM EDTA, 0.05% Tween 20, pH 8.0 @ 25°C)

*RNA buffer:* Citrate+Tween (1 mM trisodium citrate, 0.05% Tween 20, pH 6.4 @ 25°C)

Ingredients for 50 mL:

| *DNA buffer* | |
|---|---|
| Nuclease-free water | 48.564 mL |
| Tris base, 1 M | 0.500 mL |
| Disodium EDTA, 0.1 M | 0.500 mL |
| Tween 20, 10% (v/v) | 0.250 mL |
| HCl, 1 N | 0.186 mL |

| *RNA buffer* | |
|---|---|
| Nuclease-free water | 49.679 mL |
| Trisodium citrate, 1 M | 0.050 mL |
| Tween 20, 10% (v/v) | 0.250 mL |
| HCl, 1 N | 0.021 mL |

These solutions are used for preparing the beads before adding them to the mix. They are also useful for DNA and RNA elution and storage. It is possible to make concentrates of these solutions for convenience. Keep them refrigerated.

*Nucleic acid binding bead mixes*

*DNA mix:* 10 mM Tris base, 1 mM EDTA, 2.5 M NaCl, 20% PEG 8000, 0.05% Tween 20, pH 8.0 @ 25°C

*RNA mix:* 1 mM trisodium citrate, 2.5 M NaCl, 20% PEG 8000, 0.05% Tween 20, pH 6.4 @ 25°C

Read the mixing instructions below before starting to combine the ingredients.

Ingredients for 50 mL:

| *DNA binding bead mix* | |
|---|---|
| NaCl, 5 M | 25.000 mL |
| Nuclease-free water | 3.582 mL |
| Tris base, 1 M | 0.500 mL |
| Disodium EDTA, 0.1 M | 0.500 mL |
| HCl, 1 N | 0.168 mL |
| PEG 8000, 50% (w/v) | 20.000 mL |
| Tween 20, 10% (v/v) | 0.250 mL |
| Sera-Mag bead suspension | 1.000 mL |

| *RNA binding bead mix* | |
|---|---|
| NaCl, 5 M | 25.000 mL |
| Nuclease-free water | 4.672 mL |
| Trisodium citrate, 1 M | 0.050 mL |
| HCl, 1 N | 0.028 mL |
| PEG 8000, 50% (w/v) | 20.000 mL |
| Tween 20, 10% (v/v) | 0.250 mL |
| Sera-Mag bead suspension | 1.000 mL |

Mixing instructions:

1. Mix the Sera-Mag beads very well to resuspend.
2. Quickly transfer 1 mL to a 1.5 mL microcentrifuge tube (the beads settle quickly).
3. Place the tube on a magnet stand until the supernatant is clear, about 30 s.
4. Remove and discard the supernatant.
5. Add 1 mL of previously prepared TE+Tween (for DNA) or Citrate+Tween (for RNA) bufferto the bead pellet and close the tube.
6. Remove the tube from the magnet and resuspend the beads by vortexing for at least15 seconds. Spin down the liquid with a microcentrifuge.
7. Put the tube back on the magnet until the beads clear.
8. Remove and discard the supernatant.
9. Repeat steps 5 to 8 twice, for a total of 3 washes with the appropriate buffer, leaving the supernatant in the tube after the last wash.
10. In a new 50 mL conical tube, combine the nuclease-free water, NaCl and HCl. For DNA,also add the Tris and EDTA. For RNA, add only the trisodium citrate instead. Cap and mix well.
11. Remove the buffer supernatant from the bead tube still on the magnet.
12. Add 1 mL of incomplete binding buffer (prepared at step 10) to the bead tube on themagnet.
13. Remove the bead tube from the magnet and resuspend by vortexing for 15 s. Brieflyspin down the liquid without pelleting the beads.
14. Add the washed beads to the incomplete binding buffer. Cap and vortex for 30s.
15. With a 25 mL serological pipette, add 20 mL of 50% PEG stock. Dispense slowly and allow the viscous liquid to slide down the inside walls of the pipette to ensure an accurate volume is added.
16. Add the Tween 20.
17. Cap the tube and mix by inversion gently but thoroughly, until the color appears homogeneous.

The bead binding mix is ready to be used or validated. Store at 4 °C.

Important considerations and recommendations

*Chelating agents*

EDTA and citrate may interfere with some enzymatic reactions by sequestering divalent cations such as $Mg^{2+}$ and $Mn^{2+}$. These ions may damage nucleic acids or activate contaminating nucleases.On the other hand, sequestering these ions may interfere with downstream reactions that requirethem; if so, you can compensate by adding ions equimolar to the EDTA or citrate.

*pH*

The pH titrations for the buffers and bead mixes were calculated with the Python package ionize 0.8.0. They may be inaccurate for the bead mixes due to the very high ionic strengths of those solutions. Colour-change pH indicators will also be inaccurate for the same reason. A properly calibrated pH meter may be able to measure these solutions correctly. Keep in mind that the bead mix will be diluted during use when added to the sample to be purified, which will change the ionic strength and thus the pH.

*Tween 20*

Adding Tween 20 to the solutions described in the protocol is optional but provides multiple benefits. It reduces adhesion of nucleic acids to plastics, which is increased during SPRI due to thehigh ionic strength. This improves sample recovery. Tween 20 also reduces surface tension, whichcan pull beads off the pellet during supernatant removal. This effect becomes very useful if the pellet is very small. If Tween 20 is not compatible with your downstream processes or if foamingis a problem, replace its volume with nuclease-free water when mixing the solutions.

Validation

It is recommended to validate the bead mixes before use to ensure their effectiveness. They can be compared to AMPure XP or RNAClean XP, or to a previous batch of homemade mix. Validationcan be done with DNA or RNA that is representative of a typical usage scenario, a DNA ladder (note that NEB ladders may contain modifications that make their SPRI behaviour unrepresentative of normal DNA), fragmented DNA across a range of sizes, or an RNA standard.

*Example with DNA smear*

Fragment λ phage genomic dsDNA (e.g. Thermo Scientific #SD0011) with the NEBNext® dsDNA Fragmentase® kit (NEB #M0348S) for 20 minutes, according to the manufacturer's instructions, and purify the reaction product with 2 volumes of previously validated beads, then elute in TE+Tween. This produces a flat smear (50 to 2000 bp) that is easy to see on an agarose gel or anAgilent Bioanalyzer DNA 1000 chip.

Re-purify the DNA with the reference DNA bead mix and the new DNA bead mix side-by-

side using2 volumes of beads. After elution, test the DNA on agarose gel or Bioanalyzer to compare the

repurified samples with each other and with the original fragmented stock. If the yields and size distributions from the two bead mixes are equivalent, the new mix is ready for use.

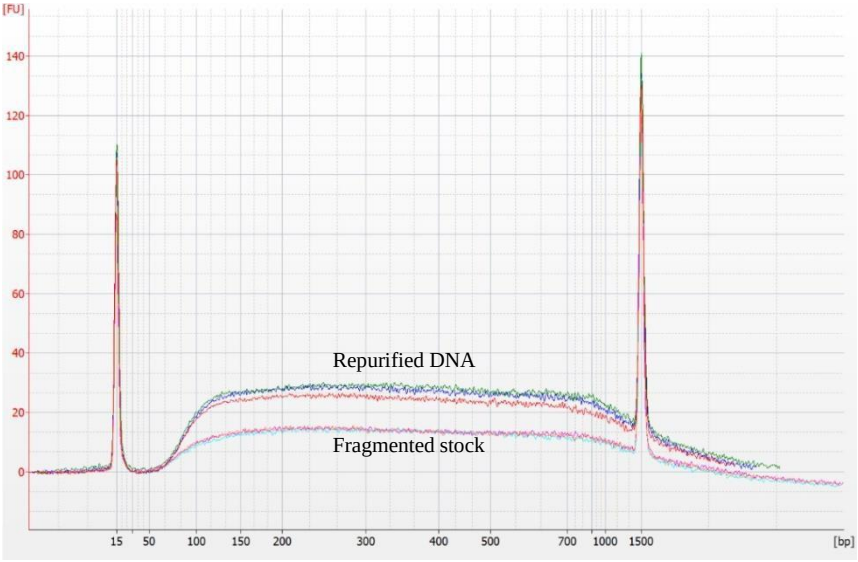


Figure 1. Example results with DNA smear. Validation done with λ phage genomic dsDNA as specified aboveand loaded on an Agilent Bioanalyzer DNA 1000 chip. Fragmented stock was purified with AMPure XP (3 measurement replicates) and the test sample with homemade bead mix (3 purification replicates) both at2 volumes of bead mix. Different amounts of DNA were loaded (approximately 25 ng stock, 40 ng repurifiedDNA) but the size distributions are very similar. This batch of mix was successfully validated.

*Example with RNA standard*

To test the RNA bead mix, dilute FirstChoice Human Brain Reference RNA (Life Tech #AM6050)100-fold to 10 ng/µL in Citrate+Tween. Purify the dilution using the RNA bead mix and elute inCitrate+Tween. Load the original input and the purified output on an Agilent Bioanalyzer RNA 6000 Pico chip.

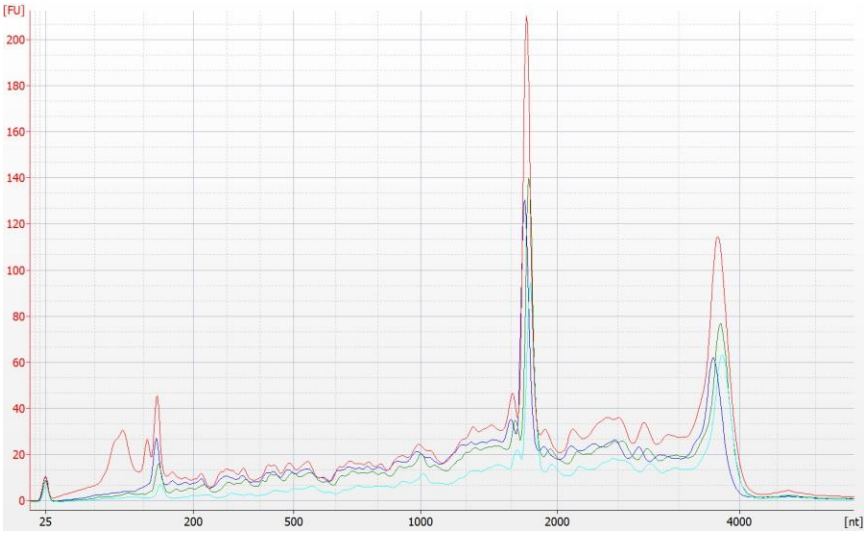


Figure 2. Example results with HBRR standard. The original diluted input (red) was loaded with samples purified using different ratios of RNA bead mix: two volumes (blue), one volume (green)

and half volume(cyan). The samples were denatured before loading.

Version history

V1.0 – January 13, 2015: First version.

V2.0 – April 17, 2015: Added support for RNA purification. Tween stock now 10%. Cosmetic changes.V2.1 – September 25, 2015: Corrected minor mistake in RNA binding bead mix recipe volumes.

V2.2 – October 21, 2015: Corrected wrong NaCl stock concentration on page 2, 3 M changed to 5 M.

## Appendix 3:

| Sample | Library | Protocol | PCR | Barcod |
|---|---|---|---|---|
| Sed_1 | Sed_1-l2 | Nextera XT v2 | none | AAGTC |
| Sed_1 | Sed_1-l1 | Nextera XT v2 | none | GTATA |
| Sed_2 | Sed_2-l2 | Nextera XT v2 | none | AAGTC |
| Sed_2 | Sed_2-l1 | Nextera XT v2 | none | GTATA |
| Sed_3 | Sed_3-l2 | Nextera XT v2 | none | AAGTC |
| Sed_3 | Sed_3-l1 | Nextera XT v2 | none | GTATA |
| Sed_4 | Sed_4-l2 | Nextera XT v2 | none | AAGTC |
| Sed_4 | Sed_4-l1 | Nextera XT v2 | none | GTATA |
| Sed_5 | Sed_5-l2 | Nextera XT v2 | none | AAGTC |
| Sed_5 | Sed_5-l1 | Nextera XT v2 | none | GTATA |
| Sed_6 | Sed_6-l2 | Nextera XT v2 | none | AAGTC |
| Sed_6 | Sed_6-l1 | Nextera XT v2 | none | GTATA |
| Lseaw_1 | Lseaw_1-l1 | Nextera XT v2 | none | CGTAG |
| Lseaw_2 | Lseaw_2-l1 | Nextera XT v2 | none | CGTAG |
| Lseaw_3 | Lseaw_3-l1 | Nextera XT v2 | none | CTCGT |
| Dseaw_1 | Dseaw_1-l1 | Nextera XT v2 | none | CTCGT |
| Dseaw_2 | Dseaw_2-l1 | Nextera XT v2 | none | CTCGT |
| Dseaw_3 | Dseaw_3-l1 | Nextera XT v2 | none | CTCGT |
| Lmex_1 | Lmex_1-l1 | Nextera XT v2 | none | CTCGT |
| Lmex_2 | Lmex_2-l1 | Nextera XT v2 | none | CTCGT |
| Lmex_3 | Lmex_3-l1 | Nextera XT v2 | none | CTCGT |

| | | | | |
|---|---|---|---|---|
| Dmex_1 | Dmex_1-l1 | Nextera XT v2 | none | TACGA |
| Dmex_2 | Dmex_2-l1 | Nextera XT v2 | none | TACGA |
| Dmex_3 | Dmex_3-l1 | Nextera XT v2 | none | TACGA |
| LMT_1 | LMT_1-l1 | Nextera XT v2 | none | TACGA |
| LMT_2 | LMT_2-l1 | Nextera XT v2 | none | TACGA |
| LMT_3 | LMT_3-l1 | Nextera XT v2 | none | TACGA |
| DMT_1 | DMT_1-l1 | Nextera XT v2 | none | TACGA |
| DMT_2 | DMT_2-l1 | Nextera XT v2 | none | TACGA |
| DMT_3 | DMT_3-l1 | Nextera XT v2 | none | TCAGC |
| Sed_Con | Sed_Con-l1 | Nextera XT v2 | none | CTACT |
| Seaw_con | Seaw_con-l1 | Nextera XT v2 | none | CGTTA |
| Clean_Seaw_con | Clean_Seaw_con-l1 | Nextera XT v2 | none | AGAGT |
| positive | positive-l1 | Nextera XT v2 | none | AAGTC |
| negative | negative-l1 | Nextera XT v2 | none | AAGTC |

Appendix 4:

Trimming the forward and reverse reads at 240 and 200 bp before merging

```
out <- filterAndTrim(fnFs, filtFs, fnRs, filtRs, truncLen=c(240,200),maxN=0, maxEE=c(2,2), truncQ=2,
        rm.phix=TRUE, compress=TRUE, multithread=FALSE)
```

Samples were pooled during sample inference steps, instead ofindependently or pseudo pooling.

dadaFs <- dada(derepsFs, err=errF, multithread = TRUE, pool = TRUE) dadaRs <- dada(derepsRs, err=errR, multithread = TRUE, pool = TRUE)

Samples which contained between 240 and 258 base pairs were kept forthe final ASV table.

seqtab2 <- seqtab[,nchar(colnames(seqtab)) %in% seq(240:258)])

Classification of amplicons was performed using DADA2`s naive Bayesian classifier and with Silva taxonomic training data formatted for DADA2 (Silvaversion 132). Additionally, exact reference matches of the ASV`s were assigned species level taxonomy.

taxa <- assignTaxonomy(seqtab.nochim, "~/tax/silva_nr_v132_train_set.fa.gz", multithread=TRUE)

taxa <- addSpecies(taxa, "~/tax/silva_species_assignment_v132.fa.gz")

appendix 5:

```
install.packages("remotes")
remotes::install_github("umerijaz/microbiomeSe
q")

install.packages("remotes")
remotes::install_github("TBrach/MicrobiomeX")

if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")

BiocManager::install("phyloseq")
BiocManager::install("microbiome
")
```

```
BiocManager::install("decontam")
library("reshape") ;
library("DESeq2")
library("phyloseq") ; library("ggplot2") ; library("vegan") ;
library("dada2")library ("xlsx")
library("eule
rr")
library("deco
ntam")
library("micr
obiome")

library("knitr")
library("ggpubr")
library("reshape2")
library("RColorBrew
er")
library("microbiome
utilities")
library("viridis")
library("tibble")
# pruning and making sample data




sampledatalist2 <- c("Dead_Maerl_Exterior", "Dead_Maerl_Exterior",
                     "Dead_Maerl_Exterior","Dead_maerl_Total" , "Dead_maerl_Total" ,
                     "Dead_maerl_Total" ,
                     "Dead_Seawater" , "Dead_Seawater" , "Live_Maerl_Exterior"
,"Live_Maerl_Exterior",
                    "Live_Maerl_Exterior" , "Live_Maerl_Total" , "Live_Maerl_Total" ,
```

```
"Live_Maerl_Total",
                "Live_Seawater"    , "Live_Seawater" ,
                 "Live_Seawater","Seawater_Control"  ,
                 "Sediment"   ,     "Sediment",
                "Sediment"      ,   "Sediment"    ,    "Sediment"   ,
                                  "Sediment","Sediment_Control")


sample.names <-
sample_names(otu_table(newps.ass))
sample.names
sampledatalist2
?sample_data
mk2sampledata = sample_data(data.frame(
 SampleType = sampledatalist2, size = nsamples(newps.ass),
 replace=TRUE,row.names = sample.names,
 stringsAsFactor
s=FALSE))
mk2sampledata


library("ape")


random_tree = rtree(ntaxa(newps.ass), rooted=TRUE,
tip.label=taxa_names(newps.ass))plot(random_tree)


mergednewps.ass <- merge_phyloseq(newps.ass, mk2sampledata , random_tree)


## pruning the controls



controlsremovedps <- subset_samples(mergednewps.ass,
SampleType %in%c("Dead_Maerl_Exterior",
"Dead_Maerl_Exterior", "Dead_Maerl_Exterior",
                                    "Dead_maerl_Total"  ,  "Dead_maerl_Total"  ,
"Dead_maerl_Total"  ,
```

```
                                        "Dead_Seawater" , "Dead_Seawater" ,
"Live_Maerl_Exterior" ,"Live_Maerl_Exterior",
                                        "Live_Maerl_Exterior" ,
"Live_Maerl_Total" ,"Live_Maerl_Total" ,  "Live_Maerl_Total",
                                        "Live_Seawater"    ,  "Live_Seawater"   ,
"Live_Seawater",
                                        "Sediment"   ,      "Sediment",
                                        "Sediment"      ,   "Sediment"    ,     "Sediment"  ,
"Sediment"
))



##ignore


pvalue.ttest
<- c()
pvalue.wilco
xon <- c()


for (taxa in test.taxa) {
 # Create a new data frame for each taxonomic group
 df <- data.frame(Abundance = abundances(controlsremovedps)[taxa,],
            Log10_Abundance = log10(1 +
            abundances(controlsremovedps)[taxa,]),Group =
            meta(controlsremovedps)$SampleType)


 pvalue.ttest[[taxa]] <- t.test(Log10_Abundance ~ Group, data = df)$p.value
 pvalue.wilcoxon[[taxa]] <- wilcox.test(Abundance ~ Group, data = df)$p.value
}
# Arrange the results in a
data.frame pvalues <-
data.frame(taxon = test.taxa,
              pvalue.ttest =
              pvalue.ttest,
```

```
                pvalue.wilcoxon =
                pvalue.wilcoxon)


## continue




##continue
## converting to centered log-ratios
(ps_clr <-
microbiome::transform(controlsremovedps, "clr"))
phyloseq::otu_table(ps_clr)[1:5, 1:5]

## creating a PCA plot using 3 pc`s

ord_clr <-
phyloseq::ordinate(ps_clr, "RDA")
phyloseq::plot_scree(ord_clr) +
geom_bar(stat="identity", fill =
"blue") +
  labs(x = "\nAxis", y = "Proportion of Variance\n")

clr1      <-      ord_clr$CA$eig[1]      /
sum(ord_clr$CA$eig)       clr2       <-
ord_clr$CA$eig[2]                    /
sum(ord_clr$CA$eig)       clr3       <-
ord_clr$CA$eig[3]                    /
sum(ord_clr$CA$eig)

pcacontrolsremoved <- phyloseq::plot_ordination(controlsremovedps, ord_clr, color
="SampleType") +
```

```
geom_point(size =

2) + coord_fixed(

clr2 / clr1 )

pcacontrolsremoved


## creating NMDS


ord_clrNMDS <- phyloseq::ordinate(ps_clr, method = "NMDS", distance = "unifrac")

clrnmds <- plot_ordination(ps_clr, ord_clrNMDS, color = "SampleType", title = "CLR
unifrac") +geom_line() + geom_point(size = 4)

clrnmds


## running stats



clr_dist_matrix <- phyloseq::distance(ps_clr, method = "jaccard")

vegan::adonis(clr_dist_matrix ~ phyloseq::sample_data(ps_clr)$SampleType)
dispr <- vegan::betadisper(clr_dist_matrix,

phyloseq::sample_data(ps_clr)$SampleType)permutest(dispr, pairwise = TRUE)



## Top 25 taxa from all the data

top20 <- names(sort(taxa_sums(finalps.ass), decreasing=TRUE))[1:25]

finallps.top20 <- transform_sample_counts(finalps.ass, function(OTU)

OTU/sum(OTU))finallps.top20 <- prune_taxa(top20, finalps.top20)



## dual bar chart of class and order top 25

plot_bar(FinalAllps.top25, "Class", fill="Order", facet_grid=~SampleType)


##Dendrogram chart

d <- distance(ps_rel_abund, method="unifrac",

type="samples")dendrogram <- hclust(d,
```

```
method="complete") plot(dendrogram)

Newdistance <- distance(controlsremovedps, method="unifrac",
type="samples")newdendrogram <- hclust(Newdistance , method =
"complete") plot(newdendrogram)



## testing for evenness and alpha normality

mergednewps.ass.even <- evenness(mergednewps.ass,
index = "all")kable(head(mergednewps.ass.even))

mergednewps.ass.meta <- meta(mergednewps.ass)
mergednewps.ass.meta$simpson <-
mergednewps.ass.even$simpson
hist(mergednewps.ass.meta$simpson)
shapiro.test(mergednewps.ass.meta$simpson)
qqnorm(mergednewps.ass.meta$simpson)

## richness and anova

richnessplot <- plot_richness(mergednewps.ass, color = "SampleType", x =
"SampleType",measures =
                c("Shannon"))
richnessplot <- richnessplot + geom_boxplot(aes(fill =SampleType), alpha=0.2)
richnessplot + theme(panel.grid.major = element_blank(), panel.grid.minor =
        element_blank(),panel.background = element_blank(), axis.line =
        element_line(colour = "black"))
plot(richnessplot)
alpha.diversity <- estimate_richness(finalps.ass, measures = c("InvSimpson",
                                "Shannon", "Chao1"))
head(alpha.diversity)
```

```
##shannon


GP <- prune_species(speciesSums(mergednewps.ass) > 0, mergednewps.ass)


shannonrichnessplot <- plot_richness(controlsremovedps, color =
"SampleType", x ="SampleType", measures =

                        c("Shannon"

                         ))

shannonrichnessplot <- shannonrichnessplot + geom_boxplot(aes( fill =
SampleType),alpha=0.2

                                        ) + stat_summary(fun.y=mean, geom="point", shape=20,
size=14, fill="SampleType")


plot(shannonrichnessplot)

alpha.diversity <- estimate_richness(GP, measures = c(

                                    "Shannon"))

alpha.diversity
```

```
#shannon anova

data <- cbind(sample_data(mergednewps.ass),

alpha.diversity)Shannon.anova <- aov(Simpson ~

SampleType, data) summary(Shannon.anova)

shannontukeyhsd <-

TukeyHSD(Shannon.anova)

shannontukeyhsd

#InvertSimpsons anova

InvertSimp.anova <- aov(InvSimpson ~

SampleType, data)summary(InvertSimp.anova)

## chao1 anova

Chao1.anova <- aov(Chao1 ~

SampleType, data)

summary(Chao1.anova)


## NMDS plots


# maerl NMDS with anova

MaerlNDMSordinate <- ordinate(maerlps, method = "NMDS", distance = "jaccard",
trymax =100)

MaerlNMDS <- plot_ordination(pslucky, MaerlNDMSordinate, color =

"SampleType")MaerlNMDS <- MaerlNMDS + geom_line() + geom_point(size

= 4)

MaerlNMDS


# maerl pca

MaerlPCASordinate <- ordinate(maerlps, method = "PCoA", distance =

"jaccard") MaerlPCA <- plot_ordination(pslucky , MaerlPCASordinate,

color = "SampleType")MaerlPCA <- MaerlPCA + geom_line() +

geom_point(size = 4)

MaerlPCA
```

```
#stats for dispersion and for
centroid Maerldis <-
distance(maerlps, "jaccard")
Maerldf <- as(sample_data(maerlps), "data.frame")
Maerlgroups <-
Maerldf[["SampleType"]]
Maerlgroups
Maerlmod <- betadisper(Maerldis,
Maerlgroups )anova(Maerlmod)
Maerlmod.HSD <-
TukeyHSD(Maerlmod)
Maerlmod.HSD
Maerl_adonis <- adonis(Maerldis ~ SampleType, data = Maerldf)
Maerl_adonis

#stats for dispersion and for centroid
BRAY curtaisMaerldisbray <-
distance(maerlps, "bray")
Maerldfbray <- as(sample_data(maerlps),
"data.frame")Maerlgroupsbray <-
Maerldf[["SampleType"]] Maerlgroupsbray
Maerlmodbray <- betadisper(Maerldisbray,
Maerlgroupsbray )anova(Maerlmodbray)
Maerlmod.HSDbray <-
TukeyHSD(Maerlmodbray)
Maerlmod.HSDbray
Maerl_adonisbray <- adonis(Maerldisbray ~ SampleType, data = Maerldfbray)
Maerl_adonisbray

#All NMDs with anova
NMDSordinate <- ordinate(finalps.ass, method = "NMDS", distance = "jaccard", try
= 10000)NMDS <- plot_ordination(pslucky, NMDSordinate, color = "SampleType",
```

```
title = "jaccard") NMDS <- NMDS + geom_line() + geom_point(size = 4)

NMDS



RealNMDSordinate <- ordinate(ps_rel_abund, method = "NMDS", distance = "unifrac")

RealNMDS <- plot_ordination(ps_rel_abund, RealNMDSordinate, color =
"SampleType", title ="unifrac")

RealNMDS <- RealNMDS + geom_line() + geom_point(size = 4)
RealNMDS



newPCAordinate <- ordinate(mergednewps.ass, method = "NMDS", distance = "unifrac")

newPCA <- plot_ordination(mergednewps.ass, newPCAordinate, color =
"SampleType", title ="unifrac")

newPCA <- newPCA + geom_line() +

geom_point(size = 4)newPCA



newjacNMDSordinate <- ordinate(mergednewps.ass, method = "NMDS", distance = "jaccard")

newNMDSjac <- plot_ordination(mergednewps.ass,
newjacNMDSordinate, color ="SampleType", title = "jaccard")

newNMDSjac <- newNMDSjac + geom_line() +

geom_point(size = 4)newNMDSjac



#stats for dispersion and for

centroid Alldis <-

distance(ps_rel_abund,

"unifrac")

Alldf <- as(sample_data(ps_rel_abund),

"data.frame")Allgroups <-

Alldf[["SampleType"]]

Allgroups

Allmod <- betadisper(Alldis, Allgroups,bias.adjust = TRUE, type = "centroid")
```

```
Allmo
d
anov
a(All
mod)
permutest(Allmod,
        pairwise = TRUE,
        permutations = 999)
Allmod.HSD <-
TukeyHSD(Allmod)
Allmod.HSD
All_adonis <- adonis(Alldis ~ SampleType, data = Alldf)
All_adonis
## controls removed
conAlldis <- distance(controlsremovedps, "unifrac")
conAlldf <- as(sample_data(controlsremovedps),
"data.frame")conAllgroups <-
conAlldf[["SampleType"]]
conAllgroups
conAllmod <- betadisper(conAlldis, conAllgroups, bias.adjust
= TRUE)
conAllmo
d
anova(co
nAllmod)
conAllmod.HSD <-
TukeyHSD(conAllmod)
conAllmod.HSD
conAll_adonis <- adonis(conAlldis ~ SampleType, data =
conAlldf)conAll_adonis
```

```
##
betastats_unifrac <- phyloseq::distance(controlsremovedps, method =
"unifrac")sampledf <- data.frame(sample_data(controlsremovedps))
adonis(betastats_unifrac ~ SampleType, data = sampledf)
TukeyHSD(betastats_unifrac)

###



scoresTukey <-
scores(Allmod.HSD)
scoresTukey
scoreAll <-
scores(Allmod)
score
  MDSrotate(NMDS)
## non-weighted unifrac
newAlldis <- distance(mergednewps.ass, "unifrac")
newAlldf <- as(sample_data(mergednewps.ass),
"data.frame")newAllgroups <-
newAlldf[["SampleType"]]
newAllgroups
newAllmod <- betadisper(newAlldis,
newAllgroups )anova(newAllmod)
newAllmod.HSD <-
TukeyHSD(newAllmod)
newAllmod.HSD
newAll_adonis <- adonis(newAlldis ~ SampleType, data =
newAlldf)newAll_adonis
```

```
##weight unifrac

newAlldisweighted <- distance(mergednewps.ass, "wunifrac")

newAlldfweighted <- as(sample_data(mergednewps.ass),

"data.frame")newAllgroupsweighted <-

newAlldfweighted[["SampleType"]] newAllgroupsweighted

newAllmodweighted <- betadisper(newAlldisweighted,

newAllgroupsweighted )anova(newAllmodweighted)

newAllmod.HSDweighted <-

TukeyHSD(newAllmodweighted)

newAllmod.HSDweighted

newAll_adonisweighted <- adonis(newAlldisweighted ~ SampleType, data =

newAlldfweighted)newAll_adonisweighted



## using deseq2



controlsremoveddeseq2 <- phyloseq_to_deseq2(controlsremovedps, ~

SampleType)dds <- DESeq(controlsremoveddeseq2)

dds

resultsdeseq2 <- results(dds) ## i think this is the problem
head(resultsdeseq2)


res = results(deseqtested)

res = res[order(res$padj,

na.last=NA), ]alpha = 0.01

sigtab = res[(res$padj < alpha), ]

sigtab = cbind(as(sigtab, "data.frame"),
as(tax_table(controlsremovedps)[rownames(sigtab), ],"matrix"))

head(sigtab)
```

```
##heat map

names <- sample_names(ps_rel_abund)

gpt1 <- prune_taxa(names(sort(taxa_sums(ps_rel_abund),TRUE)[1:50]), ps_rel_abund)

plot_heatmap(gpt1, sample.label = "SampleType", taxa.label = "Family",
low="#000033",high="#66CCFF")

plot_bar(gpt1, "Family", fill="Genus", facet_grid=~SampleType)

 gpt2 <- prune_taxa(names(sort(taxa_sums(ps_rel_abund),TRUE)[66:120]), ps_rel_abund)

plot_heatmap(gpt2, "NMDS", "bray", "SampleType", "Family",
low="#000033",high="#66CCFF", sample.order = names)


plot_heatmap(gpt, sample.order = "SampleType" ,sample.label = "SampleType",
taxa.label ="Order")

plot_heatmap(gpt, method = "NMDS", distance = "unifrac", sample.order = "SampleType"
,sample.label = "SampleType", taxa.label = "Family",)

plot_heatmap(gpt, sample.order = "SampleType" ,sample.label = "SampleType",
taxa.label ="Genus")

## making the Summarise_taxa function

phyloseq_summarize_taxa <- function(psdata,

taxonomic.ranks =

                          rank_names(psdata)) {

 if(length(taxonomic.ranks) > 1) {

 names(taxonomic.ranks) <- taxonomic.ranks

  llply(taxonomic.ranks, phyloseq_summarize_taxa, psdata = psdata)
 } else {

  taxa <- as(tax_table(psdata)[, taxonomic.ranks], 'character')

  sum_tax_table <- summarize_taxa(as(otu_table(psdata),

  'matrix'), taxa)phyloseq(otu_table(sum_tax_table,

  taxa_are_rows = TRUE),

        sample_data(psdata, FALSE))

 }

}


## making and exporting the rank table

Phylumglom <- tax_glom(finalps.ass, taxrank =
```

```
'Phylum')Phylumdat <- psmelt(Phylumglom)
Phylumdat$Phylum <- as.character(Phylumdat$Phylum)
Phylum_abundance <- aggregate(Abundance~Sample+Phylum, Phylumdat,
FUN=sum)Phylum_abundance <- cast(Phylum_abundance, Sample ~ Phylum)
Phylum_abundance

write.xlsx(Phylum_abundance , "c:/Phylum_abundance.xlsx")

##making proteobacteria order rank table
proOrderps <- subset_taxa(finalps.ass, Phylum == "Proteobacteria")

proOrderglom <- tax_glom(proOrderps, taxrank =
'Order')proOrderdat <- psmelt(proOrderglom)
proOrderdat$Phylum <-
as.character(proOrderdat$Order)
proOrder_abundance <- aggregate(Abundance~Sample+Order, proOrderdat,
FUN=sum)proOrder_abundance <- cast(proOrder_abundance, Sample ~ Order)
proOrder_abundance

write.xlsx(proOrder_abundance , "c:/proOrder_abundance.xlsx")

##making all class rank tables
Classglom <- tax_glom(finalps.ass, taxrank
= 'Class')Classdat <- psmelt(Classglom)
Classdat$Class <- as.character(Classdat$Class)
Class_abundance <- aggregate(Abundance~Sample+Class, Classdat,
FUN=sum)Class_abundance <- cast(Class_abundance, Sample ~ Class)
Class_abundance

write.xlsx(Class_abundance , "c:/Class_abundance.xlsx")

##making all orders rank table
```

```
Orderglom <- tax_glom(finalps.ass, taxrank
= 'Order')Orderdat <- psmelt(Orderglom)
Orderdat$Order <- as.character(Orderdat$Order)
Order_abundance <- aggregate(Abundance~Sample+Order, Orderdat,
FUN=sum)Order_abundance <- cast(Order_abundance, Sample ~ Order)
Order_abundance

write.xlsx(Order_abundance ,
"c:/Order_abundance.xlsx")##making all genus
rank table

Genusglom <- tax_glom(finalps.ass, taxrank =
'Genus')Genusdat <- psmelt(Genusglom)
Genusdat$Genus <- as.character(Genusdat$Genus)
Genus_abundance <- aggregate(Abundance~Sample+Genus, Genusdat,
FUN=sum)Genus_abundance <- cast(Genus_abundance, Sample ~ Genus)
Genus_abundance

write.xlsx(Genus_abundance , "c:/Genus_abundancemk2.xlsx")



## making all Family rank table
Familyglom <- tax_glom(finalps.ass, taxrank =
'Family')Familydat <- psmelt(Familyglom)
Familydat$Family <- as.character(Familydat$Family)
Family_abundance <- aggregate(Abundance~Sample+Family, Familydat,
FUN=sum)Family_abundance <- cast(Family_abundance, Sample ~ Family)
Family_abundance
write.xlsx(Family_abundance ,
"c:/Family_abundance.xlsx")##Eular diagram for
P-valus
```

```
set.seed(2)



con <- c(A = 1, B = 1, C = 1, D = 1, E = 1, F = 1, G = 1, H = 1,

       "A&B" = 0.2, "B&C" = 0.2, "C&D" = 0.2, "D&E" = 0.2, "E&F" = 0.2,

       "F&G" = 0.2, "G&H" = 0.2)

plot(euler(con))



try <- c(DMT = 1, DMX = 1, LMT = 1, LMX = 1, Sediment = 1, Deadseawater = 1,
Liveseawater =1,

       SeawaterControl = 1, SedimentControl = 1,

      "LMX&DMX" = 0.1561544, "LMX&DMT" = 0.0046314, "LMX&Deadseawater" = 0.0157490,
"LMX&LMT" = 0.0232319,

      "LMX&Liveseawater" = 0.0512665, "LMX&SeawaterControl" = 0.0007677,
"LMX&Sediment" = 0.0002774,

      "LMX&SedimentControl" =

      0.0007677, )

plot(euler(try))



#stats taken from all sample tukey HSD test (mean distance-to-centroid)

try2 <- c(DMT = 1, DMX = 1, LMT = 1, LMX = 1, Sediment = 1, Deadseawater = 1,
Liveseawater =1,

       SeawaterControl = 1, SedimentControl = 1,

       "DMT&DMX" = 0.6346832,"LMX&DMX" = 0.1561544, "LMT&DMX" = 0.9745387,
       "Sediment&DMX" = 0.1525230, "LMT&DMT" = 0.9936452,
"Sediment&DMT" =0.9965959,

       "Liveseawater&Deadseawater" = 0.9809884,
"SeawaterControl&Deadseawater" =0.3896788,

       "SedimentControl&Deadseawater" = 0.3896788,"Liveseawater&LMX" =
0.0512665,"Sediment&LMT" = 0.7251771,

       "SeawaterControl&Liveseawater" = 0.896154,

"SedimentControl&SeawaterControl" = 1)plot(euler(try2))
```

```
newtry <- c(DMT = 1.5, DMX = 1.5, LMT = 1.5, LMX = 1.5, Sediment = 1.5,
Deadseawater = 1.5,Liveseawater = 1.5,

        SeawaterControl = 1.5, SedimentControl = 1.5,

        "DMT&DMX" = 0.1540409, "Deadseawater&DMX" = 0.2895668,
"LMX&DMX" =0.6077737,

        "LMT&DMX" = 0.9999781, "Liveseawater&DMX" = 0.6557489,
"Sediment&DMX" =0.7414697,

        "LMT&DMT" = 0.2810123, "Sediment&DMT" = 0.7481562,
"LMX&Deadseawater" =0.9951121,

        "LMT&Deadseawater" = 0.1703367, "Liveseawater&Deadseawater" =

        0.9914262, "LMT&LMX" = 0.3924130, "Liveseawater&LMX" = 1.0000000,

        "Liveseawater&LMT" =
0.4352241,

        "Sediment&LMT" = 0.9304507, "SedimentControl&SeawaterControl" = 1.0000000)



plot(euler

(newtry))

sfusf

save.image(file =  "new_data.R")


#######


otu <-

abundances(controlsremoved

ps)dist <- vegdist(t(otu))

anova(betadisper(dist,

meta$group))

sort(phyloseq::sample_sums(controlsremovedps))




#### fresh start
```

```
##running stats for beta diversity

metadf <- data.frame(sample_data(ps1.rel))

unifrac.dist <- UniFrac(ps1.rel, normalized = TRUE)

permanova <- adonis(unifrac.dist ~ SampleType, data
= metadf)permanova

ps.disper <- betadisper(unifrac.dist, metadf$SampleType)
permutest(ps.disper, pairwise = TRUE, permutations = how(nperm
= 99999))

hist(unifrac.dist)
shapiro.test(unifrac.dist) ## proof that my data is not normally distributed meaning
i can notuse permtests


### creating NMDS plot

ps1.rel <- microbiome::transform(controlsremovedps, "compositional") ## makes it
relativityabundance

ps1ord <- ordinate(ps1.rel, method = "NMDS", distance = "unifrac")

p <- plot_ordination(ps1.rel, ps1ord, color = "SampleType")
p <- p + geom_line() +
geom_point(size = 4)p

##rarfing my data - Seed
711
sample_sums(controlsre
```

```
movedps)

rarepscontrols <- rarefy_even_depth(mergednewps.ass,
sample.size =min(sample_sums(mergednewps.ass)))

rareps <- rarefy_even_depth(controlsremovedps,
sample.size =min(sample_sums(controlsremovedps)))

sample_sums(rareps)


#NMDS

rarpsord <- ordinate(rareps, method = "NMDS", distance =

"unifrac") rarepsnmdsplot <- plot_ordination(rareps, rarpsord,

color = "SampleType")rarepsnmdsplot <- rarepsnmdsplot +

geom_line() + geom_point(size = 4) rarepsnmdsplot

#stats for NMDS

raredis <- distance(rarepscontrols, "unifrac")

raredf <- as(sample_data(rarepscontrols),

"data.frame")raregroups <-

raredf[["SampleType"]]

raregroups

raremod <- betadisper(raredis,

raregroups )anova(raremod)

raremod.HSD <-

TukeyHSD(raremod)

raremod.HSD

rare_adonis <- adonis(raredis ~ SampleType, data

= raredf)rare_adonis


(rare_clr <-

microbiome::transform(rareps, "clr"))

phyloseq::otu_table(rare_clr)[1:5, 1:5]


raredisclr <- distance(rare_clr, "unifrac")
hist(raredisclr

)
```

```
shapiro.test(r
arediscIr)

##trying deseq2 again

diagdds = phyloseq_to_deseq2(controlsremovedps, ~
SampleType)diagdds = DESeq(diagdds, test="Wald",
fitType="parametric")
res = results(diagdds, cooksCutoff
= FALSE)res
alpha = 0.01
sigtab = res[which(res$padj < alpha), ]
sigtab = cbind(as(sigtab, "data.frame"),
as(tax_table(controlsremovedps)[rownames(sigtab), ],"matrix"))
head(sigtab)



### decontamination

FSr = transform_sample_counts(mergednewps.ass, function(x) x /
sum(x) )FSfr = filter_taxa(FSr, function(x) sum(x) > .0005, TRUE)

fsfrord <- ordinate(FSfr, method = "NMDS", distance =
"unifrac") nmdsfsfr <- plot_ordination(FSfr, fsfrord,
color = "SampleType") nmdsfsfr <- nmdsfsfr +
geom_point(size = 4)
nmdsfsfr

FSfrunifrac <- phyloseq::distance(FSfr, method = "unifrac")
vegan::adonis(FSfrunifrac ~
phyloseq::sample_data(FSfr)$SampleType)
dispr <- vegan::betadisper(FSfrunifrac,
```

```
phyloseq::sample_data(FSfr)$SampleType)permutest(dispr, pairwise =
TRUE)
```

```
sample_names(mergednewps.ass)

controlsps <- subset_samples(mergednewps.ass, SampleType %in% c("Sediment_Control",
"Seawater_Control"))


gpt <- prune_taxa(names(sort(taxa_sums(ps_rel_abund),TRUE)[1:65]), ps_rel_abund)

con <- prune_taxa(names(sort(taxa_sums(controlsps),TRUE)[1:65]), controlsps)

plot_heatmap(con, distance = "unifrac", sample.label = "SampleType", "Family",
low="#000033", high="#66CCFF")
```

Appendix 6:

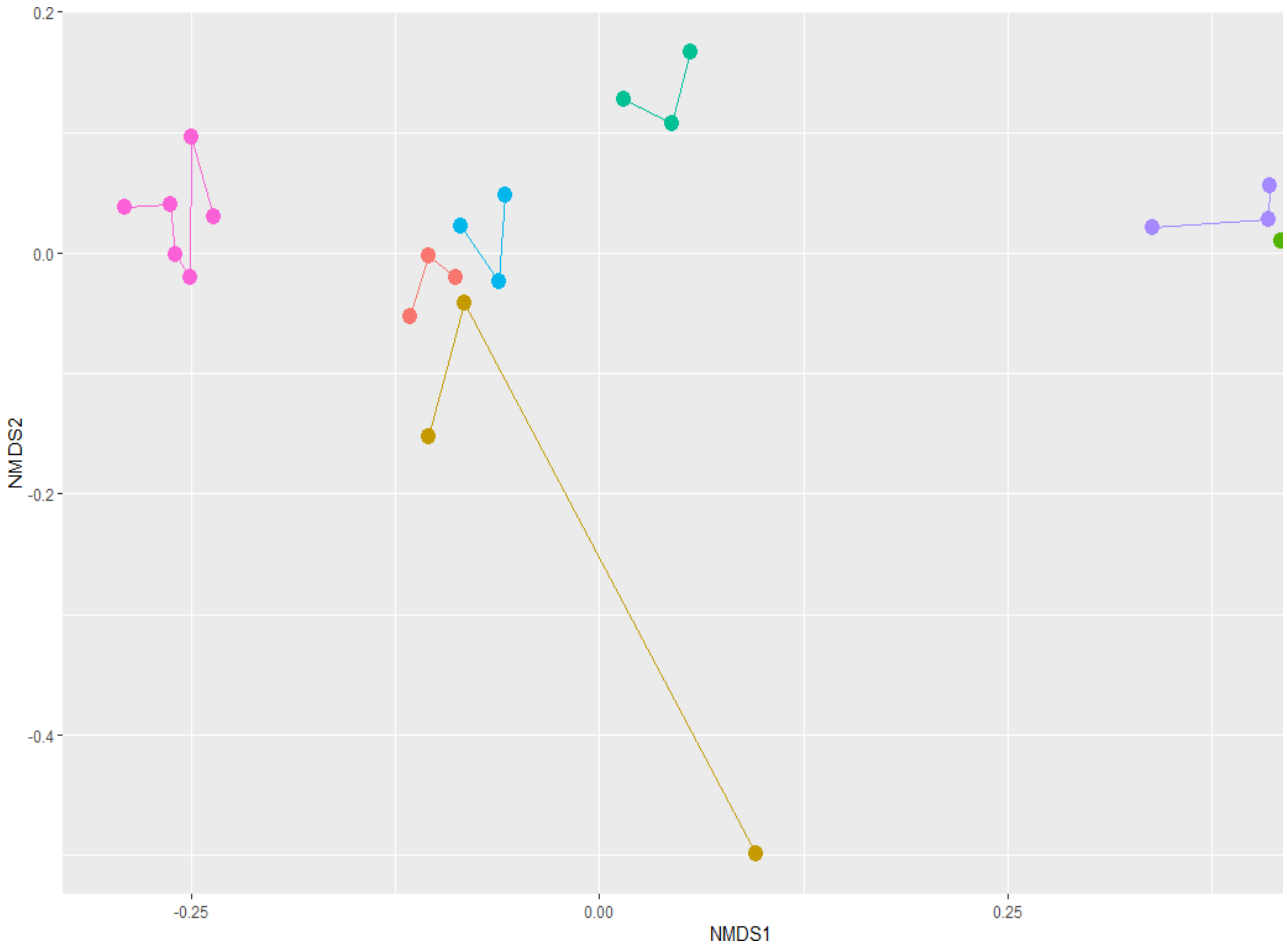


*Figure 7 NMDS with non weight unifrac distances – stress value = 0.0054*