Coursework Declaration and Feedback Form

The Student should complete and sign this part

Student Number: 2638310S	Student Name: Chenyao Sun
Programme of Study (e.g. MSc in Electronics and Electrical Engineering): MSc in Computer System Engineering	
Course Code: ENG5059P	Course Name: MSc Project
Name of **First** Supervisor: Umer Zeeshan Ijaz	Name of **Second** Supervisor: Stephen Thoms
Title of Project: Understanding microbial activity of biological and chemical dispersants of crude oil	

Declaration of Originality and Submission Information

I affirm that this submission is all my own work in accordance with the University of Glasgow Regulations and the School of Engineering requirements Signed (Student) : Chenyao Sun	
Date of Submission : 18/08/2022	

Feedback from Lecturer to Student – to be completed by Lecturer or Demonstrator	
Grade Awarded: Feedback (as appropriate to the coursework which was assessed):	
Lecturer/Demonstrator:	Date returned to the Teaching Office:

Understanding microbial activity of biological and chemical dispersants of crude oil

Chenyao Sun

2638310

Supervised by Dr Umer Zeeshan Ijaz & Dr Stephen Thoms

MSc Computer System Engineering

School of Engineering

Contents

Abstract

Many serious oil spills have occurred in frequent offshore oil transportation and exploitation. The oil can travel hundreds of miles, killing seabirds, mammals and other creatures on them. In addition, oil films formed on the sea surface can hinder gas exchange between the atmosphere and seawater, potentially impacting long-term climate change (Nancy et al., 2012). Oil is typically degraded by the activity of hydrocarbon-degrading (hydrocarbonoclastic) bacteria that exist in the oceans. Recently, understanding the role of these bacteria, particularly in response to the use of chemical surfactants, has gained popularity. Chemical surfactants are practical but not ecologically compatible and pose significant environmental risks. For example, petroleum hydrocarbons dominated by aromatic hydrocarbons, cannot be decomposed by aquatic organisms to form carcinogens after being enriched in the food chain. This project compares the performance of Finasol, a chemical dispersant used to treat offshore oil spills, and rhamnolipid, a natural biosurfactant that is more environmentally friendly, to reveal how microorganisms interact with the environment in these different treatments and how they affect the reaction of oil-degrading bacteria. Microbial community surveys through 16S rRNA amplicons are available for both types of surfactants and analyzed the dominant colonies under different treatment methods. Results found the dominant bacteria affecting the microbial community in different environments, which promoted understanding of the effects of rhamnolipid and Finasol on the natural marine microbial community. (The project is guided by the recent technological advancements in microbial in situ omics data analytics.)

Acknowledgements

I would like to thank my mentor Umer Zeeshan Ijaz for his guidance and technical support for the project. Thanks for him spending a lot of time in the weekly meeting explaining the theory and algorithm of microorganisms to me.

1 Introduction

1.1Background

Crude oil has driven the rapid development of the industry since its generation, but the environmental problems caused by it are also very serious. The offshore oil spill has the most serious and lasting harm to biology and the environment. In the 2010 oil drilling explosion in the Gulf of Mexico, about 1000 barrels of crude oil were leaked daily (Fredy et al., 2018), which also became the largest environmental disaster in the United States. The harm of oil pollution covers a wide range, and it harms birds, marine animals and plants, human beings and the marine environment. According to statistics, from 1952 to 1962, a total of 450000 seabirds died in the North Atlantic and North sea areas due to oil pollution (Sargent et al., 2011). The toxic components in the oil will kill the eggs and young fish that are heavily stuck by the oil film. Venezuela's oil spill killed over 95% of the shrimp larvae in the sea area within 24 hours. And the harmful substances in the oil will stay in the marine animals for a long time, and people who eat these marine products will also be affected.

The microbial method can be used as an effective method to control oil leakage. Some oil-degraded microorganisms decompose alkanes, aromatic hydrocarbons and other substances in the environment and convert them into bacterial cells, water and carbon dioxide. The chemical dispersant is the primary reaction tool to accelerate petroleum decomposition and improve microorganisms' degradation potential. Since the 1970s, Oil Spill Dispersants have developed rapidly. Hundreds of Oil Spill Dispersants exist worldwide (Kleindienst et al., 2015). The petroleum hydrocarbons, dominated by aromatic hydrocarbons in the early stage, could not be decomposed in aquatic organisms and had significant toxic effects on fish and shellfish aquatic organisms. They have gradually been replaced by natural surfactants extracted from raw materials such as vegetable oil and glycolipid.

This project mainly studies the dynamic changes and diversity of microbial communities under the action of chemical dispersant Finasol and biosurfactant rhamnolipid (a biological metabolite produced by Pseudomonas) and found the dominant bacterial species under different environmental influences. The microbial species and abundance information in these environmental samples can be obtained by 16S rRNA sequencing. Then annotated the sequenced sequence with the database to obtain the species information corresponding to the sequence.

1.2 Amplicon Sequence Variants

In principle, the sequence obtained by sequencing the full-length segment or part of the marker gene (16s rRNA in bacteria) in the microbial genome can be annotated to obtain the corresponding species information(Schlomann et al., 2019). However, there are tens of thousands of sequences per sample obtained by sequencing, and the workload is enormous. However, there are more than tens of thousands of sequences

per sample obtained by sequencing, and the calculation amount is tremendous. Moreover, there are a small number of probability sequencing errors in the amplification and sequencing process of the marker gene, which will reduce the accuracy of subsequent analysis. Therefore, OTU (Operational Taxonomic Unit) is introduced in diversity analysis to avoid the above problems effectively. In the early stage of OTU clustering, the UPARSE algorithm was used to extract non-repetitive sequences, which reduced the redundant calculation amount in the analysis process(Booth et al., 2020). All sequences were divided according to different similarity levels in the OTU method. After data extraction, the sequences are clustered according to the similarity between the sequences.

In recent years, more studies have adopted ASV to reduce the impact of sequencing errors and replace OTU. The default step of OTU clustering with 97% sequence similarity in the UPARSE algorithm will cover up some sequences with sequencing errors, resulting in an inaccurate abundance of some OTUs (Alisa et al., 2022). In addition, the actual variation information of sequences will also be covered by the relatively wide similarity threshold. DADA2 algorithm combines the sequencing accuracy, uses the divisive partitioning algorithm for final clustering, and calculates the p value. This clustering result is called ASV, similar to OTU clustering with 100% similarity, which improves the clustering accuracy (Guckenheimer et al., 2013). In this project, millions of sequences obtained by 16S rRNA sequencing were corrected and filtered based on the DADA2 algorithm. The statistical software R-Studio was used to perform diversity analysis, null model test, differential analysis and regression analysis on these ASV samples.

1.3Problem definition

The factors affecting the change of microbial community are very complex, including pH, temperature, time, and the richness of sampling. Therefore, it is difficult to judge whether the change of some bacterial species in the sample is related to the action of the two surfactants. This work analyzes and compares microbial community samples in multiple dimensions, such as time and environmental selection .

1.4Case Study: The bacterial community in the Faroe-Shetland Channel

Faroe Shetland channel (FSC) is a marine nature reserve in the North Atlantic between the Shetland Islands and the Faroe Islands. The channel covers an area of 5278 km^2 and has a rich history of oil exploitation(Logan et al., 2018). The relatively warm North Atlantic seawater is mixed with the subzero deep water from the Norwegian Sea, causing the region's rich biodiversity and complex ecological cycle (Corvec, 2018). This means that if an oil spill occurs here, it will seriously impact the ecological community, commerce and fishery. Moreover, the harsh and cold weather conditions in FSC significantly increase the difficulty of handling oil spill accidents.

To effectively deal with the oil spill accidents that may occur in this area in the future, the FSC ecological community, under the action of two oil spill response tools, Finasol and rhamnolipid, can be simulated based on the ecological null model to reveal the key taxa that affect the microbial community changes, to find an effective solution.

1.5Aims and Objectives

The general objective of this study is to compare the effects of chemical dispersants Finasol and biosurfactants rhamnolipid on the microbial community in the Faroe Shetland Channel, reveal the interaction between microorganisms and the environment in different treatments and investigate:

1. How Finasol and rhamnolipid affect the response and community diversity of degrading bacteria
2. Dynamic changes of microbial community with time in different treatments
3. The dominant bacterial taxa in treatment samples

2 Method

2.1 Data collection

FSC surface seawater was used as a sample to evaluate the dynamic changes of microbial communities in the influence of Finasol and rhamnolipid. These samples are divided into six groups according to different treatment methods. First, there are three main water accommodated fractions (WAF, CEWAF, BEWAF), which are prepared by mixing a specified amount of seawater, crude oil and biological or chemical surfactants in a container and have the same volume of seawater and crude oil (seawater: 1500ml, crude oil: 120ml). WAF only contains seawater and crude oil. CEWAF is a chemically enhanced water accommodated fraction containing Finasol, seawater and crude oil. The chemical surfactant and crude oil are prepared in a ratio of 1:20. BEWAF is a biological enhanced water accommodated fraction prepared from seawater, crude oil and rhamnolipid. Then a control group was set up for the two surfactants to observe their effects on the microbial community. One is SWD containing seawater and Finasol. The other is SWBS containing seawater and biosurfactant rhamnolipid. All the experimental groups should be mixed for more than 48 hours to allow the oil to be dispersed entirely and precipitated to establish the microenvironment. Finally, a sample only containing seawater without any reagent addition should be set as a control. For the six treatments of the sample (WAF, CEWAF, BEWAF, SWD, SWBS, SW), each group should be prepared twice to analyze the changes in hydrocarbon composition caused by biodegradation. These experimental groups were sampled on days 3, 7, 14 and 28 of culture

2.2 High-Throughput Sequencing

The collected DNA extracts were sequenced using the second-generation sequencing technology Illumina, and the two-step amplification procedure amplified the 16S rRNA sequence. High throughput sequencing was performed on the microbial population to analyze the gene composition and diversity of the microbial population in the environment.

2.3 DADA2 algorithm with the bioinformatics pipeline QIIME2

In order to obtain detailed biological information on microbial community species, the 16srna gene sequence obtained by sequencing was processed through the QIIME2 bioinformatics pipeline, and the DADA2 algorithm was used as a plug-in unit to correct the duplication and error in the gene sequence (Schlomann et al., 2019). Divisive Amplicon Denoising Algorithm 2 is an R package for modelling and correcting amplicon sequencing errors on multiple sequencing platforms. In amplicon analysis, the algorithm can accurately infer the sample sequence and find the difference between single nucleotides (Phadnis et al., 2018). It speculates whether the

amplicon sequence is from the template by constructing an error rate model, takes the error model of its own data as a parameter, and calculates the probability of various transpositions.

Firstly, cut each sequence to the specified length according to the approximate sequence length, the minimum quality score and other standards, detach the sequences with shorter lengths, and then remove the duplication of all sequences of all samples, preserve the unique sequence set, and calculate the average quality score of each base and the abundance of each unique sequence. Due to sequencing errors, one OTU sequence may detect multiple sequences composed of a correct sequence set with high abundance and some wrong sequences with low abundance (Dizay et al.,2017). DADA2 uses the information of sequence abundance, quality score, and the relationship between sequences to correct the wrong base of sequencing, speculate the actual sequence, and achieve the purpose of denoising. In addition, it is also necessary to remove chimeras for the detection sequence. If a sequence has a small abundance and is similar to multiple sequences, it is considered chimeras in the algorithm(Aubin et al., 2020). After removing the influence of these wrong sequences, it is spliced with a 20bp overlap to generate the ASV table.

2.4 Statistical analysis

The statistical software R-Studio 4.2.0 was used for analysis. The genes of the collected samples were sequenced based on the DADA2 algorithm, and a comprehensive evaluation was carried out regarding diversity, environmental setting and time. Using microbiome package implement diversity, subset and core microbiome analysis. The R script obtained by analysis can be obtained in the attached file.

2.4.1 Diversity analysis: Alpha Diversity

The diversity analysis of microbial communities mainly includes alpha diversity, beta diversity, species composition and relatedness index(Sari, 2013). Alpha diversity analysis is an essential part of ecological biodiversity. Alpha diversity is the diversity in a specific region or ecosystem, which refers to how many different sequences exist in a single sample(Kamzolova et al., 2014). It is a comprehensive indicator of richness and evenness. It is mainly related to two factors: the number of species (richness) and the diversity and the uniformity of individual distribution in the community. Common alpha diversity indexes include richness, Shannon, Simpson and Pielou.

Richness is the number of species in the community. Shannon index comprehensively considered the richness and evenness of the community. The higher the Shannon index value, the higher the diversity of the community(Gao et al., 2019). Simpson is one of the indices used to estimate the diversity of microorganisms in samples. It was proposed by Edward Hugh Simpson and is often used to describe a region's biodiversity in ecology quantitatively. Simpson diversity index refers to the probability that two species sampled randomly in a community are different. The greater the Simpson index, the lower the community diversity. Pielou index is an

estimate of uniformity. In addition, some indexes related to community richness, such as the ACE index, which is used to estimate the number of OTUs in the community, are one of the most commonly used indexes in ecology to estimate the total number of species(Araújo et al., 2017). It included all OTUs with sequence quantity less than 10 to estimate the number of species in the community. The higher the ACE index, the higher the richness of the community.

2.4.2 Diversity analysis: NRI and NTI

The evolutionary relationship of microorganisms in the sample can be analyzed based on the phylogenetic alpha diversity measurements. NRI is defined as the nearest relatedness index. NTI is defined as the nearest taxon index. These two indexes can be calculated by MPD (Mean Phylogenetic Distance), the average evolutionary distance between all species pairs in the community and MNPD (Mean Nearest Phylogenetic Distance), the distance between any species in the community and the species closest to itself in the community(Reiniati et al.,2017). If there are n ASV species, the n MNTD values will be calculated. The smaller index values indicated the closer the genetic relationship between species.

When the NRI and NTI index is greater than 0, the species in the sample have a higher pedigree aggregation degree than the communities randomly sampled from the species pool. It indicates that the community structure tends to gather species with similar relatives. When the NRI and NTI indexes are less than 0, the relationships between sample species are more dispersed than the communities randomly sampled from the species pool. It indicates that the community structure tends to gather the species with distant relationships (Booth& Smith, 2020). NRI tends to calculate the overall similarity of all species, while NTI pays more attention to the influence between the nearest species.

2.4.3 Diversity analysis: Beta Diversity

Different from alpha diversity in analyzing a single community structure, beta diversity is often used to compare the differences between different ecosystems, reflecting the heterogeneity of biological species caused by the environment(Lee et al., 2014). The microbial community structure changes under different treatments, and environmental factors were analyzed. Beta diversity can be revealed by calculating the distance index. The typical distance indexes are Jaccard, Bray Curtis and UniFrac.

Jaccard (Jaccard similarity index) is calculated by dividing the number of species shared by two samples by the sum of all species in the sample(Abdellatif et al., 2016). Traditionally, Euclidean distance is often used in multivariate analysis. However, in the analysis of species data, Euclidean distance does not perform well because it treats the double zero phenomena as the same existence, which will narrow the distance between two communities sharing few species, resulting in the inaccurate judgment of community diversity(Pauly et al., 2016). Double zero refers to the absence of some species in the two plots compared when calculating community similarity. Specifically, the simultaneous loss of a species in two quadrats in the community cannot be the basis for the similar composition of these two quadrats because the

reasons for the loss may be completely different. Secondly, in the species matrix, the number of unexplained double zeros depends on the number of species. Therefore it will also increase significantly with the number of rare species detected. In order to effectively reduce the error caused by the double zero problems, this paper chooses to use the asymmetric Bray Curtis distance and UniFrac distance. Among them, Bray Curtis distance only considers species abundance and evenness. The algorithm is relatively simple but does not consider the evolutionary relationship between OTUs. For the UniFrac distance, compares OTUs according to the phylogenetic tree and classifies OTUs according to 16S sequence information. The values of these distances range from 0 to 1. The value is 0, which indicates no diversity difference between the two samples. The value close to 1 indicates a greater diversity difference between the two samples.

These distance algorithms can be divided into weighted and unweighted. The unweighted distance only considers the change of species, so in the results, 0 only indicates that the OTU species between the two microbial communities are consistent(Corvec, 2018). While the weighted distance considers the changes in species' existence and abundance simultaneously. 0 in the result indicates that the species and quantity of OTU are consistent among communities(William et al., 2012). In the application of the actual environment, the influence factors are more complex. Suppose the relationship between the control and experimental treatment groups is to be studied. In that case, the composition of the community generally does not change significantly after treatment, but the abundance of the community may change greatly. Therefore, the weighted method that considers both the richness and number of species is more suitable.

Based on the clustering relationships obtained by these distance algorithms, images can be constructed using principal coordinate analysis (PCoA). It is a non-constrained data dimension reduction analysis method to study the similarity or difference of samples (Adamberg et al., 2015). Firstly, the collected eigenvalues and eigenvectors are sorted. Then the essential eigenvalues ranked at the top are selected and expressed in the coordinate system to realize the quantitative transformation of qualitative data. Its coordinate axis indicates the probability that this factor affects the microbial community. The closer the distance between the data, the smaller difference in the community composition.

2.4.4 Null Model Method: Calculating Quantitative Process Estimate and Normalized Stochasticity Ratio

The null modelling method was used to find the ecological driving factors that affect the marine microbial community composition(Cirio et al., 2018). Null model is a method for quantitative comparison of assembly processes based on beta diversity algorithm, calculated by β MNTD, β NTI, Bray Curtis and other indexes and displays a visualization result, such as the beta diversity (β RC) used to distinguish the system deterministic and stochastic occurrence rate (Raup Crick). The project adopts two methods to conduct null model analysis on marine microbial communities. The first method is quantitative process estimate (QPE). QPE aims to find the relative

proportion of different assembly processes in the system, mainly including five indicators, homogeneous selection, variable selection， homogeneous dispersal, dispersal limitation, and undominated process (Jabbarzadeh et al., 2018). Homogeneous selection and variable selection are called selection pressure. They belong to deterministic processes, where the presence of one or more conditions in the environment affects the community structure of microorganisms. Homogeneous selection implies a single environmental influence community, and variable selection means that there are multiple environmental influence community assemblies. In the stochastic process, homogeneous dispersal refers to microbes from one sample are continuously appearing in another sample due to human or environmental reasons.
In contrast, Dispersal limitation is the proportion in which no diffusion occurs. In addition, there are some unexplained factors (Undominated). That means the environment doesn't have a role to play, now if the environment doesn't have a role to play.
At the data level, QPE can observe the degree to which βMNTD (β-mean-nearest taxon distance) deviates from the mean value of the null distribution after 999 randomizations and is evaluated using βNTI (β-Nearest Taxon Index). If the observed βMNTD value is significantly greater than the null expectation (βNTI > 2 or βNTI < −2), the assembly process of the community is more deterministic, dominated by selection pressure. βNTI greater than 2 represents variable selection, and less than -2 is homogeneous selection (Johnke et al., 2018). If the observed results are not significantly different from the null distribution, it means that the environment is more inclined to randomness. If the value of βNTI is greater than 0.95, there is a dispersal limitation, and if the value is less than -0.95, it means homogenizing dispersal. If it is between 0.95 and -0.95, it represents some random processes (Johnke et al., 2018). The second method is to calculate the system's normalized stochasticity ratio (NST) to judge whether the community is prone to the competitive exclusion or environmental Filtering. NST uses the incidence-based Jaccard, which computes the actual contribution of determinism to randomness, and the abundance-based similarity and Ruzicka metrics, using the proportional-proportional (PP) and proportional-fixed (PF) algorithms. Where the value of NST is greater than 0.5, the system is stochastic, and less than 0.5 is deterministic.

2.4.5 The top 25 most abundant taxa of the sample

The top 25 most abundant bacterial species were calculated at the genus level. The different bacterial species that dominated BEWAF, CEWAF and pure oil treatments and their abundance differences could be visually compared.

2.4.6 Determine the key factors affecting ecological driving: DESeq2 Differential analysis

After obtaining samples with a high abundance of bacterial species, differences in microbial community composition can be investigated in detail through differential analysis based on DESeq2. Difference analysis uses hypothesis testing to determine whether there is a significant difference between two groups of data(Kiørboe et al.,

2014). For data with a known distribution, use parametric tests to make the results more precise. Firstly, the reads count matrix was established to standardize the data obtained by sequencing analysis, and then the discrete degree of genes was estimated by logarithmic transformation. Then the final difference analysis results, including p-value and log2Foldchange could be obtained. The log2FC reflects the expression differences between different groups. It consists of two parts, one is the difference between the samples themselves, and the gene expression itself between repeated samples has a certain degree of difference. The other part is the difference caused by different groups or experimental conditions (Kamzolova et al., 2014). After removing self-differences, the data that compares only conditional and grouped differences could be obtained. The project uses the DESeq software package to perform differential analysis of data across different conditions, including comparisons in sampling time and treatments. For the final result, plot the average on the x-axis and plot the difference on the y-axis. If the expression level has a minimum of 2 log-fold change from the mean, then the microbe between the two conditions is significantly different.

2.4.7 Determine the key factors affecting the ecological driving: Investigating the core microbiome in Finasol and Rhamnolipid treatment

To specifically analyze the microbial community structure and dominant bacterial species in different treatments, all WAFs can be analyzed by the microbiome software package. The project sets the minimum prevalence value to 0.85, meaning that microorganisms with a probability of more than 85% in the sample belong to the core microorganisms.

2.4.8 Determine the key factors affecting ecological driving: Subset regression

Previous methods have assessed the dynamics of microbial communities from multiple dimensions (Diversity, Richness, NRI and NTI). Subset regression can detect whether these different treatment methods have a positive or negative effect on the kinship and diversity of the system. The regression model is fitted into a linear equation through the data set, which has a dependent variable and several independent variables, assigns weights to each variable, obtains the beta coefficient, and judges the relationship between the diversity index, which has multiple Regression analysis of independent variables is called multiple regression analysis((Reiniati et al.,2017)). A good regression model needs to control the number of independent variables. In multivariate regression analysis, in order to obtain a concise and effective model, the variable screening will be performed. Subset regression is a class of methods for independent variable selection for multiple linear regression equations. Consider all the subsets in the total independent variables, find the optimal model among them, for example, n independent variables will fit 2n-1 subset regression equations, and then use the statistics of the regression equations as a criterion (Cross-validation Errors) to select them.

The data obtained by regression analysis are not all reliable. Only the part with a p-value < 0.05 needs to be retained. Using the same data set for both training and

model error estimation, the result of error estimation is wildly inaccurate, which can be solved by Cross-validation. Cross-validation is a method for model selection by estimating the model's generalization error, including leave-one-out and leave-P-out. Leave-one-out is to take one sample from the sample set of N each time as the verification set, and the remaining N-1 samples as the training set, and repeat N times(Eurico et al., 2016). Finally, the N results are averaged as the generalization error estimate. Leave-P-out is similar to leave-one-out but leaves P samples at a time. Each time, from the N sample set, P samples are taken as the validation set, the remaining N-P samples are used as the training set, and finally, the average N results are used as the generalization error estimation.

2.4.9 Determine the key factors affecting ecological driving: CODA GLMNET

A LASSO regression model was established using the Glmnet software package to analyze the increase and decrease of taxa in different treatments. LASSO regression performs variable selection and regularization while fitting a generalized linear model(Eurico et al., 2016). Therefore, whether the dependent variable is continuous or discrete, it can be modelled and predicted by LASSO regression. The algorithm avoids overfitting by controlling the complexity of the model. The parameter λ controls the degree of LASSO regression complexity adjustment. The larger the λ, the greater the penalty for the linear model with more variables so that a more representative variable combination with fewer variables is finally obtained. Glmnet is very efficient at dealing with the likelihood function of the penalty term that can use the sparsity of the X matrix well.

3 Results

3.1 Diversity Analysis

3.1.1 Alpha Diversity

The investigation of alpha diversity revealed the dynamic transformation of diversity within these different treatment samples. The richness, Shannon and Simpson indexes are mainly used to calculate the relationship with the sampling time regarding community richness and diversity. It can be seen that the results of the three index treatments have a similar trend of change (the diversity decreases to the lowest on the third day and then gradually rises), and the reaction abundances in different treatment samples are the same. It proves that the index adopted in the calculation has high accuracy and reliability. There was a significant difference between the two treatments of BEWAF and Finasol (CEWAF, SWD). BEWAF has the highest community diversity, while the diversity level of the two chemical treatments is lower than that of pure oil treatment. It is expected because chemical surfactants are not as environmentally friendly as biosurfactants. It will have a specific negative impact on the community. In terms of the dynamic changes with the sampling time, although these treatment samples have a similar trend and gradually rise after the diversity decreases on the third day, the response results are not the same. In the pure oil and rhamnolipid treatments, the sampling on the third day is the lowest diversity in the observation process. After 28 days, the community diversity will return to a level close to the middle stage of sampling. The diversity index in the two chemical treatments will also drop significantly on the third day, but this is not the lowest diversity in the treatment process. For example, in SWD treatment, the diversity level on the 28th day is much lower than the sampling results on day3.

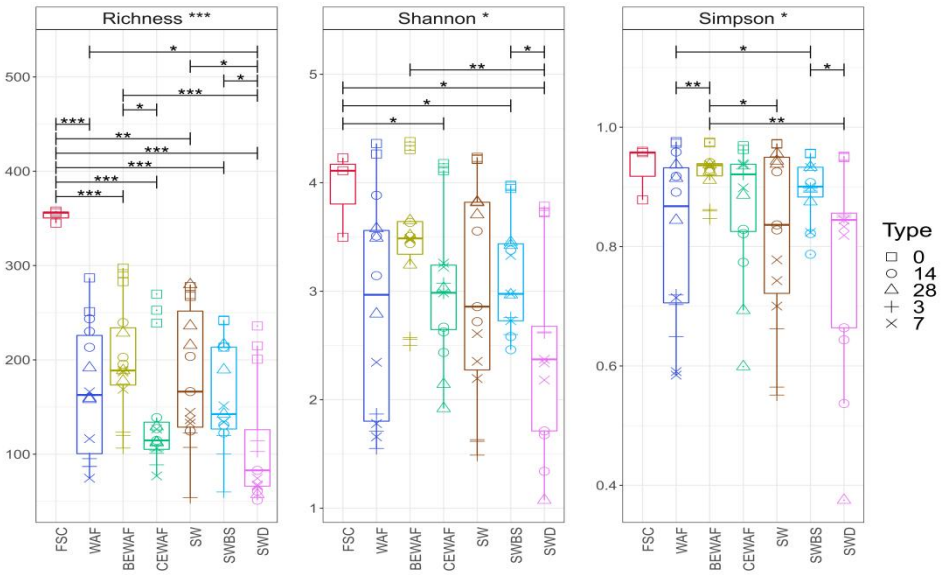


Fig 1: Alpha diversity index, Richness(Number of species), Shannon(Community diversity) and Simpson(Different probabilities)

3.1.2 NRI and NTI

NRI and NTI determine the changing trend of the system over time (Competitive exclusion or Environmental Filtering). The NRI of the original community FSC is negative, indicating that this is a more randomized ecosystem. In all treatments, the value of NRI was significantly greater than 0, and the value of NTI was significantly greater than 2, which indicated that the strong environmental filtering in these treatments was more inclined to the aggregation of closely related species. The most obvious ones are SWD and CEWAF processed by Finasol, which show strong environmental filtering at the beginning of sampling. Other schemes, such as pure oil treatment and BEWAF, still had a community structure similar to that of FSC at the early stage of sampling, and some samples with NRI close to or less than 0 appeared. The parameters used in the project are samples with the same abundance. When other parameters such as frequency and taxa are used, the value of NRI in FSC does not change. The system is still competitive exclusion, which means that the composition of the original ecological community will not be affected by the environmental settings.

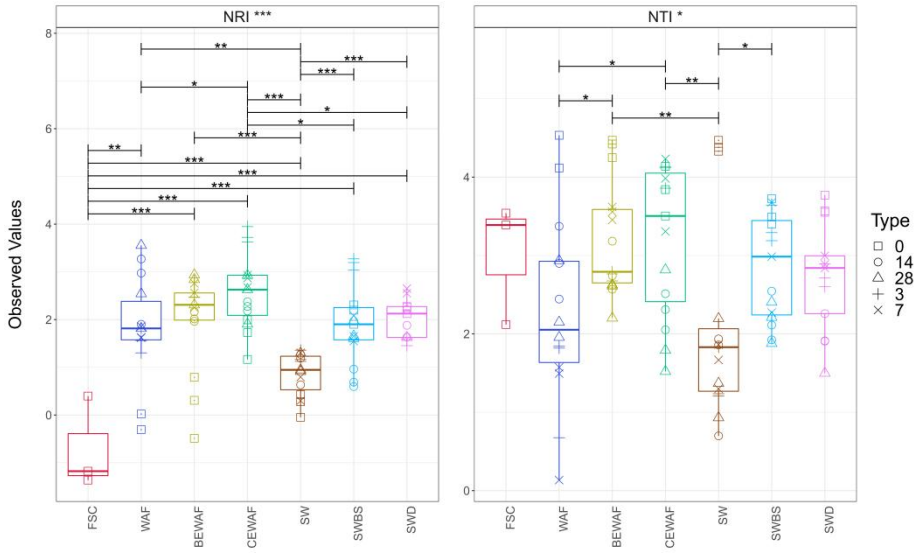
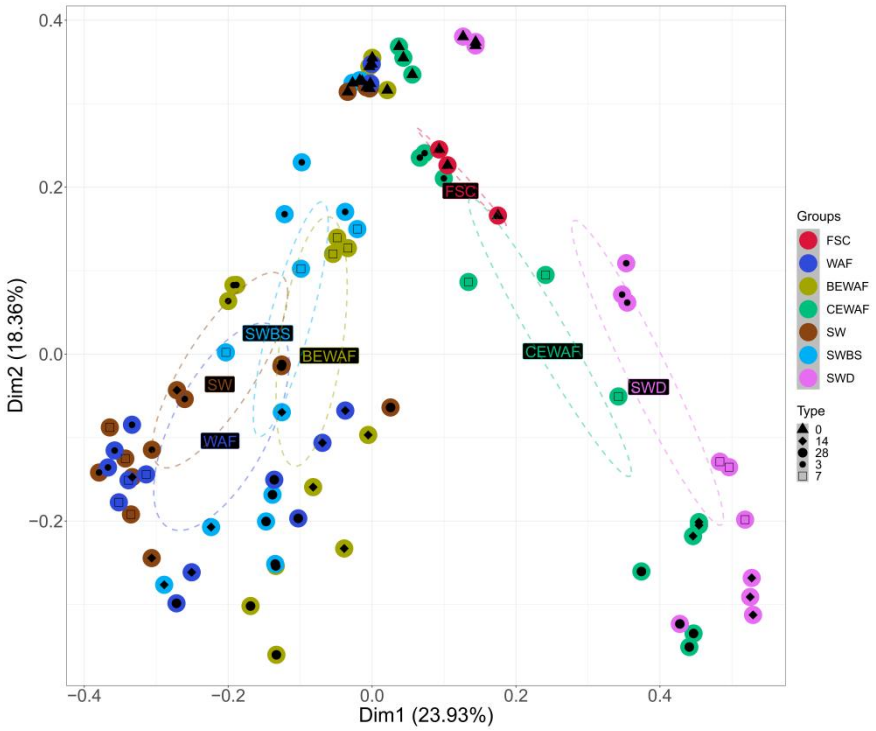


Fig 2: net relatedness index (NRI) and nearest-taxon index (NTI) phylogenetic alpha diversity measures based on MPD & NNPD

3.1.3 Beta Diversity

Beta diversity was calculated to assess community composition across treatments, and Principal coordinate analysis (PCoA) graphs were drawn(Theers et al., 2016). The project was analyzed using two distance parameters, Bray-Curtis and weighted Unifrac. The x-axis and y-axis in the table represent the two driving factors that affect the environmental ecology. According to PERMANOVA, the two influencing factors are different treatment methods and sampling time. The treatment method is the most significant factor affecting beta diversity, which can reveal up to 45% variability

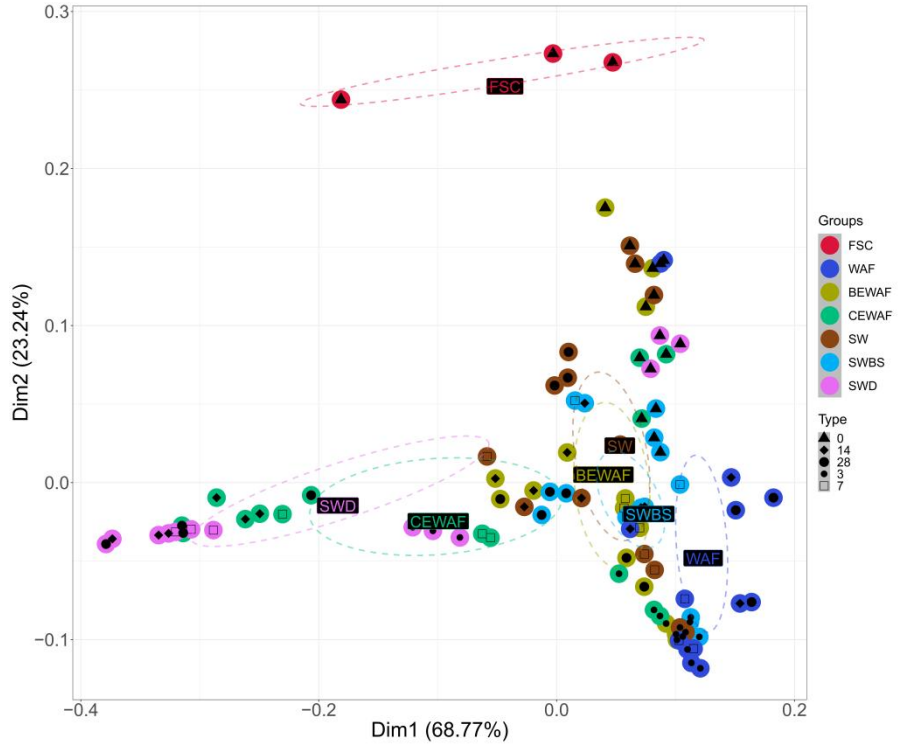
(p<0.001, R2=0.448). Sampling time explained 26% of the variability (p<0.001, R2=0.263). In the weighted Unifrac, which comprehensively considered species abundance and species, all six treatments showed a certain degree of clustering, which proved their phylogenetic similarity. Similar to the alpha diversity, the sampling was most pronounced on the third day, and these different treatments even overlapped but still showed some differences after the treatment finished. For example, the two methods, CEWAF and SWD, processed by Finasol, are far away from the overall community structure after day 28



PERMANOVA
Treatment R^2 = 0.2556 (p=0.001)
Sampling time R^2 = 0.2556 (p=0.001)
Treatment*Sampling time = 0.2516 (p=0.001)

(a)

Fig 2(a): Beat diversity- Principal coordinate analysis (PCoA) plot by Bray-Curtis distance matrices



PERMANOVA
Treatment $R^2 = 0.4482$ (p=0.001)
Sampling time $R^2 = 0.2633$ (p=0.001)
Treatment*Sampling time = 0.2245 (p=0.001)

Fig 2(b): Beat diversity -Principal coordinate analysis (PCoA) plot by weighted Unifrac distance matrices

3.2 Null Model Method: Calculating Quantitative Process Estimate and Normalized Stochasticity Ratio

QPE can estimate the relative importance of selection pressure and stochastic processes in community assembly. Among the five indices of variable selection are homogeneous selection, homogenizing dispersal, dispersal limitation and undominated. Undominated process plays an essential role in all environments. It is most obvious in the original community FSC, where the randomization process occupies all communities 66.7% of the assembly process, followed by a dispersal limitation of 33.3%. The proportion of undominated is around 50% in both biological and pure oil treatment, and the ratio is slightly lower in chemical treatment, about 28%-32%. Undominated is the primary assembly process in all treatments. Other critical assembly processes include homogeneous selection (18%-22%) and dispersal limitation, which account for 20%-30% in most samples but do not appear in SWD. In all treatments, homogenizing dispersal was not very important, only a small proportion (7.6%) in SWD, less than 2% in other environments, and not even exist in FSC. In addition, the variable selection plays an important role in chemical processing,

CEWAF (15%) and SWD (35%), but not apparent in other samples. In conclusion, selection pressure is widespread in both chemical treatment methods, while other environments are more trend to stochastic processes.

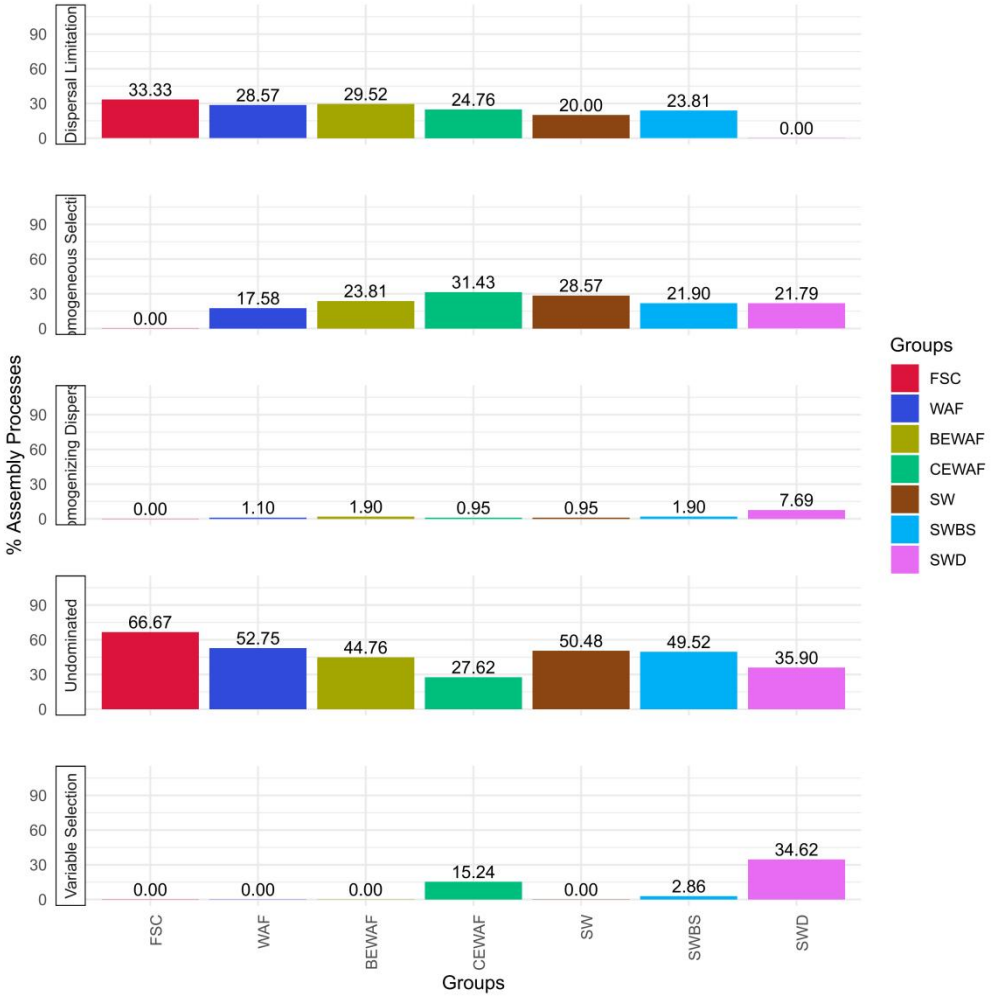
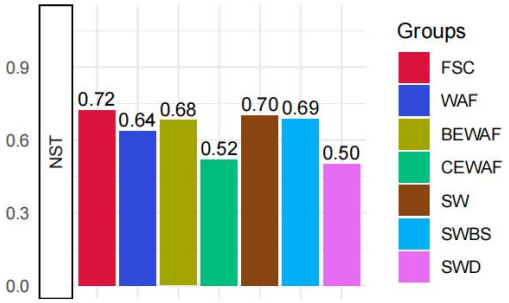


Fig 3: Quantitative process estimates (QPE) Relative importance of selection pressure, dispersal limitation or historical contingency, homogenizing dispersal or undominated community assembly processes for all treatment groups.

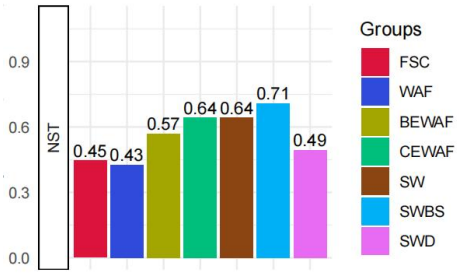
In the NST calculation, the evaluation is mainly based on proportional-fixed (P-F) and proportional-proportional modules. FSC has the highest value (81%), which indicates that the community has high randomness and is more inclined to competitive exclusion, which is consistent with the QPE results where the randomness process dominates. Other treatment methods, such as SW, WAF and rhamnolipid, all have NST values of more than 50%, indicating that these assembly processes are biased towards stochastic. It is quite different from the two chemical treatments, which in CEWAF (50%) and SWD (45%) are both deterministic processes belonging to environmental filtering. Overall, the results of these NST suggest that the microbial

community composition in the treatments studied is neither purely deterministic nor purely stochastic.

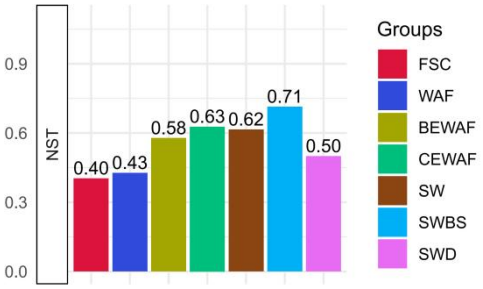
(a) Ruzicka-PF



(b) Ruzicka-PP



(c) Jaccrd-PF



(d) Jaccrd-PP

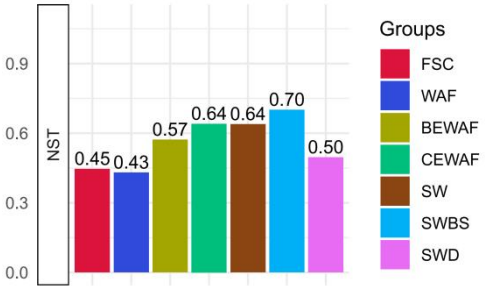


Fig 4:NST Normalized stochasticity ratio (NST) calculated based on abundance-based Ružička metric and Taxa-Richness constraints of proportional-fixed (P-F) and proportional-proportional (P-P)

3.3 Analysis of ecological driving factors

3.3.1 The top 25 most abundant taxa of each treatment

The graph shows the top 25 most abundant genera in each treatment and the change in abundance relative to sampling time, allowing for visual analysis of the composition of different communities and easy comparison of their differences. In FSC and all treatments, members of *Colwelliaceae* accounted for a significant proportion, while *Peredibacter* and *Paraglaciecola* belonged to lower levels in most treatments. In addition, there are bacteria (*Sedimentitalea*) that are only significant in specific treatments and bacteria (*Colwellia*) that show higher abundance differences in

different environments. *Colwellia* was in the top five species in abundance in the two dispersant treatments, and the enrichment levels were significantly different. Its abundance was 17% in seawater treatment and around 10% in pure oil and rhamnolipid treatments. In Finasol treatments (CEWAF, SWD), its abundance exceeded 40%. Although it is the dominant species in high abundance in many treatments, this does not mean that these dispersants positively affect *Colwellia*, promoting its growth. The abundance of these bacteria changes dynamically with time, and it is necessary to combine the sampling time to judge whether different treatments positively affect them. Taking *Colwellia* as an example, although its initial abundance is very high, it is in a continuous downward trend in the SWD treatment. It dropped to 11% on day 3, then reached 7% on day 7 and 3% on day 14, and the final content is less than 1%. Combined with the results of CEWAF, which is also treated by chemicals, although there is an increasing trend on day 3, the overall abundance still declines rapidly. Finally, it is almost unobservable on day 28. Combined with the time factor, the negative effect of Finasol on *Colwellia* can be determined, and the short upward trend in CEWAF also occurs in pure oil processing, which may be the energy in the oil.

In general, chemical or biological treatments (BEWAF, CEWAF) are compared with corresponding dispersant treatments (SWBS, SWD) and pure oil treatments (WAF) to find the effect of dispersants and oil on specific bacteria. For example, *Oleispira*, which showed an increasing trend in BEWAF, was almost negligible in WAF. On day 7, Uncultured *Micavibrionaceae* showed an increasing trend in WAF and BEWAF and peaked in the middle of the experiment (17% and 15%). On day 14, the changing trend of microorganisms under different treatments was clear, and many species with high initial abundance were reduced to extremely low levels at that time. For example, on day 14, *Colwellia* dropped to 7%, 2% and 4% for the three oil treatments (WAF, BEWAF, CEWAF). *Cycloclasticus,* which was almost negligible at the beginning of the experiment, gradually increased to 20% in WAF and was maintained until the end of the experiment. In BEWAF, *Cycloclasticus* was less than 1% on day 14 but increased to 8% and 4% by day 28. In SWD, due to the influence of Finasol, the diversity level was poor, with fewer than 25 high-abundance taxa, and only one strongly enriched bacterium (*Vibrio*) was observed. *Virbo* peaked from an initial abundance of 2% on day 3 (25%) but still decreased to a level close to the initial abundance by the end of the experiment. In the seawater treatment species, only *Colwellia* and *Amylibacter* observed strong enrichment; other species' changes were not noticeable.

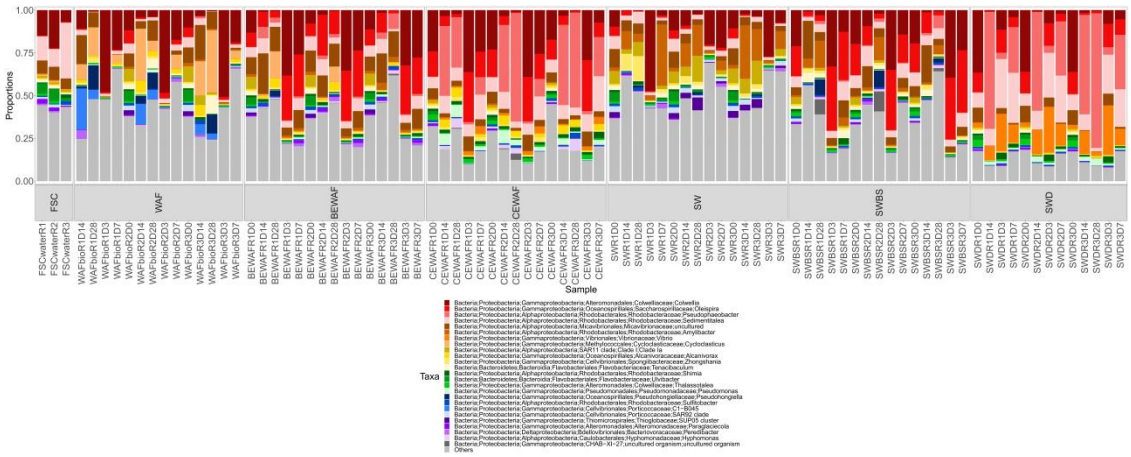


Fig 5: Top-25 most abundant taxa of microbial community

3.3.2 Core microbiome in Finasol and Rhamnolipid treatment

The core microbiome assessed the abundance and probability of analyzing the bacterial taxa that dominated the Finasol and Rhamnolipid treatments. The most abundant core microbiome in CEWAF was *Sedimentitalea* of *Rhodobacteraceae*, while *Oleispira*, *Alcanivorax* and *Cycloclasticus* played a significant role in BEWAF. *Colwellia* and *Oleispira* dominated both Finasol and Rhamnolipid treatments. In addition, some bacteria were only enriched in a single treatment. For example, Staphylococcus was only enriched in CEWAF. *Cycloclasticaceae* and *Paraglaciecola* are only abundant in BEWAF. Based on this analysis, some superficial judgments can be made about the conditions in the system that affect community change. For example, with Finasol treatment, *Vibrio* has extremely high enrichment in SWD treated only with Finasol, but relatively less in CEWAF, which may indicate that the bacteria are more prone to Finasol. Finasol is a factor that promotes its response. Overall, the negative effects of chemical dispersants on community diversity were significant. It creates a strong ambient filter.

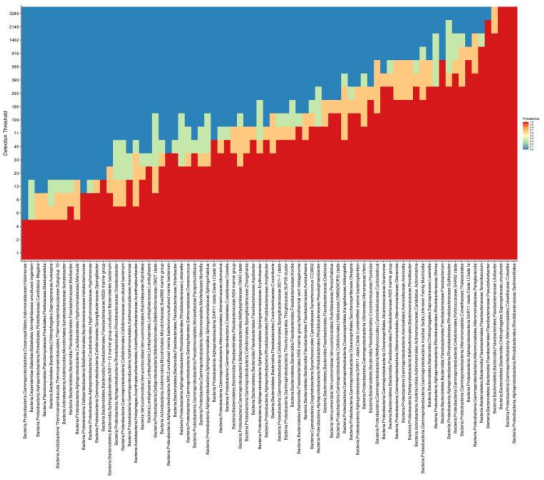


Fig 6(a): Core microbiome in FSC (Taxa with a probability of more than 85% in all samples)

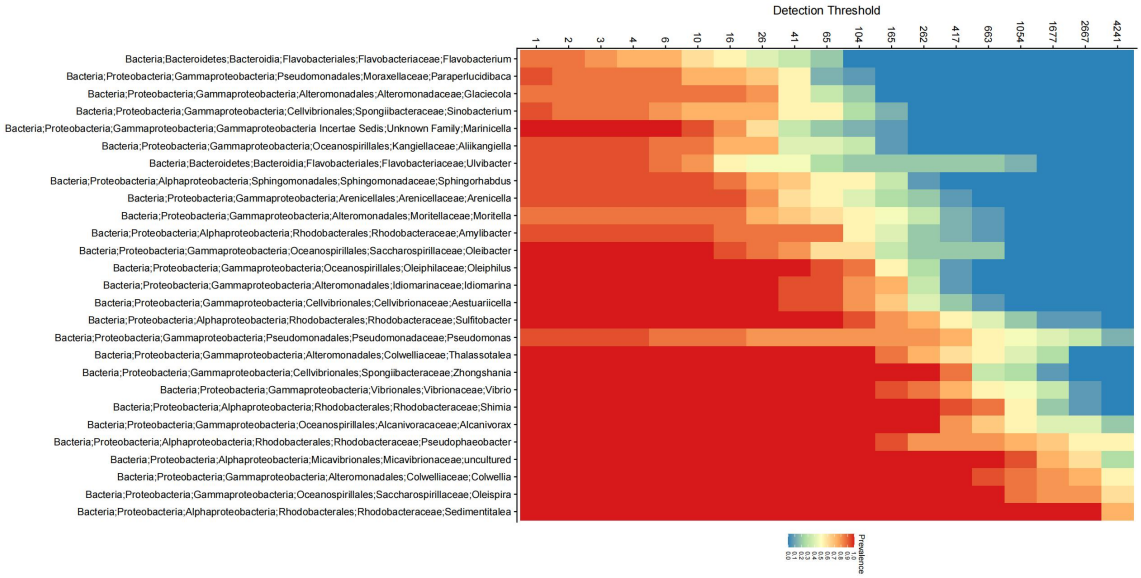


Fig 6(a): Core microbiome in BEWAF(Taxa with a probability of more than 85% in all samples)

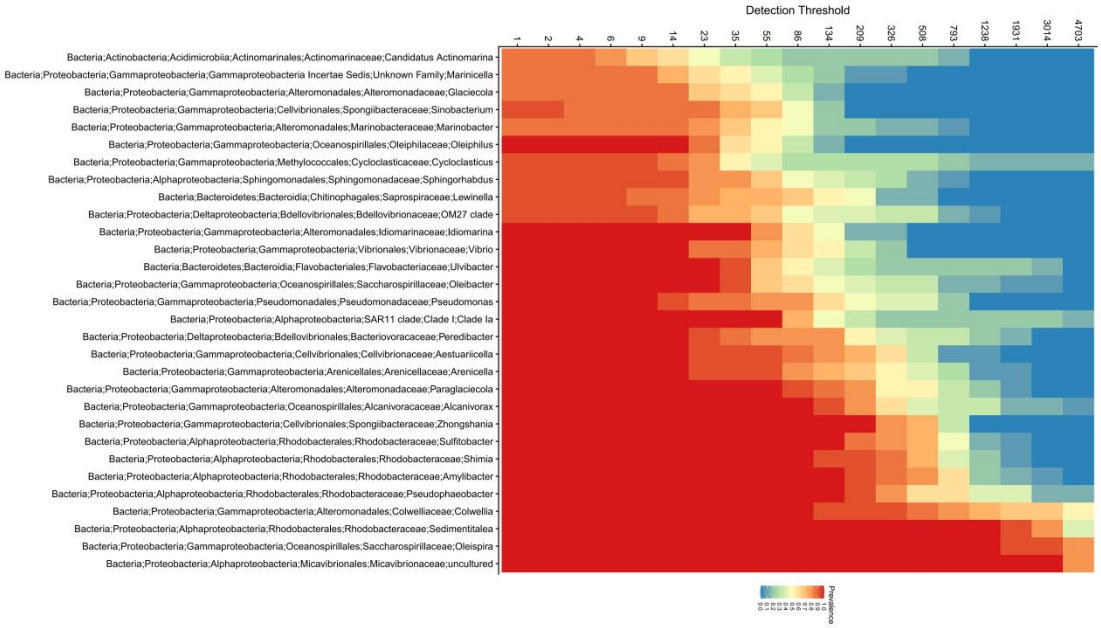


Fig 6(a): Core microbiome in CEWAF(Taxa with a probability of more than 85% in all samples)

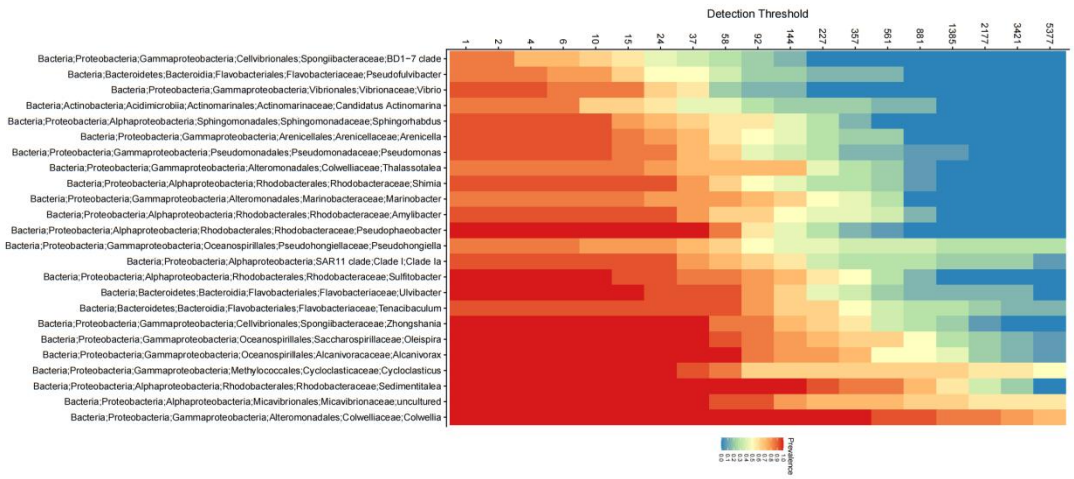


Fig 6(c): Core microbiome in WAF(Taxa with a probability of more than 85% in all samples)

3.3.3 DESeq2 Differential analysis

In order to identify the key taxa that dominate community changes during different treatments, differential analysis was performed under different conditions based on the DESeq2 algorithm. The trend of community changes can be found by comprehensively comparing the two critical factors of time and treatment methods. It can be seen that BEWAF is closer to the original community than CEWAF in the perspective of community structure. In sample time, the community composition of day0 and day7 was significantly different, and the difference tended to decrease after day14. The comparison can be applied to specific dates and processing methods by narrowing the range of conditions. In the figure, the composition difference between BEWAF and CEWAF is slight on day 7, then gradually increases, and has a significant difference on day 28. In addition, by analyzing the returned corrected p-values, significantly changed species can be identified and compared to the previous Core microbiome. For example, *Cycloclasticus*, which appeared in rhamnolipid treatment, and *Vibrio*, which was dominant in Finasol, were significantly different between the two treatments, so the differential analysis returned a minimal P value (*Cycloclasticus* padj = 2.4495e−07, *Vibrio* padj = 9.2423e−06).

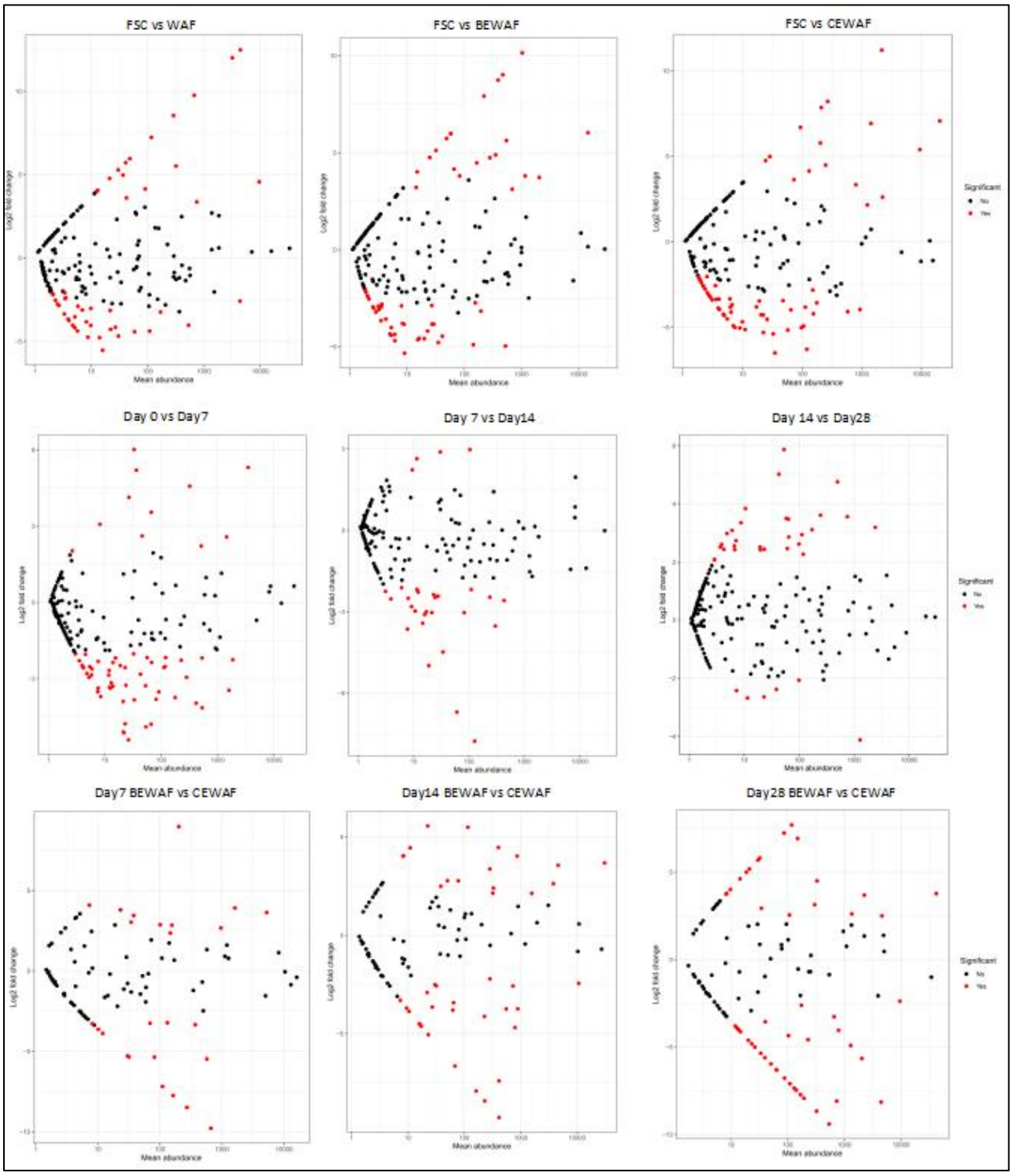


Fig 7:Differential analysis with BEWAF and CEWAF

3.3.4 Subset regression

Subset analysis was used to verify the alpha diversity results calculated by the Shannon index and evaluate the effect of chemotherapy and biological treatment on community diversity. Chemical dispersants are well known to have vital environmental filtering and significantly negatively impact community diversity. Combined with the results of alpha diversity analysis (BEWAF has the highest diversity and SWD diversity index is the lowest), the subset regression should show a negative value in the Finasol treatment and a positive value in the rhamnolipid treatment. Considering the Cross-validation model, the combination with the smallest error is BEWAF and SWD, and their p-values are less than 0.05.

	Model	Cross-validation Errors
3	Shannon ~ Treatment_BEWAF + Treatment_SWD + Treatment_FSC	0.78228
4	Shannon ~ Treatment_BEWAF + Treatment_SWD + Treatment_FSC + Incubation_day	0.79106
5	Shannon ~ Treatment_WAF + Treatment_BEWAF + Treatment_SWD + Treatment_FSC + Incubation_day	0.79121
6	Shannon ~ Treatment_WAF + Treatment_BEWAF + Treatment_SWD + Treatment_SWBS + Treatment_FSC + Incubation_day	0.79558
2	Shannon ~ Treatment_BEWAF + Treatment_SWD	0.80125
7	Shannon ~ Treatment_WAF + Treatment_BEWAF + Treatment_CEWAF + Treatment_SWD + Treatment_SWBS + Treatment_FSC + Incubation_day	0.80621
1	Shannon ~ Treatment_SWD	0.80635

Table 1: Cross-validation model finds the reliable data

	Shannon								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	3.01685 ***	0.10264	0.00000	0.09857	2.81281 – 3.22090	-0.19594 – 0.19594	29.39181	**1.251e-46**	86.00000
Treatment BEWAF	0.45496 *	0.22798	0.20222	0.10133	0.00175 – 0.90818	0.00078 – 0.40367	1.99563	**4.914e-02**	86.00000
Treatment SWD	-0.56674 *	0.24156	-0.23762	0.10128	-1.04694 – -0.08654	-0.43895 – -0.03628	-2.34618	**2.127e-02**	86.00000
Treatment FSC	0.92902 *	0.46662	0.19889	0.09990	0.00141 – 1.85663	0.00030 – 0.39749	1.99095	**4.966e-02**	86.00000
Observations	90								
R^2 / R^2 adjusted	0.155 / 0.126								
								** p<0.05*	*** p<0.01* **** p<0.001*

Table 2: The best combination model SWD & BEWAF

The data with p-values less than 0.05 in the project were considered reliable, and

these data were summarized as shown in table 3. It can be seen that the dispersant treatment of SWD has a significant negative effect on community diversity, while it is positively correlated in BEWAF, which is consistent with the previously obtained alpha diversity results. In addition, rhamnolipid-only treatment (SWBS) also had a positive effect. In contrast, sampling time had a negative effect, which may be related to the fact that many bacterial groups gradually decreased as the experiment progressed. The community diversity decreased at the end of the experiment. In addition, there is no expected data in CEWAF. The result is positive, which requires further cross-validation. Similar results have emerged for other alpha diversity such as Simpson and Richness, which will be provided in the appendix.

	Shannon								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	3.12233 ***	0.22460	0.00000	0.09928	2.67553 – 3.56912	-0.19750 – 0.19750	13.90185	**2.852e-23**	82.00000
Treatment WAF	-0.14185	0.29518	-0.06131	0.12760	-0.72905 – 0.44536	-0.31514 – 0.19251	-0.48054	6.321e-01	82.00000
Treatment BEWAF	0.46912	0.28998	0.20852	0.12889	-0.10773 – 1.04598	-0.04788 – 0.46492	1.61781	1.095e-01	82.00000
Treatment CEWAF	0.04730	0.28998	0.02102	0.12889	-0.52955 – 0.62415	-0.23538 – 0.27742	0.16312	8.708e-01	82.00000
Treatment SWD	-0.58373	0.30187	-0.24474	0.12656	-1.18424 – 0.01678	-0.49651 – 0.00704	-1.93373	5.660e-02	82.00000
Treatment SWBS	0.14876	0.28998	0.06612	0.12889	-0.42809 – 0.72561	-0.19028 – 0.32252	0.51301	6.093e-01	82.00000
Treatment FSC	0.82355	0.51055	0.17631	0.10930	-0.19209 – 1.83919	-0.04113 – 0.39375	1.61307	1.106e-01	82.00000
Incubation day	-0.01150	0.00881	-0.13342	0.10222	-0.02904 – 0.00603	-0.33677 – 0.06994	-1.30516	1.955e-01	82.00000
Observations	90								
R^2 / R^2 adjusted	0.183 / 0.113								

Table 3: Relationship between all treatments and diversity index

3.3.5 CODA GLMNET

The bacterial groups that were positively correlated with the two surfactants were counted by the GLMNET software package and verified by combining the change trends of the Top 25 taxa. In CEWAF, the positively correlated microorganisms mainly included *Marinicella*, *Aliikangiella* and *Brevundimonas*. In BEWAF, the microorganisms showing a growing trend are *Peredibacter* and *Amylibacter*. This result is close to the changing trend of the previous analysis. Although *Amylibacter* increased in the first week of CEWAF treatment, the overall trend was still decreasing. It was almost invisible at the end of the experiment. However, it has a relatively weak increasing trend in BEWAF and is maintained until the end of the experiment. In addition, some microbes, such as Alcanivorax and Sphiningomonas, do not belong to the core microbiome in the research, so their trend of increase or decrease may not be meaningful for oil processing.

Fig 8:Bacterial taxa with positive response in BEWAF and CEWAF

4 Discussion

4.1 Stochastic and deterministic process of system

From the ratios of QPE and NST, it can be found that deterministic and stochastic processes occur simultaneously in the dynamic changes of these communities. The bacterial communities in all communities studied here are neither purely stochastic nor purely deterministic. It can be seen in the results of QPE that the random process and selection pressure occupy a certain proportion in each treatment. In NST, the pure oil control treatment (WAF) and seawater control treatment (SW) had more stochastic communities. In contrast, the communities of chemical dispersant Finasol treatments (CEWAF and SWD) tended to be deterministic. It suggests that the dispersant likely elicited a deterministic-related microbial response. The relative importance of the deterministic process varied significantly over time. In biosurfactant treatments, neutral processes play a greater role in assembly treatments than chemical treatments. It suggests that biological treatment does not have a strong selection for assembling micro-communities in polluted seawater. Deterministic processes are evident in the early stages of the first three days, while stochastic processes are mainly seen in the middle and late detection stages. This result is expected because the beta diversity identified one of the factors driving community dynamics is different treatment modalities. So changing these environmental conditions in the early stage of the experiment leads to a more definitive outcome.

4.2 The dominant taxa in Petroleum degradation

The Faroe-Shetland Channel is a cold environment in the North Atlantic so the community is dominated by psychrophilic flora, including the common oil-degrading bacteria *Cyclocratics* and *Colwellia* (Lane et al., 2016). Although there is no information from the FSC to confirm the oil spill, many Crude oil degradation reactions and oil slicks surveyed by satellite indicate a certain degree of underground oil leakage in the offshore area. It may be due to the area's frequent oil extraction and transportation. Therefore, it is necessary to investigate the dominant oil-degrading bacterial groups under different environmental conditions to prevent possible large-scale oil spills in the future.

4.2.1 The core taxa in all treatments: Colwellia

Colwellia is a cold-tolerant bacterium that is widespread in deep-sea environments. It belonged to the high abundance core microbiome in all treatments. It appeared increasing trend within one week of all three oil treatments (WAF, BEWAF, CEWAF), possibly due to its metabolic diversity. Since it was one of the dominant flora in the original community FSC, this could be just the result of its higher initial abundance. Referring to the research of Kleindienst (2015), who used the chemical dispersant Corexit to treat oil spills in the Gulf of Mexico, *Colwellia* also showed similar

changes to this study, and the abundance of *Colwellia* increased to 40% in the first week. However, the change in *Colwellia* in SWD is different, decreasing significantly from the beginning of the experiment and finally almost negligible in the samples.

4.2.2 The core taxa of Finasol: Rhodobacteracaea & Vibrionaceae

Members of *Rhodobacteracaea* were the most essential taxa in Finasol treatment, including the most abundant core microbiome *Sedimentitalea* and the dominant bacterium *Pseudophaeobacter* at the end of the experiment (day 28). *Vibrionaceae* was also a core taxa that dominated only in Finasol treatment and showed strong enrichment in SWD. These *Rhodobacteracaea* and *Vibrionaceae* bacteria can produce organic intermediates through hydrocarbon degraders or consume carboxylic acids and alcohols in chemical surfactants. For example, the strong enrichment of *Virbo* in SWD, compared to the lesser content in CEWAF, suggests that the increase in vibrio is caused by Finasol, which is more inclined to metabolize dispersants rather than petroleum. Vibrio possesses high metabolic agility, which may be why they significantly increased over the initial core microbiome Colwellia in the middle and late stages of the experiment. Unlike *Vibrionaceae*, the significant increase in Rhodobacteracaea is based on their degrading effect on petroleum. *Rhodobacteracaea* possesses hydrocarbon-degrading ability. Referring to the recent study of the dispersant Corexit, *Vibrio* also showed strong enrichment, suggesting that a component in Finasol may have contributed to the continued response of *Vibrio*(Eurico et al., 2016). Some researchers believe that *Vibrio* metabolizes the metabolic by-products of other microorganisms after degrading petroleum, but this explanation does not apply to the research of this project. In this project. *Vibrio* is strongly enriched in SWD (only seawater and Finasol). There is no oil involved, so it might be the result of *Vibrio's* metabolic composition of dispersant.

4.2.3 The core taxa of BEWAF: Cycloclasticus

Similar to *Colwellia*, *Cycloclasticus* is also a taxa that was enriched in the early stage of the experiment. It occurs in rhamnolipid and pure oil treatments, showing an increasing trend within a week and maintaining it until the end of the culture. The project did not observe *Cycloclasticus* in either Finasol treatment. It is a very negative result. In oil spill research about the Gulf of Mexico, *Cycloclasticus* is a very important dominant bacterium in chemical processing for polycyclic aromatic hydrocarbons in the environment(Xuemei et al., 2016). *Cycloclasticus* did not appear in the chemical treatment of this project, indicating that Finasol had a serious negative impact on this taxa. Since it is an important degrading bacteria in the processing of aromatic hydrocarbons, the absence of *Cycloclasticus* directly affects the rate of aromatic biodegradation in chemical processing, which requires improving the composition of Finasol in the future.

4.2.4 Oil-amended treatment bacteria: Alcanivorax

Alcanivorax, first described in 1998, is an oil-degrading marine bacterium at low levels in the uncontaminated environment of the upper oceans(Emmanuel et al., 2016). When the conditions of these moderately halophilic environments are suitable, *Alcanivorax* may account for 80% of the oil-degrading microorganisms in the region. In this project, *Alcanivorax* was only observed in oil-containing treatments (WAF, BEWAF, CEWAF) because it depends on alkanes for survival and has a high degree of degradation ability. It is the main microorganism in oil. Usually, it will multiply rapidly after entering the oil into the environment, but it is not apparent in the experiment. A clear upward trend was observed after day 14 and maintained until the end of the experiment. On the one hand, it may be that the initial high-abundance group *Colwellia* in the original community also belongs to the early-stage growth bacteria occupying the living space of *Alcanivorax*. On the other hand, the FSC is a cold environment in the North Atlantic, and the low temperature may limit the reproduction of *Alcanivorax*.

4.2.5 The core taxa of FSC: Uncultured Micavibrionaceae

According to statistics, an uncultured member of the Micavibrionaceae has a very high abundance in the original community (less than *Colwellia* and *Sedimentitalea*). At the same time, it is also the most abundant core microbiome in BEWAF, showing an increasing trend in the middle of the experiment. These taxa were first described in 1982, but their role in ecosystems and oil degradation is currently not understood clearly. According to the research of Wilhelm(2014),*Micavibrionaceae* changes dynamically with seasons in the ecological community, so the observed growth trend in the samples is probably caused by the initial high abundance at the time of sampling.

4.3 Regression model analysis

As expected, regression analysis models supported the validation results for community diversity and core microbiome. Negative values in the SWD demonstrate a significant negative effect of Finasol treatment on community structure. A positive correlation trend was found in the lasso regression for members of *Rhodobacteraceae*. In addition, when using different parameters for subset analysis, the combined model of SWD and BEWAF always has smaller cross-validation errors. It might be the

apparent impact of SWD and BEWAF on ecological communities. For example, the powerful environmental filtration in SWD directly determines the composition of the microbial community. Many microorganisms decline rapidly when the dispersant enters the sample, so the judgment results are clearer and more reliable.

4.4 Comprehensive discussion with Finasol and rhamnolipid

A certain proportion of selective pressure was observed in both Finasol and rhamnolipid treatments. In general, the environmental filtering from chemical treatment is stronger than biosurfactant. Finasol was very effective in promoting biodegradation, but strong environmental selection negatively impacted *Cycloclasticus*, which affected its aromatic biodegradation rate and substantially disrupted community diversity. It was verified by regression analysis that the biosurfactant rhamnolipid could also promote the degradation of hydrocarbons and did not inhibit some important hydrocarbon-decomposing bacteria. It seems that rhamnolipid can completely replace the application of Finasol in oil spills, but there are still many difficulties in actual operation. Although rhamnolipid was described in 1947, the problem of high cost and low product output has not been solved yet. So how to commercialize it widely is still the main challenge for biosurfactants in the future.

5 Conclusion

There were differences in the response of the ecosystem to treatments with different surfactants. In these community dynamics, deterministic and stochastic processes occur simultaneously. The communities in Finasol treatment had stronger environment filtering, and the treatment with rhamnolipids tended to be more stochastic. With time, the community composition under the three oil treatments also gradually changed. The mainly dominant communities of Finasol were the member of *Rhodobacteracaea* and *Vibrionaceae*, while *Cyclocraticuss*, *Colwellia*, and *Oleispira* were increasing in the rhamnolipid treatments. Finasol stimulated the faster degradation of n-alkanes, but the strong selective pressure inhibited the reproduction of *Cyclocraticus*. It leads to a decrease in the rate of aromatic biodegradation. The rhamnolipid correction treatment maintained a high functional diversity, and the total amount of aromatic biodegradation was the highest. The experiment is simulated and cultivated in the microscopic world. The actual ecological community changes are more complex and need to refer to various factors such as temperature and season. However, the statistics of the changing trend of these core bacterial communities can still provide a reference for preventing oil spill accidents.

In conclusion, natural marine microbial communities in the Faroe-Shetland Channels responded significantly to treatment with the synthetic dispersant Finasol and the biosurfactant rhamnolipid over time. In general, Finasol has more substantial degradation effect than crude oil. The findings advance the understanding of rhamnolipid biosurfactants, and Finasol affects natural marine microbial communities in FSCs, supporting the potential application in oil spills.

6 Future work & Challenges

Oil spills are accidents that seriously impact the natural environment, so it is essential to analyze microbial taxa that can degrade oil effectively. The project selected microbial samples from the Faroe-Shetland Channel and statistics the bacteria with better adaptability under the treatment of the two dispersants. This simulation method can also be extended to other areas with potential oil spills for detection and prevention, searching for suitable dispersants in the microbial community in that area. This research also has some work to improve and future challenges:

1. The null model method used in this paper needs to be improved. The environment filtering of CEWAF is not apparent enough in the proportional-fixed module (The ratio is close to 50%). It does not demonstrate the deterministic of the system well.

2. Although the project counts these dominant bacterial groups, there are still some groups, such as *Micavibrionaceae*, whose role in the ecosystem has not been clearly defined in microbiology. It requires more research in this field in the future. In addition, many factors affect the composition of microbial communities in the virtual environment. Different temperatures and seasons will have an impact on the initial abundance of the community. The samples for this project were collected at the same temperature and season, so further validation is required in the future.

3. The project analyzes the community changes in the effect of dispersants from the perspective of understanding and observation. However, there are still many challenges to designing a dispersant suitable for most degradation bacteria to solve future oil spill accidents. As noted above, the biosurfactant rhamnolipid has not been mass-produced. There is still much groundwork to be done to prevent oil spills in the future

Reference

Alisa C Herbst, Mackenzie G Johnson, Hayley Gammons, Stephanie E Reedy, Kristine L Urschel, (2022) Patricia A Harris, Amanda A Adams. "Development and Evaluation of a Muscle Atrophy Scoring System (MASS) for Horses", Journal of Equine Veterinary Science.

Abdellatif N., Fekih-Salem R., Sari T (2016).“Competition for a single resource and coexistence of several species in the chemostat”, Mathematical Biosciences & Engineering, vol. 13, pp. 631–652

Adamberg K., Valgepea K., Vilu R (2015). “Advanced continuous cultivation methods for systems microbiology”, Microbiology-Sgm, vol. 161, pp. 1707–1719, .
Fekih-Salem R., Lobry C., Sari T., “A density-dependent model of competition for one resource in the chemostat”, Mathematical Biosciences, vol. 286, pp. 104–122, 2017.

Aubin GG, Haloun A, Treilhaud M, Reynaud A, Corvec S. (2013). Gallibacterium anatis bacteremia in a human. J Clin Microbiol 51:3897–3899

Araújo IM, et al.(2017) Hydrothermal synthesis of bacterial cellulose–copper oxide nanocomposites and evaluation of their antimicrobial activity. Carbohydr. Polym. doi: 10.1016/j.carbpol.2017.09.081.

Booth, S.C., Smith, W.P.J (2020) Light sheets unveil host–microorganism interactions. Nat Rev Microbiol 18, 65 . https://doi-org.ezproxy.lib.gla.ac.uk/10.1038

Corvec S.(2018) Clinical and Biological Features of Cutibacterium (Formerly Propionibacterium) avidum, an Underrecognized Microorganism. Clin Microbiol Rev. 30;31(3):e00064-17.

Dizay HH, Lau DG, Nottage WM. (2017). Benzoyl peroxide and clindamycin topical skin preparation decreases Propionibacterium acnes colonization in shoulder arthroscopy. J Shoulder Elbow Surg 26:1190–1195.

Eurico J. D’Sa, Edward B. Overton, Steven E. Lohrenz, Kanchan Maiti, R. Eugene Turner, and Angelina Freeman .(2016) Changing Dynamics of Dissolved Organic Matter Fluorescence in the Northern Gulf of Mexico Following the Deepwater Horizon Oil Spill. Environmental Science & Technology, 50 (10) , 4940-4950. https://doi.org/10.1021/acs.est.5b04924

Emmanuel Nyankson, Muslum Demir, Mehmet Gonen, and Ram B. Gupta . (2016) Interfacially Active Hydroxylated Soybean Lecithin Dispersant for Crude Oil Spill Remediation. ACS Sustainable Chemistry & Engineering, 4 (4), 2056-2067. https://doi.org/10.1021/acssuschemeng.5b01403

Fredy Alvarado, Edyla R. Andrade, Bráulio A. Santos, Graham Prescott, Gustavo Souza, Federico Escobar. (2018) "Forest cover is more important than farmland heterogeneity and livestock intensifification for the retention of dung beetle phylogenetic diversity," Ecological Indicators,

Gao M, Li J, Bao Z, Hu M, Nian R, Feng D, An D, Li X, Xian M, Zhang H.(2019) A natural in situ fabrication method of functional bacterial cellulose using a microorganism. Nat Commun.25;10(1):437. doi: 10.1038/s41467-018-07879-3. PMID: 30683871; PMCID: PMC6347598.

Guckenheimer J., Holmes P.J.(2013) Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer Science & Business Media .

Jabbarzadeh, M., & Fu, H. (2018). Viscous constraints on microorganism approach and interaction. Journal of Fluid Mechanics, 851, 715-738. doi:10.1017/jfm.2018.509

Johnke, J., Boenigk, J., Harms, H. & Chatzinotas, A.(2017) Killing the killer: predation between protists and predatory bacteria. FEMS Microbiol. Lett. 364 (9), fnx089.

Kiørboe, T., Jiang, H., Gonçalves, R. J., Nielsen, L. T. & Wadhwa, N. (2014) Flow disturbances generated by feeding and swimming zooplankton. Proc. Natl Acad. Sci. USA 111 (32), 11738–11743.

Kiørboe, T. (2016) Fluid dynamic constraints on resource acquisition in small pelagic organisms. European Phys. J. Special Topics 225 (4), 680–683.

Kleindienst S, Seidel M, Ziervogel K, Grim S, Loftis K, Harrison S, et al.(2015) Chemical dispersants can suppress the activity of natural oil-degrading microorganisms. Proc Natl Acad Sci.;112(48):14900–5.

Kamzolova, S. V., Yusupova, A. I., Dedyukhina, E. G., Chistyakova, T. I. et al.,(2009) Succinic acid synthesis by ethanol-grown yeasts. Food Technol. Biotechnol. 47, 144– 152.

Lee KY, Buldum G, Mantalaris A, Bismarck A.(2014) More than meets the eye in bacterial cellulose: biosynthesis, bioprocessing, and applications in advanced fiber composites. Macromol. Biosci.14:10–32. doi: 10.1002/mabi.201300298.

Lane G. Tidwell, Sarah E. Allan, Steven G. O'Connell, Kevin A.(2016) Hobbie, Brian W. Smith, and Kim A. Anderson. PAH and OPAH Flux during the Deepwater Horizon Incident. Environmental Science & Technology, 50 (14), 7489-7497. https://doi.org/10.1021/acs.est.6b02784

M. Cirio, G.K. Brennen, J. Twamley Quantum magnetomechanics: ultrahigh q levitated mechanical oscillators Phys Rev Lett, 109 (2012), Article 147206

NANCY R., MICHAELA BAYERLOVÁ, OLIVIER MISSA, and NATÁLIA MARTÍNKOVÁ (2012). "Complex patterns of host switching in New World arenaviruses : HOST SWITCHING IN NEW WORLD ARENAVIRUSES," Molecular Ecology.

Phadnis J, Gordon D, Krishnan J, Bain GI. (2016). Frequent isolation of Propionibacterium acnes from the shoulder dermis despite skin preparation and prophylactic antibiotics. J Shoulder Elbow Surg 25:304–310.

Pauly N, Pucciariello C, Mandon K, Innocenti G, Jamet A, Baudouin E, Herouart D, Frendo P, Puppo A (2006) Reactive oxygen and nitrogen species and glutathione: key players in the legume-Rhizobium symbiosis. J. Exp. Bot. 57, 1769–1776.

William W. Gilbertson, Martin Solan, James I. Prosser, (2014) Differential effects of microorganism–invertebrate interactions on benthic nitrogen cycling, FEMS Microbiology Ecology, Volume 82, Issue 1,Pages 21–22

R. D. Sargent, S. W. Kembel, N. C. Emery, E. J. Forrestel, D. D. Ackerly. (2011) "Effffect of local community phylogenetic structure on pollen limitation in an obligately insect-pollinated plant," American Journal of Botany

Schlomann, B. H., Wiles, T. J. et al. (2019) Sublethal antibiotics collapse gut bacterial populations by enhancing aggregation and expulsion. Proc. Natl Acad. Sci. USA 116, 21392–21400

Sari T.(2013), "Competitive Exclusion for Chemostat Equations with Variable Yields," Acta Applicandae Mathematicae, vol. 123, no. 1, pp. 201–219.

Theers, M., Westphal, E., Gompper, G. & Winkler, R. G. (2016) Modeling a spheroidal microswimmer and cooperative swimming in a narrow slit. Soft Matt. 12 (35), 7372–7385.

Wilhelm SW, Lecleir GR, Bullerjahn GS, Mckay RM, Saxton MA, Twiss MR, et al. (2014) Seasonal changes in microbial community structure and activity imply winter production is linked to summer hypoxia in a large lake. FEMS Microbiol Ecol. ;87(2):475–85.

Xuemei Chen, Justin A. Weibel, and Suresh V. Garimella . (2016) Continuous Oil–Water Separation Using Polydimethylsiloxane-Functionalized Melamine Sponge. Industrial & Engineering Chemistry Research, 55 (12), 3596-3602.

Appendix I: Diversity data supplement

Only three alpha parameters have been selected for evaluation. Fisher alpha and Pielou evenness can also judge community diversity. It can be seen that the data distribution is basically the same as that in richess, and SWD is still the lowest diversity among all samples.

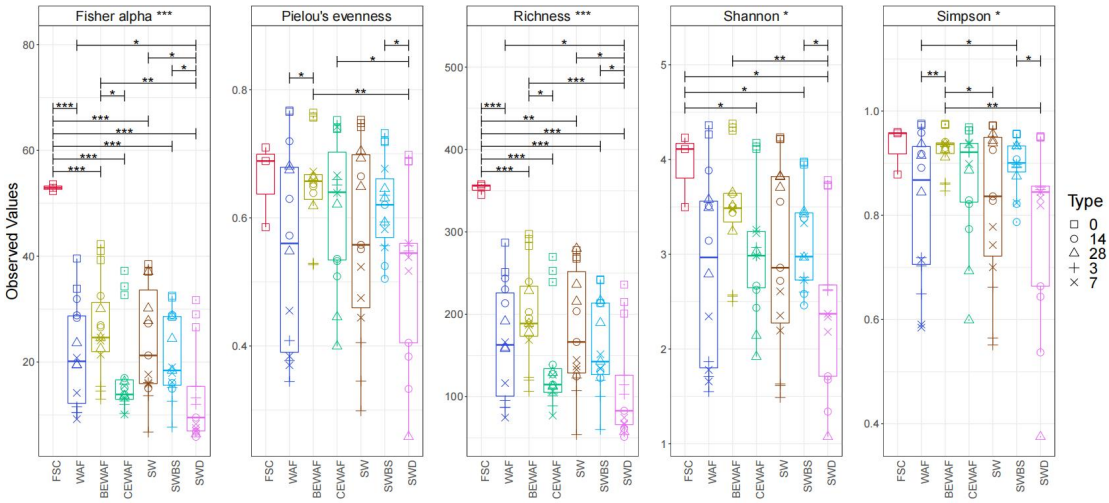


Fig 9: Alpha diversity for all index

	Richness	Shannon	Simpson	FisherAlpha	PielouEvenness
BEWAFR1D0	297.3938385	4.378843154	0.97533688	40.98486958	0.764182705
BEWAFR1D14	195.4688367	3.439255519	0.929463083	26.56360791	0.638733155
BEWAFR1D28	188.7356662	3.512877994	0.925910921	24.57849386	0.663017662
BEWAFR1D3	106.4915629	2.501645732	0.846840826	12.99199087	0.529181406
BEWAFR1D7	188.5911504	3.481882382	0.9371409	24.63067274	0.657167576
CEWAFR1D0	238.9655278	4.137043115	0.962856308	32.66571226	0.752577135
CEWAFR1D14	127.1683348	2.625025734	0.82168973	15.12339019	0.535955817
CEWAFR1D28	129.1247104	3.027333582	0.886396605	16.31195077	0.620966369
CEWAFR1D3	100.9337097	2.984387427	0.921120803	11.98822765	0.639954198
CEWAFR1D7	126.619789	3.25895768	0.93688652	15.4671230	0.666405934

	8		7	7		
WAFbioR1D3	87.09893115	1.708901038	0.702808241	10.47928874	0.377024496	
WAFbioR1D7	74.54923545	1.657998069	0.584580282	9.203884194	0.384019139	
WAFbioR2D0	286.9150563	4.361740563	0.975905983	39.56099969	0.765608037	

Table 4: Diversity data of three oil treatments

Beta diversity analyzed with unweighted UniFrac distance. It can be seen that the results are significantly different from the weighted UniFrac distance. In the weighted method, these samples appear obvious aggregation or even overlap, but it cannot be seen in this analysis. Since the weighting algorithm considers both species number and abundance, the results should be more accurate, so the project does not use this method

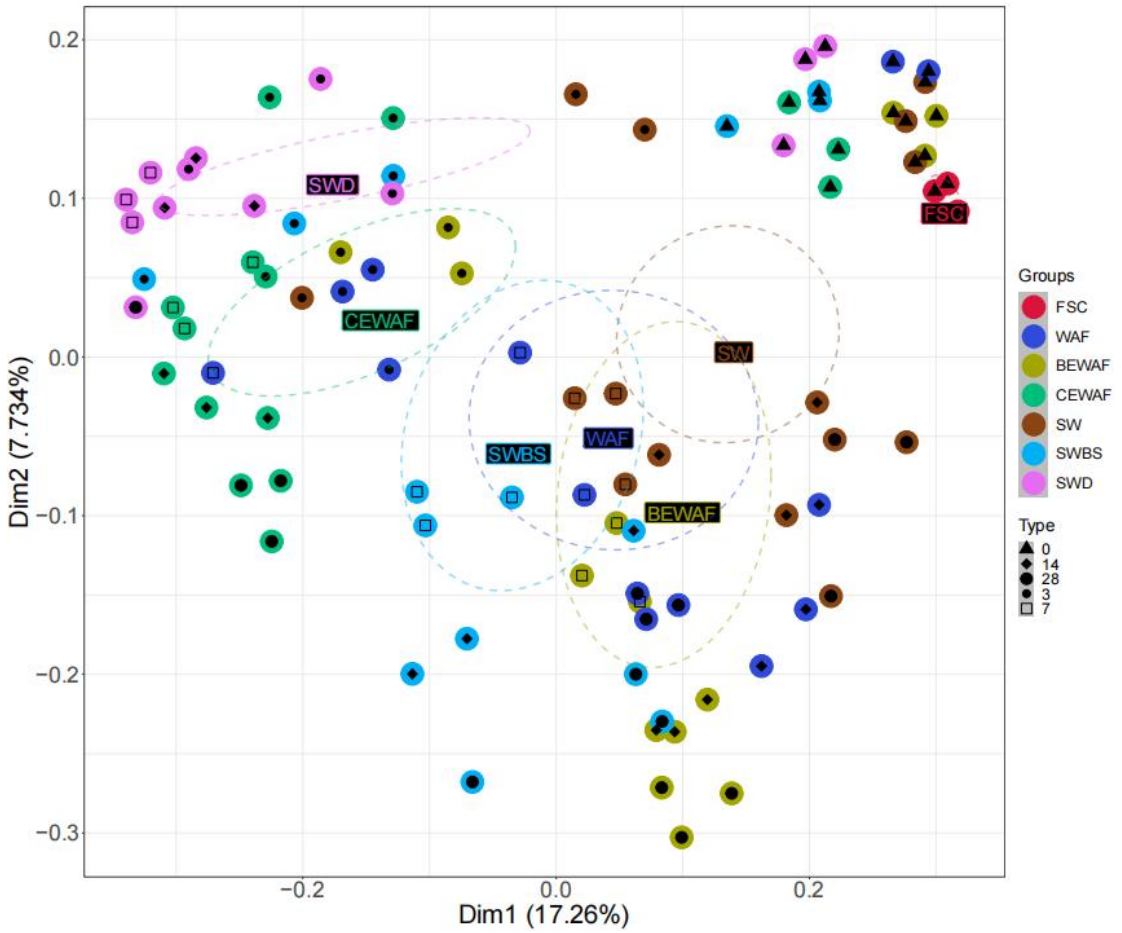


Fig10: Beat diversity -Principal coordinate analysis (PCoA) plot by unweighted Unifrac distance matrices

Considering that richness is the decisive factor affecting the composition of microbial communities, the project uses the same richness to calculate NRI and NTI. In addition, there are many parameters such as frequency and taxa that can be used as reference

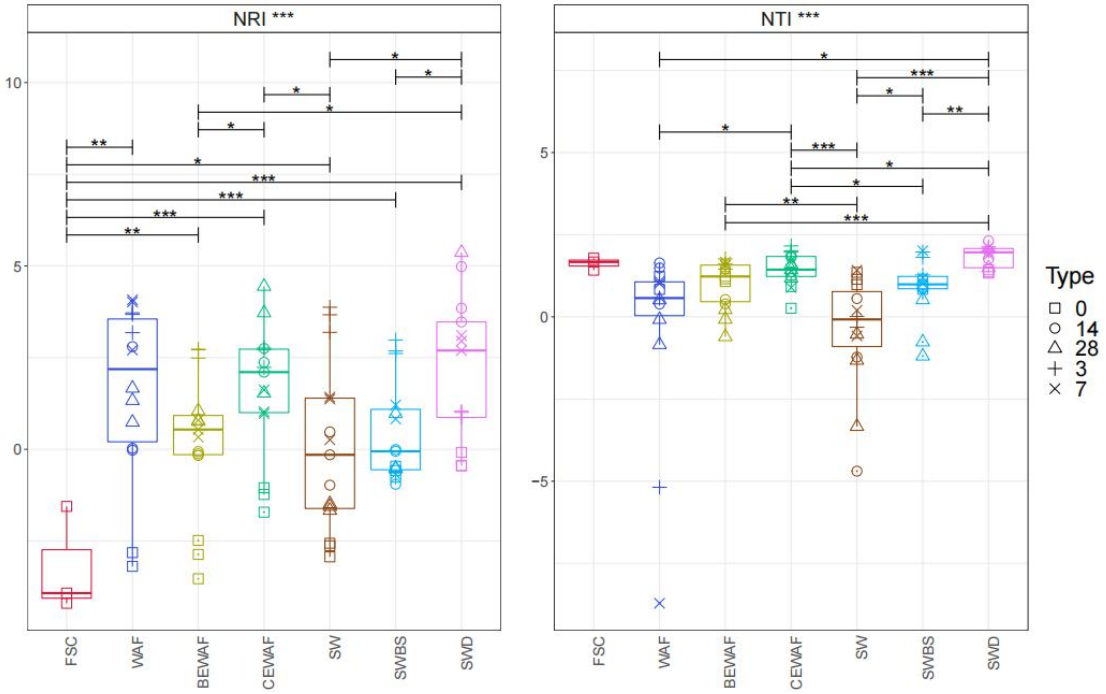


Fig11: Test NRI and NTI using frequency parameter

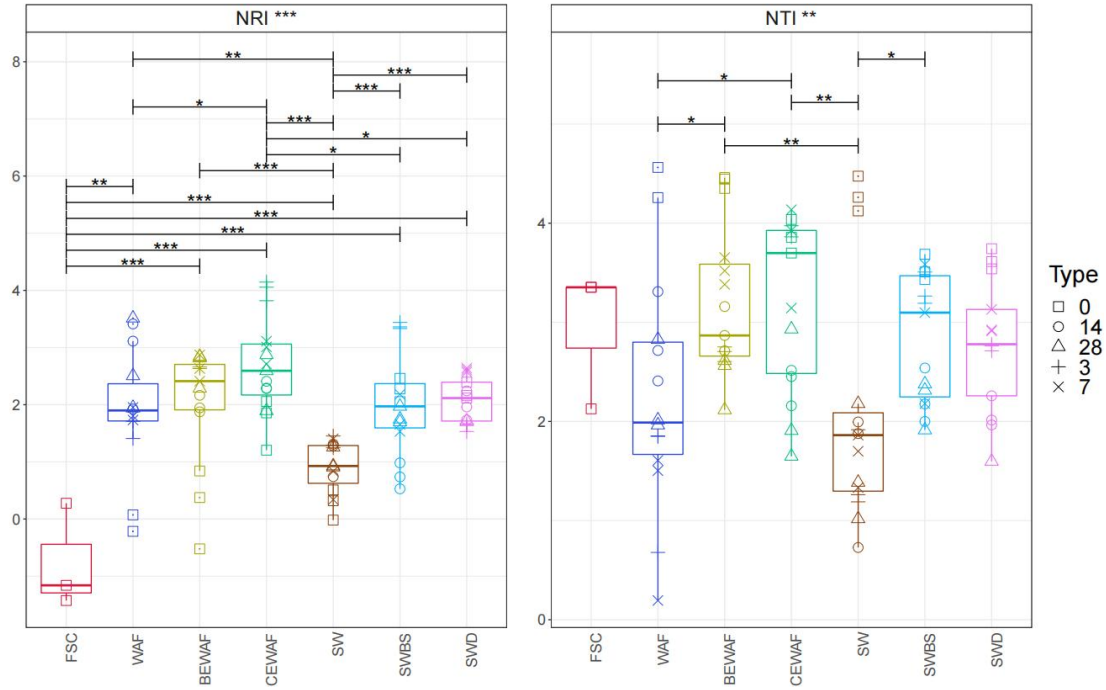


Fig12: Test NRI and NTI using taxa parameter

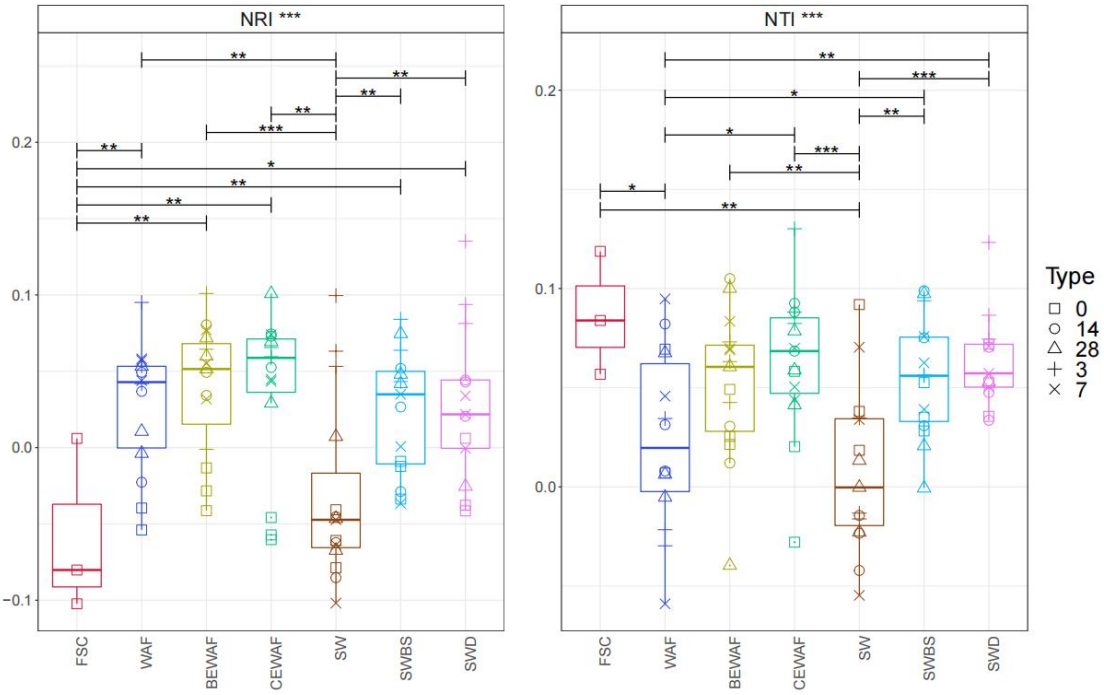


Fig13: Test NRI and NTI using trialswap parameter

It can be seen that there are problems with these results. In the calculation of NRI and NTI with richness as the parameter (Fig. 2), and returned a very ideal result. All NRI values are significantly greater than 0 and NTI values are significantly greater than 2. This shows the environmental filtration in these petroleum processing. The subsequent tests on the core microbiome also verified this result (many of the dominant bacteria in these treatments came from the same genus). This is because the initial richness is an important factor affecting the community environment, and the test results without controlling the richness are unreliable.

The null model method of this project needs to be improved. These are the results based on jaccrd distance. It can be seen that these results are not expected. A good model should show deterministic in SWD and CEWAF.

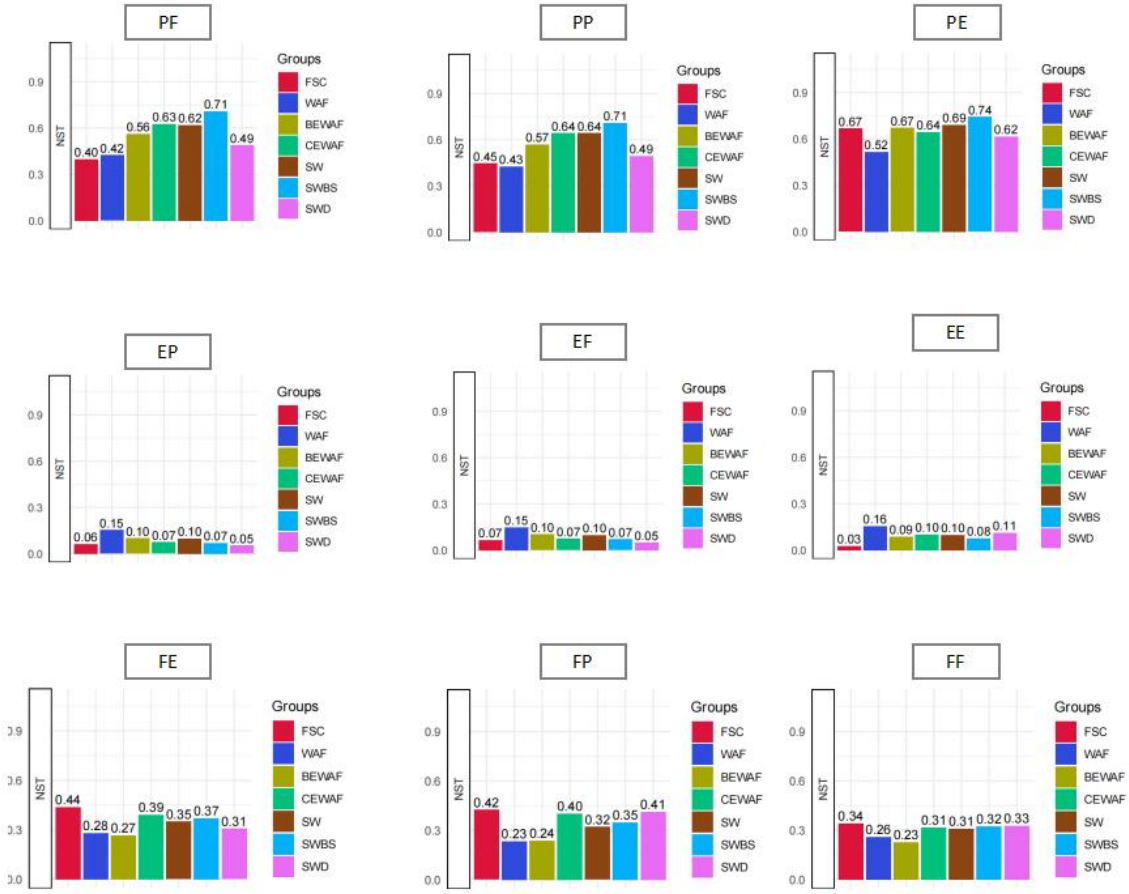


Fig 14: NST calculated by Jaccrd distance method

Appendix II: Core microbiome taxa

The project shows the difference of microbial communities under different conditions through differential analysis. And the diversity of communities was analyzed with the returned P values. The following are the P values of all bacteria in different conditions.

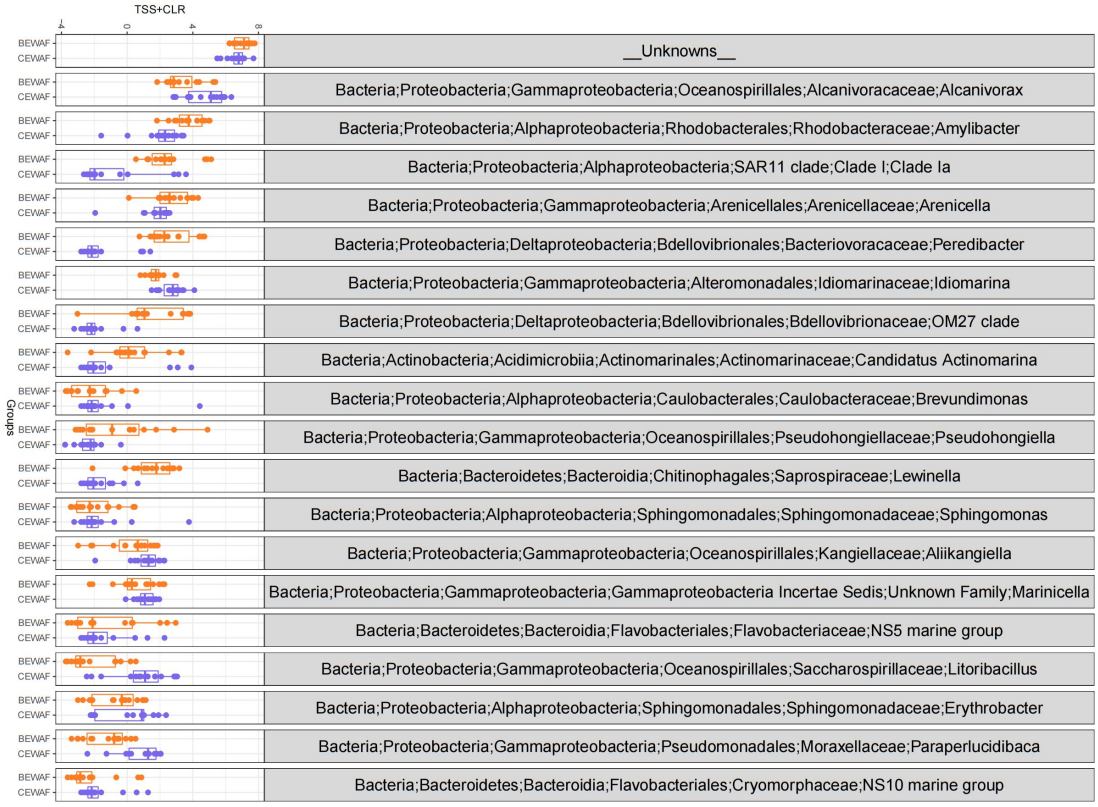


Fig15: Differences between BEWAF and CEWAF in bacterial taxa

Fig16: Differences bacterial taxa between BEWAF and CEWAF in day7

Fig17: Differences bacterial taxa between BEWAF and CEWAF in day14

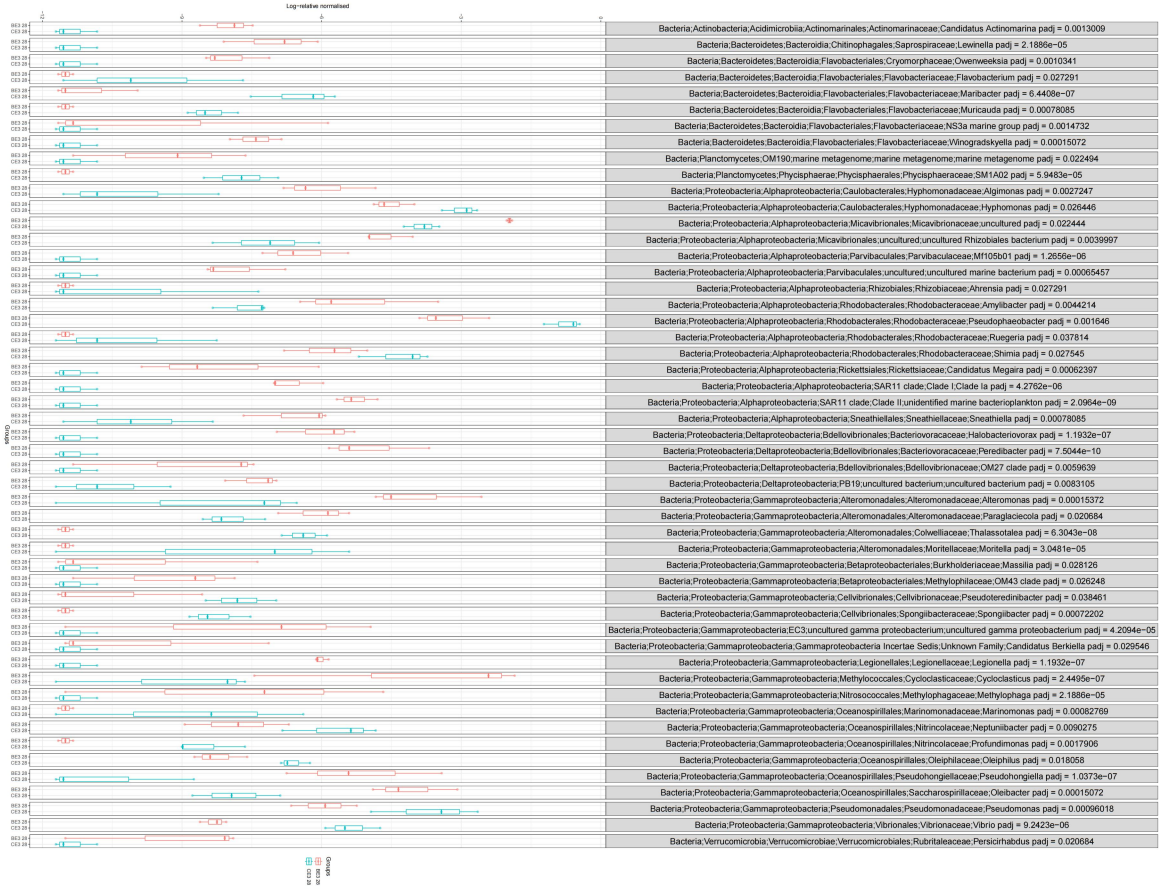


Fig18: Differences bacterial taxa between BEWAF and CEWAF in day28

	baseMean	log2FoldChange	pvalue	padj	Upregulated
Oleispira	14602.08197	6.030477915	1.52E-20	3.32E-18	BEWAF
Aestuariicella	471.4295134	9.015463045	8.88E-14	9.68E-12	BEWAF
Polaribacter	8.946599724	-5.324463998	1.74E-11	1.26E-09	FSC
Pseudomonas	390.607118	8.744056733	2.44E-11	1.33E-09	BEWAF
Lentisphaera	6.085111772	-4.681236371	1.09E-09	4.75E-08	FSC
;Alteromonas	1030.749341	10.14430254	3.84E-09	1.39E-07	BEWAF
Rubritalea	5.194989209	-4.403500229	7.87E-09	2.45E-07	FSC
Arenimonas	6.062590487	-4.356203705	3.35E-08	9.12E-07	FSC
Bacteroidetes bacterium	5.252482939	-4.003800571	2.17E-07	5.26E-06	FSC
Amylibacter	1163.084377	3.808260651	2.53E-07	5.51E-06	BEWAF
Zhongshania	688.106287	3.120763035	6.19E-07	1.23E-05	BEWAF
Marinosulfonomonas	4.974655774	-4.326198352	9.29E-07	1.56E-05	FSC
Ilumatobacter	3.478606753	-3.646542953	8.76E-07	1.56E-05	FSC
Oleibacter	545.0001299	5.626514054	3.55E-06	5.53E-05	BEWAF
Coxiella	15.32564493	-4.633522119	5.07E-06	7.37E-05	FSC
Oleiphilus	66.0719622	4.156413093	9.73E-06	0.000132577	BEWAF
Sva0996	8.791601355	-3.83230626	1.13E-05	0.000145137	FSC
Aquibacter	16.1103243	-4.557252193	1.56E-05	0.000189084	FSC
Maribacter	3.201429711	-3.073841327	2.02E-05	0.000231711	FSC
ProteoRhizobiales bacterium	220.5867978	7.919094001	2.51E-05	0.000273738	BEWAF
Synechococcus	35.03841253	-4.770782098	3.05E-05	0.000309711	FSC
Litorivivens	2.825633536	-3.204675428	3.13E-05	0.000309711	FSC
metagenome	19.90511829	-4.578513942	3.38E-05	0.000320342	BEWAF
Maritimimonas	4.226572886	-3.563606621	3.87E-05	0.000351507	FSC
Flavobacterium	143.6762465	-4.886046726	4.45E-05	0.000387768	FSC
Blastopirellula	3.125372095	-2.907813165	6.24E-05	0.00052349	FSC

Table 5: partial P values returned by BEWAF relative to the original community

	baseMean	log2FoldChange	pvalue	padj	Upregulated
Persicirhabdus	35.29754484	-6.493714184	2.78E-13	5.21E-11	FSC
Pseudomonas	2167.192037	11.22420991	8.36E-13	7.82E-11	CEWAF
Aestuariicella	269.4246058	8.215597386	1.61E-12	1.00E-10	CEWAF
Oleispira	9434.168872	5.401400016	1.83E-11	8.54E-10	CEWAF
Oleiphilus	202.9696337	5.789882583	2.46E-11	9.20E-10	CEWAF
Polaribacter	9.11268154	-5.035825253	2.24E-10	6.99E-09	FSC
Sva0996	7.508818538	-5.005230361	3.90E-10	1.04E-08	FSC
Loktanella	11.16189882	-5.130292444	1.74E-09	4.07E-08	FSC
Acanthopleuribacter	7.092665929	-4.909522679	4.18E-09	8.68E-08	FSC
Bacteroidetes bacterium	5.001451286	-4.301935281	2.25E-08	4.21E-07	FSC
Octadecabacter	5.606355895	-4.506604549	4.53E-08	7.70E-07	FSC
Neptuniibacter	210.8432981	7.861495164	6.59E-08	9.48E-07	CEWAF
Zhongshania	800.9417113	3.348492919	6.37E-08	9.48E-07	CEWAF
NS2b marine group	4.149685229	-3.955725892	1.04E-07	1.39E-06	FSC
Pseudophaeobacter	20351.05617	7.075662096	1.34E-07	1.68E-06	CEWAF
Kordia	22.56346195	-5.300834622	1.55E-07	1.81E-06	FSC
CC9902	32.92071814	-5.381804543	1.68E-07	1.84E-06	FSC
Oceanococcus	10.04601567	-4.675908121	2.30E-07	2.39E-06	FSC
Lewinella	120.4344895	-6.283016781	4.38E-07	4.31E-06	FSC

Table 6: partial P values returned by CEWAF relative to the original community

	baseMean	log2FoldChange	pvalue	padj	Upregulated
Pseudohongiella	3193.928115	12.00565786	2.85E-10	4.91E-08	WAF
hydrothermal	15.9768518	-5.528573143	6.58E-10	5.66E-08	FSC
Polaribacter	8.759412757	-4.751569409	1.82E-08	1.02E-06	FSC
Pseudomonas	288.5265863	8.544442706	2.37E-08	1.02E-06	WAF
C1-B045	4447.322535	12.47885327	5.37E-08	1.85E-06	WAF
Arenimonas	5.709656829	-4.377916829	8.92E-08	2.56E-06	FSC
Rubritalea	5.108643923	-4.176765876	1.71E-07	4.19E-06	FSC
Bacteroidetes	4.716821063	-4.034678468	2.08E-07	4.47E-06	FSC
Coxiella	14.11635439	-4.774575824	1.02E-06	1.95E-05	FSC
Porticoccus	673.4906437	9.758005818	1.58E-06	2.72E-05	WAF
Ilumatobacter	3.34500287	-3.354469838	9.63E-06	0.00015061	FSC
Marinosulfonomona	4.028214541	-3.755847764	1.08E-05	0.000154358	FSC
Rheinheimera	3.912379322	-3.663918461	2.31E-05	0.000284103	FSC
Marinobacter	320.4949313	5.512477002	2.30E-05	0.000284103	WAF
Octadecabacter	5.131471077	-3.526934829	2.55E-05	0.00029261	FSC
Loktanella	9.864990574	-4.045873607	3.44E-05	0.000369916	FSC
Synechococcus CC9902	30.45163548	-4.681671173	7.76E-05	0.000785532	FSC
Acanthopleuribacter	8.229258821	-3.828602286	0.000158733	0.001447235	FSC
Crocinitomix	22.11942024	-4.274820534	0.000159869	0.001447235	FSC
Methylophaga	116.8807017	7.231202339	0.000191096	0.001643422	WAF
Mf105b01	40.8365668	5.718465992	0.000272744	0.002233902	WAF
seawater metagenome	2.541833749	-2.778120771	0.000291339	0.002277738	FSC
Cellvibrionaceae; Aestuariicella	30.22627045	5.28437513	0.000457535	0.003421564	WAF

Table 7: partial P values returned by WAF relative to the original community

The paper mainly analyzes the core microbiome of three oil treatments, and the community changes of their control groups are shown in the figure. The pure dispersant treatment SWD had the strongest damage to the community, with less than 25 core species.In contrast, all treatment solution were not as rich as the original communities.

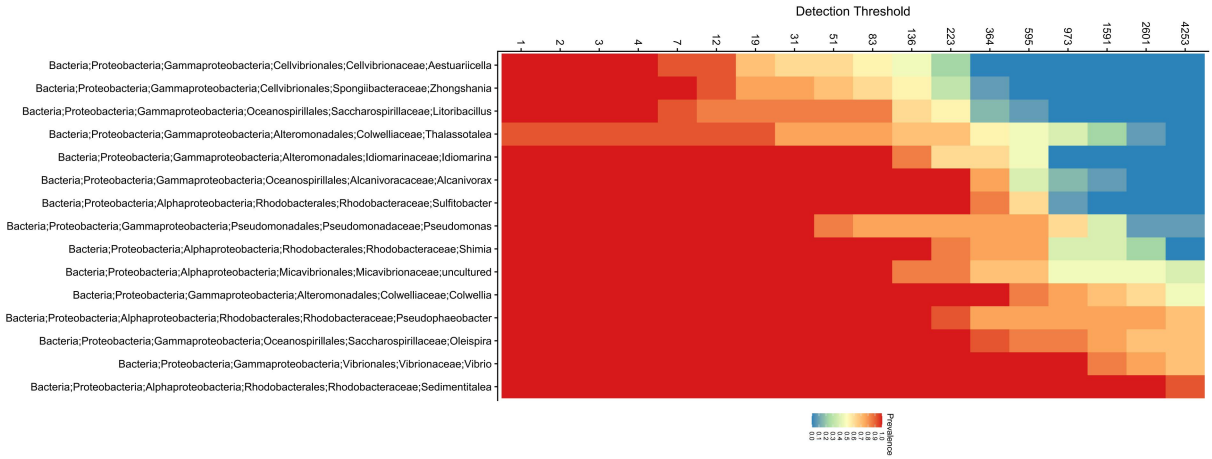


Fig 19: Core microbiome in SWD

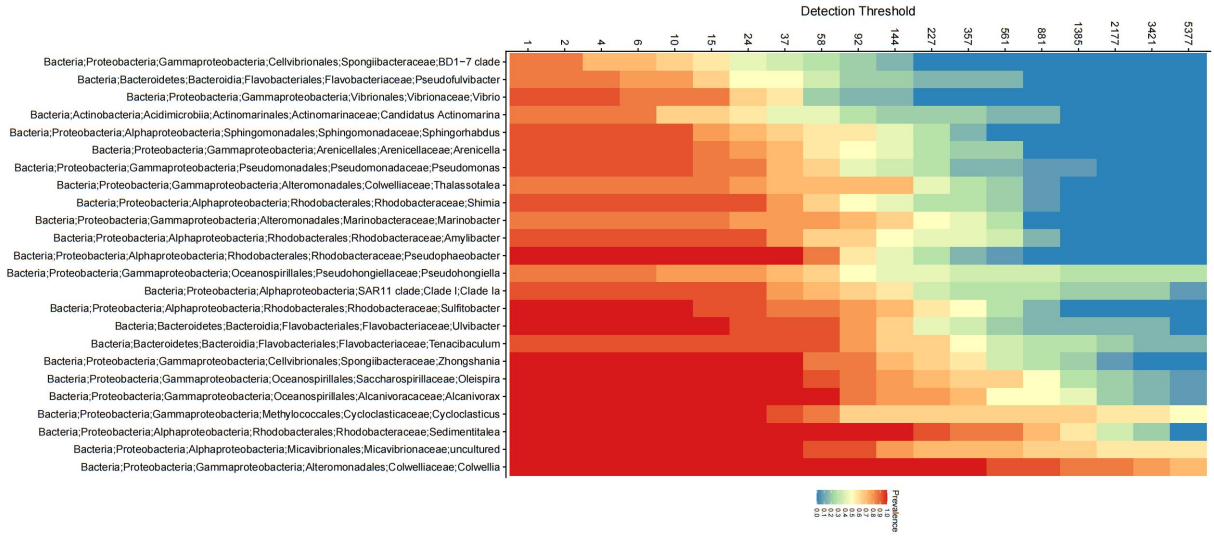


Fig 20: Core microbiome in WAF

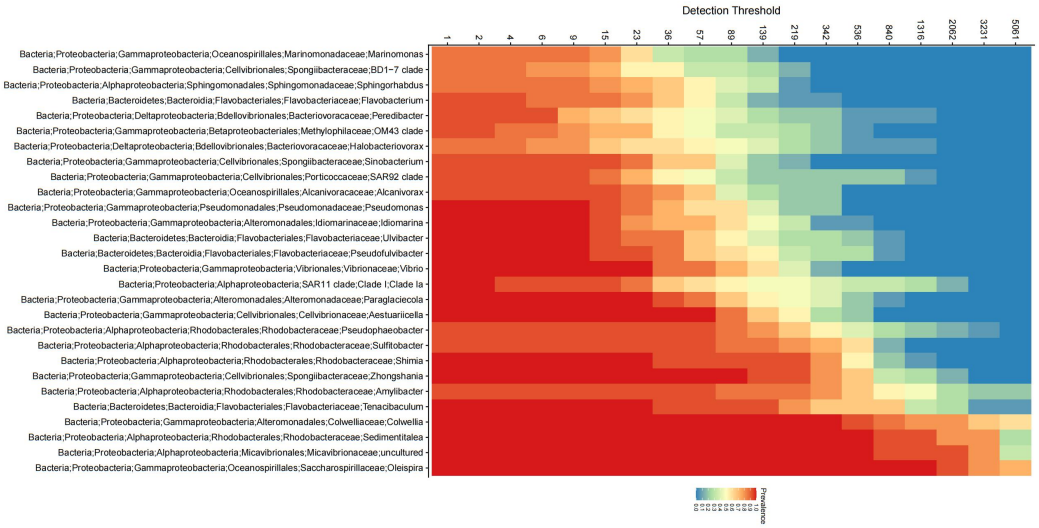


Fig 21: Core microbiome in SWBS

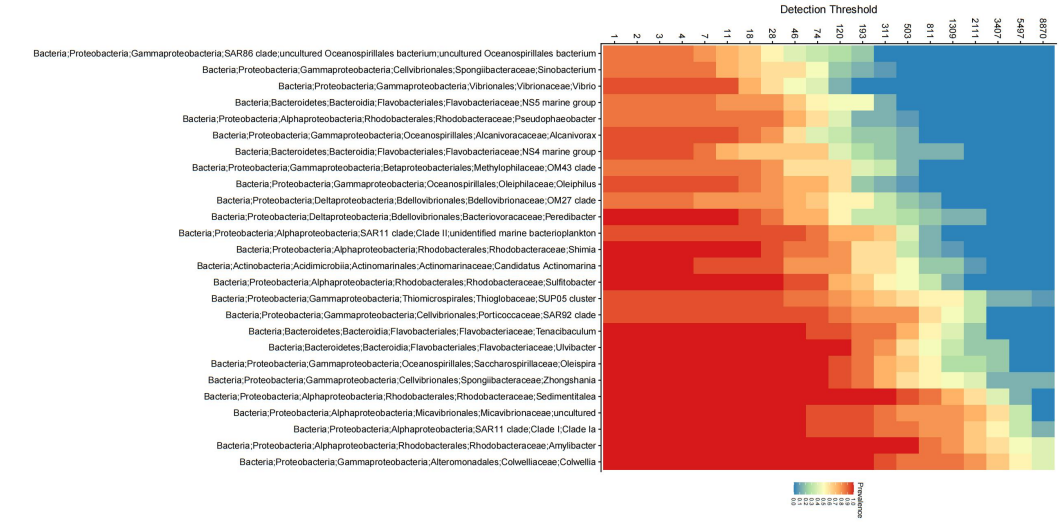


Fig 22: Core microbiome in SW

Appendix III: Cross-validation model

The project provides the best combination of cross validation SWD &BEWAF, and the data returned by other modules are as follows

	Shannon								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	3.14168 ***	0.09248	-0.00000	0.10145	2.95790 – 3.32546	-0.20162 – 0.20162	33.97159	**2.285e-52**	88.00000
Treatment SWD	-0.69157 **	0.24333	-0.28995	0.10202	-1.17513 – -0.20800	-0.49270 – -0.08721	-2.84209	**5.570e-03**	88.00000
Observations	90								
R^2 / R^2 adjusted	0.084 / 0.074								

	Shannon								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	3.06181 ***	0.10182	-0.00000	0.10023	2.85943 – 3.26419	-0.19922 – 0.19922	30.07081	**9.671e-48**	87.00000
Treatment BEWAF	0.41001	0.23069	0.18224	0.10254	-0.04851 – 0.86854	-0.02156 – 0.38605	1.77731	7.901e-02	87.00000
Treatment SWD	-0.61169 *	0.24456	-0.25646	0.10254	-1.09779 – -0.12560	-0.46027 – -0.05266	-2.50116	**1.425e-02**	87.00000
Observations	90								
R^2 / R^2 adjusted	0.116 / 0.096								

	Shannon								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	3.14065 ***	0.13753	0.00000	0.09811	2.86720 – 3.41411	-0.19506 – 0.19506	22.83542	**5.003e-38**	85.00000
Treatment BEWAF	0.45290 *	0.22692	0.20131	0.10086	0.00172 – 0.90408	0.00076 – 0.40185	1.99584	**4.915e-02**	85.00000
Treatment SWD	-0.60050 *	0.24174	-0.25177	0.10135	-1.08114 – -0.11985	-0.45329 – -0.05025	-2.48407	**1.495e-02**	85.00000
Treatment FSC	0.80522	0.47348	0.17239	0.10137	-0.13619 – 1.74663	-0.02916 – 0.37394	1.70063	9.267e-02	85.00000
Incubation day	-0.01171	0.00871	-0.13576	0.10098	-0.02902 – 0.00560	-0.33653 – 0.06501	-1.34450	1.824e-01	85.00000
Observations	90								
R^2 / R^2 adjusted	0.173 / 0.134								

	Shannon								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	3.18768 ***	0.14817	0.00000	0.09826	2.89302 – 3.48234	-0.19539 – 0.19539	21.51318	**6.371e-36**	84.00000
Treatment WAF	-0.20720	0.24060	-0.08956	0.10400	-0.68566 – 0.27126	-0.29639 – 0.11726	-0.86117	3.916e-01	84.00000
Treatment BEWAF	0.40377	0.23432	0.17947	0.10415	-0.06220 – 0.86974	-0.02765 – 0.38658	1.72316	8.854e-02	84.00000
Treatment SWD	-0.64908 *	0.24859	-0.27214	0.10423	-1.14344 – -0.15472	-0.47941 – -0.06487	-2.61101	**1.069e-02**	84.00000
Treatment FSC	0.75819	0.47734	0.16232	0.10219	-0.19105 – 1.70743	-0.04090 – 0.36555	1.58838	1.160e-01	84.00000
Incubation day	-0.01150	0.00872	-0.13342	0.10117	-0.02885 – 0.00584	-0.33460 – 0.06777	-1.31878	1.908e-01	84.00000
Observations	90								
R^2 / R^2 adjusted	0.180 / 0.131								

Table 8: Data returned by all cross validation models