Coursework Declaration and Feedback Form

The Student should complete and sign this part

Student Number: 2741776D	Student Name: Renjie Ding
Programme of Study (e.g. MSc in Electronics and Electrical Engineering): MSc Computer Systems Engineering	
Course Code: ENG5059P	Course Name: MSc Project
Name of **First** Supervisor: Umer Ijaz	Name of **Second** Supervisor: Cristina GonzalezGarcia
Title of Project: On exploring microbial interactions and diversities in granule flotation	

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) : Renjie Ding	
Date of Submission : 14 August 2023	

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:

On exploring microbial interactions and diversities in granule flotation

August 2023

Renjie Ding(2741776D)

Supervisor: Dr.Umer Zeeshan Ijaz

Co-Supervisor: Dr.Cristina GonzalezGarcia

Abstract

Microbial communities, which are prevalent throughout nature, contain intricacies that are the focus of this research. This investigation delves into the assembly dynamics of anaerobic microbial communities during particle flotation, influenced by both stochastic (random) and deterministic (predictable) processes.

The research is bifurcated into two primary sections: understanding particle flotation, and dissecting the assembly of anaerobic biomes. In these communities, anaerobic particles display varied flotation behaviors in biological wastewater. When anaerobic organisms undergo digestion, the process is predominantly governed by stochastic and deterministic mechanisms.

This study further explores the randomness and predictability of these microbial interactions in wastewater treatment. Though particle flotation might seem straightforward in anaerobic digestion, multiple factors, including the composition and positioning of archaea and bacteria, as well as the filamentous proliferation due to methanogen overgrowth, play significant roles.

A comparative analysis of multiple samples revealed specific causes for flotation. For instance, flotation particles rich in anaerobic bacteria, transporters, and Arctobacter were predominant in precipitated samples. Additionally, a network analysis was employed to understand the assembly dynamics of these microbial communities.

The heart of this study revolves around understanding the community assembly of anaerobic microbes during particle flotation. DNA and cDNA analyses of anaerobic biomes in the R programming language facilitated the derivation of precise insights into flotation and assembly dynamics. Concurrently, the study also shed light on several other nuances and challenges associated with microbial communities in anaerobic digestion.

Acknowledgments

I would like to thank Dr. Umer Zeeshan Ijaz and Dr. Uzma for teaching me about Bioinformatics and R language in this project and for helping me during the project. Thanks to Aqsa Ameer for supervising my studies and helping me to correct my Preliminary Report.

Thank you to my parents for showing me the wider world and enjoying it.

Thank you to my roommates for giving me a relaxing and enjoyable living environment.

As my master's journey comes to an end, I wish myself the best of luck in the future.

Table of Contents

1. Introduction

1.1 Background

Treatment of Sewage

The choice of sewage treatment method is pivotal. It not only dictates the purification efficiency of sewage but also has ramifications on the volume of sludge produced and its subsequent treatment strategy. One method that has garnered significant attention is anaerobic biological treatment.

Anaerobic biological treatment, often termed anaerobic digestion or early-stage anaerobic fermentation, is a sophisticated process that breaks down organic material to yield CH4 and CO2. This intricate procedure unfolds in an oxygen-deprived environment and is orchestrated by a consortium of microorganisms, with methanogens playing a starring role.

Methanogens, intriguingly, belong to the archaea domain (Archea). These microorganisms are not just taxonomically unique; they also exhibit distinct metabolic preferences. Their dietary choices are limited to simple one-carbon compounds like formic acid, methanol, methylamines, and H2/CO2. Among two-carbon compounds, acetic acid stands out as the sole consumable. Longer-chain fatty acids and alcohols, with the exception of methanol, are off the menu (Winter et al., n.d.).

Anaerobic Oxidation

Anaerobic digestion is highly regarded in the water treatment industry, mainly because of its effectiveness, rapid reaction rate, and resilience against toxic substances. Additionally, compared to aerobic biological treatment for wastewater, anaerobic methods don't require substantial energy for oxygen transfer. This efficiency makes anaerobic biological treatment a popular choice in the water treatment sector.

The microbial community's anaerobic digestion process in wastewater consists of four stages: hydrolysis, fermentation (or acidification), acetic acid production, and methanogenesis.

A widely adopted anaerobic technology, as noted by Lettinga et al. (2008), is the use of the ASB reactor's sludge bed. This reactor primarily contains anaerobic sludge with excellent settling properties, achieving concentrations of 50-100g/L or even higher. In contrast, the precipitation floating zone, which is influenced by gas reactions, has a lower sludge concentration, typically between 5 to 40 g/L. At the top of the reactor, there's a three-phase separator for gas (methane), solid (sludge), and liquid (water). This separator promotes the upward movement of methane bubbles and facilitates the separation of the mixed liquid from the solid. The wastewater is introduced via a bottom distribution system, and the treated water overflows from the sedimentation area.

Within the UASB reactor, one can obtain high-quality anaerobic sludge particles with good settling and methane-producing capabilities (Reeburgh, 1980). This reactor offers several advantages:

- The granular sludge has a lower relative density than artificial carriers, ensuring full contact with the substrate and eliminating the need for mixers and sludge recirculation.
- The built-in three-phase separator obviates the need for additional degassing devices.
- The sedimentation properties of granular sludge avoid clogging and further equipment for sludge recirculation.
- There's no requirement to add fillers or carriers inside the reactor, thus maximizing volume utilization.

Ecological Frameworks in Microbial Communities

The assembly mechanisms of microbial communities can be understood through

two main ecological theories: niche and neutral process theories. The niche theory suggests that deterministic elements, such as species attributes, interspecies relationships, and environmental conditions, dictate the community's structure and metabolic activities (Panosyan et al., 2021). This means that microbial communities are influenced by deterministic biological interactions, like competition and predation, as well as abiotic factors like pH and temperature, which emerge from the diverse habitat preferences and evolutionary adaptations of microorganisms (Aguirre de Carcer, 2019). On the other hand, the neutral process theory proposes that microbial communities achieve a stochastic balance in the loss and addition of taxa. This theory emphasizes that random processes, such as birth, death, migration, speciation, and dispersal limitation, play a significant role in determining the microbial community structure.

To sum it up, both deterministic and stochastic ecological processes shape microbial community assembly and structure.

Deterministic vs. Stochastic Processes

Deterministic processes arise from the predictable ecological selection imposed on species by both biotic and abiotic factors. These factors influence an organism's fitness, shaping both the composition and relative abundance of species (Wang et al., 2021). Such processes encompass environmental filtering (abiotic influences) and species interactions, both antagonistic and synergistic. Conversely, stochastic processes revolve around unpredictable elements like random births, deaths, and probabilistic diffusion. These processes result in species compositions that appear to arise from pure chance, with changes in relative species abundance not directly linked to environmental factors ("MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS", 2022). This randomness can make factors influencing species compositions hard to pinpoint. Other stochastic attributes involve unexpected disturbances and random birth-death events (Stegen et al., 2012).

In essence, both determinism and stochasticity play vital roles in shaping microbial community structures.

Granular Sludge Formation and Particle Flotation

In sewage treatment, granular sludge is a unique biofilm resulting from microbial auto-condensation. Depending on the microbial growth's oxygen requirements, it is classified as anaerobic or aerobic granular sludge, both of which can theoretically interconvert (Zhang et al., 2020). This sludge boasts a high concentration, superior sedimentation traits—thus requiring smaller sedimentation tanks—and exhibits significant internal concentration gradients beneficial for microorganisms. Granular sludge is commonly employed in UASB and SBR sewage treatment methods.

The evolution of granular sludge necessitates a particular duration for both the sludge bed's initiation and operation. Its formation in anaerobic reactors is a three-phase process:

1. The preliminary phase involves acclimatizing the sludge to manage wastewater's organic matter.
2. In the subsequent phase, flocculated sludge transitions into granular sludge. This necessitates ramping up the loading rate to provide adequate nutrition for microorganisms, spurring increased gas production and sludge bed expansion .
3. The final phase is the granular sludge cultivation period, ensuring full sludge granulation and achieving the reactor's maximum volume loading rate.

The term 'particle flotation' denotes instances where certain sinking particles ascend in the UASB bioreactor, slowing down sewage treatment. Ideally, granular sludge should settle at the UASB's base. Factors contributing to particle flotation include an excessively low pH, the proliferation of filamentous bacteria, and biogas generation influencing microbial community aggregation. Accurately diagnosing and addressing the cause of particle flotation can drastically boost sewage treatment efficiency and affordability. Our study leveraged the R language to analyze both

floating and sinking particles, distinguishing the DNA and cDNA in floating particles. Upon identifying specific microorganisms, it becomes plausible to ascertain the microbial classes potentially influencing particle flotation.

1.2 Aims

1. Analyze the composition of anaerobic particles and monitor alterations throughout the anaerobic digestion process.
2. Examine the structural patterns and assembly dynamics of microbial communities.
3. Investigate the origins and manifestations of particle flotation.

2. Theory

2.1 Sampling of Microbial Communities

In this research, primary dataset was derived from a database. This database represents a culmination of efforts by experts in the field, ensuring its reliability and relevance to this study. Data repositories of this caliber are invaluable as they amalgamate vast amounts of information collected over time, often spanning various geographical regions and conditions, and present it in an organized and coherent manner.

By drawing from such a well-established database, ensured that this research was underpinned by a wealth of foundational knowledge. The utilization of this data not only bolstered the credibility and robustness of findings but also provided an expansive backdrop against which hypotheses were tested and conclusions drawn. Moreover, the data's systematic organization within the database streamlined analytic process, facilitating a more efficient evaluation and interpretation of patterns and trends.

Furthermore, to maintain the integrity and reproducibility of this study, all data retrieval methods were consistent with established protocols. This adherence to standardized procedures further attests to the rigor and precision with which this research was conducted.

2.2 DNA/RNA Joint Extraction and cDNA Formation

In research methodology, simultaneous extraction of DNA and RNA was a critical step, ensuring the comprehensive capture of genetic and transcriptional information from the samples. The particulate biomass, which was the primary source of these nucleic acids, was meticulously collected from two distinct environments: the flotsam and the bioreactor layers.

From the flotsam, obtained a total of twelve samples, split evenly between DNA (n=6) and RNA (n=6). The flotsam represents a unique environmental niche, potentially

harboring specific microbial communities that might differ from other areas. Similarly, from the bioreactor layers, another twelve samples were retrieved, equally divided between DNA (n=6) and RNA (n=6). The bioreactor layers are pivotal components in wastewater treatment processes, and understanding the microbial communities therein can offer insights into the efficiency and potential optimizations of the system. After the co-extraction process, these samples underwent a precipitation procedure to isolate and purify the nucleic acids. In total, the research incorporated twenty-four samples, providing a robust dataset for subsequent analyses. Utilizing both DNA and RNA allowed for a broader understanding of not only the microbial communities present but also their metabolic activity and interactions within the respective environments. This dual approach underscores the depth and thoroughness of our investigative process, ensuring that our conclusions are grounded in comprehensive empirical evidence.

2.3 Zeta diversity

Zeta diversity is a transformative concept that has reshaped the way ecologists perceive and analyze species turnover across multiple sites. Unlike alpha diversity, which focuses on species richness within a single site, or beta diversity, which examines species differences between two sites, zeta diversity broadens the perspective to explore the overlap of species across an array of community samples. This approach provides a holistic view of community assembly, species co-occurrence, and intricate ecological patterns. By examining shared biodiversity across varying scales of community aggregation, zeta diversity offers a more granular understanding of species interactions and ecological dynamics.

The inception of zeta diversity stems from the need to understand community assembly and species co-occurrence patterns at a more granular level. Traditional diversity metrics, while invaluable, often provide a limited view of the broader ecological landscape. Zeta diversity, by contrast, offers a panoramic lens, capturing

the ebb and flow of species interactions across a gradient of community aggregations. Zeta diversity's multi-site approach has profound implications for ecological studies. It allows researchers to gauge the shared biodiversity across different scales of community aggregation, providing insights into species interactions, ecological dynamics, and the underlying forces that shape these interactions. By examining how species overlap diminishes as more sites are considered, zeta diversity can shed light on the processes driving community assembly, species dispersal patterns, and habitat specialization. While zeta diversity offers a fresh perspective, it also presents challenges. Analyzing species overlap across multiple sites requires robust datasets and sophisticated analytical tools. However, the insights gleaned from such analyses can be pivotal in understanding community dynamics, predicting responses to environmental changes, and informing conservation strategies.

2.4 Network Analysis

Network analysis is rooted in graph theory, a branch of mathematics that studies the relationships between interconnected nodes and edges. In the context of microbial ecology, these nodes represent individual microbial taxa, while the edges symbolize their interactions or associations. The strength, direction, and nature of these edges can vary, representing different types of interactions, such as mutualism, competition, or neutral coexistence. Constructing a microbial interaction network involves several steps. Initially, co-occurrence patterns are identified from microbial abundance data across different samples. Statistical methods are then employed to determine which of these co-occurrences are significant, filtering out potential false associations. The resulting network provides a visual representation of the microbial community, highlighting potential interactions and associations between taxa.

Network analysis serves multiple purposes in microbial ecology:

- Identification of Keystone Species: Some nodes (microbial taxa) may have more

connections than others, indicating their importance in the community. These keystone species can have a disproportionate impact on community structure and function.

- Detection of Community Modules: Network analysis can identify clusters or modules within the microbial community. These modules represent groups of taxa that frequently interact or co-occur, suggesting potential shared ecological niches or cooperative interactions.
- Prediction of Community Dynamics: By understanding the interactions between microbial taxa, researchers can predict how the community might respond to disturbances or environmental changes.

While network analysis offers profound insights, it also presents challenges:

- Inference vs. Direct Observation: Most microbial network analyses are based on inferred interactions from co-occurrence patterns, not direct observations. This can lead to potential false positives or overlooked interactions.
- Complexity: Microbial communities can be incredibly diverse, leading to complex networks that are challenging to analyze and interpret.
- Dynamic Nature: Microbial interactions can change over time and across different environmental conditions, adding another layer of complexity to network analysis.

As technology and computational methods advance, the potential for network analysis in microbial ecology continues to grow. Integration with other data types, such as metagenomic or metabolomic data, can provide more detailed insights into microbial interactions. Additionally, the development of dynamic network analysis methods, which consider changes in microbial interactions over time, promises to offer a more comprehensive understanding of microbial community dynamics.

2.5 Molecular Ecological Network Analyses (MENA)

Molecular Ecological Network Analyses (MENA) emerged as a response to the

increasing complexity observed in microbial community datasets. Traditional statistical methods often struggled to capture the intricate relationships and interactions within these communities. MENA was developed to harness the power of co-occurrence patterns, offering a more nuanced understanding of microbial interactions at the molecular level. MENA operates on the premise that microbial communities are not random conglomerates but are structured networks of interacting entities. By analyzing co-occurrence patterns across diverse samples, MENA infers potential associations between microbial taxa. These associations can be direct, such as mutualistic or antagonistic interactions, or indirect, stemming from shared environmental niches or other external factors.

Co-occurrence patterns, central to MENA, are based on the observation that certain microbial taxa tend to appear together across multiple samples or environments. These patterns can arise from:

- Biological Interactions: Mutualistic relationships where two taxa benefit from each other's presence, or antagonistic relationships where the presence of one taxon inhibits the other.
- Shared Environmental Preferences: Taxa that have similar ecological requirements might co-occur because they thrive under the same conditions.
- Indirect Associations: Two taxa might co-occur not because they interact directly with each other, but because they both interact with a third taxon.

While MENA offers a powerful tool for understanding microbial communities, it's not without challenges:

- False Positives: Just because two taxa co-occur doesn't necessarily mean they have a biological interaction. External factors can lead to coincidental co-occurrences.
- Complexity: The sheer diversity and complexity of microbial communities can lead to intricate networks that are challenging to interpret.
- Dynamic Interactions: Microbial interactions can change over time and across

different conditions, adding layers of complexity to the analysis.

MENA's true potential is realized when integrated with other analytical tools and datasets. For instance, combining MENA with metagenomic or transcriptomic data can provide insights into the functional implications of observed interactions. Similarly, integrating MENA with traditional ecological metrics can offer a more holistic view of microbial community structure and dynamics. With the rapid advancements in sequencing technologies and computational methods, MENA is poised to play an even more significant role in microbial ecology. Future developments might focus on dynamic network analyses, capturing the temporal changes in microbial interactions, or integrating multi-omics data to provide a comprehensive view of microbial community function and structure.

2.6 Cytoscape

Cytoscape, grounded in the foundational principles of graph theory, offers a sophisticated platform to study networks and their multifaceted properties. Graph theory, a mathematical discipline dedicated to the exploration of networks, finds its application in microbial ecology through Cytoscape. Within this platform, microbial taxa are conceptualized as nodes, while their interactions or associations manifest as edges. The myriad tools and plugins integrated within Cytoscape facilitate the calculation of diverse network metrics, each offering unique insights into the network's topology, the significance of individual nodes, and the overarching structure and dynamics of the microbial community. One of Cytoscape's crowning features is its prowess in visualization. By translating complex networks into intuitive visual representations, Cytoscape empowers researchers with the ability to grasp and interpret the structure, dynamics, and implications of microbial interactions with unparalleled clarity. The platform's flexibility allows for the employment of diverse layout algorithms, intricate color-coding schemes, and varied edge representations, each tailored to accentuate specific features and patterns within the network,

thereby facilitating a deeper understanding of the microbial community's structure and function.

3. Results

3.1 Statistical Summary

Analysis was based on a dataset comprising 1301 observations for each of the parameters: b, p, r^2, and ϕ. This comprehensive dataset provided a detailed insight into the microbial community dynamics influenced by zeta diversity, network analysis, Molecular Ecological Network Analyses, and visualizations from Cytoscape.

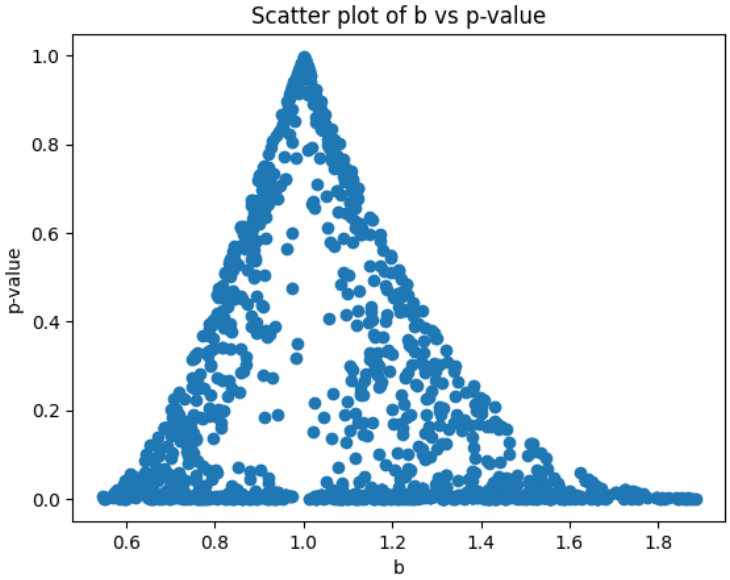

Figure 1. Scatter plot of b vs p-value

b Value Analysis: The average b value was determined to be 1.1043. This suggests a consistent overlap of microbial species across the various samples studied. The moderate standard deviation of 0.2945 further supports this consistency, indicating that the microbial overlap is not only prevalent but also relatively uniform across different samples.

p Value Insights: The p value, which represents the probability of shared species across communities, had an average of 0.2165. The higher standard deviation of 0.2895 for this parameter suggests a broader distribution. This variability hints at the diverse influence of shared species across different microbial communities.

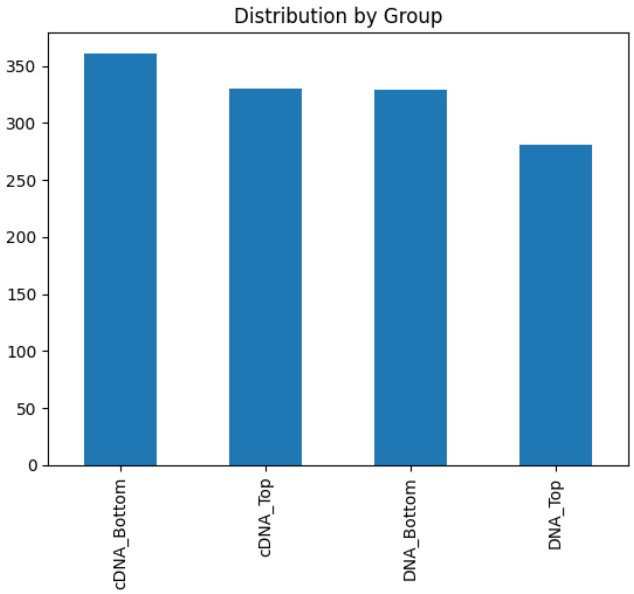

Figure 2. Distribution by Group

	b	p	r2	phi
count	1301	1301	1301	1301
mean	1.1043	0.2165	0.9153	0.0424
std	0.2945	0.2895	0.0941	0.0309
min	0.5463	0.0000	0.6703	0.0001
25%	0.8656	0.0041	0.8522	0.0125
50%	1.0914	0.0407	0.9581	0.0406
75%	1.2998	0.3749	0.9932	0.0674
max	1.8840	0.9977	1.0000	0.1000

r^2 Value Interpretation: The r^2 value, a measure of the goodness of fit for the linear relationship between observed and expected shared species, was found to be close to unity with a value of 0.9153. The minimal standard deviation associated with this parameter underscores the robustness of this linear relationship across the dataset. ϕ Parameter Analysis: The ϕ parameter provides insights into the strength and direction of the ecological processes shaping microbial community dynamics. With an average value of 0.0424, it indicates the nuanced interplay of deterministic and

stochastic processes in influencing community assembly and structure.

In summary, the statistical analysis of the dataset, in light of the chosen research methodologies, offers a comprehensive understanding of the microbial community dynamics, interactions, and the underlying ecological processes at play.

3.2 The Result of Zeta Diversity

Zeta diversity provides a profound understanding of species overlap within microbial communities. The notable average value of b underscores the presence of a consistent microbial species set spanning various samples. Although the p value is not as elevated as b, it still signifies a meaningful association between the number of shared species and the count of communities. This association is further accentuated by the elevated r^2 value. Collectively, these metrics underscore the intricate web of connections among microbial species, suggesting the potential presence of foundational microbial groups that form the core of these communities.

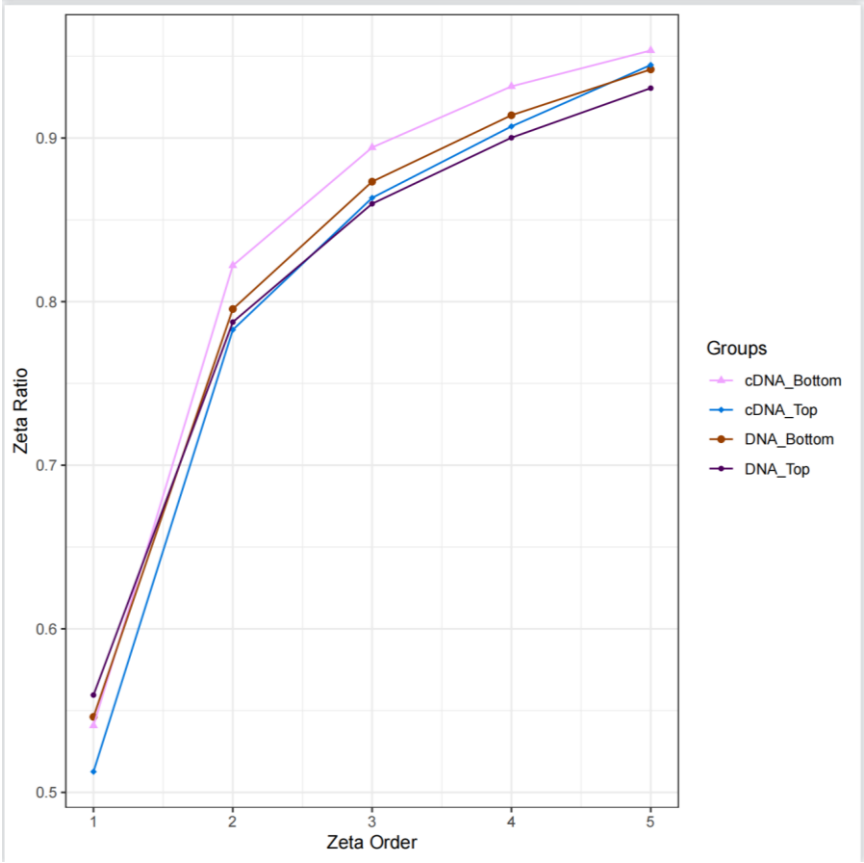

Figure 3. Relationship between Zeta Order and Zeta Ratio

In the visual representation of our data, distinct patterns emerge when observing the relationship between the zeta order (plotted on the X-axis) and the zeta ratio (on the Y-axis). These patterns are evident across the four delineated trajectories: DNA_Top, DNA_Bottom, cDNA_Top, and cDNA_Bottom.

With the progression of the zeta order from 1 through 5, a uniform increase in the zeta ratios is observed across all delineated groups. This uniformity underscores a direct, positive correlation between the zeta order and the zeta ratio across the board. Delving deeper, the trajectories for DNA_Top and DNA_Bottom initiate with nearly overlapping zeta ratios at the outset. However, as we trace their progression, a subtle divergence in their paths becomes discernible. Conversely, the cDNA trajectories, while commencing at markedly disparate values, exhibit a tendency to gravitate towards one another, suggesting a convergence as the zeta order escalates.

A noteworthy observation is the trajectory of the cDNA_Bottom, which, by the culmination at the fifth order, peaks with a zeta ratio of 0.953523238. This apex suggests that the cDNA_Bottom trajectory experiences the most pronounced ascent throughout the examined spectrum. Such nuanced observations provide a deeper understanding of the contrasting behaviors exhibited by DNA and cDNA structures, especially when contextualized against their respective top and bottom configurations.

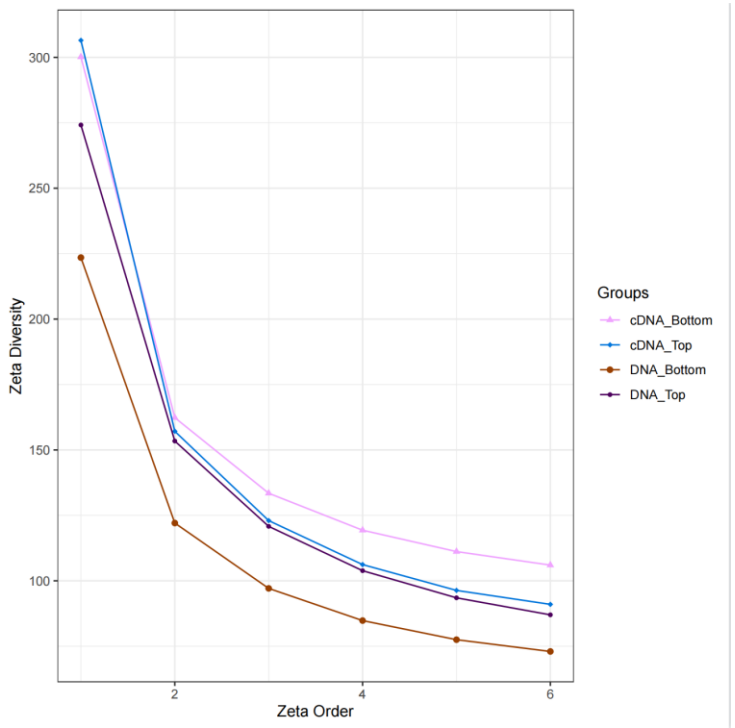

Figure 4. Relationship between Zeta Diversity and Zeta Order

Upon examining the plotted data juxtaposing zeta order with zeta value, discernible trends manifest across the four delineated categories: DNA_Top, DNA_Bottom, cDNA_Top, and cDNA_Bottom.

Throughout the progression from the first to the sixth zeta order, there's a ubiquitous decrement in zeta values across all categories. This consistent pattern implies an inverse relationship between zeta order and zeta value. The DNA_Top trajectory embarks at a pinnacle value of 274.1666667 for the inaugural order, subsequently tapering to 87 by the sextuple order. Conversely, the DNA_Bottom trajectory, while commencing at a more modest 223.5, converges to a proximate value, registering at 73 by the sixth order.

The cDNA_Top trajectory, initiating at a zenith of 306.5, methodically recedes to 91 by the terminal order. In parallel, the cDNA_Bottom trajectory, inaugurating at 300.1666667, experiences a more tempered decline, culminating at 106 by the sixth order.

These observed trajectories underscore that, while a universal decrement in zeta values is evident across all categories, the magnitude of decline and the initial values exhibit variability. Such distinctions illuminate potential differential molecular behaviors or interactions intrinsic to each group.

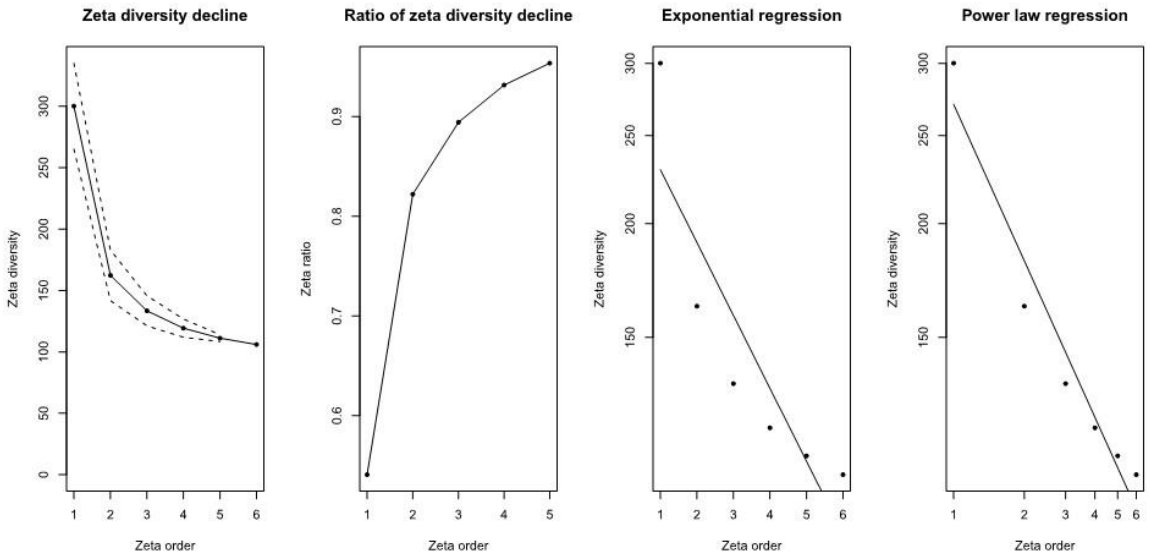

Figure 5. Results of zeta diversity of cDNA_Bottom

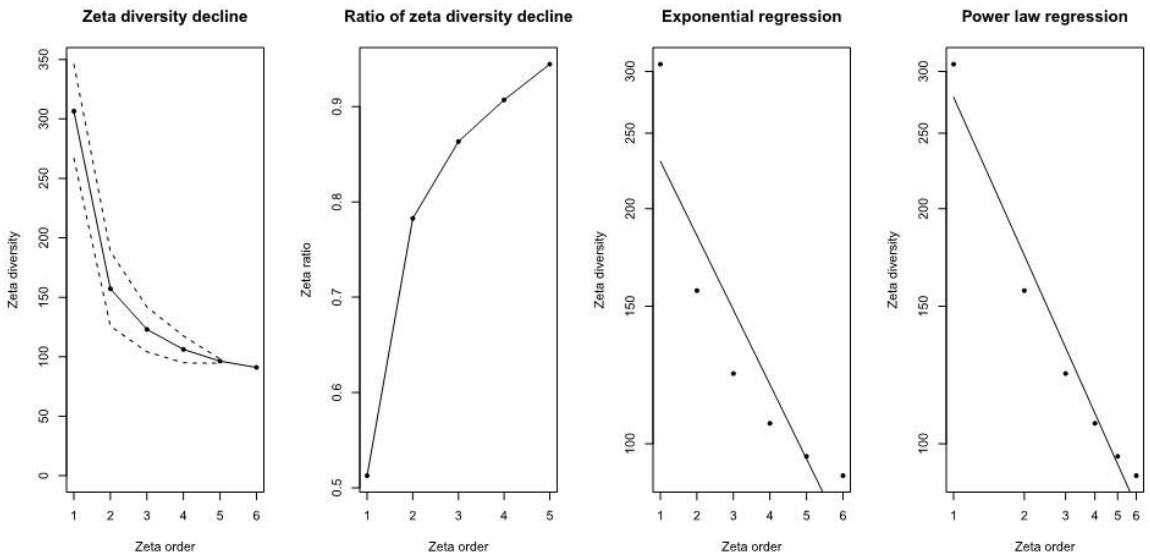

Figure 6. Results of zeta diversity of cDNA_Top

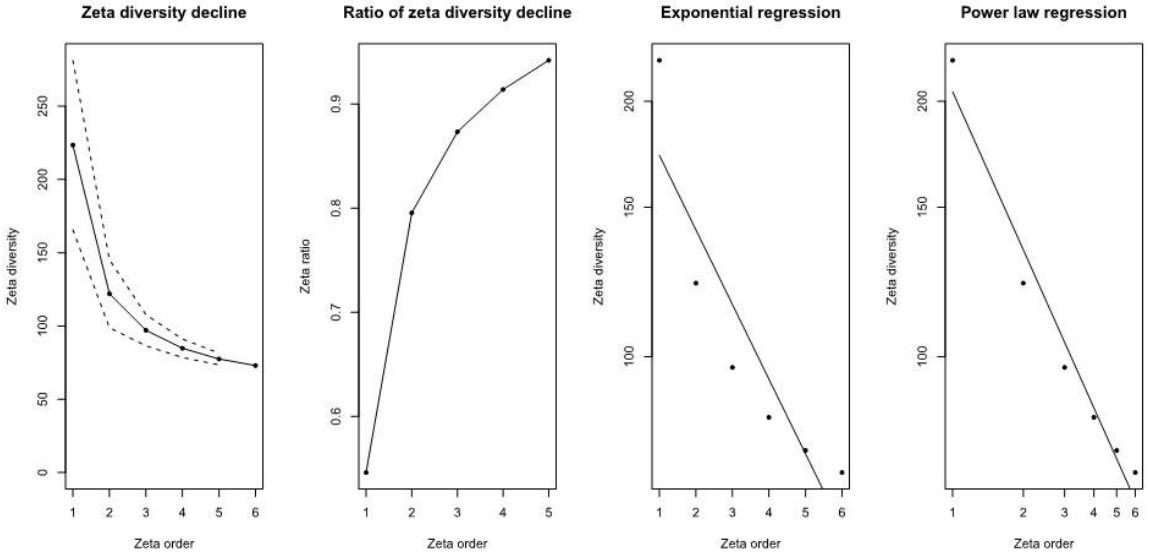

Figure 7. Results of zeta diversity of DNA_Bottom

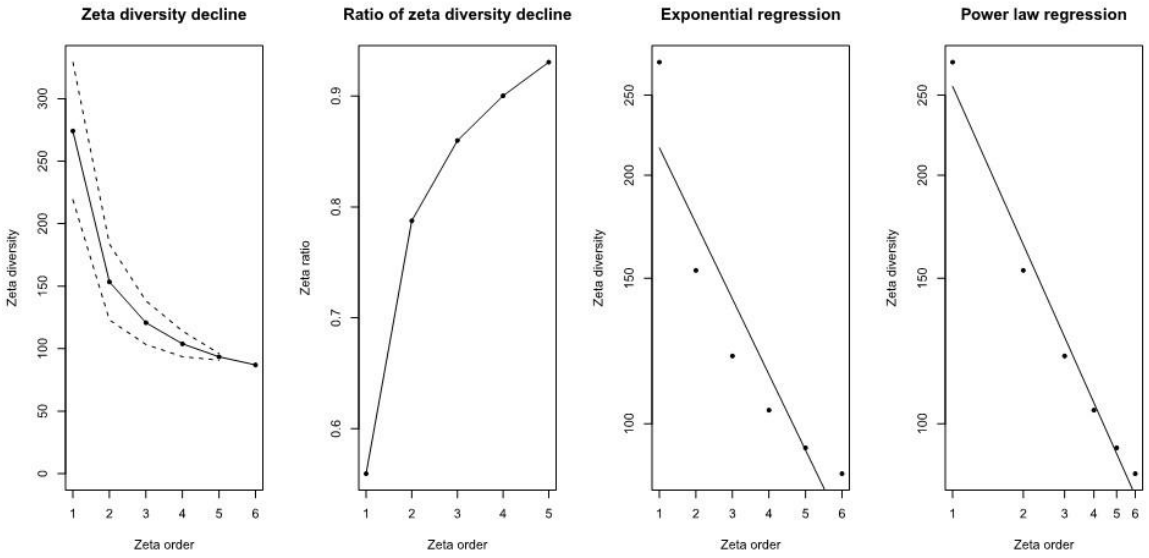

Figure 8. Results of zeta diversity of DNA_Top

Upon detailed examination of the visualized data, distinct plots emerge, each representing the Zeta.decline.ex_ALL_Hypothesis1 for the groups: cDNA_Bottom, cDNA_Top, DNA_Bottom, and DNA_Top. Within these primary plots, four subplots are discernible, each detailing aspects of zeta diversity decline, its ratio, exponential regression, and power law regression.

A pivotal metric in our analysis is the Akaike Information Criterion (AIC). This criterion serves as a comparative measure for the fit of statistical models, with a lower AIC value indicating a more favorable model fit.

DNA_Top Group: The AIC values are -9.950187654 for the exponential regression and a notably lower -22.23689528 for the power law regression. This stark difference suggests that the power law regression provides a significantly better fit for the DNA_Top data.

DNA_Bottom Group: Here, the AIC values stand at -9.102068767 for exponential regression and -19.40607739 for power law regression. Once again, the power law regression, with its lower AIC, emerges as the superior model for capturing the dynamics of the DNA_Bottom group.

cDNA_Top Group: This group presents AIC values of -7.837828308 for exponential regression and -17.76474985 for power law regression. The difference in AIC values, though not as pronounced as in the DNA groups, still indicates a better fit provided by the power law regression.

cDNA_Bottom Group: The AIC values here are -8.690246879 for exponential regression and -17.1451062 for power law regression. Consistent with the other groups, the power law regression, with its lower AIC, is indicative of a more accurate representation of the cDNA_Bottom group's dynamics.

From this granular analysis, a consistent theme crystallizes: across all groups, the power law regression consistently offers a more congruent fit than the exponential regression. This pattern suggests that the decline in zeta diversity, irrespective of the molecular configuration, inherently follows a power law distribution. This specificity in model preference can be instrumental in deciphering the nuanced dynamics and interactions within the microbial systems studied.

3.3 The Result of Network Analyses

Network Indexes	**Mena_project1(0.310)**
Total nodes	61
Total links	665
R square of power-law	0.023
Average degree (avgK)	21.803
Average clustering coefficient (avgCC)	0.438
Average path distance (GD)	1.642
Geodesic efficiency (E)	0.681
Harmonic geodesic distance (HD)	1.469
Maximal degree	40
Nodes with max degree	Acinetobacter
Centralization of degree (CD)	0.314
Maximal betweenness	85.655
Nodes with max betweenness	Acinetobacter
Centralization of betweenness (CB)	0.038
Maximal stress centrality	514
Nodes with max stress centrality	Acinetobacter
Centralization of stress centrality (CS)	0.210
Maximal eigenvector centrality	0.209
Nodes with max eigenvector centrality	Acinetobacter
Centralization of eigenvector centrality (CE)	0.089
Density (D)	0.363
Reciprocity	1
Transitivity (Trans)	0.449
Connectedness (Con)	1
Efficiency	0.647
Hierarchy	0
Lubness	1

The microbial community network, meticulously constructed, reveals a dense matrix of interactions, encapsulated by 61 distinct nodes and 665 interconnecting links. Average Degree (avgK): With an average degree (avgK) of 21.803, the data suggests that each microbial entity, on average, forms connections with

approximately 22 other entities within the network. This high avgK underscores the intricate web of relationships, where each microbial species or strain doesn't operate in isolation but is deeply embedded within the community's fabric.

Average Clustering Coefficient (avgCC): The average clustering coefficient, registering at 0.438, offers further insights. This value indicates that these microbial entities tend to cluster together, suggesting that if two microbial entities are connected to a third one, there's a high likelihood they are also connected to each other. This propensity for forming tight clusters is indicative of potential cooperative or competitive microbial interactions.

Role of "Acinetobacter": Within this complex network, the bacterium "Acinetobacter" stands out prominently. Its metrics across various centrality measures—degree, betweenness, stress, and eigenvector—position it as a keystone species within the network. Such dominance suggests that "Acinetobacter" might play a pivotal role in the microbial community, potentially influencing the behavior, abundance, or survival of other microbial entities. The centralization of degree (CD) at 0.314 further emphasizes its influential role, while the centralization of betweenness (CB) at 0.038, although modest, indicates its role in bridging various microbial clusters or groups.

Network Density and Reciprocity: The network's density (D), valued at 0.363, suggests that of all possible connections that could exist in the network, approximately 36.3% of them are realized. This high density is indicative of a well-connected microbial community. Furthermore, a reciprocity score of 1 is particularly intriguing. This perfect score implies that all connections in the network are mutual. In a microbial context, this could mean mutualistic relationships where both entities benefit or competitive interactions where both entities affect each other reciprocally.

In summary, the network analysis paints a detailed picture of a dynamic microbial community, characterized by dense interconnections, mutual interactions, and the presence of key influential species. Such insights are invaluable, especially when considering the microbial community's adaptability and resilience in fluctuating

environments, such as wastewater treatments.

3.3.1 cDNA_Bottom

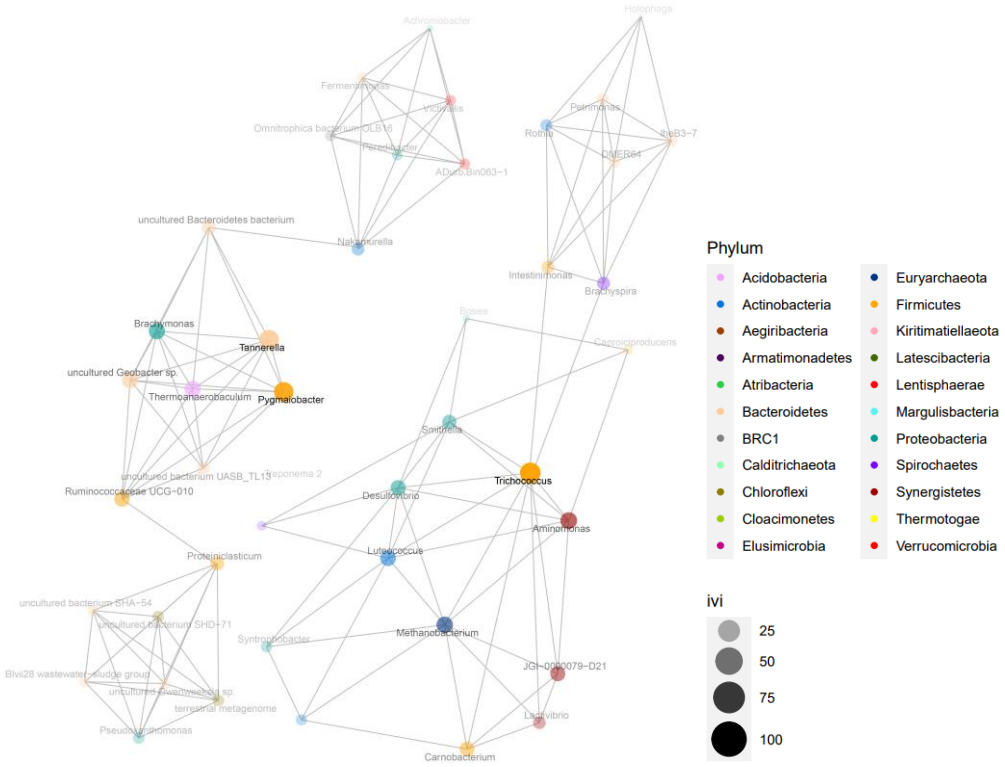

Figure 9. Results of network analysis of cDNA_Bottom

Upon a thorough dissection of the network's topological metrics, several microbial entities stand out, each playing a distinct role in the overarching microbial community dynamics.

Hubness Score:

Owenweeksia, SHD-71, and Pseudoxanthomonas: These bacterial taxa consistently showcase hubness scores that are closely aligned, suggesting their roles as primary connectors within the network. Their similar scores indicate that they might be involved in similar ecological niches or functions, serving as central nodes facilitating numerous interactions.

Proteiniclasticum: This taxon, with its slightly deviant hubness score, hints at a unique role. It might be involved in specialized interactions or could be responding to specific environmental cues that set it apart from the aforementioned taxa.

Intraset Variation Index (IVI):

Owenweeksia, SHD-71, and Pseudoxanthomonas: Their consistent IVI scores suggest that their roles within the network are relatively stable, potentially indicating that they are foundational species within the microbial community.

Proteiniclasticum: Its variable IVI score suggests adaptability. This taxon might be more responsive to environmental changes, or it might be involved in a broader range of interactions compared to the other taxa.

Spreading Score:

Owenweeksia, SHD-71, and Pseudoxanthomonas: Their analogous scores emphasize their roles as primary disseminators of information or influence within the microbial network.

Proteiniclasticum: Its distinct score suggests that it might have a unique influence on the network, potentially acting as a regulator or modulator of certain microbial interactions.

Diving deeper into individual metrics:

Degree and Betweenness Centrality: Proteiniclasticum's elevated degree suggests it has more direct connections compared to other taxa. Its heightened betweenness centrality indicates its potential role as a bridge or gatekeeper, mediating interactions between different microbial sub-communities.

Taxonomical Implications:

Bacteroidetes, Chloroflexi, and Proteobacteria: Their pronounced presence in the network reaffirms their well-documented roles in various microbial ecosystems. Their interactions and connections within the network might be reflective of their metabolic capabilities and ecological niches.

Firmicutes - Proteiniclasticum: Its standout topological metrics suggest that it might be playing a more influential role than previously understood. Given its unique position in the network, further studies could delve into its metabolic pathways, interaction partners, and potential role in ecosystem stability or adaptability.

3.3.2 cDNA_Top

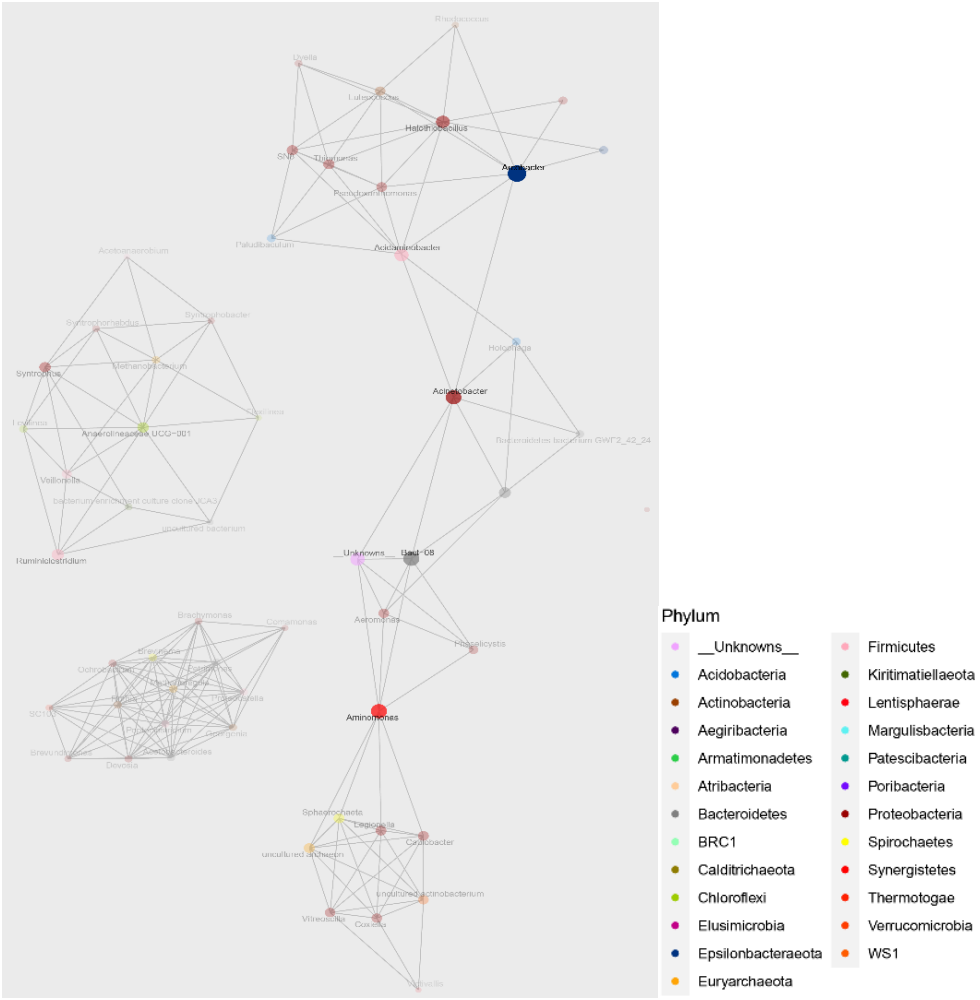

Figure 10. Results of network analysis of cDNA_Top

Within the cDNA_Top group, the microbial composition is intriguingly diverse. An uncultured archaeon stands out amidst a backdrop predominantly populated by bacteria spanning various taxonomic classes and orders.

Metric Variability: The network metrics exhibit a broad spectrum of values across different microbial entities. While certain metrics, such as the degree, maintain consistency across the initial entries, others like cluster_rank and h_index manifest pronounced variability. This suggests that while some microbial entities might share similar connectivity patterns, their roles and influence within the network can differ significantly.

Notable Outliers: Bacteriovorax emerges as a distinct outlier within the network.

With a degree of merely 1 and the lowest centrality metrics, its peripheral role in the network is evident. This might indicate its specialized or niche interactions within the microbial community.

Key Influencers: Both Peptoclostridium and Methanoregula exhibit elevated eigen_centrality, underscoring their pivotal roles within the network. Their influence is further mirrored across metrics like subgraph_centrality, suggesting their potential as central nodes orchestrating myriad interactions.

Recurring Metric Patterns: Interestingly, despite taxonomic differences, certain species such as the uncultured actinobacterium, Coxiella, and Vitreoscilla manifest identical values across most metrics. This might hint at functional redundancies or similar ecological roles despite their distinct taxonomic identities.

3.3.3 DNA_Bottom

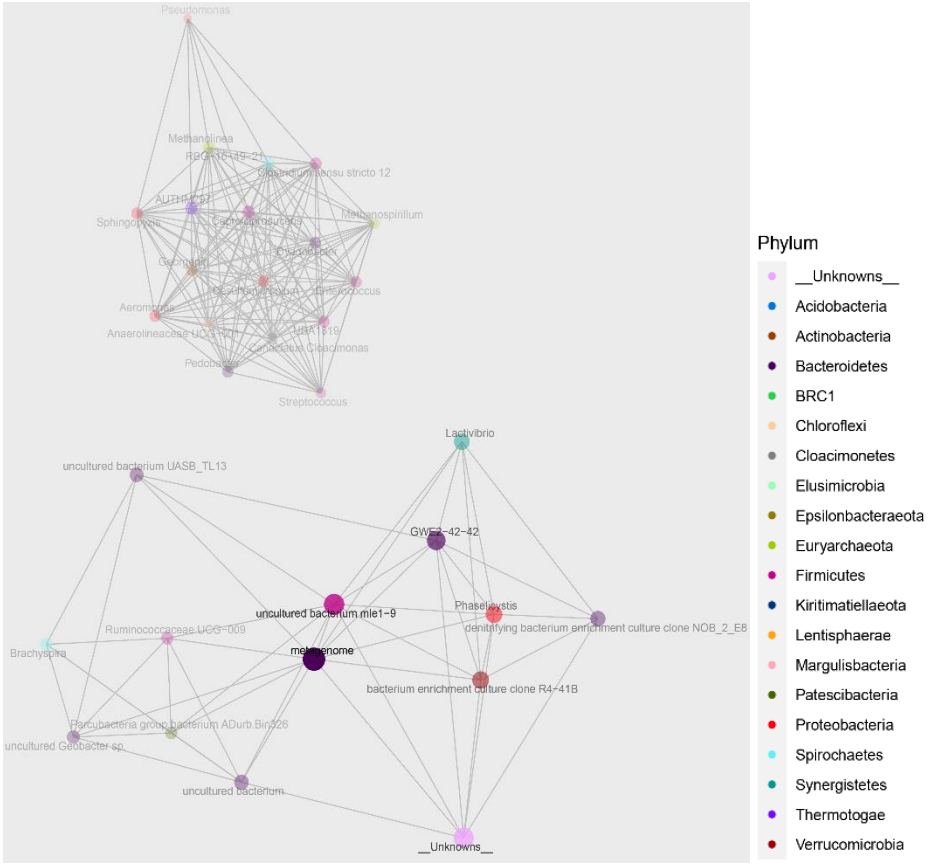

Figure 11. Results of network analysis of DNA_Bottom

The Intraset Variation Index (IVI) measures the distribution and variability of node

metrics within a specific set. The highest IVI value for "metagenome" suggests that it has a diverse range of interactions and roles within the network. This could indicate that the metagenome represents a broad spectrum of genetic material, capturing a wide array of functionalities and interactions. Its elevated IVI might imply that it's central to the network's dynamics, potentially influencing various microbial interactions and processes.

Prominent Hubness Score for AUTHM297, RBG-16-49-21, and Caproiciproducens: A high hubness score typically indicates a node's centrality and prominence within the network. The fact that AUTHM297, RBG-16-49-21, and Caproiciproducens share this elevated hubness score suggests that they are key players in the DNA_Bottom group. They likely serve as primary connectors or hubs, facilitating numerous interactions and potentially playing pivotal roles in the microbial community's stability and functionality.

A high spreading score denotes a node's potential to disseminate information or influence across the network. The fact that "Unknowns" holds the highest spreading score is particularly intriguing. It suggests that while the specific identities or functionalities of these entities might not be well-defined, they play a significant role in the network's dynamics. They might act as influencers or regulators, modulating the flow of interactions and potentially serving as bridges between known and unknown microbial entities.

In summary, these metrics highlight the intricate balance of known and unknown entities in shaping the microbial network's dynamics.

3.3.4 DNA_Top

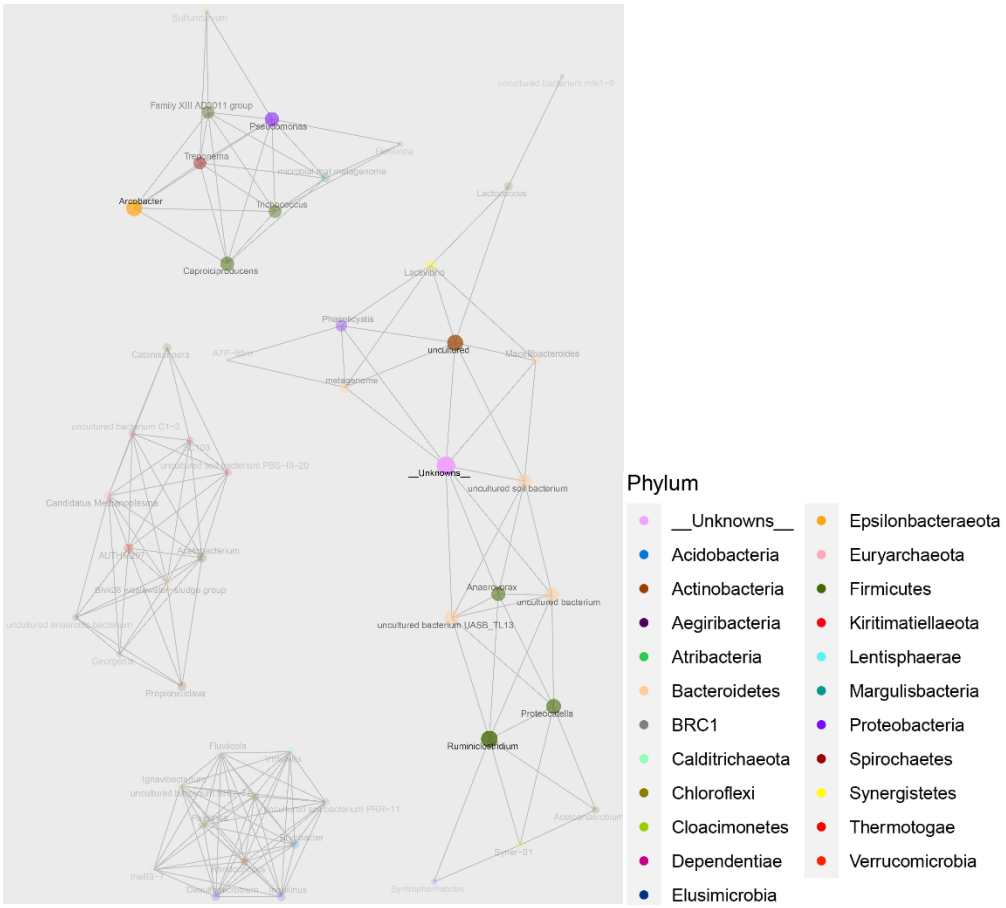

Figure 12. Results of network analysis of DNA_Top

The Intraset Variation Index (IVI) gauges the distribution and variability of node metrics within a specific set. The dominant IVI value for "Unknowns" suggests that these unidentified entities have a diverse range of interactions and roles within the network. This could imply that while their specific identities or functionalities might not be currently characterized, they are central to the network's dynamics. Their elevated IVI indicates that they might be influencing a wide array of microbial interactions and processes, potentially serving as pivotal nodes in the microbial community's structure.

A high hubness score typically signifies a node's centrality and importance within the network. The fact that Fluvicola, Inquilinus, unultured bacterium SHD-71, Ignavibacterium, Rhodococcus, Bryobacter, Desulfomicrobium, and Pelolinea all

exhibit elevated hubness scores suggests that they are integral players in the DNA_Top group. As primary connectors or hubs, they likely facilitate a multitude of interactions, playing crucial roles in maintaining the microbial community's stability and functionality. Their prominence in the network might be indicative of their metabolic capabilities, ecological niches, or adaptability within the environment.

A high spreading score denotes a node's potential to disseminate information or influence across the network. The fact that "Unknowns" holds the highest spreading score is particularly intriguing. It suggests that these unidentified entities, despite their ambiguous nature, play a significant role in the network's dynamics. They might act as influencers, regulators, or even keystone species, modulating the flow of interactions and potentially serving as bridges or connectors between known and unknown microbial entities.

3.3.5 Summaries

Taxonomy Distribution: A granular breakdown, spanning from Kingdom to Class, can elucidate the taxonomic diversity within the dataset, offering insights into the ecological breadth and potential functional capabilities of the microbial community.

Correlation Exploration: Investigating potential correlations between network metrics can unveil underlying patterns. For instance, discerning if nodes with augmented degrees consistently manifest heightened betweenness or eigen centrality can provide insights into network dynamics.

Outlier Profiling: Singular entries, such as Bacteriovorax, warrant deeper exploration. Understanding the reasons behind their metric deviations can shed light on unique ecological roles or interactions.

Hierarchical Clustering: Leveraging taxonomic data, hierarchical clustering can offer a visual representation of species relationships, juxtaposed against their network metrics. This can illuminate clusters of functionally or ecologically similar species within the network.

4. Discussion

4.1 Factors Impacting Sludge Flotation in Sewage Treatment

Particle flotation serves as a primary focus of this paper. When an overabundance of surfactants and lipid compounds infiltrate standard sewage during the aeration tank process, specific surfactants trigger a partial degradation of these materials, leading to rapid foam formation. Typically, these foams exhibit a white, lightweight appearance, dissipating once the activated sludge attains maturity. A surfeit of surfactants in the sewage can compromise the cytoplasmic membrane's stability and permeability, leading to the expulsion of crucial cellular elements and culminating in the halt of microbial growth or their demise. In the aeration stage, myriad bubbles form, readily latching onto micelles, causing the activated sludge's specific gravity to diminish, resulting in flotation. Moreover, when water contains excessive oil content, post aeration and mixing, the oil tends to bind to micelle surfaces. This binding leads to a low oxygen environment around the micelle, thereby decreasing its specific gravity and causing it to float.

Significant deviations in pH, either too high or too low, influence the catalytic functionalities of extracellular enzymes as well as those housed within the cytoplasm and cell walls of activated sludge microorganisms. Such deviations also impact nutrient absorption by these microorganisms. For instance, when the pH in a continuous flow aeration reaction tank drops below 4.0 or rises above 11.0, the microorganisms' activity within the activated sludge is typically suppressed, becomes dormant, or the organisms may perish, leading to sludge flotation ("Factors Affecting Performance of Sewage Treatment Plant - UTPedia", 2017).

Water temperature and salinity play a pivotal role in the health and function of the microorganisms found in activated sludge. Generally, these microorganisms thrive within a temperature range of 15-35°C. Surpassing 45°C, however, induces mortality in most of these microorganisms, causing them to float, save for a few that

are either long-term acclimated or uniquely resilient. While adjusting influent pH values can mitigate alkalinity's impact on activated sludge, it simultaneously results in salt production. Salt concentrations dictate osmotic pressure, a crucial determinant of microbial viability. Variations in this osmotic balance can be fatal to the microorganisms.

Toxic agents present a significant threat to aerobic activated sludge microorganisms. Primary culprits include elevated COD levels, compounds such as phenols, alcohols, aldehydes, certain acids, sulfides, heavy metals, and halides. High substrate concentrations can interact with cellular enzyme centers, impeding their access to the matrix, blocking degradation, or even causing toxic cell death. Once ingested by cells, heavy metals typically bind to -SH groups on proteins or enzymes, rendering them inactive or denatured. Trace heavy metal accumulations within cells over time can also exert a toxic influence on microorganisms. Among halides, iodine and chlorine stand out. Iodine forms a permanent bond with bacterial protein (or enzyme) tyrosine, producing diiodotyrosine and disabling the bacteria. Chlorine reacts with water to form secondary acid, which breaks down to release potent oxidants. Furthermore, material mutations in wastewater can reduce or eliminate microorganisms previously adapted to degrade specific toxins.

4.2 Anaerobic Digestion: Benefits and Challenges in Wastewater Treatment

Benefits

Energy efficiency is substantially improved in the process under consideration, and there's an added benefit: biogas, a byproduct, can be captured and reused, thereby bolstering its environmental credentials.

Another distinct advantage lies in the significantly minimized production of sludge. In fact, as demonstrated by Molinuevo-Salces et al. (2019), the volume of generated sludge is markedly low.

Considering the growth dynamics, anaerobic microorganisms exhibit a much more restrained proliferation rate compared to their aerobic counterparts. Specifically, acid-producing bacteria have a yield (Y) situated between 0.15 and 0.34 kgVSS/kgCOD. In contrast, methanogenic bacteria present a more conservative yield of around 0.03kgVSS/kgCOD. Aerobic microorganisms, for their part, fall within the range of 0.25 to 0.6kgVSS/kgCOD.

It's noteworthy to mention that certain organic materials, which are resilient and remain undegraded in the presence of aerobic microorganisms, can indeed be broken down or at least partially processed by anaerobic microorganisms. This underscores the versatility and unique capabilities of anaerobic microbes in treating diverse waste materials.

Delving deeper into the mechanics, anaerobic digestion stands out as a sophisticated and intricate microbial process. It's characterized by a sequential and interdependent collaboration between a spectrum of microorganisms, each endowed with distinct characteristics and roles. This continuous microbial synergy ensures the effective breakdown and transformation of organic matter, underpinning the robustness and efficacy of the anaerobic digestion process.

Challenges

Temperature, pH, and various other environmental factors play a critical role in determining the efficiency of the process under discussion. The system exhibits heightened sensitivity to these parameters, necessitating rigorous monitoring and control to ensure optimal performance.

Regarding the effluent quality post-treatment, there's room for improvement. While the treatment does mitigate certain contaminants, the resultant water quality often does not meet the requisite standards for direct discharge or reuse. Consequently, an additional aerobic treatment step is typically required to enhance the purity of the effluent and ensure it complies with environmental guidelines.

A notable challenge associated with this treatment process is the pronounced odor it emits. The distinct, often unpleasant, smell can pose issues, especially if the treatment facility is in close proximity to residential or commercial zones. Effective odor management strategies are essential to minimize the impact on surrounding areas.

Moreover, when evaluating the treatment's proficiency in removing specific contaminants, it's evident that its efficacy in reducing ammonia nitrogen concentrations is suboptimal. Ammonia nitrogen, a key parameter in wastewater quality, remains relatively high post-treatment, emphasizing the need for supplementary treatment or alternative strategies to address this shortcoming more effectively.

4.3 Microbial Dynamics in Anaerobic Sludge Digestion

Anaerobic digestion of sludge is a microbial-driven process chiefly governed by hydrolytic acidifying bacteria and methanogenic archaea. The significance of hydrolytic acidifying bacteria in this context is profound. These microorganisms have the capability to transform intricate compounds in the sludge, such as carbohydrates, proteins, and lipids, into more straightforward soluble monomers. Subsequently, acidifying bacteria metamorphose these hydrolyzed components into volatile fatty acids (VFAs), serving as an essential carbon reservoir that fuels further microbial activity. Therefore, it's evident that bacteria shoulder critical responsibilities, not only in the stages of hydrolysis and acidification of the sludge's organic matter but also in determining the overall efficacy of the anaerobic digestion process.

In the broader microbial community of anaerobic systems, research indicates that archaea constitute roughly 10% of the total population. This suggests that bacteria vastly outnumber archaea in these systems, reinforcing the fact that the bacterial community's dynamics and structure significantly influence the archaeal population and its function. Ahring's work further delineates the microbial diversity within

anaerobic systems, highlighting the presence of at least 20 distinct bacterial phyla. Some of the eminent phyla encompass Proteobacteria, Firmicutes, Chloroflexi, Spirochaetes, Bacteroidetes, and Actinobacteria.

Focusing on these phyla, it becomes evident that shifts within the Chloroflexi, Proteobacteria, Bacteroidetes, and Firmicutes communities have consequential impacts on the stability and efficiency of the anaerobic digestion system. This is primarily due to their dominant and indispensable roles throughout the anaerobic digestion process. In summary, understanding the intricate interplay between these bacterial phyla and archaea is vital in comprehending the intricacies of the anaerobic digestion of sludge.

Within the diverse realm of microbial communities responsible for the anaerobic digestion of organic matter, the phylum Bacteroidetes stands out for its unique capability to degrade intricate carbon-based compounds. Specifically, Bacteroidetes have shown a proficiency in breaking down complex polysaccharides such as cellulose and hemicellulose. Detailed investigations, as exemplified by the work of Berman (2019), reveal that Bacteroidetes efficiently transforms cellulose into simpler compounds, primarily monosaccharides like glucose, alongside various organic acids. Similarly, their action on hemicellulose results in the production of D-xylan and glucose.

Moreover, a distinguishing characteristic of Bacteroidetes is their predilection for further metabolizing glucose after its initial derivation from hemicellulose. This is attributed to their dominant role as principal bacteria specializing in the deglycation of sugars. Such metabolic activities by Bacteroidetes inevitably influence the concentration of volatile fatty acids (VFAs) in the system. It is imperative to understand that the abundance and activity of Bacteroidetes have a direct correlation with VFA concentrations. Alterations in Bacteroidetes populations during anaerobic digestion can consequently modulate VFA levels, which, in turn, can influence the pH balance within the system, potentially affecting the overall efficiency

of the digestion process.

To provide a holistic perspective on anaerobic digestion, it's noteworthy to mention that while bacteria, including Bacteroidetes, predominantly contribute to the hydrolysis and acidification phases, the methanogenesis stage, characterized by methane production, is exclusively mediated by archaeal communities. This demarcation of roles underscores the complex yet harmonized interactions between these microbial communities, ensuring the seamless progression of the anaerobic digestion process.

In the intricate world of microbial ecology, methanogens represent a specialized group of microorganisms that produce methane. These methanogens predominantly operate through three primary metabolic pathways: acetic acid, hydrogen, and methyl pathways. Notably, methane generation is largely attributed to methanogens that utilize acetic acid and hydrogen pathways.

Prominent research has spotlighted certain key methanogenic microorganisms, including Methanobacterium, Methanosarcina, Methanobrevibacter, Methanosaeta, and Methanomicrobium. Extending beyond these, the families Methanobacteriaceae, Methanospirillaceae, and Methanomicrobiaceae, along with the genera Methanothermobacter, Methanospirillum, Methanoculleus, and Methanomassiliicoccus, have been recognized for their critical roles in methane production.

Diving deeper into their metabolic capabilities, Methanosarcina and Methanosaeta are classified as acetic acid methanogens. Among them, Methanosarcina is exceptionally versatile. It not only functions as an acetic acid methanogen but also exhibits the characteristics of both hydrogen and methyl methanogens. As evidenced by De Vrieze et al. (2012), Methanosarcina can metabolize diverse substrates like acetic acid, methanol, methylamine, dimethylamine, and H_2/CO_2 for methane production. Remarkably, Methanosarcina is the sole strain with the prowess to produce methane through all three aforementioned pathways.

Furthermore, its resilience, marked by a high tolerance to volatile fatty acids (VFAs) and organic loading rates (OLR), underscores its significance in anaerobic environments.

Intriguingly, the population dynamics of Methanosarcina and Methanosaeta within anaerobic environments are sensitive to acetic acid concentrations. Prior research indicates that Methanosarcina thrives at elevated acetic acid concentrations, while Methanosaeta prefers environments with lower acetic acid levels. More recent findings have pinpointed Methanosarcina as the predominant strain in environments with acetic acid concentrations ranging from 250 to 500 mg COD/L.

Steering the focus away from acetic acid methanogens, other detected species like Methanosarcina, Methanothermobacter, Methanoculleus, Methanospirillum, Methanobacterium, Methanobacteriaceae, Methanospirillaceae, Methanomicrobiaceae, and Methanobrevibacter predominantly function as hydrogen-type methanogens. As elucidated by Liu et al. (2011), these organisms primarily exploit H2/CO2 or formic acid to yield CH4 gas. Additionally, Methanomassiliicoccus, identified as a methyl methanogen, employs methanol or methylamine, in conjunction with hydrogen, as electron donors to produce methane.

In summary, while the diverse archaeal population in anaerobic digestion reactors is dominated by acetic acid and hydrogen methanogens, a unifying trait among them is their adeptness at using bacterial decomposition byproducts, such as acetic acid and hydrogen, to synthesize methane. This cooperative microbial interaction underscores the intricate balance and efficiency of anaerobic digestion processes.

4.4 Key Parameters in Wastewater Treatment Optimization

Dissolved Oxygen (DO) and its Importance in Wastewater Treatment

Dissolved oxygen (DO) levels in wastewater treatment processes are typically maintained around 1-2 mg/L. Adequate DO levels are paramount for ensuring the effective functioning of aerobic microorganisms that play a pivotal role in waste

degradation.

Influence of Water Temperature on Biochemical Reactions

Water temperature significantly influences the efficiency of wastewater treatment. Within an optimal range, as the temperature increases, both the biochemical reaction rate and the proliferation of microorganisms are accelerated. This can expedite the breakdown of organic pollutants. However, cellular components, notably proteins and nucleic acids, exhibit temperature sensitivity. A sudden deviation in temperature, especially beyond the threshold of 40°C, can cause irreversible damage to these biomolecules, compromising the efficiency of the treatment process (as mentioned in "Factors Affecting Anaerobic Digestion", n.d.).

Nutrient Composition: A Balancing Act

Nutrients, primarily carbon (C), hydrogen (H), oxygen (O), and nitrogen (N), constitute around 90-97% of microbial cellular composition. The remaining 3-10% is composed of inorganic elements, phosphorus (P) being predominant. For domestic sewage treatment, there is generally no requirement to supplement with additional nutrients. However, specific industrial wastewaters may lack essential nutrients. In such cases, especially for aerobic biological treatments, nutrients are added in a recommended ratio of BOD:N:P = 100:5:1 to achieve the desired treatment efficiency. Additionally, certain inorganic nutrients like potassium (K), magnesium (Mg), calcium (Ca), sulfur (S), and sodium (Na), and trace elements including iron (Fe), copper (Cu), manganese (Mn), molybdenum (Mo), silicon (Si), and boron play essential roles in microbial metabolism.

pH: The Acidic-Alkaline Balance

The pH level of the wastewater is a crucial determinant of microbial activity. Most aerobic microorganisms thrive in a pH range of 6.5 to 8.5. Maintaining pH within this optimal range is essential to facilitate efficient organic matter degradation and ensure the overall stability of the treatment process.

Toxic Substances: Challenges in Wastewater Treatment

Certain compounds present in wastewater can act as suppressors, hindering the activity of beneficial microorganisms. These toxicants include heavy metals, cyanide, hydrogen sulfide (H2S), halogenated compounds, phenols, alcohols, and aldehydes. Monitoring and managing the concentration of these toxic substances are essential to maintain the efficiency of the treatment process.

Organic Load Rate: Striking the Right Balance

While organic matter present in sewage acts as a primary nutrient source for microorganisms, an excessive organic load can overwhelm the microbial ecosystem. An overabundance can lead to imbalances and reduced efficiency in waste degradation, emphasizing the importance of maintaining an optimal organic load rate in wastewater treatment systems ("What are the influencing factors of anaerobic biological treatment?", n.d.).

In conclusion, ensuring optimal conditions, such as appropriate DO levels, nutrient composition, pH, and organic load rate, while simultaneously managing potential challenges posed by temperature fluctuations and toxic substances, is fundamental for the effective and efficient functioning of wastewater treatment processes.

5. Conclusions

The primary focus of this research paper delves into the assembly dynamics of microbial communities within sewage treatment processes. A well-functioning model predicts that granular sludge, resulting from the aggregation of these microbial communities, should not float, thereby eliminating the production of floating particles. Yet, both experimental data and real-world observations indicate that particle floating is a commonplace occurrence.

The analytical approach adopted in this research is segmented into four distinct sections:

Microbial Community Assembly Analysis: This section primarily investigates the assembly of the microbial community, emphasizing its diversity and richness. The assembly process is governed by a combination of deterministic and stochastic factors. By understanding the interplay between these deterministic (certainty) and stochastic (randomness) elements, one can infer the diversity and richness of microbial communities in different groupings, namely DNA and cDNA. These groupings symbolize intact and active microbial communities, respectively. The insights derived from this section will elucidate whether the microbial community responsible for anaerobic digestion significantly influences the overall microbial community assembly.

Zeta Diversity Analysis: This segment delves deeper into the zeta diversity of the microbial community. The primary objective is to discern the role of determinism and randomness in shaping the microbial community. By doing so, the study aims to deduce the diversity and richness of microbial communities within DNA and cDNA groupings, as well as floating and settling groupings.

Microbial Community Differential Analysis: This section explores the variances within microbial communities and the manifestation of specific differences. Preliminary findings suggest that the microbiome could be an intrinsic factor

contributing to particle floating, which in turn may limit the diversity and richness of microbial communities.

Core Microbial Community Network Analysis: The final section zeroes in on the core microbial community, identifying specific bacterial or archaeal types that might be influential. This analysis aims to pinpoint which microbial species predominantly affect the diversity and richness of the overall community. Additionally, this section also examines the roles of determinacy and randomness in microbial community development. One of the overarching goals of this research is to ascertain which of these two elements—determinacy or randomness—exerts a more profound influence on microbial community assembly. Current evidence leans towards deterministic factors playing a more dominant role in shaping microbial community assembly.

In summary, this research offers a comprehensive examination of microbial community assembly in sewage treatment, shedding light on the factors that influence its diversity, richness, and functionality.

5.1 Microbial Dynamics: Key Intrinsic Factors in Particle Floating

In the broader discourse of our research, while the discussion section extensively covers the extrinsic factors influencing particle floating, the conclusion seeks to emphasize the intrinsic determinants that have emerged as significant contributors. These internal factors, often overlooked in conventional studies, offer a nuanced understanding of the complexities inherent in sewage treatment processes.

One of the primary internal contributors identified is the role of Archaea. Drawing from our accumulated data, there's a compelling indication that Archaea exerts a notable influence on particle floating. This discovery underscores the necessity to delve deeper into specific microbial entities, understanding their behavior and interactions within the sewage treatment environment, and how they might be influencing outcomes in ways previously unanticipated.

Furthermore, our exploration into microbial communities has revealed that cDNAs, which represent the active microbiome, appear to have a more pronounced impact on particle floating than initially presumed. This insight not only emphasizes the importance of the active microbial community but also suggests that the dynamism and activity levels of these communities might be more critical to particle floating than their mere presence.

Lastly, aligning with the findings of Evans et al. (2016), our research reaffirms the significant role of the diffusion process in the emergence of floating particles. This reiteration accentuates the need for a comprehensive understanding of both biological and physical interactions at play, and how these processes intertwine to influence the overall efficacy of sewage treatment systems.

In wrapping up, it's evident that while external factors provide a part of the puzzle, a holistic understanding of particle floating necessitates a deeper exploration of these intrinsic elements. Recognizing and addressing these internal determinants will be instrumental in shaping more effective and efficient sewage treatment strategies in the future.

5.2 On exploring microbial interactions and diversities in granule flotation

The intricate assembly of microbial communities is dissected in this research through two primary lenses, as previously alluded to.

Diversity and Richness: An in-depth analysis of the findings from the preceding four sections reveals intriguing patterns in microbial community structures. The total DNA community, representing the entire spectrum of microorganisms, exhibits greater diversity compared to the cDNA community, which signifies the active microbial entities. However, when it comes to richness, the cDNA community surpasses its total DNA counterpart. Delving deeper into specific microbial taxa, Bacteroidetes emerges as a positive influencer, enhancing the richness of the microbial community.

In contrast, Actinobacteria appears to exert a suppressive effect, diminishing the community's richness.

Certainty and Randomness: The dynamics of microbial community assembly are also governed by deterministic (certainty) and stochastic (randomness) forces. Our findings indicate a clear demarcation between the two community groupings in terms of these forces. The DNA community, representing the total microbial population, is predominantly shaped by stochastic processes. On the other hand, the cDNA community, symbolizing the active microbial fraction, is more influenced by deterministic factors. Zooming into specific microorganisms, it's noteworthy that the assembly of Archaea is significantly swayed by the stochastic elements of microbial community formation.

In summary, the assembly of microbial communities is a complex interplay of various factors, ranging from the inherent diversity and richness of the communities to the deterministic and stochastic forces that shape them. Recognizing these nuances is pivotal for a comprehensive understanding of microbial dynamics in any given environment.

6. Future Considerations

- Expanding the Sample Base: While the current study offers valuable insights into the assembly of anaerobic microbiomes, future research could benefit from a broader and more diverse sample base. This would allow for a more comprehensive understanding of the microbial communities across different wastewater environments.
- Technological Advancements: As sequencing and analytical technologies continue to evolve, it would be beneficial to revisit the study using more advanced tools. This could provide deeper insights into the microbial community structures and their interactions.
- Longitudinal Studies: A longitudinal study observing the changes in microbial communities over extended periods could provide insights into the stability, resilience, and adaptability of these communities in wastewater environments.
- Functional Analysis: Beyond community assembly, a deeper dive into the functional roles of these microbial communities in wastewater treatment processes would be invaluable. This could lead to optimized wastewater treatment strategies.
- Interactions with Other Microbial Communities: Exploring how anaerobic microbiomes interact with other microbial communities in wastewater could shed light on synergistic or antagonistic relationships that impact the treatment process.
- Environmental Impact: Future studies could delve deeper into the environmental implications of particle flotation and the release of specific microbial communities into natural water bodies.
- Economic Considerations: As the world moves towards more sustainable wastewater treatment solutions, understanding the economic implications of optimizing anaerobic microbial communities could be beneficial for policymakers

and industry stakeholders.

- Real-world Applications: Translating the findings of this study into real-world applications, such as designing more efficient wastewater treatment plants or developing microbial supplements for enhanced treatment, would be a logical next step.
- Climate Change Impact: With changing global climate patterns, understanding how these shifts might impact anaerobic microbial communities in wastewater environments could be crucial for future wastewater management strategies.

Reference

[1] Ali, M., Wang, Z., Salam, K., Hari, A. R., Pronk, M., Van Loosdrecht, M. C. M., et al. (2019). Importance of species sorting and immigration on the bacterial assembly of different-sized aggregates in a full-scale aerobic granular sludge plant. *Environ. Sci. Technol.* 53, 8291–8301. doi: 10.1021/acs.est.8b07303

[2] Winter, F., Agarwal, R., Hrdlicka, J., & Varjani, S. CO2 Separation, purification and conversion to chemicals and fuels.

[3] Alphenaar, P. A. (1994). *Anaerobic granular sludge: characterization, and factors affecting its functioning.* Wageningen: Landbouwuniversiteit te Wageningen.

[4] Zhang, B., Ning, D., Van Nostrand, J., Sun, C., Yang, Y., Zhou, J., & Wen, X. (2020). Biogeography and Assembly of Microbial Communities in Wastewater Treatment Plants in China. Environmental Science &Amp; Technology, 54(9), 5884-5892. https://doi.org/10.1021/acs.est.9b07950

[5] Campanaro, S., Treu, L., Rodriguez-R, L. M., Kovalovszki, A., Ziels, R. M., Maus, I., et al. (2019). The anaerobic digestion microbiome: a collection of 1600 metagenome-assembled genomes shows high species diversity related to methane production. *bioRxiv* 2019:680553. doi: 10.1101/680553

[6] Wang, L., Han, M., Li, X., Yu, B., Wang, H., & Ginawi, A. et al. (2021). Mechanisms of niche-neutrality balancing can drive the assembling of microbial community. Molecular Ecology, 30(6), 1492-1504.https://doi.org/10.1111/mec.15825

[7] Presley, S. J., Higgins, C. L., and Willig, M. R. (2010). A comprehensive framework for the evaluation of metacommunity structure. *Oikos* 119, 908–917. doi: 10.1111/j.1600-0706.2010.18544.x

[8] Molinuevo-Salces, B., Riaño, B., Hernández, D., & Cruz García- González, M. (2019). Microalgae and Wastewater Treatment: Advantages and Disadvantages. Microalgae Biotechnology For Development Of Biofuel And Wastewater Treatment, 505-533. https://doi.org/10.1007/978-981-13- 2264-8_20

[9] Singh, S., Rinta-Kanto, J. M., Kettunen, R., Tolvanen, H., Lens, P., Collins, G., et

al. (2019). Anaerobic treatment of LCFA-containing synthetic dairy wastewater at 20°C: Process performance and microbial community dynamics. *Sci. Total Environ.* 691, 960–968. doi: 10.1016/j.scitotenv.2019.07.136

[10] Lettinga, G., Pol, L., Van Lier, J., & Zeeman, G. (2008). Possibilities and Potential of Anaerobic Wastewater Treatment Using Anaerobic Sludge Bed (ASB) Reactors. Biotechnology, 517-526.https://doi.org/10.1002/9783527620944.ch25

[11] Trego, A. C., Mills, S., and Collins, G. (2020). Granular biofilms: function, application, and new trends as model microbial communities. *Crit. Rev. Environ. Sci. Technol.* 2020, 1–24. doi: 10.1080/10643389.2020.1769433

[12] Berman, J. (2019). Bacteria. Taxonomic Guide To Infectious Diseases, 39-119. https://doi.org/10.1016/b978-0-12-817576-7.00003-1

[13] McAteer, P. G., Trego, A. C., Thorn, C., Mahony, T., Abram, F., and O'Flaherty, V. (2020). Reactor configuration influences microbial community structure during high-rate, low-temperature anaerobic treatment of dairy wastewater. *Bioresour. Technol.* 2020:123221. doi: 10.1016/j.biortech.2020.123221

[14] Aguirre de Cárcer, D. (2019). A conceptual framework for the phylogenetically constrained assembly of microbial communities. Microbiome, 7(1). https://doi.org/10.1186/s40168-019-0754-y

[15] De Vrieze, J., Hennebel, T., Boon, N., & Verstraete, W. (2012). Methanosarcina: The rediscovered methanogen for heavy duty biomethanation. Bioresource Technology, 112, 1-9. https://doi.org/10.1016/j.biortech.2012.02.079

[16] MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS. (2022), 75(1), 1-13. https://doi.org/10.1128/mmbr.00038-10

[17] Reeburgh, W. (1980). Anaerobic methane oxidation: Rate depth distributions in Skan Bay sediments. Earth And Planetary Science Letters, 47(3), 345-352. https://doi.org/10.1016/0012-821x(80)90021-7

[18] Stegen, J., Lin, X., Konopka, A., & Fredrickson, J. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. The ISME

Journal, 6(9), 1653-1664.https://doi.org/10.1038/ismej.2012.22

[19] Factors Affecting Performance of Sewage Treatment Plant - UTPedia. Utpedia.utp.edu.my. (2017). Retrieved 18 August 2022, from http://utpedia.utp.edu.my/7250/#:~:text=Although%20there%20are%20many%20other%20factors%20that%20can,national%20company%20is%20identified%20as%20my%20case%20study.

[20] Panosyan, H., Bhattacharjee, K., Manchanda, G., & Pratap Singh, R. (2021). SPRINGER-NATURE (pp. 3-6). Springer.

[21] Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857.

[22] Bottos, E. M., Kennedy, D. W., Romero, E. B., Fansler, S. J., Brown, J. M., Bramer, L. M., et al. (2018). Dispersal limitation and thermodynamic constraints govern spatial structure of permafrost microbial communities. *FEMS Microbiol. Ecol.* 94:fiy110. doi: 10.1093/femsec/fiy110

[23] Campos, J. L., del Río, A. V., Pedrouso, A., Raux, P., Giustinianovichc, E. A., and Mosquera-Corral, A. (2017). Granular biomass floatation: a simple kinetic/stoichiometric explanation. *Chem. Eng. J.* 311, 63–71.

[24] Connelly, S., Shin, S. G., Dillon, R. J., Ijaz, U. Z., Quince, C., Sloan, W. T., et al. (2017). Bioreactor scalability: laboratory-scale bioreactor design influences performance, ecology, and community physiology in expanded granular sludge bed bioreactors. *Front. Microbiol.* 8:664. doi: 10.3389/fmicb.2017.00664

[25] Douglas, G. M., Maffei, V. J., Zaneveld, J., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2019). PICRUSt2: an improved and extensible approach for metagenome inference. *BioRxiv* 2019:672295.

[26] Eftaxias, A., Diamantis, V., Michailidis, C., Stamatelatou, K., and Aivasidis, A. (2020). Comparison of anaerobic digesters performance treating palmitic, stearic and oleic acid: determination of the LCFA kinetic constants using ADM1. *Bioprocess*

Biosyst. Eng. 43, 1329–1338.

[27] Kazieva, E., Yamamoto, Y., Tajima, Y., Yokoyama, K., Katashkina, J., and Nishio, Y. (2017). Characterization of feedback-resistant mevalonate kinases from the methanogenic archaeons *Methanosaeta concilii* and *Methanocella paludicola*. *Microbiology* 163, 1283–1291. doi: 10.1099/mic.0.000510

[28] Keating, C., Hughes, D., Mahony, T., Cysneiros, D., Ijaz, U. Z., Smith, C. J., et al. (2018). Cold adaptation and replicable microbial community development during long-term low-temperature anaerobic digestion treatment of synthetic sewage. *FEMS Microbiol. Ecol.* 94:fiy095. doi: 10.1093/femsec/fiy095

[29] Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

[30] Li, W., Zheng, P., Ji, J., Zhang, M., Guo, J., Zhang, J., et al. (2014). Floatation of granular sludge and its mechanism: a key approach for high-rate denitrifying reactor. *Bioresour. Technol.* 152, 414–419. doi: 10.1016/j.biortech.2013.11.056

[31] Locey, K. J., and Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci.* 113, 5970–5975. doi: 10.1073/pnas.1521291113

[32] Lu, X., Zhen, G., Estrada, A. L., Chen, M., Ni, J., Hojo, T., et al. (2015). Operation performance and granule characterization of upflow anaerobic sludge blanket (UASB) reactor treating wastewater with starch as the sole carbon source. *Bioresour. Technol.* 180, 264–273. doi: 10.1016/j.biortech.2015.01.010

[33] Zhu, X., Campanaro, S., Treu, L., Seshadri, R., Ivanova, N., Kougias, P. G., et al. (2020). Metabolic dependencies govern microbial syntrophies during methanogenesis in an anaerobic digestion ecosystem. *Microbiome* 8:22. doi: 10.1186/s40168-019-0780-9

[34] Zhang, X., Johnston, E. R., Liu, W., Li, L., and Han, X. (2016). Environmental changes affect the assembly of soil bacterial community primarily by mediating stochastic processes. *Glob. Chang. Biol.* 22, 198–207. Doi: 10.1111/gcb.13080

Appendix

```
library(phyloseq)
library(zetadiv)
library(ggplot2)
physeq<-import_biom("../../Data/feature_w_tax.biom")
meta_table<-read.csv("../../Data/Paul_metadata2.csv",header=T,row.names=1)
abund_table<-otu_table(physeq)
abund_table<-t(abund_table)
#Uncomment if you'd like to get rid of samples below a certain library size
abund_table<-abund_table[rowSums(abund_table)>=5000,]
OTU_taxonomy<-data.frame(as(tax_table(physeq),"matrix"))
colnames(OTU_taxonomy)<-
c("Kingdom","Phylum","Class","Order","Family","Genus","Otus")
#Ensure that all columns of OTU_taxonomy are character and not factors
OTU_taxonomy[] <- lapply(OTU_taxonomy, function(x) as.character(x))
OTU_taxonomy[is.na(OTU_taxonomy)]<-""
OTU_taxonomy$Otus<-gsub("D_6__|s__","",OTU_taxonomy$Otus)
OTU_taxonomy$Genus<-gsub("D_5__|g__","",OTU_taxonomy$Genus)
OTU_taxonomy$Family<-gsub("D_4__|f__","",OTU_taxonomy$Family)
OTU_taxonomy$Order<-gsub("D_3__|o__","",OTU_taxonomy$Order)
OTU_taxonomy$Class<-gsub("D_2__|c__","",OTU_taxonomy$Class)
OTU_taxonomy$Phylum<-gsub("D_1__|p__","",OTU_taxonomy$Phylum)
OTU_taxonomy$Kingdom<-gsub("D_0__|d__","",OTU_taxonomy$Kingdom)
#Remove singletons and adjust OTU_taxonomy
abund_table<-abund_table[,colSums(abund_table)>1]
OTU_taxonomy<-OTU_taxonomy[colnames(abund_table),]
```

```
#get rid of contaminants with "Unassigned", "Chloroplast" and "Mitochondria"
assignment", and "non classified" at Phylum level
abund_table<-abund_table[,!(OTU_taxonomy$Kingdom %in% c("Unassigned") |
OTU_taxonomy$Phylum=="" | OTU_taxonomy$Order %in% c("Chloroplast") |
OTU_taxonomy$Family %in% c("Mitochondria"))]
#extract subset of abund_table for which samples also exists in meta_table
abund_table<-abund_table[rownames(abund_table) %in%
rownames(meta_table),]
#when reducing the abund_table, there is a high likelihood that an OTU was only
present in a sample that is removed, so we shrink
#the abund_table to get rid of empty columns
abund_table<-abund_table[,colSums(abund_table)>0]
#make your meta_table smaller by only considering samples that appear in
abund_table
meta_table<-meta_table[rownames(abund_table),]
#make OTU_taxonomy smaller by only considering OTUs that appear in
abund_table
OTU_taxonomy<-OTU_taxonomy[colnames(abund_table),]
#At this point we have abund_table, meta_table, and OTU_taxonomy are ready
and their dimensions should match
#PARAMETERS CHANGE THE GROUPING COLUMN AS YOU
DESIRE##############################
#In the hypothesis space, all you need is to select the rows in meta_table you are
interested in
#and then allocate a column to meta_table$Groups that you want to use.
# label="Hypothesis1_Healthy_Province_BIRTH_Male"
# meta_table<-meta_table[meta_table$Gender %in% c("Male") &
meta_table$Condition %in%
```

```
c("Healthy_Baseline_T1","Ramadan_Baseline_T1","Intermittent_Baseline_T1"),]
# meta_table$Groups<-factor(as.character(meta_table$Province_birth),levels=c(
#     "ICT",
#     "Balochistan",
#     "KPK",
#     "Punjab",
#     "Sindh",
#     "AJK"))
# colours <- c(
#     "#F0A3FF",
#     "#0075DC",
#     "#993F00",
#     "#4C005C",
#     "#2BCE48",
#     "#FFCC99",
#     #Next colors are for lines mainly used in the PCoA script
#
"#000080","#4876FF","#CAE1FF","#9FB6CD","#1E90FF","#00F5FF","#00C957",g
rey.colors(1000));
# #Provide xy.coord, either it is not available then xy.coord=NULL otherwise make
a dataframe of the coordinates
# xy.coord<-NULL
# #xy.coord<-
data.frame(row.names=rownames(meta_table),x=0,y=meta_table$Day)
# Zeta_method<-"ALL" #ALL DNN FPO
# label="Hypothesis1_Healthy_Province_BIRTH_Female"
# meta_table<-meta_table[meta_table$Gender %in% c("Female") &
meta_table$Condition %in%
```

```
c("Healthy_Baseline_T1","Ramadan_Baseline_T1","Intermittent_Baseline_T1"),]
# meta_table$Groups<-factor(as.character(meta_table$Province_birth),levels=c(
#     "ICT",
#     "Balochistan",
#     "KPK",
#     "Punjab",
#     "Sindh",
#     "AJK"))
# colours <- c(
#     "#F0A3FF",
#     "#0075DC",
#     "#993F00",
#     "#4C005C",
#     "#2BCE48",
#     "#FFCC99",
#     #Next colors are for lines mainly used in the PCoA script
#
"#000080","#4876FF","#CAE1FF","#9FB6CD","#1E90FF","#00F5FF","#00C957",g
rey.colors(1000));
# #Provide xy.coord, either it is not available then xy.coord=NULL otherwise make
a dataframe of the coordinates
# xy.coord<-NULL
# #xy.coord<-
data.frame(row.names=rownames(meta_table),x=0,y=meta_table$Day)
# Zeta_method<-"ALL" #ALL DNN FPO
# label="Hypothesis1_Healthy_Province_RESIDENCE_Male"
# meta_table<-meta_table[meta_table$Gender %in% c("Male") &
meta_table$Condition %in%
```

```
c("Healthy_Baseline_T1","Ramadan_Baseline_T1","Intermittent_Baseline_T1"),]
# #exclude groups where you don't have atleast 3 samples
# meta_table<-meta_table[!meta_table$Province_residence %in%
c("KPK","Balochistan"),,drop=FALSE]
# meta_table$Groups<-
factor(as.character(meta_table$Province_residence),levels=c(
#     "ICT",
#     "Punjab",
#     "Sindh"))
# colours <- c(
#     "#F0A3FF",
#     "#4C005C",
#     "#2BCE48",
#     #Next colors are for lines mainly used in the PCoA script
#
"#000080","#4876FF","#CAE1FF","#9FB6CD","#1E90FF","#00F5FF","#00C957",g
rey.colors(1000));
# #Provide xy.coord, either it is not available then xy.coord=NULL otherwise make
a dataframe of the coordinates
# xy.coord<-NULL
# #xy.coord<-
data.frame(row.names=rownames(meta_table),x=0,y=meta_table$Day)
# Zeta_method<-"ALL" #ALL DNN FPO
label="Hypothesis1"
meta_table<-meta_table[meta_table$Position_Type %in%
c("cDNA_Top","DNA_Bottom","cDNA_Bottom","DNA_Top" ),]
#First provide grouping column
meta_table$Groups<-as.character(meta_table$Position_Type)
```

```
#The colours in the the next instruction match the factors for meta_table$Groups
meta_table$Groups<-factor(meta_table$Groups,c(
   "DNA_Top",
   "DNA_Bottom",
   "cDNA_Top",
   "cDNA_Bottom"
))
colours <- c(
   "#F0A3FF",
   "#0075DC",
   "#993F00",
   "#4C005C",
   #Next colors are for lines mainly used in the PCoA script

"#000080","#4876FF","#CAE1FF","#9FB6CD","#1E90FF","#00F5FF","#00C957",g
rey.colors(1000));
#Provide xy.coord, either it is not available then xy.coord=NULL otherwise make a
dataframe of the coordinates
xy.coord<-NULL
#xy.coord<-
data.frame(row.names=rownames(meta_table),x=0,y=meta_table$Day)
Zeta_method<-"ALL" #ALL DNN FPO
#PARAMETERS CHANGE THE GROUPING COLUMN AS YOU
DESIRE##############################
#Adjust abund_table to contain only those rows that got selected in the Hypothesis
space
abund_table<-abund_table[rownames(meta_table),]
#After adjustment, get rid of OTUs that are all empty
```

```
abund_table<-abund_table[,colSums(abund_table)>0]
#Adjust OTU taxonomy
OTU_taxonomy<-OTU_taxonomy[colnames(abund_table),]
#Calculate zeta diversity
#Very crude way of converting abund_table to incidence_table
abund_table[abund_table>1]=1
#Convert this to data.frame from otu_table
abund_table<-data.frame(as(abund_table,"matrix"))
collated_zeta.val<-NULL
collated_zeta.ratio<-NULL
collated_AIC<-NULL
#Now iterate through all the groups
for (i in levels(meta_table$Groups)){
   #Extract at/mt for those groups
   at<-abund_table[meta_table$Groups==i,,drop=FALSE]
   #readjust at with only those features (OTUs/ASVs) that exist in atleast one
sample
   at<-at[,colSums(at)>0,drop=FALSE]
   mt<-meta_table[rownames(at),,drop=FALSE]
   #Calculate expectation of zeta diversity decline
   #As per authors, the function below computes the expectation of zeta diversity,
   #the number of species shared by multiple assemblages for a range of orders
   #(number of assemblages or sites), using a formula based on the occupancy of
   #the species, and fits the decline to an exponential and a power law
relationship.
   res<-NULL
   if(is.null(xy.coord)){
      if(Zeta_method=="ALL"){
```

```
pdf(paste("Zeta.decline.ex_",Zeta_method,"_",label,"_",i,".pdf",sep=""),height=5,wi
dth=10)
        res<-Zeta.decline.ex(at,orders=1:nrow(at),plot=TRUE)
        dev.off()
      }
   } else{
      xy=xy.coord[rownames(mt),]
      if(Zeta_method=="DNN"){

pdf(paste("Zeta.decline.mc_",Zeta_method,"_",label,"_",i,".pdf",sep=""),height=5,wi
dth=10)
        res<-
Zeta.decline.mc(at,orders=1:nrow(at),xy=xy.coord[rownames(mt),],plot=TRUE,FP
O=c(min(xy[1]),min(xy[2])), DIR = TRUE)
        dev.off()
      } else if(Zeta_method=="FPO") {

pdf(paste("Zeta.decline.mc_",Zeta_method,"_",label,"_",i,".pdf",sep=""),height=5,wi
dth=10)
        res<-
Zeta.decline.mc(at,orders=1:nrow(at),xy=xy.coord[rownames(mt),],plot=TRUE,FP
O=c(min(xy[1]),min(xy[2])))
        dev.off()
      }

   }
```

```
   #Extract the values and collate them together
   tmp<-
data.frame(zeta.order=res$zeta.order,zeta.val=res$zeta.val,Groups=rep(i,length(r
es$zeta.val)))
   if (is.null(collated_zeta.val)){collated_zeta.val<-tmp} else {collated_zeta.val<-
rbind(collated_zeta.val,tmp)}
   tmp<-data.frame(zeta.order=res$zeta.order[-
length(res$zeta.order)],zeta.ratio=res$ratio,Groups=rep(i,length(res$ratio)))
   if (is.null(collated_zeta.ratio)){collated_zeta.ratio<-tmp} else
{collated_zeta.ratio<-rbind(collated_zeta.ratio,tmp)}
   tmp<-
data.frame(AIC.exp=res$aic["zeta$zeta.exp",2],AIC.pl=res$aic["zeta$zeta.pl",2],Gr
oups=i)
   if (is.null(collated_AIC)){collated_AIC<-tmp} else {collated_AIC<-
rbind(collated_AIC,tmp)}
}
#Plot collated_zeta.val
p<-
ggplot(aes(x=zeta.order,y=zeta.val,group=Groups,colour=Groups,shape=Groups),
data=collated_zeta.val)
p<-p+geom_point()
p<-p+geom_line()
p<-p+scale_color_manual("Groups",values=colours)
p<-p+scale_shape_manual("Groups",values=c(c(17:25),c(33:127)))
p<-p+theme_bw()
p<-p+ylab("Zeta Diversity")
p<-p+xlab("Zeta Order")
pdf(paste("COLLATED_ZETA_VAL_",Zeta_method,"_",label,".pdf",sep=""))
```

```
print(p)
dev.off()

#Plot collated_zeta.ratio
p<-
ggplot(aes(x=zeta.order,y=zeta.ratio,group=Groups,colour=Groups,shape=Groups
),data=collated_zeta.ratio)
p<-p+geom_point()
p<-p+geom_line()
p<-p+scale_color_manual("Groups",values=colours)
p<-p+scale_shape_manual("Groups",values=c(c(17:25),c(33:127)))
p<-p+theme_bw()
p<-p+ylab("Zeta Ratio")
p<-p+xlab("Zeta Order")
pdf(paste("COLLATED_ZETA_RATIO_",Zeta_method,"_",label,".pdf",sep=""))
print(p)
dev.off()

#Now save all the files generated for further usage
write.csv(collated_zeta.val,paste("COLLATED_ZETA_VAL_",Zeta_method,"_",labe
l,".csv",sep=""))
write.csv(collated_zeta.ratio,paste("COLLATED_ZETA_RATIO_",Zeta_method,"_",
label,".csv",sep=""))
write.csv(collated_AIC,paste("COLLATED_AIC_",label,".csv",sep=""))
```