Microbial Visualization Research Based on Biosensing Technology

Ramyan Radhakrishnan[*]

Islamic Azad University, Iran

[*]*corresponding author*

Abstract: Microorganisms are the most diverse organisms in the world. Microorganisms live together in a certain environment and gradually form a stable collective through interaction, which is a microbial community. So far, many researchers have devoted themselves to the study of microbial community structure to explore microbial diversity, but it is difficult to present the research results in a simple form, and the characteristics of information transmission of biosensor technology can be used to present the research results to researchers. Therefore, this paper A microbial visualization platform based on biosensors is built. After sequencing, OTU clustering and other operations, high-definition images are generated by visualization tools, and researchers can analyze the α-diversity and β-diversity of microbial communities based on the images.

1. Introduction

The development of bioinformatics has brought new ideas to the research on the diversity of microbial communities today, how to complete gene sequencing more accurately in biology, how to mine and analyze sequencing data more efficiently and scientifically in computer science, and analyze the The integration and friendly display of the result information to the researchers is the current hot spot of microbial community diversity analysis. Based on this, this paper conducts microbial diversity analysis and displays the microbial community structure in a visual way.

So far, there are many techniques for analyzing microbial community structure. For example, some scholars have clearly demonstrated the differences between the phylogenetic tree constructed based on pan-genome and the phylogenetic tree constructed based on core gene variation of the studied species, indicating that horizontal gene transfer has largely mediated the evolution of species [1]. Some researchers have used the 16S rRNA of microorganisms to predict the function of microbial communities. After comparing with whole genome sequencing, the effect is very good. Then use a brand new microbial niche module to perform data analysis on the biological community, avoiding the limitations of OTU [2]. In the study of microbial communities, some scholars designed

a method based on graph theory, embedded OTUs into a graph through the similarity between OTUs, and then used spectral clustering and other methods to cluster OTUs into multiple modules. Get a module-based expression matrix. This will be a powerful tool for researchers to further understand the ecosystem function of microbial communities [3]. Sequence-based metagenomics has captured a wealth of information in microbial community studies, such as the metagenome study of microbial populations in the Sargasso Sea yielding sequences of about one million genes and revealing entire gene classes, diversity alignments Diversification expected in research of cultured organisms [4]. Nonetheless, the techniques for microbial analysis have yet to be further refined.

This paper first introduces the types of biosensors, then proposes Alpha analysis and Beta analysis for microbial community diversity analysis, and designs a visualization platform based on biosensors to display the results of α-diversity and β-diversity analysis through visualization. Demonstrating microbial diversity is conducive to improving the efficiency of people's research on microbial structure.

2. Related Technologies

2.1. Types of Biosensors

(1) Optical biosensor

Optical biosensors are technologies that use the interaction between light and matter to analyze the presence of a measured object. Optical sensors are widely used in various fields because of their simple structure, convenient detection and good detection sensitivity [5].

(2) Electrochemical biosensors

Electrochemical sensors typically detect analytes by recording current, potential, impedance, or capacitance signals. Electrochemical sensors have the advantages of high detection sensitivity, fast signal response, good selectivity, and require a small amount of sample [6]. Reagent-free and ready-to-use testing requirements can be achieved, electroactive labeling substances are more stable, and electroactive contaminants are relatively few.

2.2. Comparative Analysis of Microbial Community Diversity

There are two main directions of diversity-based comparative analysis: Alpha analysis and Beta analysis, both of which belong to the category of ecological statistical analysis of microbial communities, but represent two different spatial dimensions [7].

(1) Alpha analysis

Alpha diversity analysis mainly evaluates the distribution of species in any community through the calculation of diversity index, such as Observed species, Simpson, Chao, Ace, etc. [8].

The observed species index represents the number of divided OTUs, that is, the number of species [9]. In the analysis of microbial community diversity, it is necessary to evaluate whether the sequencing depth is sufficient to judge the accuracy of the sequencing results. In this case, the sequence can be sampled in different numbers, and then the expected value of the number of OTUs corresponding to the number of reads in the sample is calculated, and the final statistical results are diluted. sex curve (abundance curve) to make predictions [10].

The Simpson index evaluates the probability that two individuals randomly selected from a community belong to the same species. According to this definition, the calculation of the Simpson index can be used as one of the basis for estimating the diversity of microbial communities [11]. Its

calculation method is as follows:

$$Simpson = 1 - \sum_{i=1}^{G} p_i^2 \tag{1}$$

where G is the total number of species and p_i is the proportion of species i in the community. Through formula deduction, it can be concluded that when all individuals belong to the same species, the Simpson index reaches the lowest value of 0; and when each individual corresponds to a different species, the index has the highest value of 1-1/G.

Rank-Abundance (relative abundance) is often used to simultaneously explain two important aspects of microbial community diversity in a sample: species abundance and uniformity [12]. Its calculation method is as follows:

$$Rank - Abundance = 1 - 2\frac{\sum \min(G_{A,i}, G_{B,i})}{\sum G_{A,i} + \sum G_{B,i}} \tag{2}$$

Among them, $G_{A,i}$ and $G_{B,i}$ respectively represent the number of sequences corresponding to the i-th OTU, taking samples A and B as examples.

(2) Beta analysis

Beta diversity analysis is a general direction and concept, which is specifically used in the research field of microbial community diversity, and has formed some more valuable analysis methods, focusing on similarity clustering tree, PCoA (principal component analysis) and NMDS (Non-metric multidimensional scaling analysis) et al [13]. The basis of these analyses is the Beta diversity numerical matrix, that is, the Beta diversity distance matrix is calculated according to the distance numerical estimation algorithm representing the degree of difference between different samples. As shown in Table 1, the Beta diversity difference of the three samples is shown, and the numerical value reflects the difference in diversity between samples. The difference between sample 1 and sample 2 is 0.4657, the difference between sample 1 and sample 3 is 0.8132, and the difference between sample 2 and sample 3 is 0.9715.

Table 1. Beta diversity differences

	Sample 1	Sample 2	Sample 3
Sample 1	0	0.4657	0.8132
Sample 2	0.4657	0	0.9715
Sample 3	0.8132	0.9715	0

3. Construction of a Microbial Visualization Platform Based on Biosensors

3.1. Visualization Platform

The microbiological visualization platform adopts the server-client mode. As shown in Figure 1, the user only needs to log in to the website, upload the microbiome data to be analyzed to the website, and select the required analysis items. The platform will be used in the background. Machine learning, as well as statistical knowledge, analyze the data, visualize the results, and return them to the user [15].

Build an interactive visualization platform based on biosensors. The type of biosensor is optical

biosensor. As long as the user uploads the relevant data (evolutionary tree, OTU-table, etc.), the relevant visualization results can be returned, so that the user can easily get the visualization. As a result, compared to traditional visualization software, installation and other troubles are saved [16]. The basic part of the visualization platform includes four functions: microbial community annotation, evolutionary tree, cluster tree visualization, data matrix visualization, and related statistical charts. The visualization of phylogenetic tree clustering and related statistical charts are the core, which can link the annotation results and data matrix visualization, and can clearly observe the annotation of the most interesting microorganisms, as well as the content of the microorganisms in the environment and other information [17]].

In order to make all our analysis results more intuitive to feedback to users, the role of visualization tools is reflected. We used the plotly tool, which is a tool based on D3js. The final visualization results are interactive, and we can selectively see the samples we want to see, or the flora of interest. The required result is finally output through JavaScript syntax, and the high-definition vector image is obtained [18].

Due to the huge features of microbial data and the large number of features of samples, statistical methods are also an indispensable means in order to find some microbial groups that are highly correlated with certain features. For some continuous sample features, our tool will calculate the coefficient of linear correlation [19].
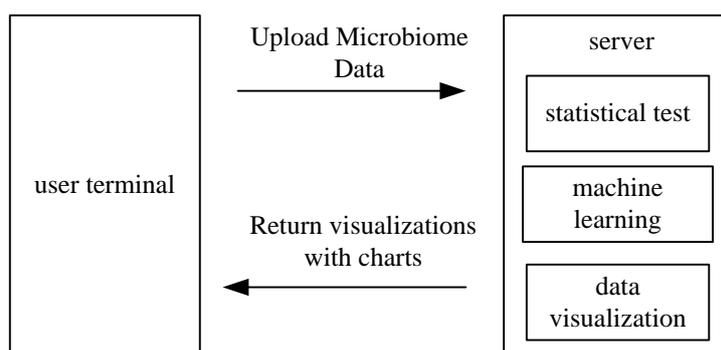
Figure 1. Microbial visualization platform architecture

3.2. Analysis of Microbial Community Diversity

(1) Screening of input sequences

The species composition in the metagenomic sequencing data of the existing environmental microbial samples is complex, and the relative species or homologous genes will seriously affect the identification accuracy of the method. To achieve accurate identification, it is necessary to ensure that the inclusions are basically correct, but also to ensure that there are not many left out, that is, to control false positives (refers to genes that do not belong to the specified strain but are mistakenly included by the method) and false negatives (refers to This gene belongs to the specified strain but was mistakenly ignored by the method). Therefore, for the case where the input sequence is a genome sequenced from a single strain, all the input sequences will be used for downstream analysis; while for the case where the input sequence is a metagenomic assembly sequence, it is necessary to screen out the sequences from the research object strains before constructing the pan-genome network.

(2) Sequencing and preprocessing

Sequencing techniques can generally be classified into three categories: whole-genome shotgun

sequencing, amplicon sequencing, and transcriptome sequencing. Whole-genome shotgun sequencing explores entire genetic data without any pre-set filters. Amplicon sequencing, on the other hand, focuses on a small specific region on the genome, starting from a sequence of conserved tags (called primers) to achieve detailed knowledge of the region. Amplicon sequencing is a powerful tool for in-depth analysis of system evolution and evolution. Transcriptome sequencing, also known as RNA-seq, uses RNA-specific primers to extract and clone the RNA components of a sample, and perform shotgun sequencing. The sequencing result is a quantitative spectrum of RNA transcription activity, and researchers can choose the appropriate sequencing according to their needs. technology.

(3) Data analysis

In data analysis, OTU (species taxonomic unit) is one of the most important concepts. Usually, researchers cluster microbial sequences together (with a similarity of more than 97%), so that each class is an OTU and can be regarded as a "species". For many years, the OTU has been the most commonly used unit of microbial diversity, especially when analyzing datasets of 16S rRNA marker gene sequences. The data analysis part includes identifying the composition of the microbial community, downstream analysis, and data visualization. After the OTU is divided, the next step is to use the OTU as the operation unit to carry out cluster analysis. The purpose of this process is to use the similarity degree to group the sequences. OTU clustering usually requires a certain similarity threshold as a reference limit, and the mainstream approach is to use 97% similarity as the identification standard. The main steps of this analysis are shown in Figure 2.
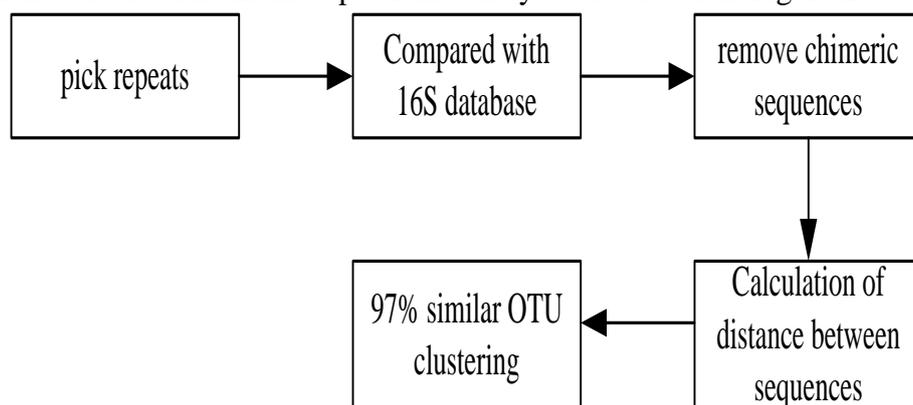

Figure 2. OTU clustering process

4. Application of Visualization Platform in Microbial Diversity Analysis

4.1. Visualization of Microbial Community Composition

The species content annotation map uses the species information of the representative sequences, and then by calculating some simple summation calculations of OTU abundance in the OTU-table, the percentage of all sequences occupied by some species in these sequences will be obtained. This allows us to clearly see how the species composition of the community is. The species composition of microbial communities is displayed from seven levels of biological taxonomic units. As shown in Figure 3, the columns of different lengths represent the percentage of species.
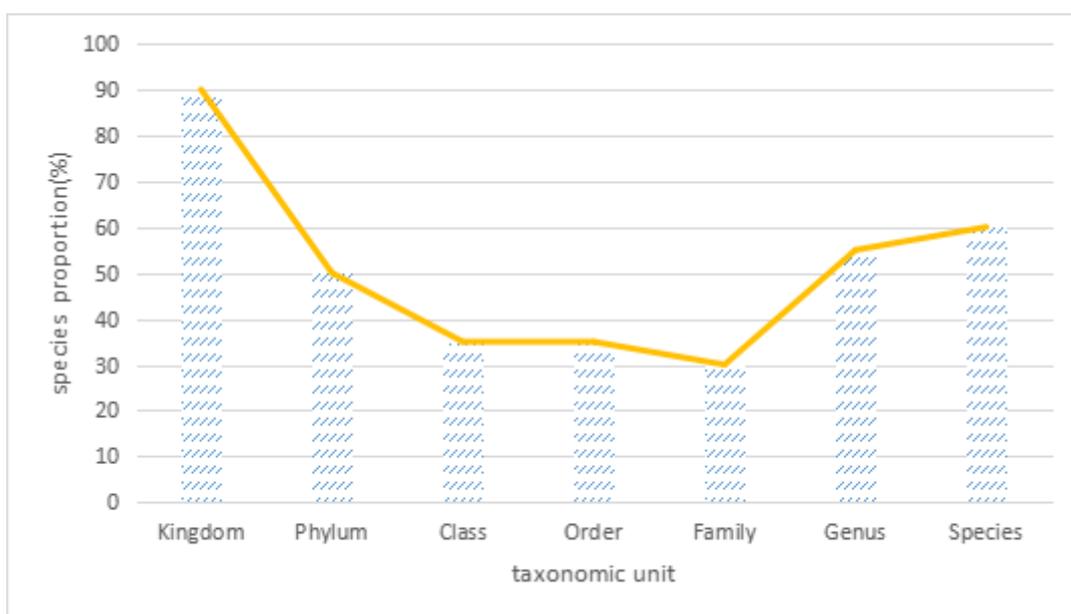
Figure 3. Species community composition

4.2. Visualization of Microbial Community α-Diversity and β-Diversity

The diversity of microbial communities includes alpha-diversity and beta-diversity, and the platform provides a visualization module for alpha rarefaction. Alpha rarefaction randomly extracts the number of sequences to be sequenced and calculates the alpha-diversity, gradually increasing the number of sequences to be sequenced. When the number of extracted sequences is large enough, the alpha-diversity will stabilize, as shown in Figure 4. That means that the number of sequences sequenced in our sample is sufficient. Here, we have a variety of common measures of alpha diversity, including Observer OTUs and Shannon indices.
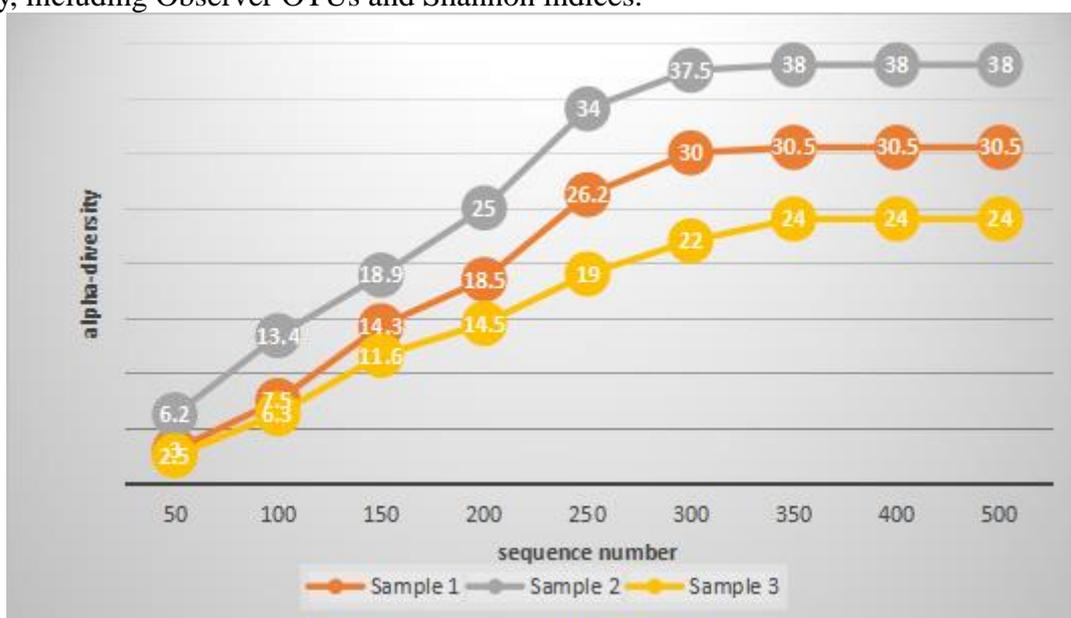

Figure 4. Schematic diagram of alpha rarefaction

The main process of Alpha analysis is: based on the calculation formulas of different indices in the statistical analysis method, call the corresponding index calculator to calculate the Alpha similarity index, obtain indices such as Chao, Ace, Shannon, etc., and save the results in the corresponding folder. middle.

The main process of beta analysis is to count the distribution of OTUs in each sample, calculate the intersection, that is, the number of OTUs shared by different samples; count the differences in the composition and abundance of OTUs, and use the UPGMA algorithm to calculate the similarity between samples., and convert the similarity relationship into a tree-structured Newick format to facilitate subsequent visualization; according to the taxonomic classification results obtained in the previous composition analysis, calculate the proportion of each OTU abundance to the total number of species in each sample; then use The Unifrac algorithm calculates the Beta diversity distance matrix reflecting the differences between samples, and finally calls the corresponding algorithm package to perform PCoA analysis or NMDS analysis based on the previously calculated distance matrix.

5. Conclusion

With the continuous breakthrough of sequencing technology and the development and popularization of biosensor technology, bioinformatics has undergone a dramatic change. The use of high-throughput sequencing data for computational analysis of microbial community diversity has become an important method that researchers rely on, and the results are also important for ecological Research has opened up new ideas. The biosensor-based microbial visualization platform developed in this paper is to meet the actual needs of microbial community diversity researchers. It displays the results of microbial community structure research on the visualization platform to help them analyze microbial structure and improve the efficiency of microbial research..

References

[1] Gil J, DAC Sánchez, Morales-Osorio J G. Edaphic factors associated with the development of avocado wilt complex and implementation of a GIS tool for risk visualization. Scientia Horticulturae, 2021, 228(110316):1-9. https://doi.org/10.1016/j.scienta.2021.110316

[2] Kumar S S, Kumar V, Kumar R, et al. Microbial fuel cells as a sustainable platform technology for bioenergy, biosensing, environmental monitoring, and other low power device applications. Fuel, 2019, 255(NOV.1):115682.1-115682.16. https://doi.org/10.1016/j.fuel.2019.115682

[3] Fsa B, Jkbc D, Lk C, et al. Life on the driveline: Molecular detection and fluorescence in situ hybridization-based visualization of microbial species in patients with left ventricular assist devices. The Journal of Heart and Lung Transplantation, 2018, 37(1):163-166.

[4] Bocobo G A, Tharan N, Choudhury N, et al. Automation of Postoperative Surgical Drain Monitoring with Novel Biosensing Technology: Proof of Concept in a Peritoneal Injury Model:. Surgical Innovation, 2021, 28(4):504-506. https://doi.org/10.1177/1553350620979819

[5] Saadati A, Hasanzadeh M, Shadjou N. Application of advanced nanomaterials in enzyme biosensing technology. Iranian Journal Of Chemistry & Chemical Engineering-International English Edition, 2020, 38(19100894):309-337.

[6] Patel S K, Parmar J, Sorathiya V, et al. Graphene-Based Plasmonic Absorber For Biosensing Applications Using Gold Split Ring Resonator Metasurfaces. Journal of Lightwave Technology, 2021, PP(99):1-1.

[7] Casquel R, Holgado M, Laguna M F, et al. Engineering vertically interrogated interferometric

sensors for optical label-free biosensing. Analytical and Bioanalytical Chemistry, 2020, 412(14):1-13. https://doi.org/10.1007/s00216-020-02411-3

[8] *Nikitin M P, Orlov A V, Znoyko S L, et al. Multiplex biosensing with highly sensitive magnetic nanoparticle quantification method. Journal of Magnetism & Magnetic Materials, 2018, 459(AUG.):260-264.*

[9] *Claire Z, Abdullah A, Yrjo G, et al. Postharvest Supply Chain with Microbial Travelers: a Farm-to-Retail Microbial Simulation and Visualization Framework. Applied and Environmental Microbiology, 2018, 84(17):AEM.00813-18. https://doi.org/10.1128/AEM.00813-18*

[10] *Lenchenko E M, Sachivkina N P, Blumenkrants D A, et al. Visualization of microbial biofilms in case of digestive disorders in lambs. Veterinary Science Today, 2021, 1(1):59-67.*

[11] *Dame R T, Lin S N, Wuite G. Direct visualization of the effect of DNA structure and ionic conditions on HU–DNA interactions. Scientific Reports, 2021, 11(1):1-10. https://doi.org/10.1038/s41598-021-97763-w*

[12] *Cecconet D, Bolognesi S, Daneshgar S, et al. Improved process understanding and optimization by multivariate statistical analysis of Microbial Fuel Cells operation. International Journal of Hydrogen Energy, 2018, 43(34):16719-16727.*

[13] *Shank, Elizabeth A. Considering the Lives of Microbes in Microbial Communities. Msystems, 2018, 3(2):e00155-17. https://doi.org/10.1128/mSystems.00155-17*

[14] *Wan K Y, Goldstein R E. Time Irreversibility and Criticality in the Motility of a Flagellate Microorganism. Physical review letters, 2018, 121(5):058103.1-058103.6.*

[15] *Singh B, Katiyar D, Tilak R, et al. Prevalence of Urinary Tract Infection Causing Microorganism and Determination of Susceptibility to Antibiotic among Slum Women of District Varanasi, India. International Journal of Current Microbiology and Applied Sciences, 2018, 7(2):3483-3497. https://doi.org/10.20546/ijcmas.2018.702.415*

[16] *Wright W, Obanion M S, Majkowwicz D C, et al. Evaluating Immersive Visualization Technology for Use in Geospatial Science Education. Surveying and Land Information Science, 2020, 79(1):15-22.*

[17] *Tinnaluri N, Devanuri J. Heatline Visualization for Thermal Transport in Complex Solid Domains with Discrete Heat Sources at the Bottom Wall. International Journal of Heat and Technology, 2019, 37(1):100-108. https://doi.org/10.20546/ijcmas.2018.702.415*

[18] *Mizukoshi K, Arakawa T, Mitsubayashi K. Convenience biosensing approach of lactic acid in stratum corneum for skin care assessment. Skin Research and Technology, 2020, 26(4):455-464. https://doi.org/10.1111/srt.12834*

[19] *Das B, Franco J L, Logan N, et al. Nanozymes in Point-of-Care Diagnosis: An Emerging Futuristic Approach for Biosensing. Nano-Micro Letters, 2021, 13(1):1-51. https://doi.org/10.1007/s40820-021-00717-0*