Comparative analysis of bacterial community diversity between soil and water of Dongzhai Harbor Mangrove reserve using 16S rRNA gene sequencing and shotgun metagenomic sequencing

Xianzi Rong
Hainan Normal University
Wenfei Zhang (wenfei2007@163.com)
Hainan Normal University
Bo Xu
Hainan Normal University
Peng Li
Hainan Normal University
Hongping Wu
Hainan Normal University
Yinghong Jin
Hainan Normal University
Ruiping Wang
Hainan Normal University

Research Article

Keywords: bacterial community diversity, soil, water, 16S rRNA gene sequencing, shotgun metagenomic sequencing

Posted Date: December 11th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-114741/v1

Abstract

The mangrove ecosystem has rich biological resources and is the fourth highest service value ecosystem in the world. The Dongzhai Harbor Mangroves (DHM) reserve is located in the intertidal zone where the freshwater and seawater are dynamically interlaced, and its unique habitat may hide special microbial resources. In this study, we analyzed and compared bacterial community composition and diversity between the DHM soil and water by 16S rRNA gene sequencing and shotgun metagenomic sequencing. We found that the dominant species in both soil and water of DHM were Proteobacteria and Actinobacteria, while the most differentially abundant species were Chloroflexi and Bacteroidetes. However, shotgun metagenomic sequencing analysis identified more highly abundant species in the water than in the soil, and identified more species with significant differences between the soil and the water ($P \leq 0.0001$). Network analysis identified more co-occurring species in the soil and water. Kyoto Encyclopedia of Genes and Genomes (KEGG) functional analysis identified three relatively abundant pathways: metabolism (accounting for more than 50%), genetic information processing, and environment information processing. Our results increase our understanding of bacterial community diversity in the water and soil of the DHM. Additional information that is hidden in the environment may be obtained by shotgun metagenomic sequencing; this technique can be used to mine more microbial resources from various environments.

Introduction

Mangrove forests as special coastal ecosystems are found in tropical and subtropical coastal intertidal zones. They act as a transitional zone among land, sea, and rivers, and are the final land environment before the land-based pollutants into the sea, which has been shown to purify pollutants[1] and is periodically submerged by the tide under certain conditions. The properties of these environments determine the strength and characteristics of microbial ecological functions[2]. High tide brings in salt water, and when the tide recedes, solar evaporation of the seawater leads to further increases in salinity. In addition, mangrove trees shed a great deal of organic matter, particularly the fallen leaf covered with bacteria, protozoans, and fungi into water. As a result that the microbial community in mangrove waters is more pronounced than the nearshore coastal environment. Mangrove soils have an acidic pH, are high in organic matter and salt content, and are mainly anaerobic (although surface soils are aerobic)[3,4]. The unique characteristics of this habitat have led to the development of a rich and unique microbial community[5].

Microbes play an important role in material circulation, energy flow, and ecosystem stability[6]. In recent years, microorganisms as important members of mangrove ecosystems have received increasing attention[7,8]. Microbial diversity is an important part of biodiversity, and understanding the microbial diversity is the basis of microbial research in mangrove wetlands, which is of great value for developing microbial resources. But mangrove microbial communities are still little understood, primarily due to the limitations of research methods[9].

The Dongzhai Harbor Mangroves (DHM) reserve, which is currently the largest coastal mudflat in China, is located in Yanfeng Town, Haikou City, China, northeast of Hainan Island (110° 32'-110° 37' E and 19° 51'-20° 1'N). There are three economically important tree species in the mangrove family including Kandelia candel, Bruguiera gymnorrhiza, and Rhizophora stylosa[10]. In recent years, numerous studies addressing the protection of plant and animal biodiversity in the DHM reserve have been published[11,12]. According to our knowledge, there is no report of comprehensive comparing soil and water microbial diversity study and developing microbial resources in DHM. Therefore, it is necessary to found more hidden information from microorganism in DHM. In this study, combining 16S rRNA gene and shotgun metagenomic sequencing, we comprehensively understand microbial community composition and microbial diversity in the soil and water of DHM.

Results

Sequencing statistics. The 16S rRNA gene sequencing results for the soil and water samples are shown in Table S1. Across all of the samples, the total number of sequences was 321,131; the total number of bases (bp) was 134,242,228; and the average sequence length was 418.04 bp. The quality control results for the shotgun metagenomic sequences are shown in Table S2. The average numbers of raw reads in the soil and water samples were 99,317,796.6667 and 98,984,089.3333, respectively, while the average numbers of clean reads in the soil and water samples were 98,645,909 and 97,975,913, respectively. After quality control, the average proportion of raw reads for both the soil samples and the water samples was 99%, and the average proportion of raw bases for both the soil samples and the water samples was 98%.

The Shannon curve shown as Fig. S2, was one of the indexes that was used to estimate the microbial diversity of each sample. The result showed that our sequencing data were sufficient to reflect the vast majority of the bacterial diversity in the soil and water samples. The larger Shannon index for the soil samples indicated that bacterial diversity in soil was higher than that in the water in the DHM. The soil in DHM wetland as a highly heterogeneous medium are of marine alluvium, which is rich in organic matter. It provides a variety of suitable habitat and environment conditions for microorganism, supporting the high soil microbial diversity and a variety of different microbial13.

Species composition and difference analysis. As shown in the species community bar chart (Fig. 1a, b), Archaea and bacteria remained after the removal of Eukaryota and viruses. Most of the microorganisms detected in the soil and water samples from the DHM were bacteria (98.21% and 98.61%, respectively). Only a small proportion of the samples were Archaea (1.60% and 1.18%, respectively). The 16S rRNA gene analysis identified the dominant species in the DHM water samples as Proteobacteria (54.1%), Actinobacteria (13.0%), and Bacteroidetes (26.7%). However, the dominant species in the DHM soil samples were Proteobacteria (47.9%), Actinobacteria (7.04%), and Chloroflexi (16.5%) (Fig. 1a). The shotgun metagenomic analysis identified the dominant species in the water samples as Proteobacteria (65.55%), Actinobacteria (11.99%), and Bacteroidetes (16.44%), and identified the dominant species in the soil samples as Proteobacteria (62.32%), Actinobacteria (9.31%), and Chloroflexi (7.65%) (Fig. 1b). However, bacterial community composition and structure in soil or water were affected by different

environmental factors. As a results that bacterial communities abundance in water were different from those in soil, and the major reason for this discrepancy might be related to different habitats.

Species heatmap clustering is based on similarities in relative abundance among species and samples, and aggregates species with high abundance and low abundance in separate block[14]. The bacterial community heatmap of the 30 most abundant species based on 16S rRNA gene sequences and shotgun metagenomic sequences showed the certain species that were highly abundant in the water samples were moderately abundant or uncommon in the soil samples, such as Pseudarcicella. However, other species had similar levels of relative abundance in both soil and water, such as the highly abundant species of unclassified Gemmatimonadetes (Fig. 1c,d)

The phylum Bacteroidetes was significantly different between soil and water ($P \leq 0.0001$Fig. 2a), as showed in the bar chart of species differences based on 16S rRNA gene analysis. However, shotgun metagenomic analysis identified additional species that were significantly different between soil and water ($P \leq 0.0001$), as shown in the bar plot of Welch's *t*-test (Fig. 2b), including Bacteroidetes, Firmicutes, Cyanobacteria, Planctomycetes, Acidobacteria, Spirochaetes, Thaumarchaeota, and unclassified bacteria. Due to unequal amplication 16S rRNA gene sequence of species, it may be biased. However, shotgun metagenomes cover a widespread microbial community and generate huge number of reads with various length by using this sequencing technologies, as a result that covering more species with significantly difference and genetic information.

Network analysis. The co-occurrence network based on 16S rRNA gene analysis indicated that the species Chloroflexi, Acidobacteria, Planctomycetes, Proteobacteria, Gemmatimonadetes, Cyanobacteria, Ignavibacteriae, Bacteroidetes, Actinobacteria, Firmicutes, and Verrucomicrobia co-occurred in the soil and water samples (Fig. 3a). Shotgun metagenomics analysis identified additional species that co-occurred in the soil and water (Fig. 3b). The species correlation network of the top 35 phyla indicated that most species correlated with others (Fig. 4), including Euryarchaeota, Firmicutes, Verrucomicrobia, Chlorobi, and Tenericutes. Based on the clustering coefficient, these taxa play an important role in the species correlation network. Shotgun metagenomic sequencing not only elucidates species structure and systematic community evolution, but also supports the genetic analysis of functional metabolic networks within the microbial community.

Functional diversity analysis. 16S rDNA functional predictions indicated that three main categories, including metabolism, genetic information processing, and environment information processing, which have relatively high abundance in both the soil and the water (Fig. 5a). In the shotgun metagenomic functional analysis, we annotated six types of functional genes, and were also found to be enrich in three main categories of metabolism, environment information processing, and genetic information processing (Fig. 5b). Of these, metabolism represents more than 50% of all functional classifications at KEGG pathway level 1. This suggested that microbiome of DHM display a relatively high level of metabolism activities and genetic stability.

Fig. 5 Functional prediction at KEGG pathway level 1orthologies **a** 16S rDNA functional prediction; **b** shotgun metagenomic functional annotation.

Discussion

The microbial diversity research is of great value for many aspects, which includes 1) the development of biological resources, 2) accelerating the discovery process of new functional genes and bioactive substances, 3) the clarification of the relationship between microbial communities and habitat, 4) the elucidation of the association between the community structure and function[15]. Previous studies have suggested that bacteria are the most dominant group of microorganisms in this environment[7]. Bacterial communities play an important role in material transformation in mangrove ecosystems[16]. A study has found that Proteobacteria and Chloroflexi were ubiquitous and dominant in the DHM soil[17], and our result is consistent with it (Fig. 1a,b).

In this study, we compared bacterial community structure and diversity between the soil and water of the DHM using 16S rRNA gene sequencing and shotgun metagenomics sequencing. The two approaches used here can better reveal and reflect the species composition, levels of species abundance, as well as the bacterial functions in the soil and water sample of DHM. However, the two methods generated slightly different results for the analyses of microbial community and structure. The shotgun metagenomic sequencing identified more similar, different, and associated species, and retrieved more hidden information than 16S rDNA sequencing. Shotgun metagenomic sequencing technique can thus be used to mine additional microbial resources from various environments.

The wetland ecosystem in DHM is open and the microorganisms in soil and water always keep contacting and communicating. Both the soil and water samples shared certain microbial community structures. Within the bacteria, Proteobacteria was the dominant phylum, followed by Bacteriodes, Chloroflexi, and Actinobacteria. The similar bacterial phylum dominancy was detected like those observed in the Brazilian mangrove metagenomes and soil microbiome from a managed mangrove in Malaysia[18]. It also has been reported that members of the Proteobacteria are the most abundant group of bacteria in soil and are known to harbor a diverse group of metabolic enzymes[19].

The species symbiosis and intercorrelation in soil and water sample could better explain that the habitat in DHM wetland ecosystem is not completely closed and isolated from each other. Functional profiles were analyzed and revealed the similar basic functional categories between the soil and water in DHM through the 16S rRNA gene and shotgun metagenomics sequences. As a result the high abundance of sequences assigned to metabolic pathways, which is also commonly found in other mangrove metagenomes[19].

Our results increased our comprehensive understanding of microbial community diversity in DHM water and soil. However, the relationship between mangrove microbial diversity and the maintenance of environmental stability, as well as the connection between mangrove microbial diversity and

microorganism functionality, remain to be clarified. Future studies are necessary to investigate microbial diversity in response to human activities and environmental change. In particular, it is important to mine additional genetic resources using the metagenomics sequencing technique.

Methods

Sample collection and DNA extraction. Soil and water samples were collected in the DHM and the Yanfeng River basin during ebbing tides in July 2018. We set up five sampling quadrats (A, B, C, D, and E) in the upper, middle and lower reaches of Yanfeng River. The sampling map and the latitude and longitude information as listed in Fig. S1. A composite sample is made by combining five subsamples from the same area in a site. The soil samples were collected from the topsoil layer (5–10 cm deep) before the leaves and grass covering the topsoil at each sampling point were cleared. Soil samples had four replicates for each quadrats and twenty soil samples were taken from five quadrats. We fully mixed all of the samples, removed debris, and divided the mixed sample into three equal parts, which were then stored in sterile plastic bags.

The water samples were collected simultaneously at each sampling quadrats. All of the water samples (20 L in total) were fully mixed, and then divided into three equal subsamples. Microorganisms in each subsample were collected into sterile tubs by centrifugation. Total DNA was extracted by Zhou's In-situ Pyrolysis[20]. The extracted DNA samples were quantified using a NanoDrop 2000 (Thermo Fisher Scientific, USA) and DNA quality was determined by 1% agarose gel electrophoresis.

16S rRNA gene and shotgun metagenomic sequencing. Extracted DNA was sent to Shanghai Majorbio Bio-pharm Technology Co., Ltd for 16S rRNA gene sequencing and shotgun metagenomic sequencing. The V3–V4 hypervariable regions of the bacterial 16S rRNA gene were amplified with primers 338F (5' – ACTCCTACGGGAGGCAGCAG - 3') and 806R (5' - GGACTACHVGGGTWTCTAAT - 3'). PCR products were quality checked by 1% agarose gel electrophoresis and purified using AxyPrep DNA Gel Extraction Kits (Axygen Biosciences, Union City, CA, USA). Purified amplicons were quantified using a QuantiFluorTM-ST (Promega, USA), according to the manufacturer's protocol.

In order to perform the shotgun sequencing, the extracted DNA was fragmented to an average size of about 300 bp using Covaris M220 (Gene Company Limited, China), and paired-end libraries were constructed using TruSeq^{TM} DNA Sample Prep Kits (Illumina, San Diego, CA, USA). Adapters containing the full complement of sequencing primer hybridization sites were ligated to the blunt-ends of all fragments. The mixed PCR products of 16S rRNA genes and paired-end sequencing library were sequenced using an *Illumina* HiSeq PE 4000 platform at Shanghai Majorbio Bio-pharm Technology Co., Ltd. (Shanghai, China) with paired-end sequencing technology following the manufacture's instruction (www.illumina.com).

Comparative analyses of the soil and water bacterial communities. The 16S rDNA and metagenomic data were analyzed using the Majorbio I-Sanger Cloud Platform (www.majorbio.com). Sequences with

≥97% similarity were clustered into Operational taxonomic units (OTUs) using U Search (V Session 7.0; http://drive5.com/uparse/). To identify the species corresponding to each OTU, the Bayesian algorithm of the Ribosomal Database Project (RDP)classifier was used to carry out taxonomic analysis against the Silva database. Based on the species identification corresponding to each OTU cluster, all samples were flattened based on the minimum number of sample sequences and further analyzed.

Representative sequences from the non-redundant gene catalog were aligned to the NCBI NR database using BLASTP (Version 2.2.28+, http://blast.ncbi.nlm.nih.gov/Blast.cgi) with an e-value cutoff of $1e^{-5}$ to obtain taxonomic annotation information for each sampled species. We visualized the composition of the bacterial community at the phylum level using a bar chart.

Species abundance and composition at the genus level (for the 30 most abundant genera) were compared between the soil and water communities using community heatmaps. Species hierarchal clustering methods and sample-level clustering methods were average. Welch's *t* -tests were used to identify species with significantly different numbers of reads between the soil and water communities.

Network analysis. The phylum co-occurrence networks were evaluated (for the 30 most abundant phyla) to explore the co-existence relationships among species in soil and water samples. To evaluate these networks, we used 16S rDNA network analysis and metagenomic-based species distribution network analysis. The phylum correlation networks were also evaluated to explore the interactions among phyla, based on the metagenomic univariate correlation network of the 35 most abundant phyla. For these analyses, the Spearman correlation coefficient model was used with a cutoff of 0.5 and a p-value of 0.05.

Functional annotation in the soil and water. To perform functional predictions based on 16S rDNA, we first standardized the OTU abundance table using PICRUSt[21], and then obtained the KEGG Ortholog (KO) of each OTU based on the greengene id. KEGG annotations were performed using BLASTP (Version 2.2.28+) against the KEGG database[22] (http://www.genome.jp/keeg/), with an e-value cutoff of $1e^{-5}$. We identified level 1 metabolic pathways in the soil and water samples based on 16S rDNA functional predictions and metagenomic-based taxonomic annotations.

Data Availability

16S rRNA gene sequences have been deposited in the NCBI Short Read Archive database under accession number SRP249356; And shotgun metagenomic sequences have been deposited in the NCBI Short Read Archive database under accession number SRP249901.

Declarations

Acknowledgments

This work was supported by the Education Bureau of Hainan Province (No. HNKY2017ZD-13), and

the National Natural Science Foundation of China (No.41366001).We acknowledge Hainan Dongzhaigang Nature Reserve Authority for assistance with soil and water sample collection.

Authors contributions

R.X. Z.and Z.W.F. designed the experiments; R.X. Z.performed the experiments; X.B., J.Y.H.,W.H.P.,W.R.P.and L.P. contributed reagents/materials/analysis tools, all authors analyzed the data; R.X. Z.and Z.W.F. wrote the manuscript. All authors read the final manuscript.

Additional information

Supplementary information accompanies this paper at https://orcid.org/0000-0002-4115-268X.

Competing Interests The authors declare no competing interests.

References

1. Zhu, Y. H., Liu, L., Liu, K., Wang, S. G. & Ai, B. Progress in researches on plant biomass of mangrove forests. Wetland ence 12, 515-526 (2014).
2. Gebarowska, E., Pusz, W., Kucinska, J. & Kita, W. Comparative analysis of airborne bacteria and fungi in two salt mines in Poland. Aerobiologia 34, 127-138 (2018).
3. Navodha, D. & Upali, C. Effects of Mangrove Zonation and the Physicochemical Parameters of Soil on the Distribution of Macrobenthic Fauna in Kadolkele Mangrove Forest, a Tropical Mangrove Forest in Sri Lanka. Advances in Ecology,2014,(2014-12-16) 2014, 13 pages (2014).
4. Duarte, R. T. D., Nóbrega, F., Nakayama, C. R. & Pellizari, V. H. Brazilian research on extremophiles in the context of astrobiology. International Journal of Astrobiology 11, 325-333 (2012).
5. Jiang, Y. X. Establishment and Application of Metagenomics Platform Technology Based on Mangrove Soil Microbial Resources Research and Development, Xiamen University, (2007).
6. Jing, L. I., Yurong, L., Jizheng, H. E. & Yuanming, Z. Insights into the responses of soil microbial community to the environmental disturbances. Acta Scientiae Circumstantiae 33, 959-967 (2013).
7. Fitriyah, D., Arimurti, S. & Senjarini, K. Physiological and Molecular Characteristics of Bacterial Isolates from Bandealit Coastal Area Jember, East Java, Indonesia. HAYATI Journal of Bioences 20, 89-93 (2013).
8. Marko, M. & Pawliczak, R. The role of microbiota in allergy development. Alergologia Polska - Polish Journal of Allergology (2017).
9. Jiang, Y. X., Zheng, T. L. & Tian, Y. [Research on mangrove soil microorganisms: past, present and future]. ACTA MICROBIOLOGICA SINICA 46, 848-851 (2006).
10. Islam, M. S., Lian, C., Geng, Q., Kameyama, N. & Hogetsu, T. Chloroplast microsatellite markers for the mangrove tree species Bruguiera gymnorrhiza, Kandelia candel, and Rhizophora stylosa, and cross-amplification in other mangrove species. Conservation Genetics 9, 989-993 (2008).

11. Liao, B., Li, M., Zheng, S., Chen, Y. & Huang, Z. [Niches of several mangrove species in Dongzhai Harbor of Hainan Island]. Chinese Journal of Applied Ecology 16, 403-407 (2005).
12. Feng, Q. & Lin, Z. Study on the diversity of benthic arthropod in mangrove area of Dongzhaigang, Hainan. Marine ences (2013).
13. Hong-Jie, C. Research Advances in Soil Microbial Diversity and its Impact Factors. Territory & Natural Resources Study (2015).
14. CHEN Mei-chun, Z. Y.-j., LIU Bo, WANG Jie-ping, LIU Xiao-gang, YANG Wen-jing. . Analysis of Endophytes Diversity in Jasminum sambac (L.) Ait by Metagenomic Technology Journal of Tropical and Subtropical Botany 26, 77-87 (2018).
15. Zhong, W. Application of nucleic acid-based methods in the study of soil microbial diversity. Acta Pedologica Sinica 46, 334-341 (2009).
16. Holguin, G., Vazquez, P. & Bashan, Y. The role of sediment microorganisms in the productivity, conservation, and rehabilitation of mangrove ecosystems: an overview. Biology and Fertility of Soils 33, 265-278 (2001).
17. Peng Li, S. W., Haoneng Yin, Hongping Wu, Qin Peng, Wenfei Zhang, Ruiping Wang. Bacterial community diversity and dynamics of Dongzhai harbor mangrove soils in China. Canadian Journal of Microbiology (2019).
18. Spain, A. M., Krumholz, L. R. & Elshahed, M. S. Abundance, composition, diversity and novelty of soil Proteobacteria. Isme Journal 3, 992-1000 (2009).
19. Priya, G. et al. Metagenomic insights into the phylogenetic and functional profiles of soil microbiome from a managed mangrove in Malaysia. Agri Gene 9, 5-15, doi:10.1016/j.aggene.2018.07.001 (2018).
20. ZHOU, J. DNA recovery from soils of diverse composition. Applied and environmental microbiology 62, 316-322 (1996).
21. Aßhauer, K. P., Bernd, W., Rolf, D. & Peter, M. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. Bioinformatics, 2882-2884 (2015).
22. Gerlich, M. & Neumann, S. KEGG: kyoto encyclopedia of genes and genomes. Nuclc Acids Research 28, 27-30 (2000).

Figures

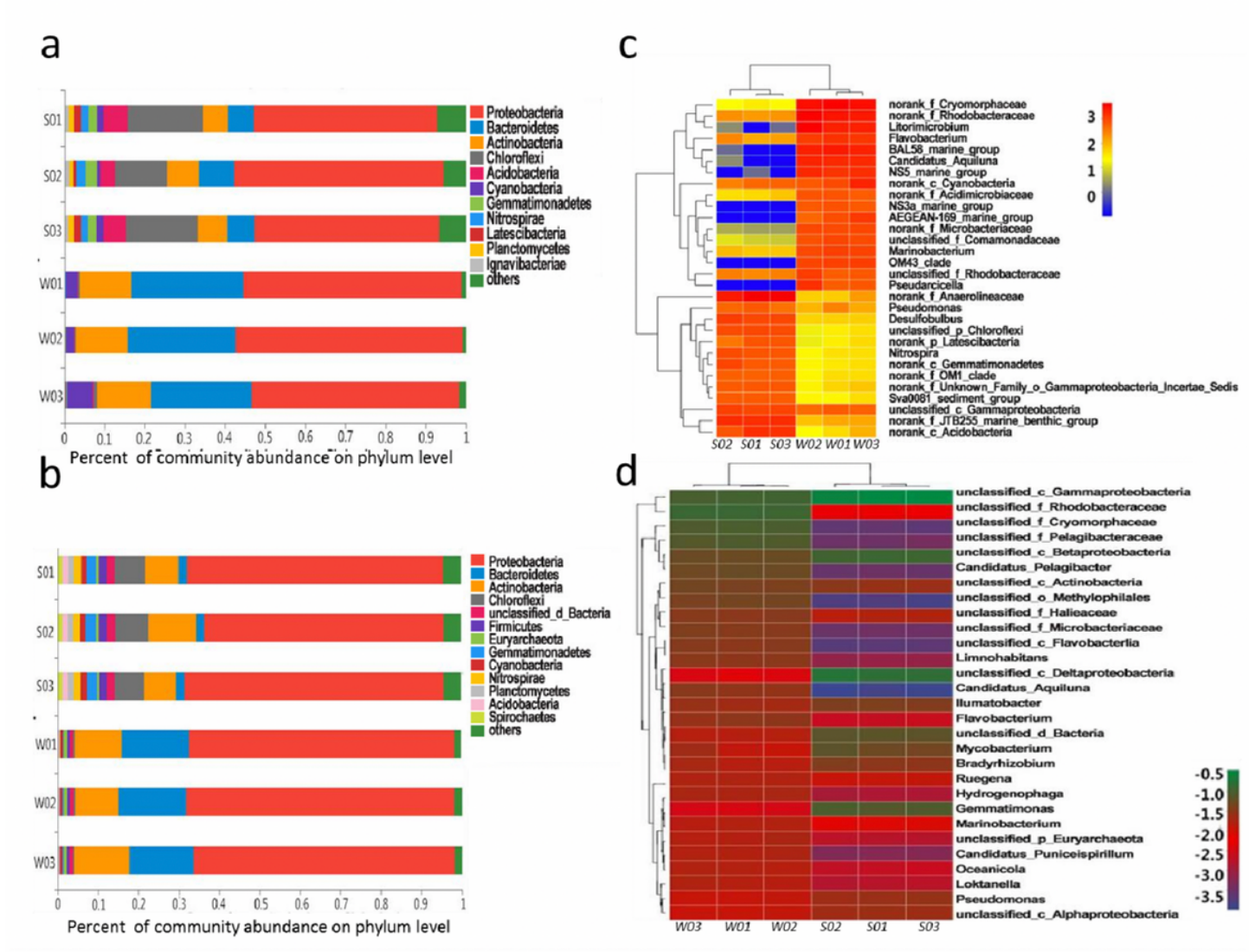


Figure 1

a and c Species composition and relative abundance in the soil and water samples based on 16S rRNA gene sequences; b and d species composition and relative abundance in the soil and water samples based on shotgun metagenomic data.

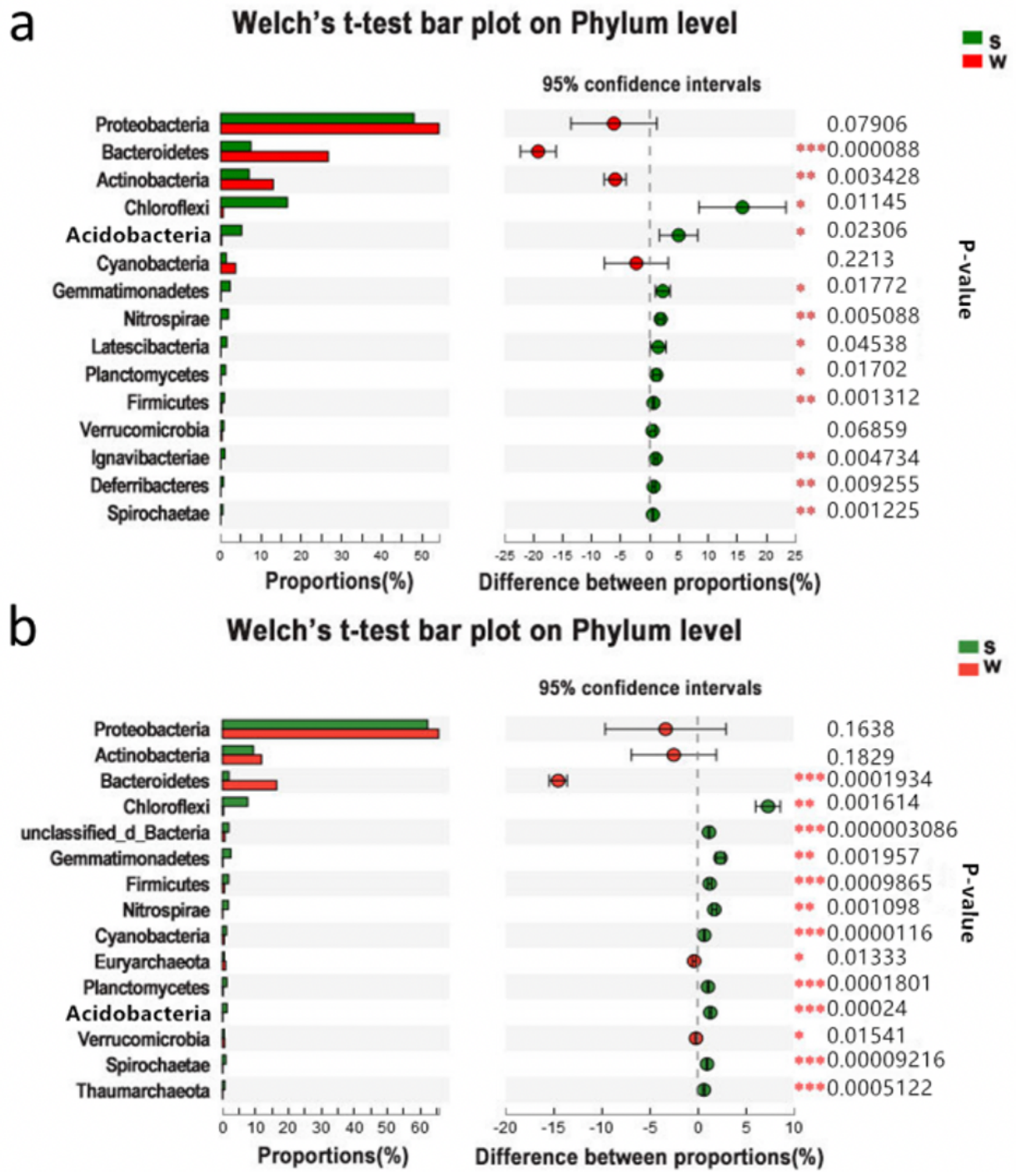


Figure 2

Analysis of the significantly different phyla between the soil and water. Statistical differences were identified using Welch's t test. *P < 0.05, **P < 0.001, ***P < 0.0001. a 16S rRNA gene analysis and b shotgun metagenomic analysis.

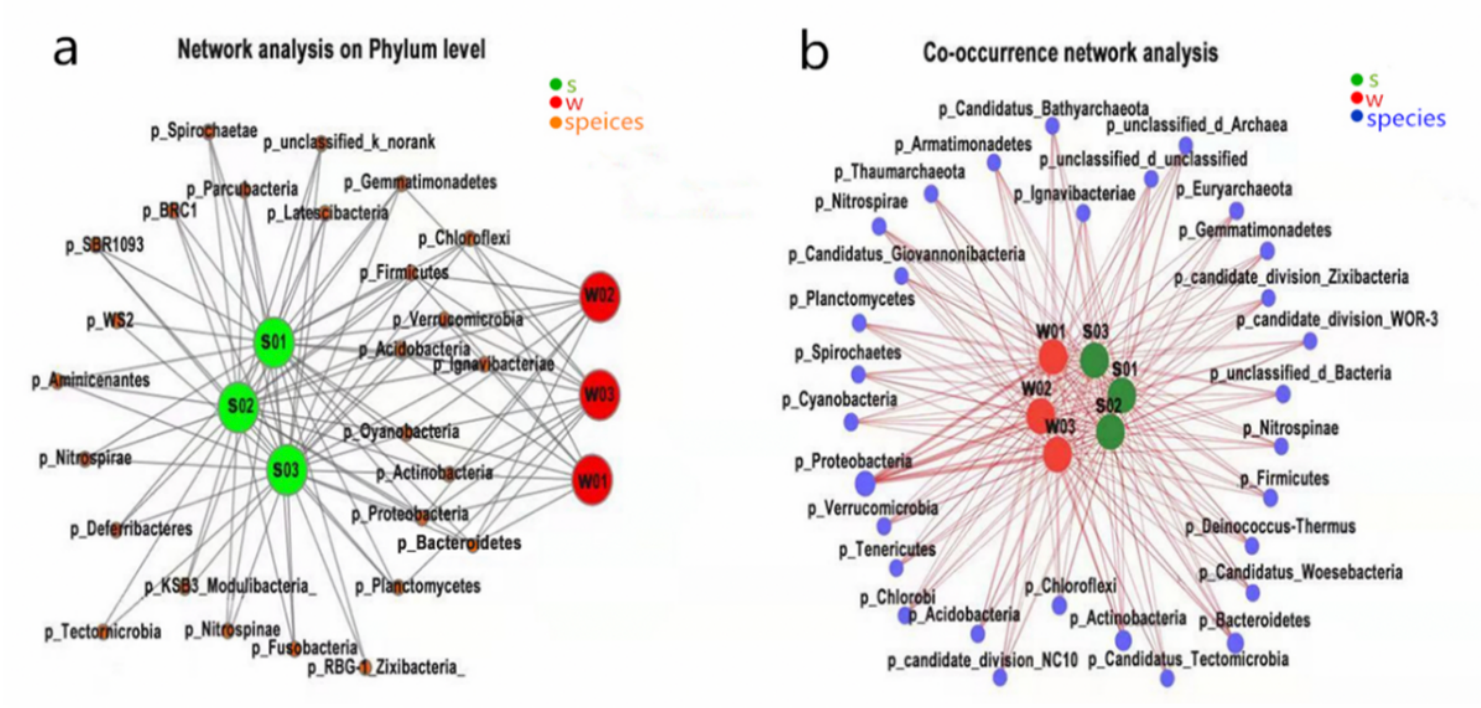
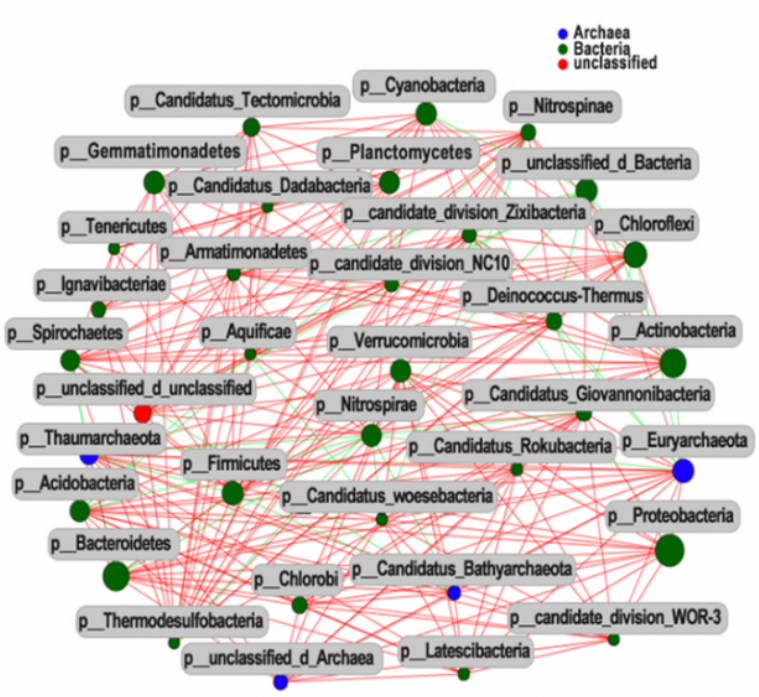


Figure 3

Co-occurrence species network analysis. a 16S rRNA gene analysis and b shotgun metagenomic analysis.



Appendix

Node_Name	Degree	Clustering
p__Euryarchaeota	13	1
p__Firmicutes	13	1
p__Verrucomicrobia	13	1
p__Chlorobi	13	1
p__unclassified_d_Archaea	13	1
p__Tenericutes	13	1
p__candidate_division_WOR-3	13	1
p__Planctomycetes	12	0.863636364
p__Deinococcus-Thermus	12	0.863636364
p__Armatimonadetes	12	0.863636364
p__Aquificae	12	0.863636364
p__Actinobacteria	11	0.818181818
p__Candidatus_Tectomicrobia	12	0.772727273
p__Nitrospirae	16	0.766666667
p__Spirochaetes	16	0.766666667
p__Thaumarchaeota	16	0.766666667
p__Nitrospinae	16	0.766666667
p__Proteobacteria	16	0.758333333
p__Bacteroidetes	16	0.758333333
p__Cyanobacteria	16	0.758333333

Figure 4

Species correlation networks showing the 35 most abundant phyla.

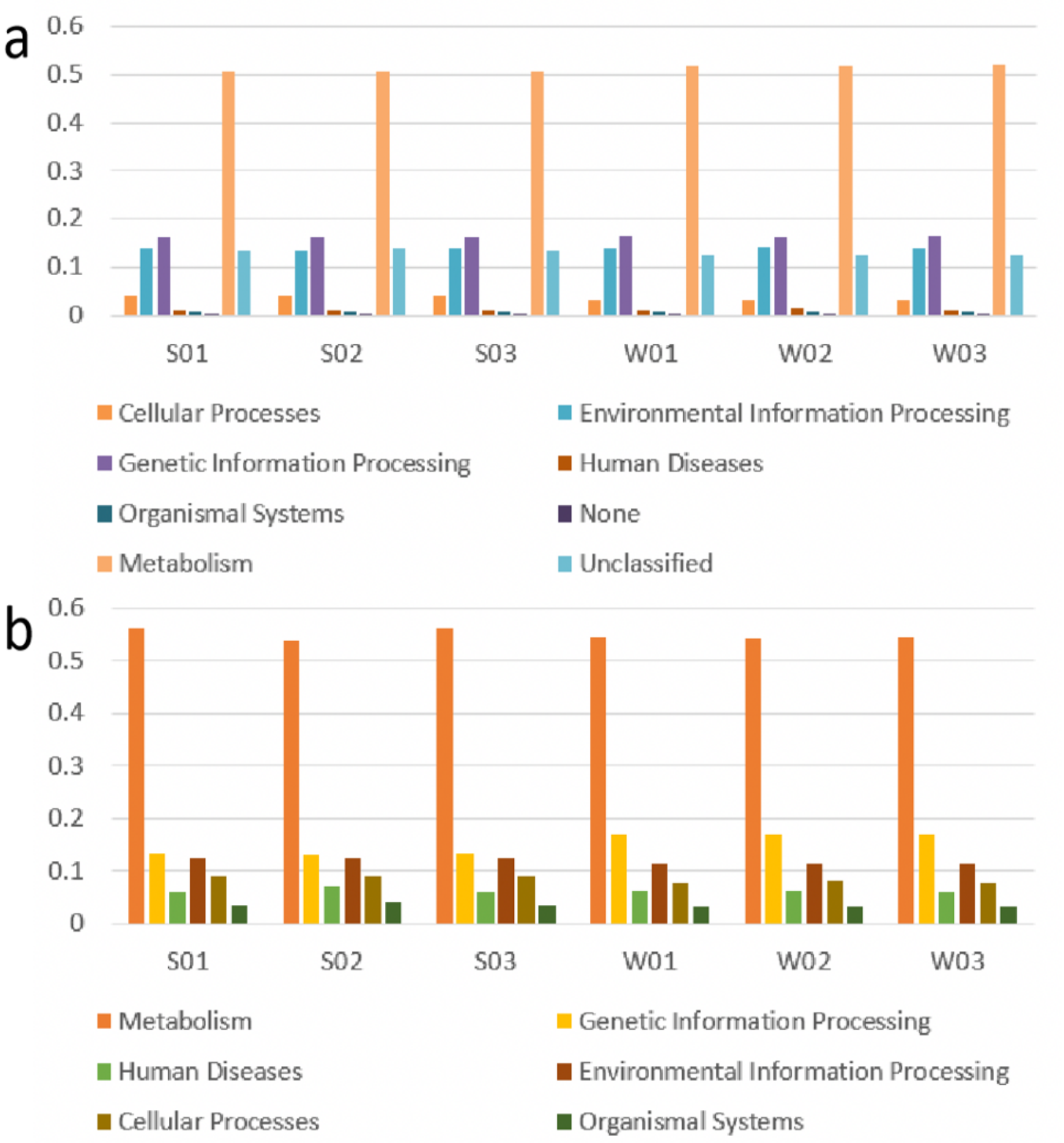


Figure 5

Functional prediction at KEGG pathway level 1orthologies a 16S rDNA functional prediction; b shotgun metagenomic functional annotation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SRSupplementaryinformation.pdf