

# Predicting Humification Degree of Organic Solid Waste During Aerobic Fermentation Using Designated Bacterial Community

**Hong-Yu Yang**

Northeast Agricultural University

**Shu-Bo Zhang**

Northeast Agricultural University

**Yue Zhao**

Northeast Agricultural University

**Xin-Yu Xie**

Northeast Agricultural University

**Zi-min Wei** (✉ [weizimin@neau.edu.cn](mailto:weizimin@neau.edu.cn))

Northeast Agricultural University <https://orcid.org/0000-0002-9511-6692>

**Xu Zhang**

Northeast Agricultural University

**Xiao-Meng Chen**

Northeast Agricultural University

**Wu-Jun Qiu**

Northeast Agricultural University

---

## Research

**Keywords:** Bacterial community, Biological indicator, Humification degree, Aerobic fermentation, Organic solid waste

**Posted Date:** April 5th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-328051/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Microbe is the driver for disposing of organic solid waste (OSW) during aerobic fermentation. Notwithstanding, the significance of microbe is always underestimated in numerous preceding studies about the aerobic fermentation products assessment.

**Results:** Here, we investigate the humification degree (HD) and humic acid content could be assessed in terms of the bacterial community. Microbial sequencing and bioinformatics approaches are combined to analyze the biological characteristics of 105 aerobic fermentation samples collected. The bacterial communities could make predictions, which even correctly determines the categories of OSWs with 94% accuracy. Furthermore, the bacterial genera screened are designated as the bacterial code, which is substituted into the Random forest model to predict HD. And the bacteria codes can also provide a better prediction of the HD.

**Conclusions:** Our result suggested that bacteria codes a reliable biological method is proposed to assess HD effectively. Our experiment not only proves that the aerobic fermentation can be revealed by biological means. The bacterial codes can also be used as an ecological and biological indicator to evaluate the quality of the aerobic fermentation of different materials. This study further provides new insights on the use of microbe to evaluate the content of various substances during the degradation process of OSWs.

## Introduction

Aerobic fermentation is a significant measure to achieve organic solid waste (OSW) resources, reduction, and stabilization[1]. The essence of aerobic fermentation is a fast process of humification. In this process, the large quantities of humus formation is dependent on the activity of microbes[2]. Microbial metabolism plays a major role in the high-efficiency, rapid humification process[2, 3]. The quantity of humus and the degree of humification(HD) affect the stability of the aerobic fermentation and determine the quality of the final product[4]. Humus substance have crucial effects on soil fertility, soil remediation, soil productivity, and soil biological characteristics[5, 6]. Therefore, HD and humus content are essential indicators for evaluating the quality of humification products.

HD evaluation is a comprehensive indicator to determine the final quality of humic acid products, and a key evaluation indicator to influence the safe application of humic acid products. Various non-biological indicators for the traditional evaluation of HD, such as physical indicators, chemical indicators, spectroscopy indicators, etc[7, 8]. However, physical and chemical indicators are limited in measure HD ,and the spectral analysis method can not provide comprehensive information, so they are not enough to quickly and accurately determine the stability of organic matter[8]. Therefore, it is crucial to find a biological index to evaluate humification products efficiently. In addition, previous studies used to explore the response relationship between microorganisms and humic acid in aerobic fermentation[9, 10]. Most of them only reveal the response mechanism of a single material, rather than predicting and evaluating the HD of different aerobic fermentation materials based on microbial community data[9, 11].

Moreover, with the development of scientific research technology in recent years, high-throughput sequencing technology has been applied more widely in aerobic fermentation research[12]. With sequencing technology development, a mushrooming number of statistical methods based on bacterial communities have emerged, especially machine learning[13]. The random forest has been verified as a better machine learning model to observe bacteria's response to environmental factors and offers a straightforward and well-documented approach for creating predictive tools[14]. It has become the current research trend. Importantly, there is ample evidence that bacterial communities directly or indirectly respond to changes in the environment<sup>[15]</sup>. The bacterial data did not always classify a site as outside of target ranges when the chemical data would indicate that the chemistry changes are not always affecting the biological communities in the same way.

This study investigates the bioindicator ability of bacteria during the aerobic fermentation by high throughput sequencing for five kinds of OSW(chicken manure(CM), cow dung(CD), sewage sludge(SS), garden waste(GW) and rice straw(RS)). Bacteria diversity analysis suggests that there are similarities in bacterial communities during the aerobic fermentation, which we call consistent bacteria. Consistent bacteria and the key divergent bacteria obtained by statistical screening constitute microbe indicators that can characterize the aerobic fermentation of different OSWs. We use the random forest to prove a significant relationship between the bacterial community and the types of OSW in the aerobic fermentation. And predict the collected humification indicators. Our study establishes bacteria codes that have a better performance in predicting humification and proposes a bioindicator to evaluate HD.

## Results

### Analysis of bacterial diversity.

We compared the richness and diversity of the bacteria from different OSWs during aerobic fermentation. Four kinds of alpha diversity were analyzed with Kruskal- Wallis test (Fig.1a). Shannon index and Simpson index of bacteria were significantly different between these OSWs. The alpha diversity of RS significantly differed from the CM and SS, and the alpha diversity in CD and SS were very notable differences. The ACE index and Chao1 index analysis results showed that the richness of bacteria in CD is significantly different. In contrast, the alpha diversity of bacteria in GW relatively are similar to other OSWs.

Non-metric multidimensional scaling (NMDS) showed that the differences of SS were smaller than other OSWs, whereas the differences of CD were larger. There were significant differences ( $p < 0.001$ ) in the bacterial communities between OSWs (Supplementary Table.1, 2, Fig.1b). It was also illustrated a effect on the microbial community structure during aerobic fermentation due to OSW's heterogeneity. What is more, bacterial community structure has more significant differences between SS and the other OSWs, with  $R^2$  (0.244 -0.387). In contrast, the community structure was relatively similar between GW and other OSWs. Some similarities can still be revealed various OSWs at locations of overlapping projections in NMDS (Figure 1.b).

Our study indicated that the diversity and community structure of bacteria was significantly different for various OSWs. However, NMDS analysis also showed that regardless of the plane, the projections of the bacterial communities have a certain overlap. Therefore, if the purpose of evaluating aerobic fermentation products could be accomplished, both divergent bacteria and consistent bacteria should be considered during the aerobic fermentation of all OSWs.

### **Consistent bacteria.**

To explore the Consistent bacteria in all OSWs, OTUs were classified into 988 genera. Proteobacteria, Bacteroidetes, Firmicutes, and Actinobacteria were found (Supplementary Fig.1 a, c, e, g, i) with high relative abundance in all OSWs at the phylum. The samples of all OSWs were classified similarly at the phylum level. However, the relative abundance of phylum level varied greatly.

Both the aerobic fermentation of RS and CD, the relative abundance of dominant genera (except Others) was around 20%. Among the aerobic fermentation process of SS, the diversity of bacteria was lower. Moreover, the relative abundance of microorganisms in the same sample changed little, and the organisms are relatively stable at the genus level (Supplementary Fig.1, b, d, f, h, j). Ten genera of bacteria were present in 70% of the samples with high relative abundance (Fig.2). We named them as consistent bacteria (CB).

### **Divergent bacteria.**

To identify the divergent bacteria, some bacteria genera were selected with LEfSe statistical analysis (Fig.3, Supplementary Tables 3-7). We summarize the bacteria belonging to the different groups in various OSWs (Supplementary Figure 2). Too many biomarkers obtained from the LEfSe analysis were not conducive to subsequent research. Thereby, we conducted a further classification for divergent bacteria so that we could find the critical divergent bacteria in various bodies of OSWs.

### **Divergent hub bacteria.**

To determine which genus significantly interacts with other bacteria, we construct five networks for bacterial communities of different OSW respectively (Fig. 4). Overall, positive correlations were more remarkable than negative correlations among all OSWs (Table 1). Compared to other materials, the network density of CD was 0.169 with a higher number of nodes and edges (184 and 2852, respectively), indicating that the bacterial communities of CD were the most complex and structurally stable. The nodes' number of SS was only 47, with fewer species at phylum than others (Fig.4c). Still, its network density was 0.101 and close to GW (0.105), larger than RS and CM. These results indicate that SS's bacterial community composition was more straightforward than others. But the bacteria community was more tightly linked, which was consistent with the NMDS analysis. Although the number of RS nodes was 99, the lowest network density was only 0.054, implying that RS's bacterial community was looser.

The hub bacteria were selected based on the top 10% of degree ranking at the genus level for various OSWs. The hub bacteria may have an enormous impact on bacterial community composition. For

example, *Haliangium* was ranked in the top 10% of degree in CD, considered one of the hub bacteria in the CD bacterial community. The hub bacteria were combined with the LEfse analysis to simplify divergent hub bacteria (DHB) (Fig.5c). Moreover, DHB differentiate significantly between the various OSWs, suggesting that they may influence bacterial community structure differences.

### **Divergent functional bacteria.**

FAPROTAX was applied to predict the functions of bacterial communities. We found that the bacterial communities in every OSWs covered roughly the same functions with 30 basic functions. Nevertheless, there are also slight differences between OSWs. Bacterial communities in GW cover more functions, followed by RS, CD, CM, and SS, respectively (Supplementary Fig.3). We observed that the aerobic fermentation process mainly contained three major functional groups: C metabolism (including lignin decomposition, cellulose decomposition, etc.), and N metabolism (including nitrification, denitrification, etc.), and fermentation. Therefore, the FAPROTAX results are simplified into three aspects. (Supplementary Table 8).

Our results indicated that the bacteria community also demonstrate functional diversity (Fig.5a). The CM and RS emphasized fermentation more than carbon and nitrogen metabolism. In contrast, the bacterial community of CD invests more in carbon and nitrogen metabolism capacity. For SS, carbon metabolism capacity was stronger than aerobic fermentation and nitrogen metabolism, and nitrogen metabolism was weaker. Moreover, GW was more balanced overall, with carbon and nitrogen metabolism are roughly the same, and fermentation slightly higher than carbon and nitrogen metabolism.

The results of FAPROTAX analysis were combined with the results of LEfse analysis to screen out the divergent functional bacteria (DFB) for each of OSWs. These genera may contribute to the significant differences among OSWs bacterial communities' function.

### **Constructing bacterial codes.**

Firstly, The DFB and DHB were united as key divergent bacteria (KDB). Then, KDB and CB were combined form bacteria codes for various OSWs (Fig. 6). Bacterial codes contain the CB that are stably present during aerobic fermentation, the DHB that occupy critical positions in the community structure, and the DFB that are functionally different for different OSWs. That is, we obtained all bacteria capable of characterizing the overall bacterial profile of aerobic fermentation.

**Random forest projections.** After obtained bacterial codes of various OSWs in this experiment, the Random forest model is applied to verify the feasibility of bacterial codes. All microbial OTUs at the genus level were input into the random forest model (Supplement Fig.4). The results showed that the bacterial communities could correctly determine the categories of OSWs with more than 94% accuracy. It is also indicated that the bacteria community can be used to predict and assess aerobic fermentation process with random forest model.

The bacterial codes were input into the random forest model to predict some fundamental humification indicators (Fig.7). Furthermore, the standard regression curves were plotted based on the prediction results. And when the slope is closer to 1, the prediction results are more accurate. Among the four model scenarios that fit the linear regression analysis, the prediction of humification indicators slope between 0.46-0.78. the prediction result is relatively assuring. Except for the samples of CM, the prediction accuracy of HA prediction results is similar, and the slopes are all about 0.7 (Fig.6a). The slopes are all greater than 0.6. the slope value of the CD is even 0.78, which is the highest value among all FA prediction results (Fig.6b). For the prediction of the humification index (Fig.6c), it is suitable for CD. On the contrary, the CM humification index prediction result is the minimum of all the predicted slope values. (Fig.6d). The accuracy of the model HD prediction in the microbe varied across genus taxa between the five OSW (Supplement Fig.5). Though, all the results suggested bacterial codes can response to the HD of aerobic fermentation from the prediction results. It is also indicted that the potential ability of bacteria codes to predict aerobic fermentation's HD.

## Discussion

Determining and detecting the HD of products has been a major concern for researchers in the aerobic fermentation[16-18]. In general, researchers use fundamental physicochemical indicators to determine the HD of aerobic fermentation's products[19]. Hence, the introduction of biological indicators is indispensable for the evaluation of HD of fermented products. Here, bacterial codes are developed for bacterial communities in different materials. Bacterial codes can evaluate the HD of different materials more accurately in this study. Moreover, a random forest model is used to predict the quality of the aerobic fermentation's products. Collectively, the whole bacterial code recognition and prediction system can be used as a feasible biological tool to evaluate the quality of fermentation of different materials and control fermentation cycling. Moreover, this system can also be used for other biological processes related to microbial reaction processes. For example, it can evaluate the hazard level of pollutants in soil or water and assess the risk of intestinal microbial diseases.

We introduce the microbial community as an unknown parameter to investigate the HD. To fully characterize the microbial community during the aerobic fermentation, the functions and composition sequences of different microbes in OSW from various sources were summarized in a multi-dimensional analysis. We suppose that attention should be paid to the similarities of bacterial communities in experiments. Consistent bacteria are a specific group of bacteria in aerobic fermentation. Although the abundance is different, bacteria preserved under different nutrition and physical and chemical factor may have better characteristics under the specific aerobic fermentation conditions. Common microbes may also be the most stable community in response to environmental perturbations. The predominance of aerobic thermophiles and thermophilic bacteria (e.g., *Truepera*) is probably due to the aerobic fermentation overall warming[20, 21]. Moreover, *Trueper*, *Bacillus*, and *Enterococcus* are associated with the decomposition of substances such as cellulose[22, 23]. Also, *Devosia* and *Flavobacterium* may be involved in functions related to nitrogen metabolism during the aerobic fermentation[24]. This means that these genera are involved in the aerobic fermentation of most samples. Therefore, to predict the HD of

samples more accurately, attention must be paid to these bacteria. They have great potential in monitoring the aerobic fermentation and humification quality of different OSWs.

Also, divergent bacteria populations can be classified into multiple categories from multiple perspectives. First, core microbial taxa can be determined based on network analysis parameters (such as degree, betweenness, centrality, etc.)[25]. Core groups of microbes might be the dominant microbes in higher abundances, or they might be the prominent conductors in a microbial network[26]. Conductors can direct minor microbes to carry out a series of life activities. These conductors may not be abundant, but the metabolites they secrete are the nutritional basis for the life activities of other microbes[27]. Secondly, researchers select specified microbes to study specified metabolic pathways and microbial functions, such as N and C metabolism[28]. Microbes are classified according to their differences in function to investigate the metabolic potential of different taxonomic groups on specific substrates or specific life activities[29]. At present, microflora, microfluidics, and flow cytometry can be used to classify or separate microbes with different functions[30, 31]. Third, to better describe and predict the evolution of the whole structure and function of the community. Researchers are now making a variety of predictions using mathematical models and statistical methods. Data from random forests show that bacterial communities are instrumental in classifying and predicting the HD parameters. They compare the predicted value of the random forest time series of bacterial codes with the actual value of the decay index, as shown below. (Supplement Fig.6). Except that the predicted value of FA of bacterial codes is different from the actual value, the predicted values of other decay values are closer to the actual value, which indicates the feasibility of characterizing the decay degree by bacterial codes.

Our experiments have verified the bacterial community as a biological indicator to evaluate the quality of aerobic fermentation better. However, it still needs many experiments to prove it in practical application. First, the raw materials for the aerobic fermentation are collected around Harbin, which may be limited by region. Also, the types of materials selected are limited, and many follow-up experiments are needed to verify the experimental results. In the future, the universality of this method can be verified by multi-point and multi-material joint tests. The results verified by multiple experiments can be used to build accurate mathematical models, and the humic acid process evaluation database can be constructed. Second, since the composition of the microbial community was characterized only by 16S v3-v4 amplicon sequencing, only the relationship between the bacterial community composition and HD was considered during the fermentation process. Also, the relationship between fungi, archaea, or protozoa and HD was not discussed. Furthermore, with modern life science and technology development, multivariate analysis can be applied in this study, such as quantitative proteomic analysis, metabolomics analysis, microfluidic technology, etc. Also, the quality of aerobic fermentation's products is limited only by humic acid concentration. External factors, such as different materials and environmental factors, can also affect the degree of humification. It is worth noting that microbes can characterize these external factors. That is to say, the use of biological indicators to characterize environmental factors has its unique advantages compared to traditional methods. Our experiment not only proves that the aerobic fermentation can be revealed by biological means. The bacterial codes can also be used as an ecological and biological indicator to evaluate the quality of the aerobic fermentation among different materials.

# Conclusions

Aerobic fermentation is considered an essential component of sustainable development as it can reduce environmental problems caused by organic wastes. Microbes are the engine of aerobic fermentation, given the importance of bacterial communities to ensure harmless products after aerobic fermentation, it is time that monitoring efforts better account for changes by biological method, rather than relying on abiotic changes to determine the quality of aerobic fermentation. We evaluated bacterial communities of 5 types of OSWs during aerobic fermentation and the HD of host materials using 105 samples. The research analyzed the bacterial community from multiple dimensions and constructed bacterial code that can better predict HD. A greater use of the bacterial communities as indicators in degradation of OSWs will not only improve our ability to manage aerobic fermentation but also further provides new insights on the use of microbe to evaluate the content of various substances during the degradation process of OSWs.

# Methods

## The samples of humification.

A total of 105 humus samples of five OSWs were obtained from the Laboratory of Environmental Microbiology of Northeast Agricultural University, China, in this study. The CM and the CD were collected from the College of Animal Science, Northeast Agricultural University, SS from Harbin sewage Treatment plant, dried, and processed into grain sizes ranging from 10mm-20mm. The RS was taken from Xiangyang farm, and the GW was taken from the mixed forest of Northeast Agricultural University and cut into fragments of about 20mm. These materials were aerobically fermented in a separate reactor[32]. For the sufficient humification, the solid organic waste's initial moisture content was maintained in the range of 60–65%, the C/N ratio was 25:1, and the reactor's ventilation rate was 0.5 L/min[33]. The piles were turned over each sampling time to ensure a steady supply of oxygen. Some samples were frozen at -20°C for microbiological analysis, while others were air-dried for the determination of humic acids(HA) and fulvic acids(FA) organic matter(OM), and total organic carbon(TOC). TOC was estimated using the potassium dichromate volumetric method.

## Humus analysis.

Dry samples(10g) were evenly mingled by a mixed solution of 0.1M  $\text{Na}_4\text{P}_2\text{O}_7 \cdot 10\text{H}_2\text{O}$  and NaOH at the ratio of 1:10(W: V), the mixture was shaken at room temperature for 24h, centrifuge the suspension at 10,000 RPM at 4°C for 20 minutes. Thereafter, the obtained supernatant was filtered through a 0.45µm Millipore membrane, the limpid liquid obtained is the humus solution. Adjusting the pH of the solution to 1 by 6M HCL and standing at 4°C for 12h. Centrifuge the solution at 10,000 RPM for 10 min. The FA solution was obtained by purifying the supernatant with XAD-7 exchange resin. After repeated washing with 0.1mol/L HCL solution three times, the precipitation was dissolved in 0.05mol/L  $\text{NaHCO}_3$  to obtain HA solution. All samples were performed in triplicate and calculated the mean to reduce the error.

SHIMADZU TOC-Vcph analyzer was applied to detect the TOC, HA, and FA[34]. The humification indices were calculated according to the following formula[35].

Humification Index(HI)=  $C_{HA}/TOC \times 100$ ;

Degree of Polymerization (DP)= $C_{HA}/C_{FA}$ .

### **High-throughput sequencing of 16S rRNA.**

Sample DNA was extracted with a DNA kit (Omega Biotek, Inc.). Using a NanoDrop Spectrophotometer(Gene Co. Ltd, shanghai,China) to evaluate DNA quantity and quality To investigate the bacterial structure of various organic solid wastes, the 16S primers 341F(CCTAYGGGRBGCASCAG) and 806R (GGACTACNNGGTATCTAAT) were used to amplify the V3-V4 regions of bacterial 16S rRNA gene. The sample is finally sequenced on the Illumina HiSeq platform (Novogene Bioinformatics Technology Inc, Beijing, China)[36].

### **Network analysis.**

To determine the relationship between the bacteria of different solid organic wastes, using network analysis to describe the underlying co-occurrences among bacteria[37, 38]. Network analysis is performed at the genus level and finally exhibited at the phylum level. Data filtering was performed before network construction with the 'Hmisc' package of the R software. The genus whose relative abundance sum is less than 0.005 is removed to reduce the computational complexity and ensure the accuracy of the calculation, avoid the error calculation caused by the null value[39]. Spearman correlation was used to determine all the pairwise associations[40].

To obtain the genus with a higher correlation coefficient, the genus with a Spearman coefficient lower ( $r < 0.8$ ) was eliminated. The less significant correlations were dislodged by R software, and only significant correlations ( $P < 0.05$ ) were retained for the downstream procedure[41]. The Benjamini-Hochber (BH) multiple test was used to correct the  $P$ -value, adjusted the false discovery rate (FDR), and the chance of false rejecting the null hypothesis is not more than 0.05[42]. Imported the correlation results into the Gephi (0.9.2) and visualized them with Frucherman Reingold algorithms. The network density, clustering coefficient, and other topological parameters of the network and the degree of each node in the network were calculated by Gephi. The clustering coefficient was used to reflect the degree of embedding between a node and its neighbors, that is, the degree of clustering between nodes. The network density represents the degree of aggregation and the degree of connection tightness of nodes. The degree was used to search the key hubs in the network[43, 44]. A node with a higher degree indicates that there are more nodes connected to it in the network. When the node changes, it will affect more nodes, making the network structure more likely to change, and it is also more likely to become a key hub in the network. Key hub.

### **Function prediction of bacteria.**

Functional annotation of prokaryotic taxa (FAPROTAX) was used to evaluate bacterial communities' potential functions and divide OTUs into one or more functional groups based on existing literature that contains currently culturable bacteria [45]. FAPROTAX contains more than 7600 functional annotations from more than 80 functional groups (e.g., xylanolysis, nitrification, methanol oxidation). Variations of specific strains within species non-cultivable species were ignored, for example, horizontal gene transfer led to the differentiation between strains. This may be closely related to different carbon compounds degradation, and this is why some carbon metabolism pathways were not divided (similarly for fermentation)[44]. However, other functions were more conservative, such as which were associated with nitrification, denitrification, or sulfate respiration[46]. The method is run through matching the annotation information of the OTU based on 16S rRNA sequencing with the information database through the python program and output the predicted result of the function (<http://www.zoology.ubc.ca/louca/FAPROTAX/lib/php/index.php?p=Instructions>).

### **Prediction of humification indices.**

Random forest was used to predict the humification indices in aerobic fermentation and evaluated the maturity quality of different organic solid wastes. Random forests have been confirmed by previous studies to be superior to other modeling methods in environmental bacterial communities[47]. Across all the samples contained a total of 988 genera after rarefying. To reduce the number of explanatory variables entering each random forest model, the two key categories of genera (C and KD) by multiple methods in this study were finally screened which best represented various samples. Random forest was performed by 'randomForest' package with default parameters.[48] The linear regression models were used to evaluate the accuracy of random forest prediction for the maturity of the material. The closer  $R^2$  is to 1, the higher the fit of the predicted value. The closer the slope of the equation is to 1, the closer the predicted value is to the actual value.

### **Statistics analysis.**

All statistical analyses and data visualizations in the R platform (4.0.2). Employ Canoco 5 to perform NMDS analysis basic on OTUs data of different samples and visualize them by Originpro 2019. The alpha diversity of different materials in the aerobic composting process was visualized by the 'vioplot' package in the R. Kruskal-Wallis tests were used to measure whether there are differences in alpha diversity among multiple materials[49]. Permutation multivariate analysis of variance (PERMANOVA) was calculated by the 'vegan' package of the R software. PERMANOVA was used to determine differences in bacterial communities among different organic solid wastes.[50] Bray-Curtis distance was utilized by PERMANOVA to calculate the difference in bacterial communities, whether it is the overall level or pairwise comparison. Linear discriminant analysis effect size (LEfSe)[51] was used to explore the biomarkers (according to the biosystematics, it is divided into 6 categories from the kingdom level to the genus level) of five different organic solid wastes, that is, the significant divergent bacteria between the material and others based on  $P < 0.05$  and  $LDA > 2.0$ [48, 52].

# Abbreviations

OSW: organic solid waste; HD: humification degree; SS: sewage sludge; CD: cow dung; CM: chicken manure; GW: garden waste; RS: rice straw.

# Declarations

## Acknowledgements

We wish to thank Hao Jingkun for providing chicken manure, Zhu Longji for providing cow dung, Wang Liqin for providing garden waste, Shi Mingzi for providing sewage sludge, and Wu Di for providing rice straw.

## Authors' contributions

XMC and JQW facilitated the sample collection. HYY, ZMW, XYX, and XZ designed the research. HYY and SBZ performed the research and conducted the data analyses with assistance from YZ and XYX. HYY and SBZ drafted the manuscript; all authors reviewed and edited the manuscript.

## Funding

This work was financially supported by the National Natural Science Foundation of China (No. 51778132) and Development Project (No. 2019YFC1906400).

## Availability of data and materials

All raw sequences derived from this experiment were submitted into the Short Read Archive of NCBI and can be found under the BioProject accession number PRJNA710204.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

# References

1. Shi MZ, Zhao XY, Zhu LJ, et al. Elucidating the negative effect of denitrification on aromatic humic substance formation during sludge aerobic fermentation. *J Hazard Mater.* 2020;388;

- <https://doi.org/10.1016/j.jhazmat.2020.122086>.
2. Lu LH, Zeng GM, Fan CZ, et al. Characterization of a laccase-like multicopper oxidase from newly isolated *Streptomyces* sp C1 in agricultural waste compost and enzymatic decolorization of azo dyes. *Biochem Eng J.* 2013;72:70-76; <https://doi.org/10.1016/j.bej.2013.01.004>.
  3. Chen M, Xu P, Zeng GM, et al. Bioremediation of soils contaminated with polycyclic aromatic hydrocarbons, petroleum, pesticides, chlorophenols and heavy metals by composting: Applications, microbes and future research needs. *Biotechnol Adv.* 2015;33:745-55; <https://doi.org/10.1016/j.biotechadv.2015.05.003>.
  4. Senesi N, Plaza C. Role of humification processes in recycling organic wastes of various nature and sources as soil amendments. *Clean-Soil Air Water.* 2007;35:26-41; <https://doi.org/10.1002/clen.200600018>.
  5. Pedra F, Polo A, Ribeiro A, et al. Effects of municipal solid waste compost and sewage sludge on mineralization of soil organic matter. *Soil Biology & Biochemistry.* 2007;39:1375-82; <https://doi.org/10.1016/j.soilbio.2006.12.014>.
  6. Palansooriya KN, Shaheen SM, Chen SS, et al. Soil amendments for immobilization of potentially toxic elements in contaminated soils: A critical review. *Environ Int.* 2020;134; <https://doi.org/10.1016/j.envint.2019.105046>.
  7. Ikeya K, Maie N, Han XZ, et al. Comparison of carbon skeletal structures in black humic acids from different soil origins. *Soil Sci Plant Nutr.* 2019;65:109-13; <https://doi.org/10.1080/00380768.2019.1584016>.
  8. Medina J, Monreal CM, Antilen M, et al. Influence of inorganic additives on wheat straw composting: Characterization and structural composition of organic matter derived from the process. *J Environ Manage.* 2020;260; <https://doi.org/10.1016/j.jenvman.2020.110137>.
  9. Wu D, Wei Z, Gao X, et al. Reconstruction of core microbes based on producing lignocellulolytic enzymes causing by bacterial inoculation during rice straw composting. *Bioresour Technol.* 2020;315:123849; <https://doi.org/10.1016/j.biortech.2020.123849>.
  10. Gong CM. Microbial safety control of compost material with cow dung by heat treatment. *J Environ Sci-China.* 2007;19:1014-19; [https://doi.org/10.1016/S1001-0742\(07\)60164-8](https://doi.org/10.1016/S1001-0742(07)60164-8).
  11. Zhu L, Zhao Y, Zhang W, et al. Roles of bacterial community in the transformation of organic nitrogen toward enhanced bioavailability during composting with different wastes. *Bioresour Technol.* 2019;285:121326; <https://doi.org/10.1016/j.biortech.2019.121326>.
  12. Van Rossum T, Ferretti P, Maistrenko OM, et al. Diversity within species: interpreting strains in microbiomes. *Nat Rev Microbiol.* 2020;18:491-506; <https://doi.org/10.1038/s41579-020-0368-1>.
  13. Ye L, Mei R, Liu WT, et al. Machine learning-aided analyses of thousands of draft genomes reveal specific features of activated sludge processes. *Microbiome.* 2020;8:16; <https://doi.org/10.1186/s40168-020-0794-3>.
  14. Kim H, Lee KK, Jeon J, et al. Domestication of *Oryza* species eco-evolutionarily shapes bacterial and fungal communities in rice seed. *Microbiome.* 2020;8:20; <https://doi.org/10.1186/s40168-020->

00805-0.

15. Yuan J, Wen T, Zhang H, et al. Predicting disease occurrence with high accuracy based on soil macroecological patterns of Fusarium wilt. *ISME J.* 2020;14:2936-50; <https://doi.org/10.1038/s41396-020-0720-5>.
16. Zhao Y, Wei YQ, Zhang Y, et al. Roles of composts in soil based on the assessment of humification degree of fulvic acids. *Ecol Indic.* 2017;72:473-80; <https://doi.org/10.1016/j.ecolind.2016.08.051>.
17. Guimaraes DV, Gonzaga MIS, da Silva TO, et al. Soil organic matter pools and carbon fractions in soil under different land uses. *Soil & Tillage Research.* 2013;126:177-82; <https://doi.org/10.1016/j.still.2012.07.010>.
18. Kawasaki S, Maie N, Watanabe A. Composition of humic acids with respect to the degree of humification in cultivated soils with and without manure application as assessed by fractional precipitation. *Soil Sci Plant Nutr.* 2008;54:57-61; <https://doi.org/10.1111/j.1747-0765.2007.00219.x>.
19. de Gannes V, Eudoxie G, Hickey WJ. Prokaryotic successions and diversity in composts as revealed by 454-pyrosequencing. *Bioresource Technology.* 2013;133:573-80; <https://doi.org/10.1016/j.biortech.2013.01.138>.
20. Krishnan Y, Bong CPC, Azman NF, et al. Co-composting of palm empty fruit bunch and palm oil mill effluent: Microbial diversity and potential mitigation of greenhouse gas emission. *J Clean Prod.* 2017;146:94-100.
21. Ki BM, Kim YM, Jeon JM, et al. Characterization of odor emissions and microbial community structure during degradation of pig carcasses using the soil burial-composting method. *Waste Manage.* 2018;77:30-42; <https://doi.org/10.1016/j.wasman.2018.04.043>.
22. Hemati A, Aliasgharzad N, Khakvar R, et al. Role of lignin and thermophilic lignocellulolytic bacteria in the evolution of humification indices and enzymatic activities during compost production. *Waste Manag.* 2021;119:122-34; <https://doi.org/10.1016/j.wasman.2020.09.042>.
23. Kong ZJ, Wang XQ, Wang MM, et al. Bacterial ecosystem functioning in organic matter biodegradation of different composting at the thermophilic phase. *Bioresource Technology.* 2020;317; <https://doi.org/10.1016/j.biortech.2020.123990>.
24. Zhang WM, Yu CX, Wang XJ, et al. Increased abundance of nitrogen transforming bacteria by higher C/N ratio reduces the total losses of N and C in chicken manure and corn stover mix composting. *Bioresource Technology.* 2020;297; <https://doi.org/10.1016/j.biortech.2019.122410>.
25. Zhou JZ, Deng Y, Luo F, et al. Phylogenetic Molecular Ecological Network of Soil Microbial Communities in Response to Elevated CO<sub>2</sub>. *Mbio.* 2011;2; <https://doi.org/10.1128/mBio.00122-11>.
26. Zhu LJ, Zhao Y, Zhang WS, et al. Roles of bacterial community in the transformation of organic nitrogen toward enhanced bioavailability during composting with different wastes. *Bioresource Technology.* 2019;285; <https://doi.org/10.1016/j.biortech.2019.121326>.
27. Fisher CK, Mehta P. Identifying Keystone Species in the Human Gut Microbiome from Metagenomic Timeseries Using Sparse Linear Regression. *Plos One.* 2014;9; <https://doi.org/10.1371/journal.pone.0102451>.

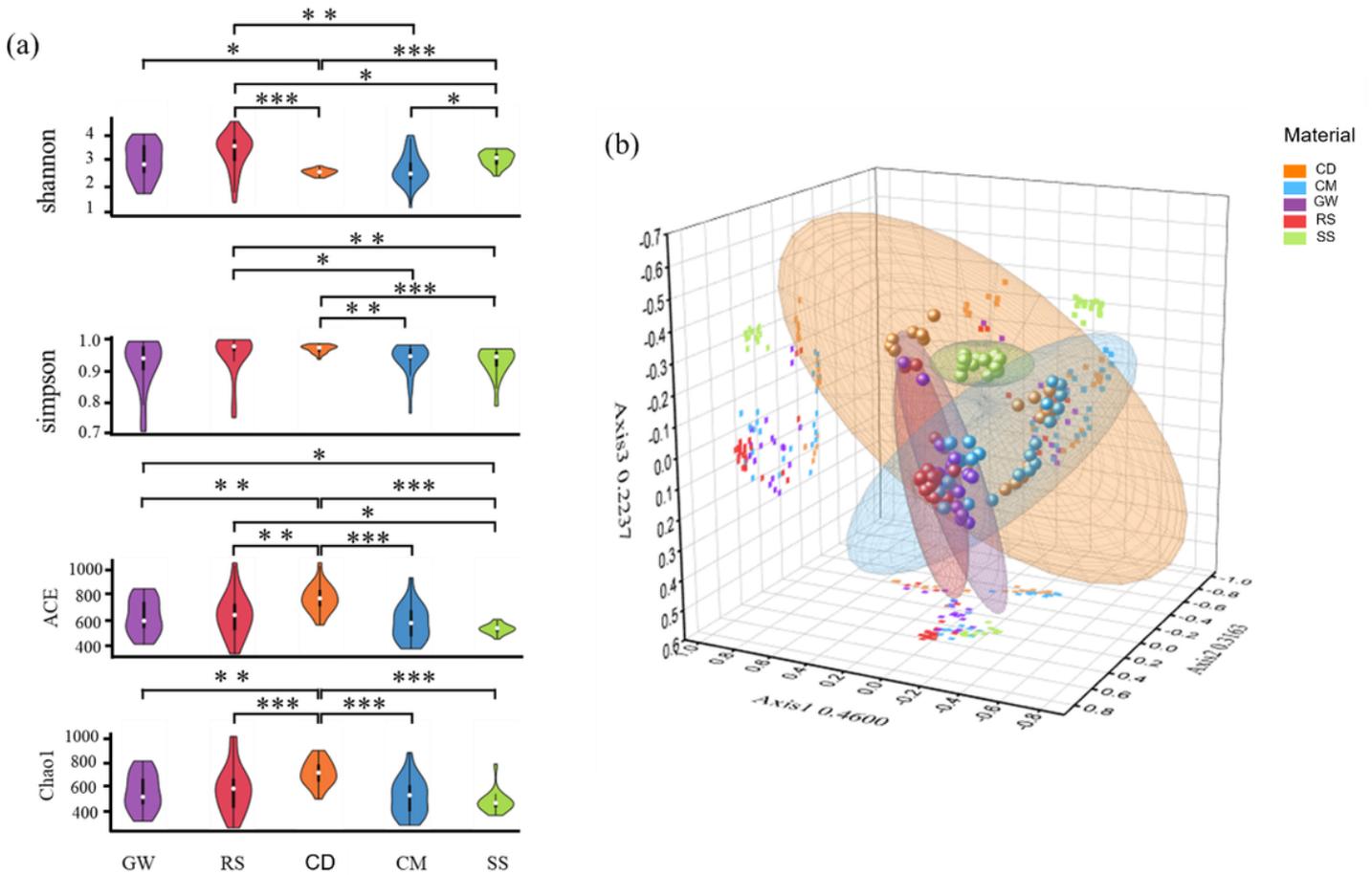
28. Dijkstra P, LaViolette CM, Coyle JS, et al. N-15 enrichment as an integrator of the effects of C and N on microbial metabolism and ecosystem function. *Ecol Lett*. 2008;11:389-97; <https://doi.org/10.1111/j.1461-0248.2008.01154>.
29. Qu ZL, Liu B, Ma Y, et al. Differences in bacterial community structure and potential functions among Eucalyptus plantations with different ages and species of trees. *Applied Soil Ecology*. 2020;149; <https://doi.org/10.1016/j.apsoil.2020.103515>.
30. Maurice CF, Haiser HJ, Turnbaugh PJ. Xenobiotics Shape the Physiology and Gene Expression of the Active Human Gut Microbiome. *Cell*. 2013;152:39-50; <https://doi.org/10.1016/j.cell.2012.10.052>.
31. Kou SZ, Cheng DH, Sun F, et al. Microfluidics and microbial engineering. *Lab Chip*. 2016;16:432-46; <https://doi.org/10.1039/c5lc01039j>.
32. Zhu LJ, Wei ZM, Yang TX, et al. Core microorganisms promote the transformation of DOM fractions with different molecular weights to improve the stability during composting. *Bioresource Technology*. 2020;299; <https://doi.org/10.1016/j.biortech.2019.122575>.
33. Cui HY, Zhao Y, Chen YN, et al. Assessment of phytotoxicity grade during composting based on EEM/PARAFAC combined with projection pursuit regression. *J Hazard Mater*. 2017;326:10-17; <https://doi.org/10.1016/j.jhazmat.2016.09.059>.
34. Wu JQ, Wei ZM, Zhu ZC, et al. Humus formation driven by ammonia-oxidizing bacteria during mixed materials composting. *Bioresource Technology*. 2020;311; <https://doi.org/10.1016/j.biortech.2020.123500>.
35. Baffi C, Dell'Abate MT, Nassisi A, et al. Determination of biological stability in compost: A comparison of methodologies. *Soil Biology & Biochemistry*. 2007;39:1284-93; <https://doi.org/10.1016/j.soilbio.2006.12.004>.
36. Bullerjahn GS, McKay RML, Bernat G, et al. Community dynamics and function of algae and bacteria during winter in central European great lakes. *J Great Lakes Res*. 2020;46:732-40; <https://doi.org/10.1016/j.jglr.2019.07.002>.
37. Cardona C, Weisenhorn P, Henry C, et al. Network-based metabolic analysis and microbial community modeling. *Curr Opin Microbiol*. 2016;31:124-31; <https://doi.org/10.1016/j.mib.2016.03.008>.
38. Rottjers L, Faust K. From hairballs to hypotheses-biological insights from microbial networks. *FEMS Microbiol Rev*. 2018;42:761-80; <https://doi.org/10.1093/femsre/fuy030>.
39. Barberan A, Bates ST, Casamayor EO, et al. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME Journal*. 2012;6:343-51; <https://doi.org/10.1038/ismej.2011.119>.
40. Brisson V, Schmidt J, Northen TR, et al. A New Method to Correct for Habitat Filtering in Microbial Correlation Networks. *Front Microbiol*. 2019;10:585; <https://doi.org/10.3389/fmicb.2019.00585>.
41. Soffer N, Zaneveld J, Vega Thurber R. Phage-bacteria network analysis and its implication for the understanding of coral disease. *Environ Microbiol*. 2015;17:1203-18; <https://doi.org/10.1111/1462-2920.12553>.
42. Faust K, Sathirapongsasuti JF, Izard J, et al. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol*. 2012;8:e1002606; <https://doi.org/10.1371/journal.pcbi.1002606>.

43. Sporns O, Honey CJ, Kotter R. Identification and classification of hubs in brain networks. *Plos One*. 2007;2:e1049; <https://doi.org/10.1371/journal.pone.0001049>.
44. Louca S, Parfrey LW, Doebeli M. Decoupling function and taxonomy in the global ocean microbiome. *Science*. 2016;353:1272-77; <https://doi.org/10.1126/science.aaf4507>.
45. Liang SC, Deng JJ, Jiang Y, et al. Functional Distribution of Bacterial Community under Different Land Use Patterns Based on FaProTax Function Prediction. *Pol J Environ Stud*. 2020;29:1245-61; <https://doi.org/10.15244/pjoes/108510>.
46. Ma J, Liu H, Zhang C, et al. Joint response of chemistry and functional microbial community to oxygenation of the reductive confined aquifer. *Science of the Total Environment*. 2020;720.
47. Smith MB, Rocha AM, Smillie CS, et al. Natural Bacterial Communities Serve as Quantitative Geochemical Biosensors. *Mbio*. 2015;6; <https://doi.org/10.1128/mBio.00326-15>.
48. Hermans SM, Buckley HL, Case BS, et al. Using soil bacterial communities to predict physico-chemical variables and soil quality. *Microbiome*. 2020;8; <https://doi.org/10.1186/s40168-020-00858-1>.
49. Coelho FF, Giasson E, Campos AR, et al. Digital soil class mapping in Brazil: a systematic review. *Sci Agr*. 2021;78; <https://doi.org/10.1590/1678-992X-2019-0227>.
50. Guo QX, Yan LJ, Korpelainen H, et al. Plant-plant interactions and N fertilization shape soil bacterial and fungal communities. *Soil Biology & Biochemistry*. 2019;128:127-38; <https://doi.org/10.1016/j.soilbio.2018.10.018>.
51. Chong J, Liu P, Zhou GY, et al. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nat Protoc*. 2020;15:799-821; <https://doi.org/10.1038/s41596-019-0264-1>.
52. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12; <https://doi.org/10.1186/gb-2011-12-6-r60>.

## Tables

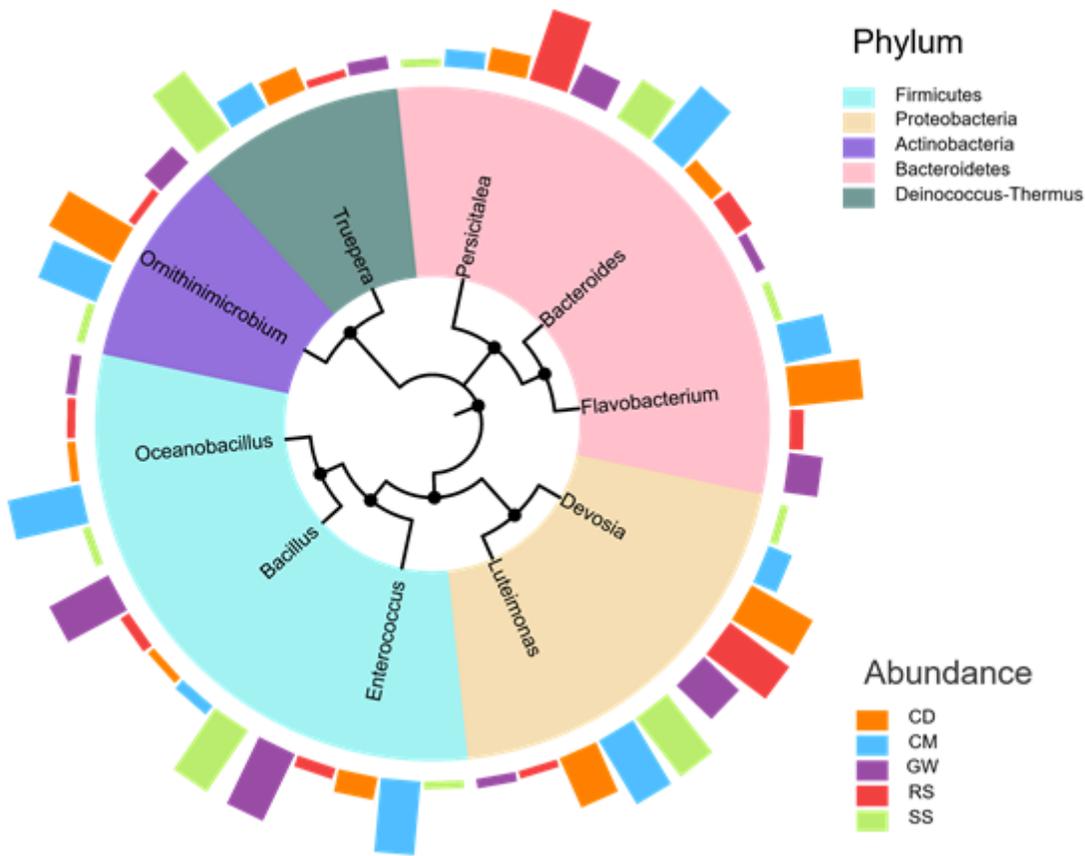
Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

## Figures



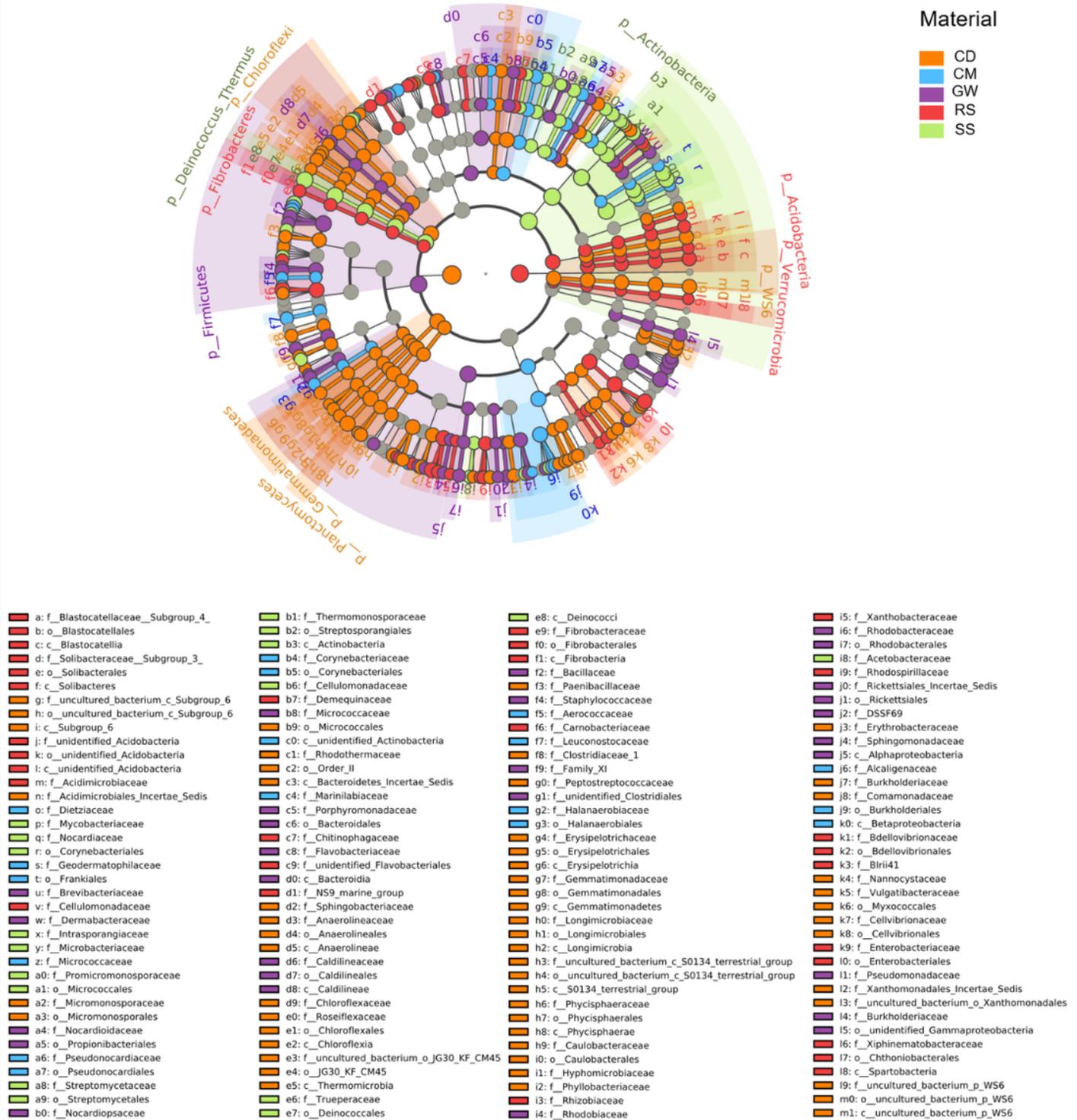
**Figure 1**

Alpha and beta diversity of bacteria in humification process based on high-throughput sequencing of 16S rRNA. The differences of bacterial diversity (Shannon index, Simpson index, ACE, and Chao1) displayed in violin plot were calculated by the Kruskal-Wallis test. a Where \* indicates the magnitude of differences, and there are significant differences (\*  $p < 0.05$ , \*\*  $p < 0.01$ ), highly significant differences (\*\*\*)  $p < 0.001$ . b shows NMDS analysis results for five aerobic fermentation samples of bacterial communities, with different color coverage indicating five other organic solid wastes in a three-dimensional scatter diagram.



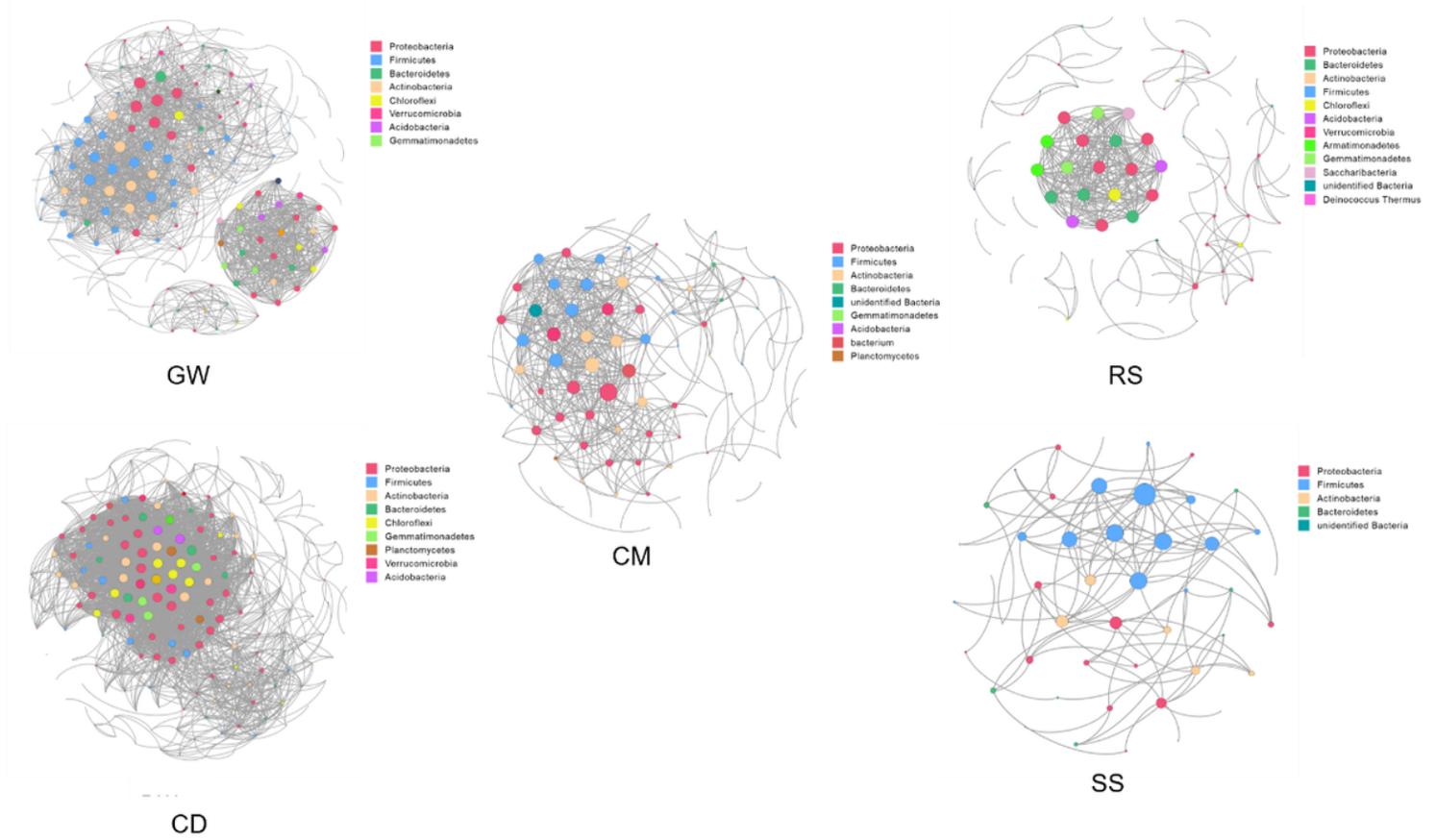
**Figure 2**

Phylogenetic tree of consistent bacteria. According to OTUs' abundance analysis, Consistent bacteria was selected, existed in no less than 70% samples of each aerobic fermentation material treatment. The histogram shows the proportion of this bacterial genus in the various organic solid waste after abundance normalization. The background color of the phylogenetic tree indicates to which phylum bacteria belong.



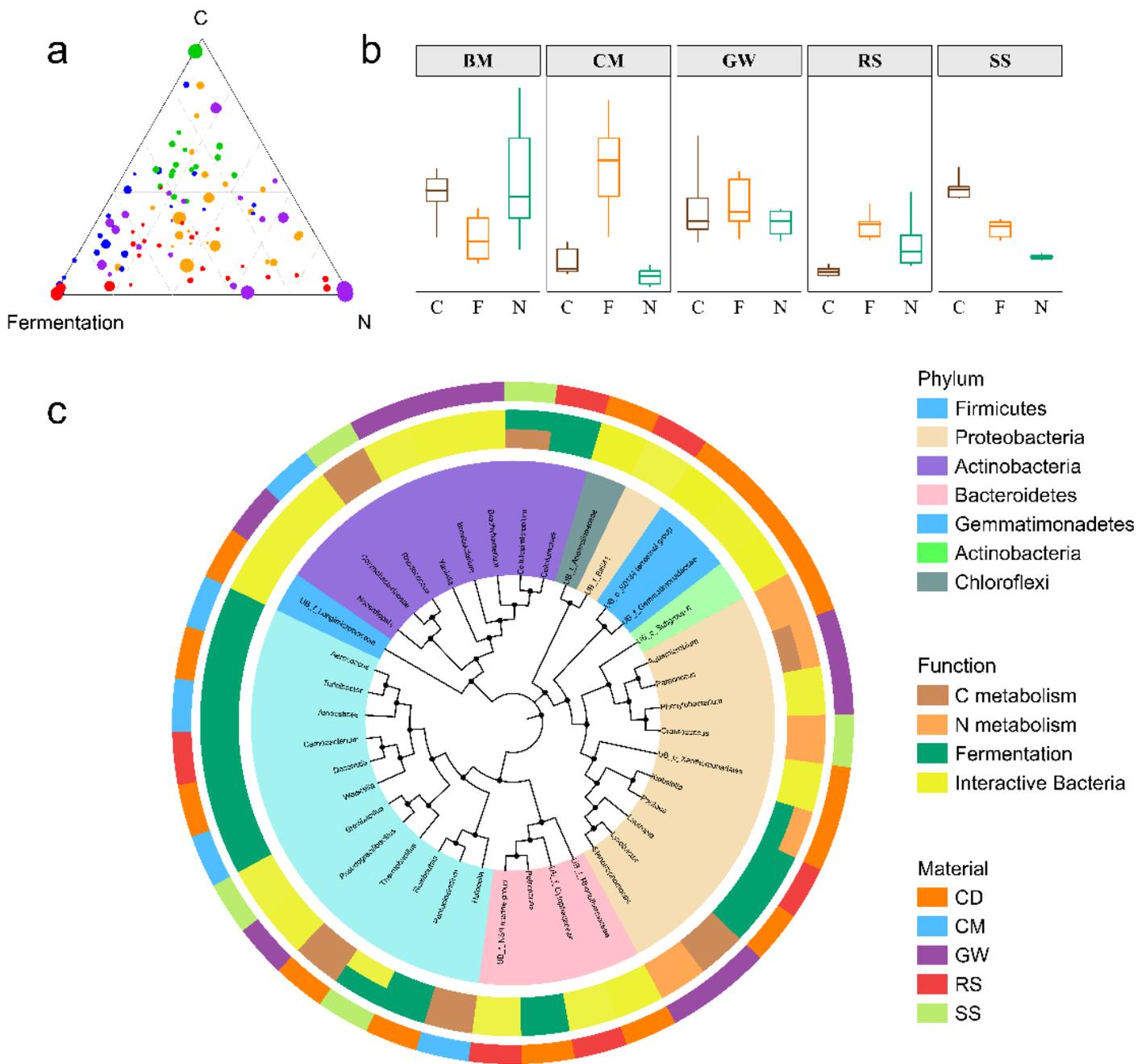
**Figure 3**

LEfSe results show that bacterial biomarkers are sensitive to different organic solid waste. The branching diagram has six layers of circles. Each circle ring represents from inside to outside, kingdom, phylum, class, order, family, and genus. A high abundance of species is labeled with color by color to distinguish statistical biomarker in various organic solid waste, where gray nodes represent taxonomically insignificant microorganisms.



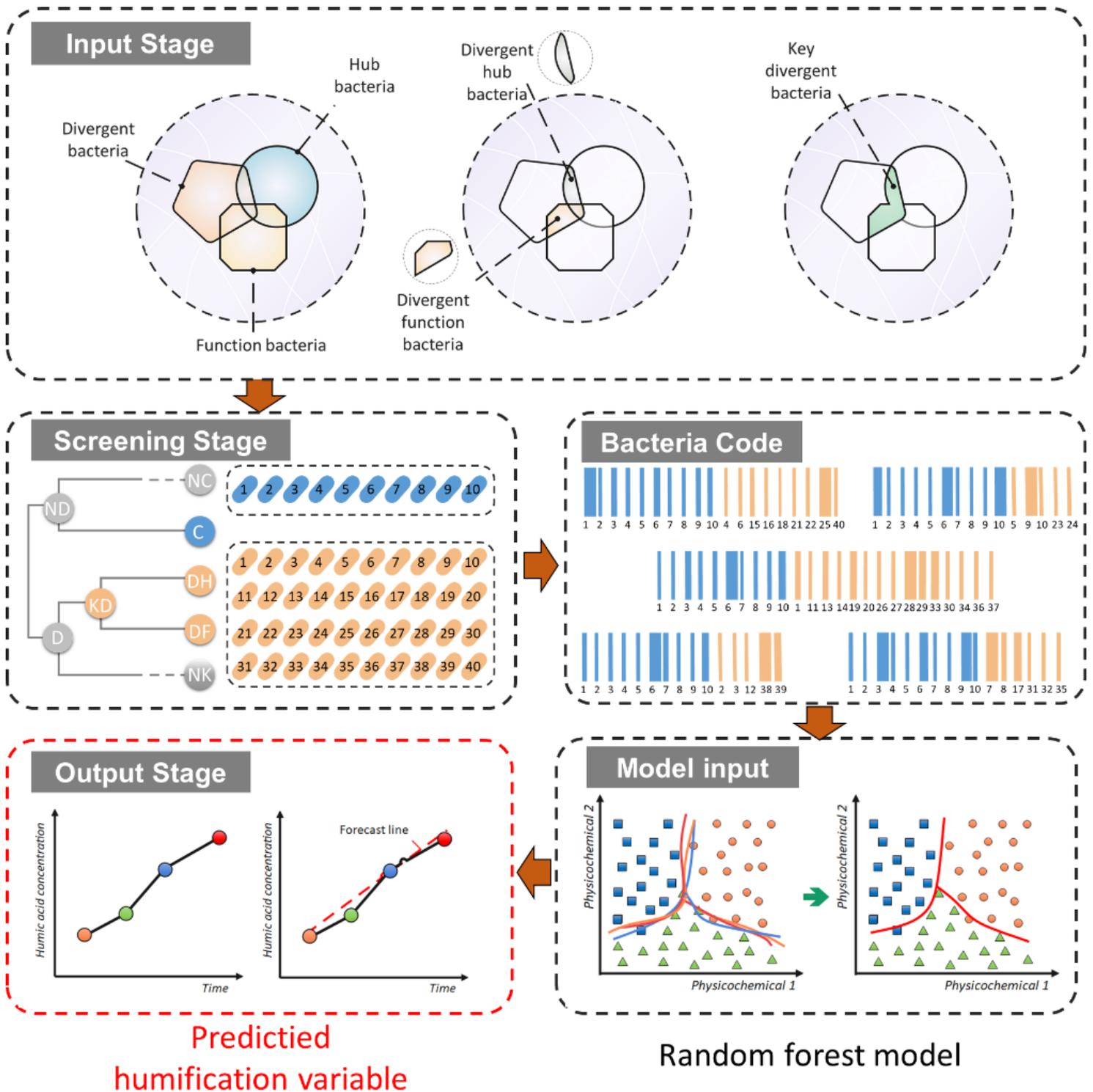
**Figure 4**

The network was constructed based on different organic solid wastes—including chicken manure [CM], cow dung [CD], garden waste [GW], rice straw [RS], and sewage sludge [SS]—on the co-occurrence between bacteria taxa at the genus level. The node size symbolizes the degree of each bacteria, and the node filled in various colors is phylum taxa.



**Figure 5**

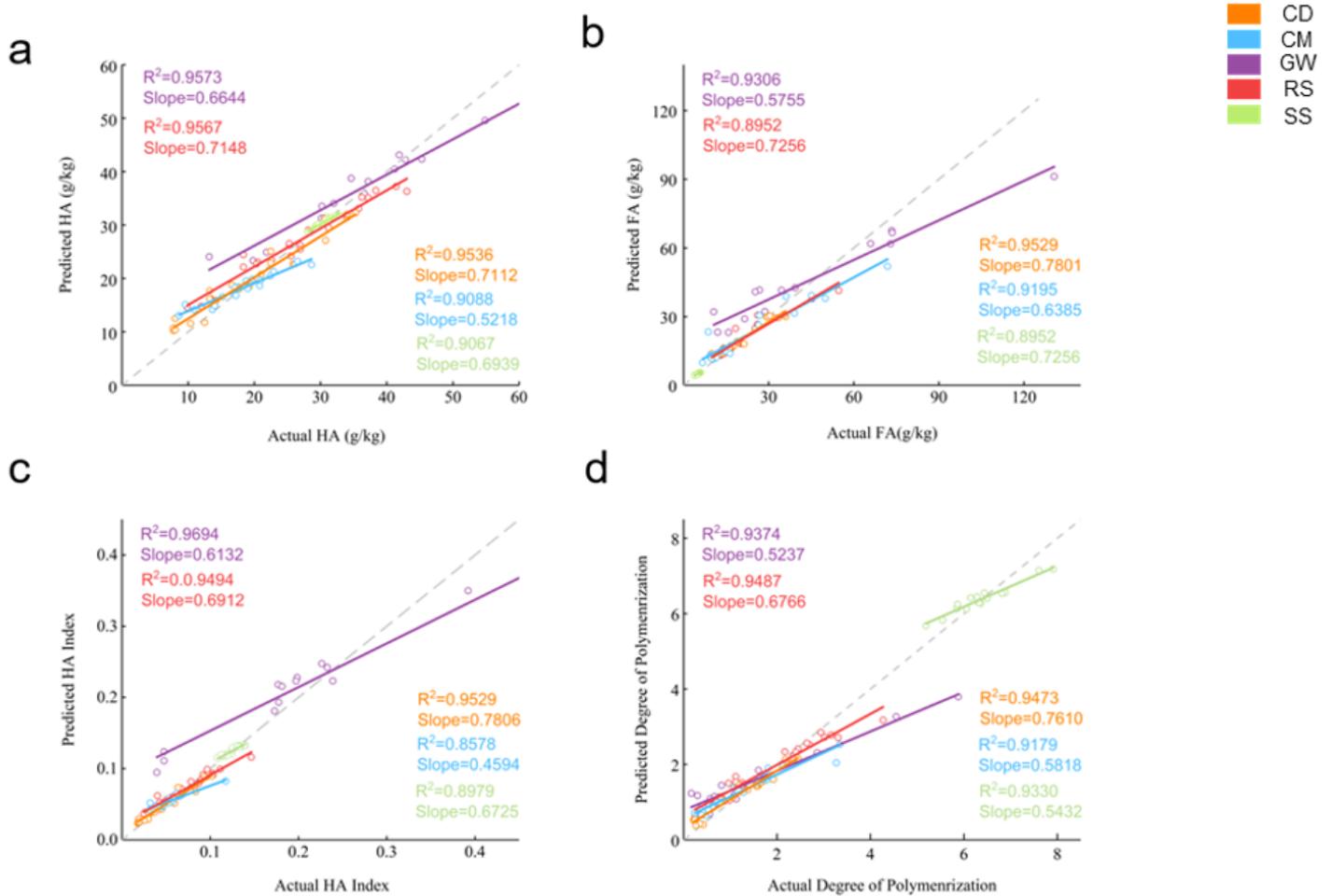
FAPROTAX functional prediction results and phylogenetic tree of essential differential microorganisms. a Ternary plots based on FAPROTAX functional prediction results different colors represent different organic solid waste samples, and node sizes indicate the relative abundance of functional groups in the samples. b Box plots of function prediction results, with box colors representing different functional groups (C-metabolism, N-metabolism, and F-fermentation). c Function prediction, network analysis, and LEfse analysis, with three circles outside the phylogenetic tree, the outermost circle represents the material to which the bacterial genus belongs, and the two inner circles were indicating the function of the bacterial genus or whether it is a crucial node microorganism. The color range of the evolutionary tree represents gate-level taxa.



**Figure 6**

Construction of bacterial codes for various organic solid waste. Input stage: bacterial community data ; Screening stage: non- Divergent bacteria (ND), non-Consistent bacteria (NC), Consistent bacteria (C), Divergent bacteria (D), Key divergent bacteria (KD); Divergent hub bacteria (DH), Divergent Function bacteria (DF), non-Key divergent bacteria (NK) ; Consistent bacteria (C) and key divergent (KD) make up the bacterial code. Bacteria code: the specific bar code of different OSW. Two types of bacteria were colored

with blue and orange. Each number has a corresponding bacteria. The width of each bar is normalized bacteria abundance in the bacteria code. Model input: The bacterial code is substituted into the random forest model to make a prediction Output stage: the value of predict humification.



**Figure 7**

Stochastic forest predictions of four humification indicators during aerobic fermentation. a-d represents the stochastic forest predictions of four humification indicators(HA,FA,HI,DP) during aerobic fermentation, respectively. The purple dots represent garden waste, the red dots represent rice straw, the orange dots represent cow manure, the blue dots represent chicken manure, and the green dots represent sludge. The grey dashed line represents the best-predicted case. The adjusted R<sup>2</sup> and slope are indicated in the figure

## Supplementary Files

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

- [Supplementmaterial.doc](#)
- [Table1.jpg](#)