



# Gradient Internal Standard Method for Absolute Quantification of Microbial Amplicon Sequencing Data

Shilei Wang,<sup>a</sup>  Qun Wu,<sup>a</sup> Ying Han,<sup>b</sup> Rubing Du,<sup>a</sup> Xiaoyong Wang,<sup>b</sup>  Yao Nie,<sup>a</sup> Xiaowei Du,<sup>b</sup> Yan Xu<sup>a</sup>

<sup>a</sup>State Key Laboratory of Food Science and Technology, Synergetic Innovation Center of Food Safety and Nutrition, School of Biotechnology, Jiangnan University, Wuxi, Jiangsu, China

<sup>b</sup>Technical Center, Xinghuacun Fenjiu Distillery Co. Ltd., Fenyang, Shanxi, China

**ABSTRACT** High-throughput amplicon sequencing is a critical tool for studying microbiota; however, it results only in relative abundance data. Thus, changes in absolute abundance of microbiota cannot be determined, which hinders further microbiology research. We have therefore established a gradient internal standard absolute quantification (GIS-AQ) method to overcome this issue, which can simultaneously obtain the absolute abundances of bacteria and fungi. Deviations from the quantitative equations of microbes and internal standards were eliminated through calibration. Compared with traditional quantitative real-time PCR and microscopy quantifications, this method is reliable ( $R^2_{\text{average}} = 0.998$ ;  $P < 0.001$ ) and accurate ( $P_{\text{internals versus microscopy}} > 0.05$ ). The GIS-AQ method can be adapted to any amplicon primer choice (e.g., 336F/806R and ITS3/ITS4), rendering it applicable to ecosystem studies including food, soil, and water samples. Crucially, when using solid-state fermentation samples from various temporal dimensions, the results obtained from the relative and absolute abundance are different. The absolute abundance can be used to study the difference in communities between different samples, and the GIS-AQ method allows this to be done rapidly. Therefore, combining the absolute abundance with relative abundance can accurately reflect the microbiota composition.

**IMPORTANCE** To solve the problem of amplicon sequencing cannot discern the microbiota absolute abundance, we proposed a gradient internal standard absolute quantification method. We used Chinese liquor fermentation as a model system to demonstrate the reliability and accuracy of the method. By comparing the relative and absolute abundances of microbiota in various temporal dimensions, we found dynamic changes in the absolute abundance of communities under various temporal dimensions from the relative abundance. Based on its design principle, this method can be widely applied to different ecosystems. Therefore, we believe that the GIS-AQ method can play an immeasurably useful role in microbiological research.

**KEYWORDS** internal standards, absolute quantification, microbiota, solid-state fermentation

The study of microbiota such as those in soil (1), diseases (2), fermented food (3), and sites of human activity (4, 5) has paid greater attention to various temporal dimensions, such as seasonal (6) and age (7) differences. Comparing microbial abundances between different samples in various temporal dimensions is important for expanding the breadth and depth of the research.

Omics techniques have been utilized as tools in microbiology (8, 9). Amplicon sequencing is the most common method for analyzing microbial community structures (10); however, this method can generate spurious results (11). The relative abundance does not reflect the variation in absolute abundances of the individual microbes (12).

**Citation** Wang S, Wu Q, Han Y, Du R, Wang X, Nie Y, Du X, Xu Y. 2021. Gradient internal standard method for absolute quantification of microbial amplicon sequencing data. *mSystems* 6:e00964-20. <https://doi.org/10.1128/mSystems.00964-20>.

**Editor** Stephen R. Lindemann, Purdue University

**Copyright** © 2021 Wang et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Qun Wu, wuq@jiangnan.edu.cn, Yao Nie, ynie@jiangnan.edu.cn, or Yan Xu, yxu@jiangnan.edu.cn.

**Received** 6 October 2020

**Accepted** 3 December 2020

**Published** 12 January 2021

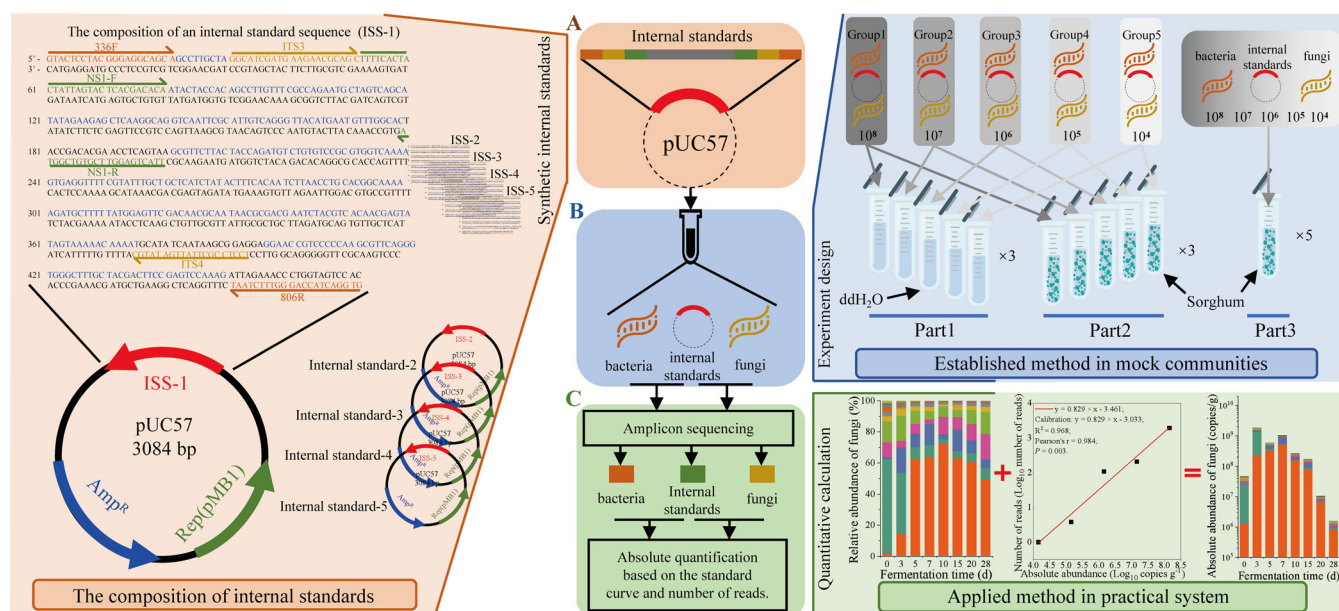
Thus, to compare samples across temporal and spatial dimensions, we must use microbiota absolute quantitative results (13).

The internal standard can be utilized to solve problems with absolute microbial quantification (12). Recently, researchers have conducted quantitative analyses of microbiota abundances through the synthetic internal standards of 16S rRNA, 18S rRNA, and internal transcribed spacer (ITS) (14) using a standard external strain, such as a fluorescently labeled strain (15). Presently, only one internal standard can be added to ecosystems with only one concentration. However, the different individual microbes in the microbiota cover a wide range of concentrations in the actual samples (16, 17). For example, the concentrations of different bacteria ranged from  $6.9 \times 10^4$  CFU  $g^{-1}$  to  $6.5 \times 10^8$  CFU  $g^{-1}$  in compost (18). Typically, the closer the internal standard is to the microbial concentration, the more accurate the quantitative result is. Studies have indicated that different results are obtained using internal standards with various concentrations, and only 20 to 80% of microbes can be accurately quantified (14). Hence, to quantify different microbial absolute abundances, it was suggested to determine the same sample multiple times by adding different internal standard concentrations at various times, which is time-consuming. As a result, it remains challenging to accurately determine microbes with different concentrations by measuring a sample only once. Therefore, it is important to establish an efficient method to accurately quantify microbes with large concentration ranges in complex microbiota.

In this study, we established a gradient internal standard absolute quantification (GIS-AQ) method to absolutely quantify microbes in complex microbiota in solid-state fermentation. The gradient internal standard group concentrations were simultaneously added to the same sample to quantify microbial concentrations of different orders of magnitude. The specific primers of the 16S rRNA, ITS, and internal standards were designed on the same sequence using one sequence fragment to reduce the number of internal standards added. Moreover, the deviations from the quantitative equations of microbes and internal standards were eliminated through calibration. In addition, we built a mock community with specific primers for each organism in the communities with quantitative real-time PCR (qPCR) validation. Based on the mock community, we demonstrated the reliability and accuracy of the GIS-AQ method by microscopy quantification (the microbial gradient dilution solution was counted using optical microscopy). Finally, the GIS-AQ method was used to identify the absolute abundance of microbiota during solid-state Chinese liquor fermentation. This method is useful for quickly understanding the absolute abundance of microbial communities over different temporal dimensions and can play an immeasurable role in microbiological research.

## RESULTS

**GIS-AQ method principle.** First, we added five different internal standards (pUC-57 plasmid) with a  $10\times$  concentration gradient to a sample to construct a gradient internal standard group for absolute quantification (Fig. 1A). Each internal standard contained a specific pair of primers and recognition sequences (Table 1) that were not present in the microbial genomes based on a search of the National Center for Biotechnology Information (NCBI) database and were flanked by the universal primers of the bacterial 16S rRNA V3-V4 and fungal ITS2 regions (Fig. 1A). The concentration of plasmid was calculated by the calculation equation in Materials and Methods. Second, before genomic DNA extraction, five internal standards (plasmid [Fig. 1A]) were added to the same sample at different concentrations to construct the concentration gradient (approximately  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ , and  $10^8$  copies  $g^{-1}$ ). The concentrations were chosen based on previous studies on the microbial concentration ranges ( $10^3$  to  $10^9$  copies  $g^{-1}$ ) in the Chinese liquor solid-state fermentation ecosystem (19–21) (Fig. 1B). Due to the specific primers and recognition sequences (see Materials and Methods) of the five internal standards, the internal standard sequences (ISS) were easy to identify during data analysis. Third, there was a positive correlation (Pearson's  $r > 0.9$ ;  $P < 0.05$ )



**FIG 1** Design and application method of internal standards. (A) Synthetic internal standards. The internal standards were designed and constructed to contain different primer sequences (336F/806R, ITS3/ITS4, and IS-F/IS-R) and a specially designed gene fragment. (B) Experiment design. Three experiments were designed to demonstrate the accuracy of the internal standards. (Part 1) The internal standards and microbes were added to the ddH<sub>2</sub>O. (Part 2) The internal standards and microbes were added to the sorghum. (Part 3) Five internal standards and 10 microbes were added to the sorghum. (C) Quantitative calculation. Based on the traditional amplicon sequencing results and the quantitative equation obtained in part 3, the microbial absolute abundance in the sample was calculated.

between the copies of the internal standards added and detected (Fig. 1C). The numbers of reads ( $\log_{10}$  reads) were normalized to the concentration ( $\log_{10}$  copies per gram) based on their linear correlation (raw data converted by  $\log_{10}$ ) to complete the individual microbial quantification.

**Detection of the microbial absolute abundance in mock communities using the GIS-AQ method.** To explore the possibility of quantifying bacterial and fungal communities using internal standards, we used five bacteria and five fungi to build a mock community (see Materials and Methods). Using the mock community, we established a gradient internal standard absolute quantification (GIS-AQ) method and demonstrated its accuracy using the following five steps.

First, we tested whether the five internal standards were stable and reliable (Fig. 1B, part 1; see Materials and Methods). Based on the qPCR results, the five parallel experiments demonstrate that there were no significant differences ( $P > 0.05$ ) and the quantitative equation was stable and reliable ( $R^2 = 0.999$ ;  $P < 0.001$ ; see Fig. S1 in the supplemental material) for use in this method.

Second, we measured the recovery rates of the five internal standards, five bacteria, and five fungi of different absolute abundances during genomic DNA extraction from a solid substrate (steamed sorghum) (Fig. 1B, part 2; see Materials and Methods). The results illustrate that the recovery rates of the internal standards, bacteria, and fungi were 70.02 to 83.90%, 71.87 to 87.10%, and 70.45 to 86.12%, respectively, with no significant differences ( $P > 0.05$  [Fig. S2]).

Third, we calibrated the internal standards to eliminate the deviations between the quantitative equations of the internal standards, bacteria, and fungi, because we assumed that the internal standards cannot be completely equivalent to the microbes in the ecosystems. We built two staggered mock communities using 10 bacteria and fungi common in liquor fermentation. We then added five different absolute abundances of internal standards (approximately  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ , and  $10^8$  copies  $g^{-1}$ ) to each mock community (Fig. 1B, part 3; see Materials and Methods). Then, we extracted the internal standards (plasmids) and microbial genomic DNA for quantification (see

**TABLE 1** Internal standard sequences<sup>a</sup>

Internal standard	Sequence length (bp)	% GC	Complete internal standard sequence
IS1	472	46.40	5'-GTACTCTACGGGAGGCAGCAGCCTTGTAGGCATCGATGAAGAACGCAGCTTTTCACTACTATTAGTACTCACGACAC AATACTACCACAGCCTTGTTCGCCAGAATGCTAGTCAGCATATAGAAGAGCTCAAGCGAGGTC AATTCGCATTGTCAG GGTTACATGAATGTTTGGCACTACCGACACGAACCTCAGTAAGCGTTCCTACTACCAGATGCTGTGTCCGCGTGGTCA AAAGTGAGGTTTTCGTATTTGCTGCTCATCTATACTTTCAACAATCTTAACCTGCACGGCAAAAAGATGCTTTTATGGAGT TCGACAACGCAATAACGCGACGAATCTACGTCACAACGAGTATAGTAAAAACAAAATGCATATCAATAAGCGGAGGAG GAACCGTCCCCAAGCGTTCAGGGTGGGCTTTGCTACGACTCCGAGTCCAAAGATTAGAAACCCTGGTAGTCCAC-3'
IS2	472	46.20	5'-GTACTCTACGGGAGGCAGCAACTCCCTGTTGCATCGATGAAGAACGCAGCAAGAATGATGATAATCATGACTGT GCTAAGACGGTGTGCAAAACAAAGCGGTCTTACGGTCAGTCGATTTTCTCTCGAGTCTCGTCCAGTTGAGCGTATCACT CTCAATGTACTAGCAAGCCAAGAAGGCTGTGCTTGGAGTCAATCTGATGTAGGATGATCTCCAGACACCAGGCCACTAC TCTTCATACTAAAGCATAAACGTCGAACAGTCATGAAAGTCTTAGTACCGGACGTACCATTTTACTGTGAATATTACCT GAAGCTGTACCGTTATTGAGGAGCAAAAGATGTAGTACTGCTCTTATCATATTTGTATTGGCATATCAATAAGCGGAGGA CCCACGCACCTGATCGCTCCTGTTTGTCTTTAAGGACCGGACGAACCACAGAGCATTAGAAAACCTGGTAGTCCAC-3'
IS3	472	45.60	5'-GTACTCTACGGGAGGCAGCAGCAGTCCCTGTTGCATCGATGAAGAACGCAGCCTTAACCTACTATCAATACCCATGACT TGACTCTGCTGCAGTACGTATCGCTGAAAACAGTGTAGTGTAAAGGAATGCTCTGACCAGGACAACACACGTAGTG AAAGTTACATGTTTGGTTCCTCGACTCGGATCTGAGTTGACCAATGACTCACTTGAGATCTGAACCCTAGTGATG ATAAATATGTATCTCGTTCACGCAGATTGCCAGCACTTTCAGAATCATGATGTGCATGGTAGAATGACTCTTATAACGAA CTTCGACATGATAATATCCCCCTTCAACTTCTAGAGAAAGAAAAGTATTGACATGAGGCATATCAATAAGCGGAGGA TTCATCAGCTAACGTAACGGTTAGAGGCTCGCTAAATCGCACTGTCCGGCTCCCTATTAGAAAACCTGGTAGTCCAC-3'
IS4	472	45.40	5'-GTACTCTACGGGAGGCAGCAGTCCCTGTTGCATCGATGAAGAACGCAGCCTTCCAGCACAACAGCCAAAGAAGT TTCCAATTTTTTATTTCCGAATGACATGTGTCTCCTTGGGTAATTCGCGCAAAATCGCGCAAAACTTAGGAGCCAGGGGAAAC AGATAGGTCTAATTAACCTAAGGGAGTAAATCTTGAATCGTTGAACTATATACTTACGCTGGAACCTTCTCCGG CGAATTTTTACTGTCACCAACTACGAGATTTGAAGTAAACCAATTAAGCACATAGTCGCGCTATCCGACAATTTCCAAT TATAACATATCGTTCATGAAGGCCAGAATTACTTACCGCCCTTCCATGCGTGCAGCATATCAATAAGCGGAGGAGC TGATCCGAGTCGAGTAAAAACACCAGTACCCAAAACCAGGCGGGCTCGCCACATTAGAAAACCTGGTAGTCCAC-3'
IS5	472	46.90	5'-GTACTCTACGGGAGGCAGCAGTCCCTGTTGCATCGATGAAGAACGCAGCCTTCCAGCACAACAGCCAAAGAAGT ATCCGAAGAGAGAGTGTGCGATCCTCCGTTAAGATATTCTTACGTATGATAGCTATGATTTTGTAGAGGTAGCGAAC GCGTTAAACATTTACAGATAGTGGGATTGCGGCAAAAGGGCGTATAATTGTGGACTAACATAGTCGTAACACTACGATG GTACCAACTCAATCTCAGCTCGTGCCTAAATAACGTAATCTCAACTGATTCTTGGCAATCTACGGAGCGACTTGA TTATTAACAGTTGTCTAGCGAGTTCTAATCTTTACCAACATCGTAATAGCTCCAAAGCATATCAATAAGCGGAGGAATC CGCAGTGGCCGGTAGACACACGCTCCACCCCGTGTCTGTGACGGGGACTAAATTAGAAAACCTGGTAGTCCAC-3'

<sup>a</sup>The bold sequences are the primers 336F/806R. The italicized sequences are the primers ITS3/ITS4. The underlined sequences are the specific internal standard sequence primers.

Materials and Methods). The quantitative equation established by internal standards was as follows:

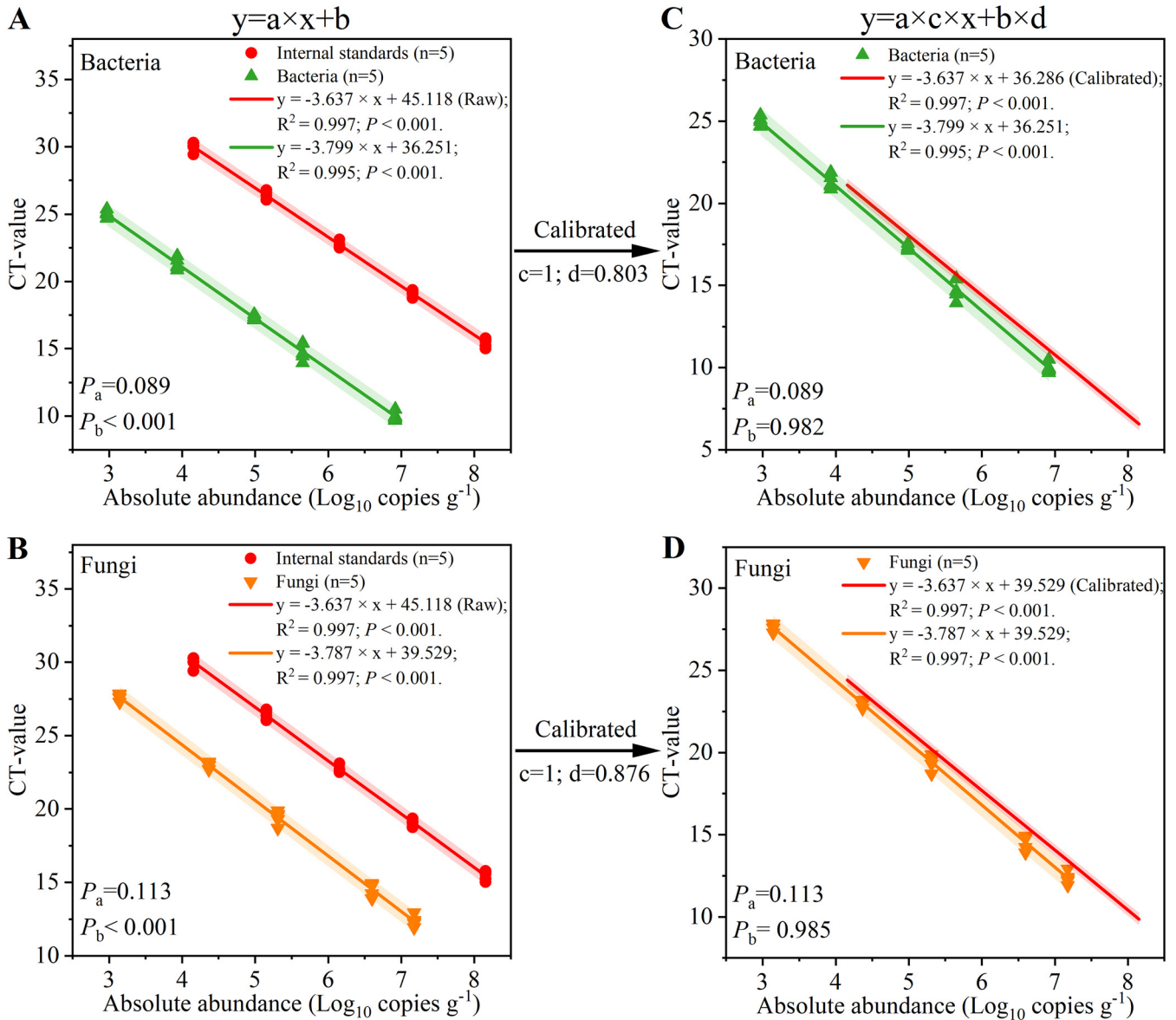
$$y_{\text{CT-value}} = a \times x_{\text{internals}} + b$$

which was calibrated to

$$y_{\text{CT-value}} = a \times c \times x_{\text{internals}} + b \times d$$

where  $a$  and  $b$  are the slope and intercept of the internal standards and microbes, respectively,  $c$  and  $d$  are the calibrated coefficients in the internal standards and microbes, respectively,  $x_{\text{internals}}$  is the concentration of the internal standards ( $\log_{10}$  copies per milliliter) or microbes ( $\log_{10}$  copies per milliliter), and  $y_{\text{CT-value}}$  is the cycle threshold value (CT-value) of qPCR results.

An analysis of variance (ANOVA) of the slopes (variable  $a$  in the equation) and intercepts (variable  $b$  in the equation) of the quantitative equations of the internal standards, bacteria ( $P_a = 0.089$ ;  $P_b < 0.001$  [Fig. 2A]), and fungi ( $P_a = 0.113$ ;  $P_b < 0.001$  [Fig. 2B]) demonstrated that there were no significant differences ( $P_a > 0.05$ ) between the slopes, but there were significant differences ( $P_b < 0.001$ ) between the intercepts. Thus, we attempted to calibrate the quantitative equations of the internal standards based on the intercepts of bacteria and fungi. Due to the lack of a significant difference between the slopes ( $P_a > 0.05$ ), the slope calibrated coefficients were 1 ( $c_{\text{bacteria}} = 1$ ;  $c_{\text{fungi}} = 1$  [Fig. 2C and D]). The intercepts of the quantitative equations of the internal standards, bacteria, and fungi were  $45.118 \pm 1.316$ ,  $36.251 \pm 0.534$ , and  $39.529 \pm 0.970$ , respectively (Fig. 2A and B). Therefore, the calibrated coefficients were 0.803 and 0.876

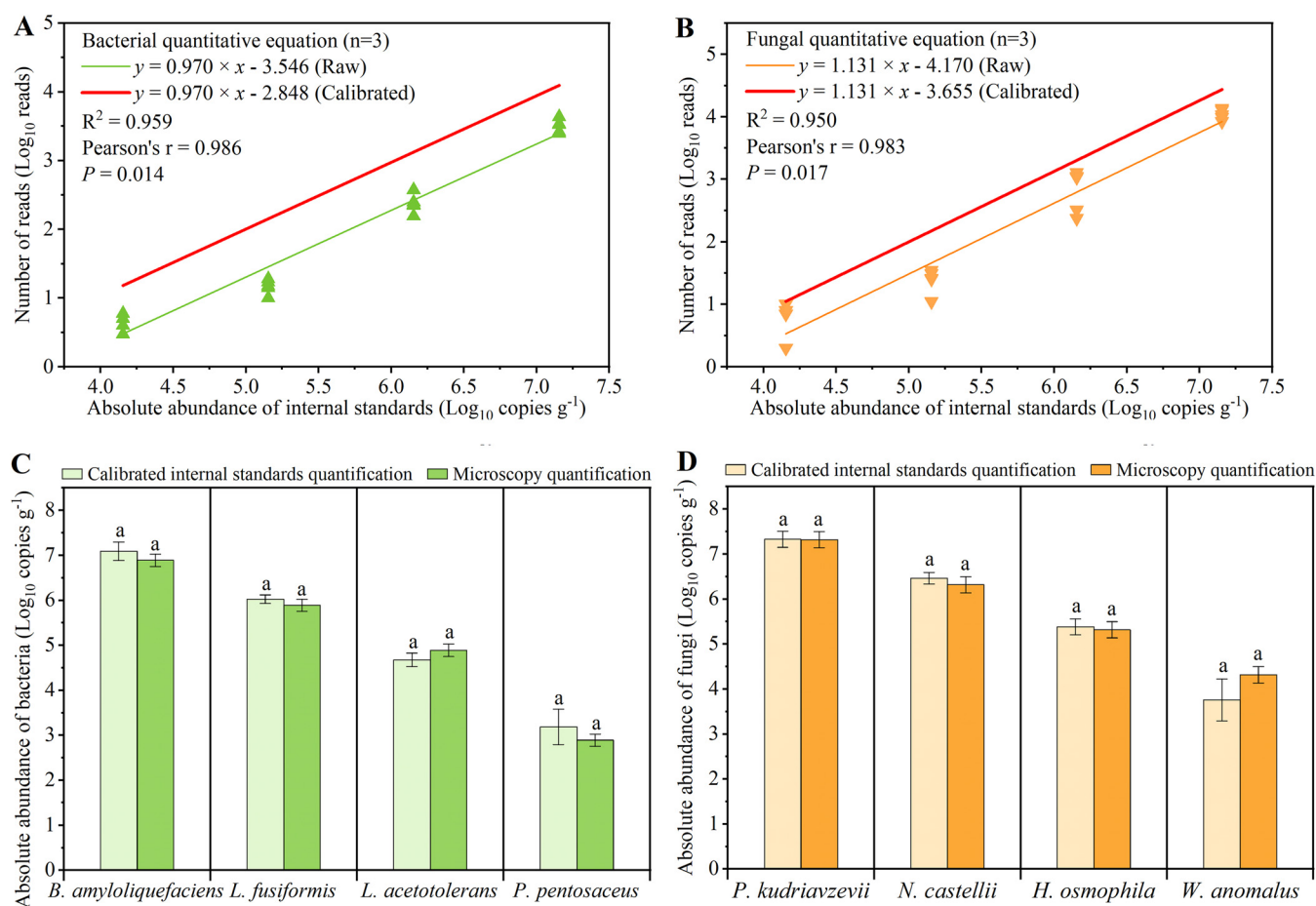


**FIG 2** Calibration of the quantitative equations of the internal standards. (A) The equations of internal standards and bacteria before calibration. (B) The equations of internal standards and fungi before calibration. (C) The equation of internal standards calibrated by the equation of bacteria. (D) The equation of internal standards calibrated by the equation of fungi.  $P_a$  and  $P_b$  were calculated using one-way ANOVA.

for bacteria and fungi, respectively ( $d_{\text{bacteria}} = 36.251/45.118 = 0.803$ ;  $d_{\text{fungi}} = 39.529/45.118 = 0.876$  [Fig. 2C and D]). When the calibrated equations were used, the deviations between the internal standards and microbes effectively disappeared (bacteria,  $P_b = 0.982$ ; fungi,  $P_b = 0.985$  [Fig. 2C and D]), producing excellent agreement between the results of the quantitative equations of the internal standards and microbes.

Fourth, amplicon sequencing was used to sequence the genome samples extracted in the third step. We assumed that the deviation between the quantitative results of the internal standards and microbes in the third step would also exist between the results of amplicon sequencing. Therefore, the calibrated coefficients ( $d_{\text{bacteria}} = 0.803$ ;  $d_{\text{fungi}} = 0.876$ ) obtained in the third step were applied to the quantitative equations of amplicon sequencing (Fig. 3A and B) (see Materials and Methods). We used microscopy quantitation methods to validate the GIS-AQ method. Microscopy quantitation is the most accurate quantitative method because it can directly count the number of added individual microbes (15, 22). Therefore, the accuracy of the GIS-AQ method can be





**FIG 3** Detection of the microbial absolute abundances in the mock communities using GIS-AQ. (A) Average calibration equation of bacteria ( $n=5$ ). (B) Average calibration equation of fungi ( $n=5$ ). (C) Comparison of two absolute quantitative methods for bacteria. (D) Comparison of two absolute quantitative methods for fungi. The same letter (a) indicates no statistically significant differences ( $P > 0.05$ ).

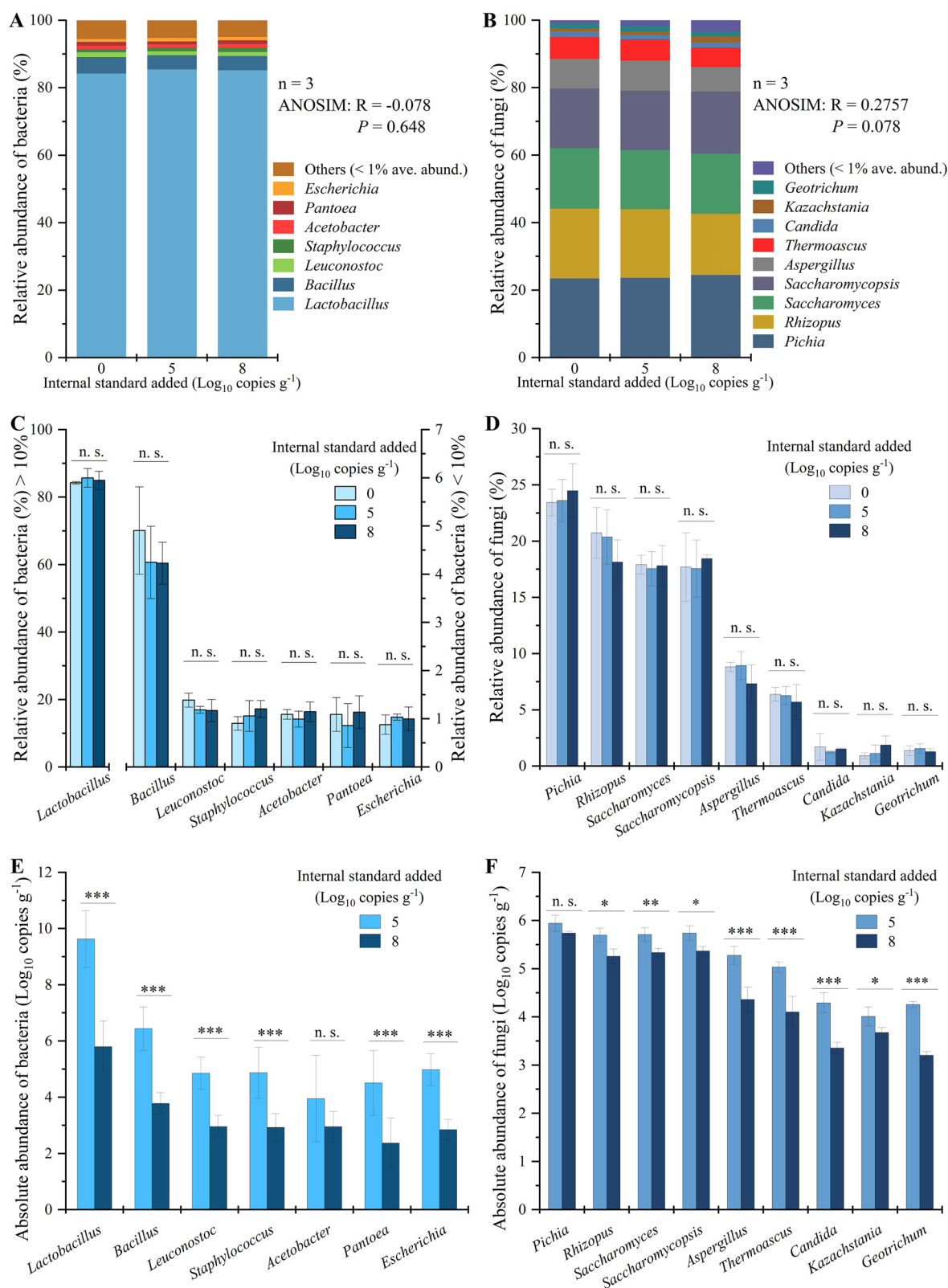
validated if there was no significant difference between the quantitative results by calibrated quantitative equations and microscopy quantitation. Based on the gradient of the internal standards and calibrated coefficients, we established calibrated quantitative equations for the bacteria and fungi (Fig. S3) in each sample.

Finally, we compared the different quantitative results (Fig. S4A and B) and found that there were no significant differences ( $P > 0.05$ ) between the quantitative results using the calibrated quantitative equations and the microscopy quantitation in the mock communities (Fig. 3C and D). These results suggest that the GIS-AQ method can accurately ( $P_{\text{internals versus microscopy}} > 0.05$ ) quantify the microbial absolute abundances.

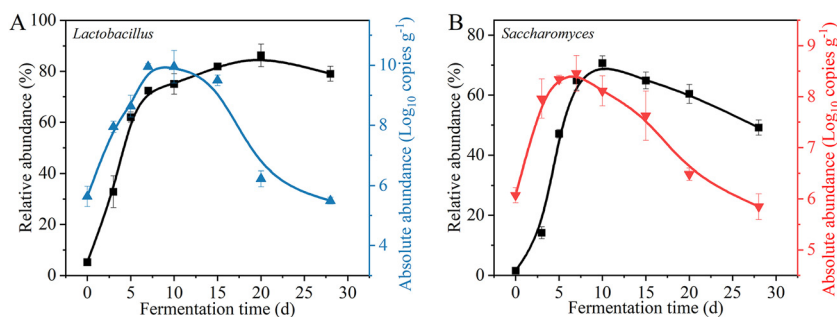
**Perturbation of quantification with single internal standards.** Before we applied the GIS-AQ method to solid-state fermentation, we sought to answer two questions. (i) In the amplification sequencing process, does the use of internal standards affect the relative abundance of microbes in solid-state fermentation? (ii) Does the use of a single internal standard perturb the absolute abundance?

To determine the effect of additional internal standards on the solid-state fermentation, we only added one internal standard to three independent solid fermentation samples with 0 copies  $\text{g}^{-1}$  (control), approximately  $10^5$  copies  $\text{g}^{-1}$ , and  $10^8$  copies  $\text{g}^{-1}$  (Fig. S5A and B) (see Materials and Methods). To facilitate comparison between the samples, we analyzed the composition of the microbial communities after the internal standards were removed.

There were no significant differences between groups with different internal standards added in the bacterial communities (analysis of similarities [ANOSIM]:  $R = -0.078$ ;  $P = 0.648$  [Fig. 4A]) and fungal communities (ANOSIM:  $R = 0.2757$ ;  $P = 0.078$  [Fig. 4B]).



**FIG 4** Perturbation of results with a single internal standard in solid-state fermentation samples ( $n=3$ ). (A and B) Relative abundances of bacteria (A) and fungi (B) after removing the internal standards. (C and D) Relative abundances of dominant bacteria (C) and fungi (D) (relative abundance  $\geq 1\%$ ) in different samples (the concentrations of internal standards are approximately 0 copies  $\text{g}^{-1}$ ,  $10^5$  copies  $\text{g}^{-1}$ , and  $10^8$  copies  $\text{g}^{-1}$ ). (E and F) Absolute quantitation comparison of dominant bacteria (E) and fungi (F) quantified by different concentrations of internal standards (approximately  $10^5$  copies  $\text{g}^{-1}$  and  $10^8$  copies  $\text{g}^{-1}$ ).



**FIG 5** Abundances of *Lactobacillus* and *Saccharomyces* in the Chinese liquor solid-state fermentation ( $n = 3$ ). (A) Relative and absolute abundances of *Lactobacillus*. (B) Relative and absolute abundances of *Saccharomyces*.

Then, genera with an average relative abundance of  $\geq 1\%$  (dominant microbiota) were selected for analysis. There were no significant differences ( $P > 0.05$ ) in the relative abundances of the seven dominant bacteria (Fig. 4C) and nine dominant fungi (Fig. 4D) in the groups with different internal standards added. These results indicate that during the amplification sequencing process, the use of internal standards did not affect the relative abundances of microbes in solid-state fermentation.

The absolute abundances of bacteria and fungi were obtained based on the concentrations of the single internal standard (approximately  $10^5$  copies  $g^{-1}$  and  $10^8$  copies  $g^{-1}$ ). The absolute abundances of all bacteria were significantly different ( $P < 0.05$ ) between the two internal standards of different concentrations (except for *Acetobacter*) (Fig. 4E). The absolute abundances of all fungi were also significantly different ( $P < 0.05$ ) between the two internal standards of different concentrations (except for *Pichia*) (Fig. 4F). These results indicate that adopting gradient internal standards is necessary because of the quantitative perturbation that results from using a single internal standard.

**Detection of the absolute abundance of microbiota in solid-state fermentation samples using the GIS-AQ method.** To explore the difference between the relative and absolute abundances, the Chinese liquor solid-state fermentation (a type of ecosystem) (see Materials and Methods) was used as a model system for quantification. Microbial compositions were constructed for the solid-state fermentation samples from the first and second batches based on 16S rRNA and ITS sequences (Fig. S6A and B). The absolute quantitative equations of bacteria and fungi were obtained based on the calibrated quantitative equations of the internal standards (Fig. S7). Then, the absolute abundances of the communities in the fermentation were calculated based on the calibrated quantitative equations (Fig. S6C and D).

For bacteria, *Lactobacillus* (absolute dominant genus;  $\geq 10\%$  average abundance) was selected as an example to illustrate the difference between relative and absolute abundance. The relative abundance of *Lactobacillus* gradually increased until day 20 of fermentation, while the absolute abundance reached its peak on day 7 (Fig. 5A). For fungi, *Saccharomyces* (absolute dominant genus) was selected as an example to illustrate the difference between relative and absolute abundance. The relative abundance of *Saccharomyces* gradually increased until day 10 of fermentation, while the absolute abundance reached its peak on day 7 (Fig. 5B).

In summary, the results obtained from the relative and absolute abundances are different. The relative abundance is more suitable for the study of the different compositions of communities in one sample. The absolute abundance can be used to study the difference in communities between different samples, and the GIS-AQ method allows this to be done swiftly.

## DISCUSSION

**Characteristics of different quantitative methods.** We established and implemented a gradient internal standard absolute quantification (GIS-AQ) method to



determine the absolute abundances of bacteria and fungi. This method establishes a reference and extends existing quantitative methods for microbial communities. Every method has its advantages and disadvantages; therefore, we compared the GIS-AQ method with existing quantitative methods to provide context and allow the reader to choose the most appropriate method for their respective research question. To resolve the shortcomings of traditional amplicon sequencing, various methods have been proposed to assess the absolute abundance of microbiota, such as flow cytometry (11, 16, 17), qPCR (17), fluorescence *in situ* hybridization coupled with catalyzed reporter deposition (CARD-FISH) (23), and spike-in (internal standard) (12, 14, 15, 22, 24–26). However, determining the absolute abundance using these methods remains challenging. For example, combining amplicon sequencing and flow cytometry can accurately quantify the number of cells in liquid samples. However, it is mainly suitable for specific labeled cells in the liquid samples. In the case of qPCR, the external standard used for qPCR quantification seldom has the same recovery rate as the ecosystem. The CARD-FISH method is complicated and time-consuming to operate. The internal standard approach is the most widely used quantitative method. For example, some researchers designed three internal standards (plasmids) for prokaryotic 16S rRNA, eukaryotic 18S rRNA, and fungal ITS in soil samples. To realize the microbial quantification, three internal standards were added before DNA extraction and repeated DNA extraction and sequencing, changing the concentration of internal standards in each repetition and establishing a quantitative equation, of the same soil sample (14). Researchers designed a synthetic full-length 16S rRNA gene spike-in standard for prokaryotic 16S rRNA in soil samples, and they also used the same method as before to achieve the quantification of prokaryotic 16S rRNA (26). Compared with the existing internal standard quantitative methods, the GIS-AQ method can be described as follows. (i) We constructed a gradient internal standards group, and one round of amplicon sequencing can achieve the determination of different concentrations of internal standards. The GIS-AQ method is also faster and more economical than performing amplification sequencing multiple times to establish a quantitative equation. (ii) We designed the primer information of bacteria and fungi on the same internal standard sequence, which facilitates the quantification of bacteria and fungi at the same time. (iii) Based on the mock communities, the quantitative equations of the internal standards were calibrated to eliminate the potential deviations between the internal standards and communities in the process of genome DNA extraction. Therefore, we believe that the GIS-AQ method has advantages for quantifying microbial communities.

**Application of the absolute abundance data.** The relative abundance of amplicon sequencing is based on the normalization or transformation of the total read number (27). Relative abundance is suitable for statistical analysis of the same samples or samples that have similar communities, such as natural soils (28) and healthy guts (29). For samples with large differences in communities, such as fertilizing soils (25) and patients' guts, the relative abundance may easily generate information that is contrary to the actual situation (11). Therefore, at present, the analysis method of combining relative and absolute abundances is primarily used for samples with large differences in microbial communities. Specific manifestations are as follows. (i) When characterizing the changes in communities, the absolute abundance is more convincing than the relative abundance because it shows the dynamic changes in communities between different samples (such as time series samples) that cannot be seen by traditional relative abundance (11). This application represents the most important use of absolute abundance (14). (ii) Regarding the replacement of relative abundance data with absolute abundance data in traditional statistical analysis, at present, no systematic comparative studies that have shown that absolute abundance data can be used in traditional statistical analysis methods are known; however, some researchers have attempted substitution in different statistical methods, such as principal-component analysis (25), network analysis (17), and cluster patterns (heat map) (15). The scope of the application of absolute abundance in the field of microbial quantification still requires further

research to support and verify current results. Thus, absolute abundance is a crucial challenge for future microbiota variance studies (24).

**Applicability of the GIS-AQ method.** The GIS-AQ method established in this study provides a new reference for the field of microbial quantification. As the microbial abundances, species, and substrates are different in various research fields, the GIS-AQ method can be applied to different research fields after appropriate adjustments. These adjustments mainly reflect the design of the internal standards, mock communities, and substrates. (i) The parameters of the internal standards should be designed according to the general primer, GC content, and the sequence length corresponding to the amplification regions. For example, when the amplification region is V4 (30), V4-V5 (31), or ITS1 (32), it is necessary to change the primer of the internal standards used. (ii) The mock communities and substrates should be the same or similar to the object of study because the quantitative principle of internal standards is to use the quantitative equations of the internal standards to quantify the absolute abundance of the microbes. At present, it is unknown whether differences exist between the internal standards and microbes in the process of genome extraction and amplification; however, the problem can be solved using mock communities and substrates. Therefore, when studying the oceans (33), the mock communities should have microbes with different abundances in the oceans, and the mock substrates chosen should be a sterile isotonic saline solution with a seawater composition or sterilized seawater.

**Matters requiring attention in the use of the GIS-AQ method.** The amount of manual labor required by the GIS-AQ method is larger than for existing quantitative research using internal standards (25). This is mainly due to the construction of the mock communities and substrates. As we assume that the internal standards cannot be completely equivalent to the microbes in the natural systems, it is necessary to eliminate the deviations between the internal standards and microbial quantitative equations. Theoretically, different mock communities and substrates need to be constructed for various research ecosystems. However, if there is no change in the ecosystem, the mock communities and substrate need only be constructed once; therefore, the extra manual labor required is limited. In the Chinese liquor solid-state fermentation ecosystem, the detection line of GIS-AQ method is  $10^4$  to  $10^5$  copies  $g^{-1}$ , which is a very low detection line compared with that for the dominant microbes ( $10^8$  to  $10^9$  copies  $g^{-1}$ ) (21, 34). Furthermore, the concentration ranges of the internal standards can be adjusted according to the microbial concentration ranges in different ecosystems, such as soils ( $10^5$  to  $10^9$  copies  $g^{-1}$ ) (25) and guts ( $10^7$  to  $10^{11}$  copies  $g^{-1}$ ) (17). For the use of the GIS-AQ method, the following points should be noted. (i) There should be uniformity between the internal standards, because different internal standards need to be used to represent various concentrations. (ii) The concentration gradients of microbes in mock and natural communities should be similar. (iii) When calibrating the quantitative equations, the difference in the calibration coefficients between bacteria and fungi cannot be ignored.

In summary, we developed a GIS-AQ method for bacteria and fungi. The GIS-AQ method expands upon traditional amplicon sequencing based on relative abundance. Based on the design principle, after proper modification, the GIS-AQ method is compatible with any application using any amplification region (such as V3, V4, V3-4, ITS1, and ITS2) and is thus applicable in a broad range of ecosystems, including soil, rhizospheres, and food. Therefore, the GIS-AQ method is a valuable addition to microbiology research.

## MATERIALS AND METHODS

**Design and construction of internal standard sequences.** The internal standard sequences (ISS) were designed based on three key elements. (i) The internal standard sequences had the same universal primer binding sites (PBSs) as those of the bacteria and fungi in the amplicon sequencing samples. (ii) The synthetic stuffer sequences had a similar length and GC% as that of the microbes in the solid-state fermentation. (iii) The specific PBSs were used to recognize the internal standard sequences (Fig. 1).

First, the primers 336F/806R of the 16S rRNA V3-V4 region (35) and ITS3/ITS4 of the ITS2 region (36) were used as universal primers for bacteria and fungi for the internal standards. Second, based on

previous research (21), we selected the average length of amplified fragments and GC content of the 10 microbes (five bacteria and five fungi [Table S1]) with the highest relative abundances in the solid-state fermentation process as the references to construct the random sequences (14). There was a set of random sequences (Table 1) that satisfied the following criteria: uniform GC content ( $\pm 1\%$ ), no homopolymers of  $>10$  bp, no repeats of  $>16$  bp (as determined by BLAST), and no self-complementary regions of  $>10$  bp (26). In addition, the optimized set of artificial sequences shared negligible identity with sequences in the National Center for Biotechnology Information (NCBI) database. Third, we designed specific primers for internal standards sequences based on the NCBI Primer-BLAST (Table S2). The specific primers satisfied the following criteria: similar length, similar melting temperature ( $T_m$ ) values, similar GC contents, no homopolymers of  $>3$  bp, and no self-complementary regions of  $>5$  bp (26). The designed primers were used as upstream and downstream primers, and BLAST was performed in the NCBI database, which demonstrated negligible identification with known sequences in the NCBI nucleotide collection (nr/nt) database (web-BLAST performed in May 2019). The specificity of these primers was demonstrated by a PCR cross-validation. The PCR conditions were as described in a previous study (37).

Complete internal standards sequences are provided in Table 1. Full-length internal standard sequences (472 bp) were chemically synthesized and inserted into the pUC-57 plasmid cloning vector by Sangon Biotech (Shanghai, China). Plasmid cloning vectors with internal standard sequence inserts were transformed into *Escherichia coli* TOP10 competent cells [genotype:  $F^-$  *mcrA*  $\Delta$ (*mrr-hsdRMS-mcrBC*)  $\phi$ 80 *lacZ* $\Delta$ M15  $\Delta$  *lacX74* *recA1* *ara* $\Delta$ 139  $\Delta$ (*ara-leu*)7697 *galU* *galk* *rpsL* (Str)*endA1* *nupG*] (Sangon Biotech) as per the manufacturer's instructions. Internal standard DNA was extracted from overnight liquid cultures using a SanPrep spin column and collection tube (Sangon Biotech) according to the manufacturer's instructions. The internal standard DNA concentrations were measured using a PicoGreen double-stranded DNA (dsDNA) assay kit (Invitrogen Life Technologies, Grand Island, NY) (26), and the internal standard concentration was calculated as

$$C_{\text{internal}} (\text{copies ml}^{-1}) = \frac{m_{\text{DNA}} (\text{ng ml}^{-1}) \times 6.022 \times 10^{23} (\text{copies mol}^{-1})}{\text{DNA length (bp)} \times 10^9 (\text{ng g}^{-1}) \times 660 (\text{Da mol}^{-1})}$$

where  $C_{\text{internal}}$  is the internal standards concentration, and  $m_{\text{DNA}}$  is the mass of the internal standards. In our study, five internal standards (IS1 to IS5) were stored at  $-80^\circ\text{C}$  for future processing at a concentration of  $1 \mu\text{g ml}^{-1}$ .

**Strains.** The microbes used in the experiment were isolated from the Chinese liquor solid-state fermentation process: *Bacillus amyloliquefaciens* LBM 10008, *Lysinibacillus fusiformis* LBM 10019, *Lactobacillus acetotolerans* LBM 10005, *Weissella paramesenteroides* LBM 10007, *Pediococcus pentosaceus* LBM 10024, *Pichia kudriavzevii* LBM 20017, *Naumovozyma castellii* LBM 20004, *Hanseniaspora osmophila* LBM 20013, *Wickerhamomyces anomalus* LBM 20006, and *Saccharomyces cerevisiae* LBM 20001. Using the internal standard-specific primer design method, we constructed specific primers (Table S2) for the microbes. The specificity of these primers was demonstrated using PCR cross-validation.

**qPCR.** The qPCR was performed on a real-time PCR system (Applied Biosystems, Foster City, CA) with a mixture of  $0.4 \mu\text{l}$  ( $20 \mu\text{M}$ ) of primers,  $10.0 \mu\text{l}$  of SYBR green supermix (SYBR Premix *Ex Taq* II; TaKaRa, Shanghai, China),  $1.0 \mu\text{l}$  of DNA templates, and  $8.2 \mu\text{l}$  of double-distilled water (ddH<sub>2</sub>O) (38). The PCR program was as follows: preheating at  $98^\circ\text{C}$  for 10 min, 40 cycles at  $98^\circ\text{C}$  for 30 s and  $60^\circ\text{C}$  for 30 s, and an increase of  $0.5^\circ\text{C}$  every 5 s from  $72^\circ\text{C}$  to  $95^\circ\text{C}$  for the melting-curve analysis to confirm the amplification specificity (39).

**The calibrated coefficients.** The recovery rates of internal standards and microbes differed in the genome extraction (14). To calibrate the differences in the recovery rate, we designed the following experiments.

First, we diluted all five internal standards (IS1 to IS5) according to the gradient of approximately  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ , and  $10^8$  copies  $\text{ml}^{-1}$ . Then, five internal standards of the same concentration were mixed to obtain the internal standard concentration gradient. qPCR was performed on the mixed liquid to determine whether there was an amplification preference among the various internal standards. Second, genomic DNA was extracted from the five groups of mixed liquid using the easy nucleic acid isolation (E.Z.N.A.) soil DNA kit (Omega Bio-Tek, Inc., Norcross, GA) according to the manufacturer's instructions to determine the internal standard recovery rate. Third, we used qPCR to establish the melting curves for five internal standards, five bacteria, and five fungi to discern the differences between their recovery rates. Finally, according to the microbe recovery rates, the quantitative equation established by internal standards, provided in Results, was determined.

**Mock communities.** To demonstrate the GIS-AQ method, we used mock communities to mimic the taxonomic diversity of the microbes in the solid-state fermentation process. Based on previous research (21), bacteria and fungi found to be dominant in the solid-state fermentation process of Chinese liquor fermentation were used as mock communities. *B. amyloliquefaciens* and *L. fusiformis* were incubated for 24 h at  $37^\circ\text{C}$  and 200 rpm in a Luria-Bertani sterile liquid medium (Sangon Biotech, Shanghai, China). *L. acetotolerans*, *W. paramesenteroides*, and *P. pentosaceus* were incubated for 48 h at  $37^\circ\text{C}$  in de Man-Rogosa-Sharp medium (Oxoid Ltd., Hants, UK). *P. kudriavzevii*, *N. castellii*, *H. osmophila*, *W. anomalus*, and *S. cerevisiae* were incubated for 48 h at  $30^\circ\text{C}$  in a yeast extract-peptone-dextrose medium (Sangon Biotech).

After incubation, the 10 cultured bacteria and fungi were eluted with phosphate-buffered saline (0.01 M, pH 7.2) at 4,000 rpm for 20 min (5804 R centrifuge; Eppendorf, Hamburg, Germany) three times. Then, the microbes and internal standards were added to the sterile sorghum substrates after

microscopy quantification. The sterile sorghum substrates were preprepared via autoclaving sorghum at 121°C for 20 min to simulate the substrates in the Chinese liquor solid-state fermentation process (21). The added concentrations of microbes were as follows: approximately  $10^8$  copies  $g^{-1}$  of *B. amyloliquefaciens*, approximately  $10^7$  copies  $g^{-1}$  of *L. fusiformis*, approximately  $10^6$  copies  $g^{-1}$  of *L. acetolerans*, approximately  $10^5$  copies  $g^{-1}$  of *W. paramesenteroides*, approximately  $10^4$  copies  $g^{-1}$  of *P. pentosaceus*, approximately  $10^8$  copies  $g^{-1}$  of *P. kudriavzevii*, approximately  $10^7$  copies  $g^{-1}$  of *N. castellii*, approximately  $10^6$  copies  $g^{-1}$  of *H. osmophila*, approximately  $10^5$  copies  $g^{-1}$  of *W. anomalus*, and approximately  $10^4$  copies  $g^{-1}$  of *S. cerevisiae*. The added concentrations of internal standards were as follows: approximately  $10^8$  copies  $g^{-1}$  of IS1, approximately  $10^7$  copies  $g^{-1}$  of IS2, approximately  $10^6$  copies  $g^{-1}$  of IS3, approximately  $10^5$  copies  $g^{-1}$  of IS4, and approximately  $10^4$  copies  $g^{-1}$  of IS5. Next, the sorghum substrates containing different absolute abundances of bacteria, fungi, and internal standards were uniformly mixed (Table S3). We performed five independent preparations for the mock communities.

**Sample collection in natural solid-state fermentation.** For the solid-state liquor fermentation, steamed grains were mixed with the starter at a ratio of 9:1 (wt/wt) and fermented in sealed jars for 28 days. Subsequently, the fermented grains were distilled to obtain the liquor. We collected a total of 24 samples (100 g per sample) from 3 jars (three independent preparations) in the layer center (0.5 m deep) on days 0, 3, 5, 7, 10, 15, 20, and 28 (Table S3). All samples were stored at  $-80^{\circ}C$  for further DNA extraction.

**Single internal standard perturbation experiment preparation.** To calibrate the influence of different concentrations of the internal standards on the results, we designed a single internal standard perturbation experiment. The perturbation experiment was designed based on the samples from day 10 of fermentation. Previous studies have demonstrated that the composition of the microbiota is stable and representative on day 10 of fermentation (21). We divided the same substrates (on day 10) into three parts (10 g each) and then added the internal standards (IS1) at 0 copies  $g^{-1}$  (control, equivalent ddH<sub>2</sub>O),  $10^5$  copies  $g^{-1}$ , and  $10^8$  copies  $g^{-1}$  (Table S3) and mixed evenly. All samples were stored at  $-80^{\circ}C$  for further DNA extraction. We conducted three independent preparations for all perturbation experiments.

**DNA extraction, qualification, and sequencing analysis.** Five-gram samples from the mock communities and natural fermentation were used to extract the total genomic DNA using an E.Z.N.A. soil DNA kit (Omega Bio-Tek, Norcross) according to the manufacturer's instructions. The V3-V4 region of the 16S rRNA gene was amplified using the universal primers 336F (5'-GTACTCCTACGGGAGGCAGCA-3') and 806R (5'-GTGGACTACHVGGGTWTCTAAT-3') (35). For the fungi, the ITS2 region was amplified using the primers ITS3 (5'-GCATCGATGAAGAACGCAGC-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') (36). A unique 8-nucleotide barcode sequence was added to the primers to differentiate between the samples. PCR mixtures contained 2.5  $\mu$ l of Pyrobest buffer (10 $\times$ ), 2  $\mu$ l of deoxynucleoside triphosphates (dNTPs; 2.5 mM), 1  $\mu$ l of each primer (10  $\mu$ M), 0.4 U of Pyrobest DNA polymerase (TaKaRa Holdings, Inc., Kusatsu, Japan), 15 ng of template DNA, and ddH<sub>2</sub>O up to 25  $\mu$ l. Amplification was performed as described previously (21). Amplicons were pooled in equimolar quantities and subjected to high-throughput sequencing using MiSeq sequencing for 2  $\times$  300-bp paired-end sequencing (Illumina, San Diego, CA) (21).

**Sequence processing.** All raw sequences were processed using Quantitative Insights into Microbial Ecology (QIIME) (version 1.8) (40). High-quality sequences were obtained by removing sequences with  $>2$  ambiguous bases,  $>10$  homopolymers, primer mismatches, average quality scores of  $<20$ , and lengths (excluding the barcode region or primer) of  $<50$  bp (21). Using USEARCH (version 10), chimeras were removed (41). UCLUST (version 1.2.22) clustered the trimmed sequences into operational taxonomic units (OTUs) with 97% sequence similarity (42). The Central Bureau of Fungal Cultures database (CBS-KNAW, [www.wi.knaw.nl](http://www.wi.knaw.nl), a public database of fungal sequences), EzBioCloud database ([www.ezbiocloud.net](http://www.ezbiocloud.net), a public database of bacterial sequences), and internal standard sequence information (Table S2) were used to align bacterial 16S rRNA gene, fungal ITS2 region, and internal standard sequences, respectively (21).

**Absolute quantitation.** To detect the absolute abundance of the microbes in the mock communities and ecosystems, we used gradient internal standards to build a linear equation between the concentration ( $\log_{10}$  copies per gram) and the number of reads ( $\log_{10}$  reads) of internal standards. The quantitative equation is as follows:

$$y_{\text{reads}} = a \times x_{\text{internals}} + b$$

which was calibrated to

$$y_{\text{reads}} = a \times c \times x_{\text{internals}} + b \times d$$

where  $a$  and  $b$  are the slope and intercept of the internal standards and microbes, respectively,  $c$  and  $d$  are the calibrated coefficients of the internal standards and microbes, respectively, and  $x_{\text{internals}}$  is the concentration of internal standards ( $\log_{10}$  copies per gram).

Based on the above equations, absolute abundances representing the microbial copy number ( $x_{\text{internals}}$ ) in the mock and natural samples were calculated based on the number of reads ( $y_{\text{reads}}$ ) in the amplicon sequencing.

**Statistical analysis.** Diversity and statistical analyses were conducted using XLSTAT (version 19.02.42992). The significant difference between recovery rates, slopes and intercepts were calculated using one-way ANOVA using SPSS Statistics (version 22). The Shannon index and Chao1 estimator were

calculated using QIIME (version 1.8) (43, 44). ANOSIM was used to test differences between samples based on the Bray-Curtis distance matrices (45).

**Data availability.** The bacterial and fungal raw sequence data were deposited in the DNA Data Bank of Japan (DDBJ) database under the accession numbers [DRA010282](https://doi.org/10.1093/dnaa/daa012) and [DRA010283](https://doi.org/10.1093/dnaa/daa013). The mock community cultures are available by request for noncommercial purposes.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, PDF file, 0.01 MB.

**FIG S2**, PDF file, 0.01 MB.

**FIG S3**, PDF file, 0.02 MB.

**FIG S4**, PDF file, 0.01 MB.

**FIG S5**, PDF file, 0.01 MB.

**FIG S6**, PDF file, 0.01 MB.

**FIG S7**, PDF file, 0.01 MB.

**TABLE S1**, XLSX file, 0.02 MB.

**TABLE S2**, XLSX file, 0.02 MB.

**TABLE S3**, XLSX file, 0.02 MB.

## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (NSFC; 31530055), National Key R&D Program of China (2018YFD0400402), Jiangsu Province Science and Technology Project (BE2017705), China Postdoctoral Science Foundation (2017M611702), and national first-class discipline program of Light Industry Technology and Engineering (LITE2018-12), the Priority Academic Program Development of Jiangsu Higher Education Institutions, the 111 Project (111-2-06), and the Collaborative Innovation Center of Jiangsu Modern Industrial Fermentation.

We declare that we have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## REFERENCES

1. Castro SP, Cleland EE, Wagner R, Sawad RA, Lipson DA. 2019. Soil microbial responses to drought and exotic plants shift carbon metabolism. *ISME J* 13:1776–1787. <https://doi.org/10.1038/s41396-019-0389-9>.
2. Cornforth DM, Diggle FL, Melvin JA, Bomberger JM, Whiteley M. 2020. Quantitative framework for model evaluation in microbiology research using *Pseudomonas aeruginosa* and cystic fibrosis infection as a test case. *mBio* 11:e03042-19. <https://doi.org/10.1128/mBio.03042-19>.
3. Wolfe BE, Button JE, Santarelli M, Dutton RJ. 2014. Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. *Cell* 158:422–433. <https://doi.org/10.1016/j.cell.2014.05.041>.
4. Sielaff AC, Urbaniak C, Mohan GBM, Stepanov VG, Tran Q, Wood JM, Minich J, McDonald D, Mayer T, Knight R, Karouia F, Fox GE, Venkateswaran K. 2019. Characterization of the total and viable bacterial and fungal communities associated with the International Space Station surfaces. *Microbiome* 7:50. <https://doi.org/10.1186/s40168-019-0666-x>.
5. Gohli J, Boifot KO, Moen LV, Pastuszek P, Skogan G, Udekwu KI, Dybdal M. 2019. The subway microbiome: seasonal dynamics and direct comparison of air and surface bacterial communities. *Microbiome* 7:160. <https://doi.org/10.1186/s40168-019-0772-9>.
6. Tong X, Leung MHY, Wilkins D, Cheung HHL, Lee PKH. 2019. Neutral processes drive seasonal assembly of the skin mycobiome. *mSystems* 4:e00004-19. <https://doi.org/10.1128/mSystems.00004-19>.
7. Koizumi T, Nara K. 2020. Ectomycorrhizal fungal communities in ice-age relict forests of *Pinus pumila* on nine mountains correspond to summer temperature. *ISME J* 14:189–201. <https://doi.org/10.1038/s41396-019-0524-7>.
8. Tringe SG, Rubin EM. 2005. Metagenomics: DNA sequencing of environmental samples. *Nat Rev Genet* 6:805–814. <https://doi.org/10.1038/nrg1709>.
9. Oulas A, Pavloudi C, Polymenakou P, Pavlopoulos GA, Papanikolaou N, Kotoulas G, Arvanitidis C, Iliopoulos I. 2015. Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinform Biol Insights* 9:75–88. <https://doi.org/10.4137/BBI.S12462>.
10. Almeida A, Shao Y. 2018. Genome watch: keeping tally in the microbiome. *Nat Rev Microbiol* 16:124. <https://doi.org/10.1038/nrmicro.2018.13>.
11. Props R, Kerckhof FM, Rubbens P, De Vrieze J, Sanabria EH, Waegeman W, Monsieurs P, Hammes F, Boon N. 2017. Absolute quantification of microbial taxon abundances. *ISME J* 11:584–587. <https://doi.org/10.1038/ismej.2016.117>.
12. Guo X, Zhang X, Qin Y, Liu Y, Zhang J, Zhang N, Wu K, Qu B, He Z, Wang X, Zhang X, Hacquard S, Fu X, Bai Y. 2020. Host-associated quantitative abundance profiling reveals the microbial load variation of root microbiome. *Plant Commun* 1:100003. <https://doi.org/10.1016/j.xplc.2019.100003>.
13. Wang X, Wang M, Xie X, Guo S, Zhou Y, Zhang X, Yu N, Wang E. 2020. An amplification-selection model for quantified rhizosphere microbiota assembly. *Science Bull* 65:983–986. <https://doi.org/10.1016/j.scib.2020.03.005>.
14. Tkacz A, Hortal M, Poole PS. 2018. Absolute quantitation of microbiota abundance in environmental samples. *Microbiome* 6:110. <https://doi.org/10.1186/s40168-018-0491-7>.
15. Yang L, Lou J, Wang H, Wu L, Xu J. 2018. Use of an improved high-throughput absolute abundance quantification method to characterize soil bacterial community and dynamics. *Sci Total Environ* 633:360–371. <https://doi.org/10.1016/j.scitotenv.2018.03.201>.
16. Vieira-Silva S, Sabino J, Valles-Colomer M, Falony G, Kathagen G, Caenepeel C, Cleyne I, van der Merwe S, Vermeire S, Raes J. 2019. Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. *Nat Microbiol* 4:1826–1831. <https://doi.org/10.1038/s41564-019-0483-9>.
17. Vandeputte D, Kathagen G, D'Hoe K, Vieira-Silva S, Valles-Colomer M, Sabino J, Wang J, Tito RY, De Commer L, Darzi Y, Ermeire SV, Falony G, Raes J. 2017. Quantitative microbiome profiling links gut community variation to microbial load. *Nature* 551:507–511. <https://doi.org/10.1038/nature24460>.
18. Awasthi MK, Duan Y, Awasthi SK, Liu T, Zhang Z. 2020. Effect of biochar



- and bacterial inoculum additions on cow dung composting. *Bioresour Technol* 297:122407. <https://doi.org/10.1016/j.biortech.2019.122407>.
19. Wang S, Xiong W, Wang Y, Nie Y, Wu Q, Xu Y, Geisen S. 2020. Temperature-induced annual variation in microbial community changes and resulting metabolome shifts in a controlled fermentation system. *mSystems* 5:e00555-20. <https://doi.org/10.1128/mSystems.00555-20>.
  20. Liu J, Wu Q, Wang P, Lin JC, Huang L, Xu Y. 2017. Synergistic effect in core microbiota associated with sulfur metabolites in spontaneous Chinese liquor fermentation. *Appl Environ Microbiol* 83:e01475-17. <https://doi.org/10.1128/AEM.01475-17>.
  21. Wang S, Wu Q, Nie Y, Wu J, Xu Y. 2019. Construction of synthetic microbiota for reproducible flavor compound metabolism in Chinese light-aroma-type liquor produced by solid-state fermentation. *Appl Environ Microbiol* 85:e03090-18. <https://doi.org/10.1128/AEM.03090-18>.
  22. Stämmler F, Gläsner J, Hiergeist A, Holler E, Weber D, Oefner PJ, Gessner A, Spang R. 2016. Adjusting microbiome profiles for differences in microbial load by spike-in bacteria. *Microbiome* 4:28. <https://doi.org/10.1186/s40168-016-0175-0>.
  23. Schmidt H, Eickhorst T. 2014. Detection and quantification of native microbial populations on soil-grown rice roots by catalyzed reporter deposition-fluorescence *in situ* hybridization. *FEMS Microbiol Ecol* 87:390–402. <https://doi.org/10.1111/1574-6941.12232>.
  24. Ji BW, Sheth RU, Dixit PD, Huang Y, Kaufman A, Wang HH, Vitkup D. 2019. Quantifying spatiotemporal variability and noise in absolute microbiota abundances using replicate sampling. *Nat Methods* 16:731–736. <https://doi.org/10.1038/s41592-019-0467-y>.
  25. Jiang S, Yu Y, Gao R, Wang H, Zhang J, Li R, Long X, Shen Q, Chen W, Cai F. 2019. High-throughput absolute quantification sequencing reveals the effect of different fertilizer applications on bacterial community in a tomato cultivated coastal saline soil. *Sci Total Environ* 687:601–609. <https://doi.org/10.1016/j.scitotenv.2019.06.105>.
  26. Tourlousse DM, Yoshiike S, Ohashi A, Matsukura S, Noda N, Sekiguchi Y. 2017. Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon sequencing. *Nucleic Acids Res* 45:e23. <https://doi.org/10.1093/nar/gkw984>.
  27. Quinn TP, Erb I, Richardson MF, Crowley TM. 2018. Understanding sequencing data as compositions: an outlook and review. *Bioinformatics* 34:2870–2878. <https://doi.org/10.1093/bioinformatics/bty175>.
  28. Li S, Wang P, Chen Y, Wilson M, Yang X, Ma C, Lu J, Chen X, Wu J, Shu W, Jiang L. 2020. Island biogeography of soil bacteria and fungi: similar patterns, but different mechanisms. *ISME J* 14:1886–1896. <https://doi.org/10.1038/s41396-020-0657-8>.
  29. Zhao S, Lieberman TD, Poyet M, Kauffman KM, Gibbons SM, Groussin M, Xavier RJ, Alm EJ. 2019. Adaptive evolution within gut microbiomes of healthy people. *Cell Host Microbe* 25:656–667.e8. <https://doi.org/10.1016/j.chom.2019.03.007>.
  30. Yu K, Yi S, Li B, Guo F, Peng X, Wang Z, Wu Y, Alvarez-Cohen L, Zhang T. 2019. An integrated meta-omics approach reveals substrates involved in synergistic interactions in a bisphenol A (BPA)-degrading microbial community. *Microbiome* 7:16. <https://doi.org/10.1186/s40168-019-0634-5>.
  31. Kling JD, Lee MD, Fu F, Phan MD, Wang X, Qu P, Hutchins DA. 2020. Transient exposure to novel high temperatures reshapes coastal phytoplankton communities. *ISME J* 14:413–424. <https://doi.org/10.1038/s41396-019-0525-6>.
  32. Karhu K, Auffret MD, Dungait JA, Hopkins DW, Prosser JI, Singh BK, Subke JA, Wookey PA, Agren GI, Sebastia MT, Gouriveau F, Bergkvist G, Meir P, Nottingham AT, Salinas N, Hartley IP. 2014. Temperature sensitivity of soil respiration rates enhanced by microbial community response. *Nature* 513:81–84. <https://doi.org/10.1038/nature13604>.
  33. Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G, Djahanschiri B, Zeller G, Mende DR, Alberti A, Cornejo-Castillo FM, Costea PI, Cruaud C, d'Ovidio F, Engelen S, Ferrera I, Gasol JM, Guidi L, Hildebrand F, Kokoszka F, Lepoivre C, Lima-Mendez G, Poulain J, Poulos BT, Royo-Llonch M, Sarmiento H, Vieira-Silva S, Dimier C, Picheral M, Searson S, Kandel-Lewis S, Tara Oceans coordinators, Bowler C, de Vargas C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Jaillon O, Not F, Ogata H, Pesant S, Speich S, Stemmann L, Sullivan MB, Weissenbach J, Wincker P, Karsenti E, Raes J, Acinas SG, Bork P. 2015. Ocean plankton. Structure and function of the global ocean microbiome. *Science* 348:1261359. <https://doi.org/10.1126/science.1261359>.
  34. Wang P, Wu Q, Jiang X, Wang Z, Tang J, Xu Y. 2017. *Bacillus licheniformis* affects the microbial community and metabolic profile in the spontaneous fermentation of *Daqu* starter for Chinese liquor making. *Int J Food Microbiol* 250:59–67. <https://doi.org/10.1016/j.jfoodmicro.2017.03.010>.
  35. Zhang X, Tian X, Ma L, Feng B, Liu Q, Yuan L, Fan C, Huang H, Huang H, Yang Q. 2015. Biodiversity of the symbiotic bacteria associated with toxic marine dinoflagellate *Alexandrium tamarense*. *J Biosci Med (Irvine)* 3:23–28. <https://doi.org/10.4236/jbm.2015.36004>.
  36. Toju H, Tanabe AS, Yamamoto S, Sato H. 2012. High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PLoS One* 7:e40863. <https://doi.org/10.1371/journal.pone.0040863>.
  37. Hu X, Wang H, Wu Q, Xu Y. 2014. Development, validation and application of specific primers for analyzing the clostridial diversity in dark fermentation pit mud by PCR-DGGE. *Bioresour Technol* 163:40–47. <https://doi.org/10.1016/j.biortech.2014.04.008>.
  38. Liu C, Feng S, Wu Q, Huang H, Chen Z, Li S, Xu Y. 2019. Raw material regulates flavor formation via driving microbiota in Chinese liquor fermentation. *Front Microbiol* 10:1520. <https://doi.org/10.3389/fmicb.2019.01520>.
  39. Jiang J, Liu Y, Li H, Yang Q, Wu Q, Chen S, Tang J, Xu Y. 2019. Modeling and regulation of higher alcohol production through the combined effects of the C/N ratio and microbial interaction. *J Agric Food Chem* 67:10694–10701. <https://doi.org/10.1021/acs.jafc.9b04545>.
  40. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley R, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunencko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f.303>.
  41. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27:2194–2200. <https://doi.org/10.1093/bioinformatics/btr381>.
  42. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>.
  43. Tao Y, Li J, Rui J, Xu Z, Zhou Y, Hu X, Wang X, Liu M, Li D, Li X. 2014. Prokaryotic communities in pit mud from different-aged cellars used for the production of Chinese strong-flavored liquor. *Appl Environ Microbiol* 80:2254–2260. <https://doi.org/10.1128/AEM.04070-13>.
  44. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998. <https://doi.org/10.1038/nmeth.2604>.
  45. Bokulich NA, Thorngate JH, Richardson PM, Mills DA. 2014. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. *Proc Natl Acad Sci U S A* 111:E139–148. <https://doi.org/10.1073/pnas.1317377110>.