Check for updates

High-throughput cultivation and identification of bacteria from the plant root microbiota

Jingying Zhang [1,2,3,7], Yong-Xin Liu [1,2,3,7], Xiaoxuan Guo[1,2,3], Yuan Qin[1,2,3,4], Ruben Garrido-Oter [5,6 ✉], Paul Schulze-Lefert [5,6 ✉] and Yang Bai [1,2,3,4 ✉]

Cultivating native bacteria from roots of plants grown in a given environment is essential for dissecting the functions of the root microbiota for plant growth and health with strain-specific resolution. In this study, we established a straightforward protocol for high-throughput bacterial isolation from fresh root samples using limiting dilution to ensure that most cultured bacteria originated from only one microorganism. This is followed by strain characterization using a two-sided barcode polymerase chain reaction system to identify pure and heterogeneous bacterial cultures. Our approach overcomes multiple difficulties of traditional bacterial isolation and identification methods, such as obtaining bacteria with diverse growth rates while greatly increasing throughput. To facilitate data processing, we developed an easy-to-use bioinformatic pipeline called 'Culturome' (https://github.com/YongxinLiu/Culturome) and a graphical user interface web server (http://bailab.genetics.ac.cn/culturome/). This protocol allows any research group (two or three lab members without expertise in bioinformatics) to systematically cultivate root-associated bacteria within 8–9 weeks.

Introduction

Plant roots assemble a taxonomically structured microbial community known as the root microbiota[1]. Marker gene-based amplicon profiling and metagenomic sequencing have been used to describe the taxonomic composition and gene content of the root microbiota of a broad diversity of plants, including model and crop species[2–6]. Recent functional and mechanistic studies have shed light on services provided by root commensal microorganisms to the host[7–11]. For example, the root bacterial microbiota protect plants from soil-borne fungal and oomycete pathogens[7,8] and participate in nutrient use[12]. The host plant selectively regulates its root microbiota, which, in turn, influences diverse plant physiological processes[8,10,13–18]. Cultivated microorganisms from the root microbiota have served as a core resource for these functional and mechanistic studies.

To investigate the functions of the root microbiota in a given environment, it is necessary to isolate and culture microorganisms from root samples of plants grown in the same environment (e.g., the same soil type). Although thousands of cultivated microbes have been isolated and are stored in international stock centers, these microbes come from a wide variety of environments and host species. Usually there is not enough information to determine whether a cultivated set of microbes is a representative cross-section of the microbial communities that colonize a given niche. In addition, microbial adaptation is expected to contribute to genotypic and functional differentiation among microbes sampled from different habitats and hosts. For these reasons, it is often necessary to create new stocks of bacteria from a given natural soil and host species combination to dissect the interactions between plants and the root microbiota in this system[19–22]. The lack of a rapid, efficient tool for microbial cultivation has made it difficult for individual research groups to obtain comprehensive culture collections from their root microbiota samples of plants grown in different environmental conditions.

When isolating and identifying root-associated bacteria with traditional methods, such as colony picking and Sanger sequencing, many research groups face difficulties, including: (i) obtaining bacteria with diverse growth rates using reasonable time and cost. For instance, fast-growing bacteria
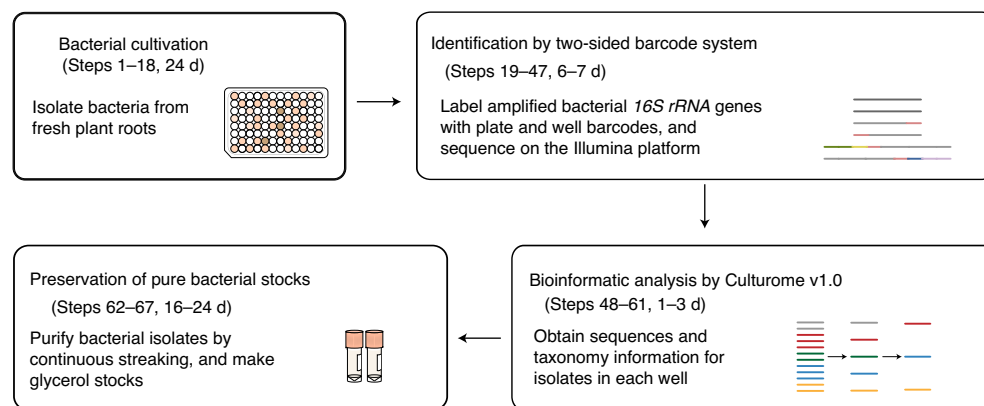
Fig. 1 | Overview of the high-throughput bacterial cultivation and identification system.

might inhibit the growth of slow-growing ones, often resulting in preferential isolation of fast growers. (ii) High-throughput identification of cultivated bacteria that are pure or associate with other bacteria. Sanger sequencing requires pure DNA templates of clonal cultures and is inefficient for obtaining sequence information from groups of bacteria that intimately associate with each other. (iii) Efficient processing and evaluation of bacterial identification data. To overcome these difficulties and obtain individual bacterial cultures from a given root sample, we established an integrated protocol for high-throughput bacterial cultivation, identification and bioinformatic data analysis (Fig. 1). First, to streamline the tedious process of colony picking and repeatedly obtaining fast-growing bacteria, we isolated the bacteria in 96-well cell culture plates by performing a limiting dilution of the root microbiota sample with liquid media. These dilutions were performed under the assumption that bacteria present in the dilution are separated into different wells following a Poisson distribution. The dilution factors were chosen in such a way that ~30% of the wells showed visible bacterial growth after an incubation period, a condition predicted to maximize the total number of wells containing clonal cultures[23]. Second, we adapted a one-sided barcode polymerase chain reaction (PCR) system targeting bacterial *16S rRNA* gene sequences based on 454 sequencing and cultivated 54–65% of bacterial taxa that reproducibly associated with roots of *Arabidopsis* plants grown in natural soil[9]. We recently reported an improved workflow using a two-sided barcode PCR system based on Illumina technology to increase sequencing depth and accuracy by avoiding the formation of chimeric sequences obtained in one-sided barcode design. Using the updated protocol, we cultivated 69–72% of root bacterial taxa in field-grown rice and found that rice plants coordinate root microbiota to use soil nitrogen[12]. Finally, to facilitate analyzing cultivation data, we developed an easy-to-use bioinformatic pipeline and a graphical user interface web server to evaluate the cultivation outcome and cross-reference cultivated bacteria with the corresponding root microbiota profiling. This pipeline is called 'Culturome' (https://github.com/YongxinLiu/Culturome); it can be used from the local computer or via a graphical user interface web server (http://bailab.genetics.ac.cn/culturome/). Our protocol allows any research group to cultivate bacteria from local samples at an affordable cost and in a short amount of time (8–9 weeks; two or three lab members without expertise in bioinformatics).

Applications of the method

The protocol presented here can be used to establish a taxonomically comprehensive bacterial culture collection, including beneficial, commensal and pathogenic bacteria, from a native root microbiota sample from a given environment and plant species. The cultured bacteria can be used to study the functions and assembly of the root microbiota under normal or stressed conditions with various gnotobiotic plant growth systems. After evaluating the bacterial diversity of a root sample by *16S rRNA* gene amplicon sequencing, cultured bacteria with corresponding *16S rRNA* gene sequences can be selected. Co-cultivation of such synthetic bacterial communities of varying complexity (i.e., different numbers of cultured bacterial strains) with axenic plants in germ-free soil, clay or agar matrices can reveal physiological functions of the root microbiota for the plant host. For example, the phenotypic outputs of axenic plants and plants inoculated with synthetic bacterial communities can be compared if the plants are cultivated under defined abiotic or biotic stress and standard growth conditions. Standardized synthetic community approaches provide opportunities to reproduce root

microbiota patterns and to investigate molecular mechanisms underlying plant–microbiota interactions[19]. Our protocol can be modified to cultivate members of the bacterial microbiota from other plant tissues (e.g., leaves and seeds) as well as other ecological systems, such as soil or food.

Comparison to other methods

Approaches such as colony picking[24–26], cell sorting[27,28] and in situ cultivation[29,30] that have been previously used to culture bacteria from a variety of environmental sources are also suitable for the isolation of bacterial strains, although the latter two approaches demand specialized equipment. Our protocol has the following advantages compared to other approaches. First, it can be performed using common laboratory equipment and does not require the use of expensive equipment, such as fluorescence-activated cell sorting instruments. Second, our protocol has a higher throughput than colony picking. The isolation procedure in the protocol can be completed in a single day, whereas colony picking can require several weeks of continuous colony isolation, depending on colony growth speed. Third, this protocol avoids the problem of repeatedly obtaining fast-growing bacteria using the colony picking method. Bacteria are separated in different wells in our protocol, thereby preventing fast-growing bacteria from inhibiting the growth of slow-growing bacteria, which can occur when bacteria are cultivated on agar plates. Fourth, the two-sided barcode PCR system is based on Illumina sequencing technology, which is a more accurate, high-throughput procedure than the previous identification process using 454 pyrosequencing[9]. Finally, to facilitate summarization of the cultivation results, we provide an easy-to-use bioinformatic pipeline and a graphical user interface web server. Taken together, our protocol is simpler and more efficient than previous methods for individual research groups to study the functions and mechanisms of the root microbiota.

Limitations of this approach

This approach can be used only for bacteria and is not suitable for use with filamentous eukaryotes such as fungi or oomycetes that can be isolated using selective medium and single hyphal tip picking[7]. It is also unsuitable for use with bacteria that do not grow in liquid media. This approach is not designed to cultivate bacteria that require anaerobic conditions, such as anaerobic archaea, for which an anaerobic laminar flow hood is needed.

Experimental design

Prepare fresh root microbiota samples

To capture the taxonomic and functional diversity of root-associated bacteria in our lab, we culture bacteria from fresh roots of plants growing in natural soil (e.g., using 6-week-old *Arabidopsis thaliana* or 8-week-old rice (*Oryza sativa* L.), when the root microbiota are fully differentiated[2,3,5,11]; Fig. 2a). There are several considerations:

- It is better to use fresh root samples than frozen root samples to capture bacteria that are sensitive to freezing[31].
- Because the root microbiota vary depending on the environmental conditions (e.g., soil type and nutrient availability) and host plants features (e.g., plant genotype and developmental stage), we should culture bacteria from the roots of a plant species at a specific developmental stage and plant growth condition that is relevant to the scientific question being addressed. This maximizes the chances of detecting the functionally relevant commensal bacteria. Further, by cross-referencing the *16S rRNA* gene sequences of the cultured bacteria with the *16S rRNA* gene sequences obtained from culture-independent amplicon sequencing of root samples, we can compose synthetic bacterial communities that mimic the taxonomic bacterial diversity present in root samples.
- Because bacteria harboring the same *16S rRNA* gene sequences often have different gene repertoires[32], it is generally required to isolate multiple strains with the same *16S rRNA* gene sequence to assess the intergenomic heterogeneity of a specific bacterial taxon. The most efficient way to achieve this is to isolate bacteria with the same *16S rRNA* gene but from independent plant individuals[9,12]. Bacteria isolated in this way arise from independent root colonization events, increasing the probability of capturing intergenomic heterogeneity[9].

Selection of culture media

It is important to select a proper medium for bacterial cultivation, as different bacteria might require different nutrients for growth[33–37]. Compared to unplanted soil, where organic carbon is normally rate limiting for bacterial growth[38], exudation of soluble photoassimilates from living roots and
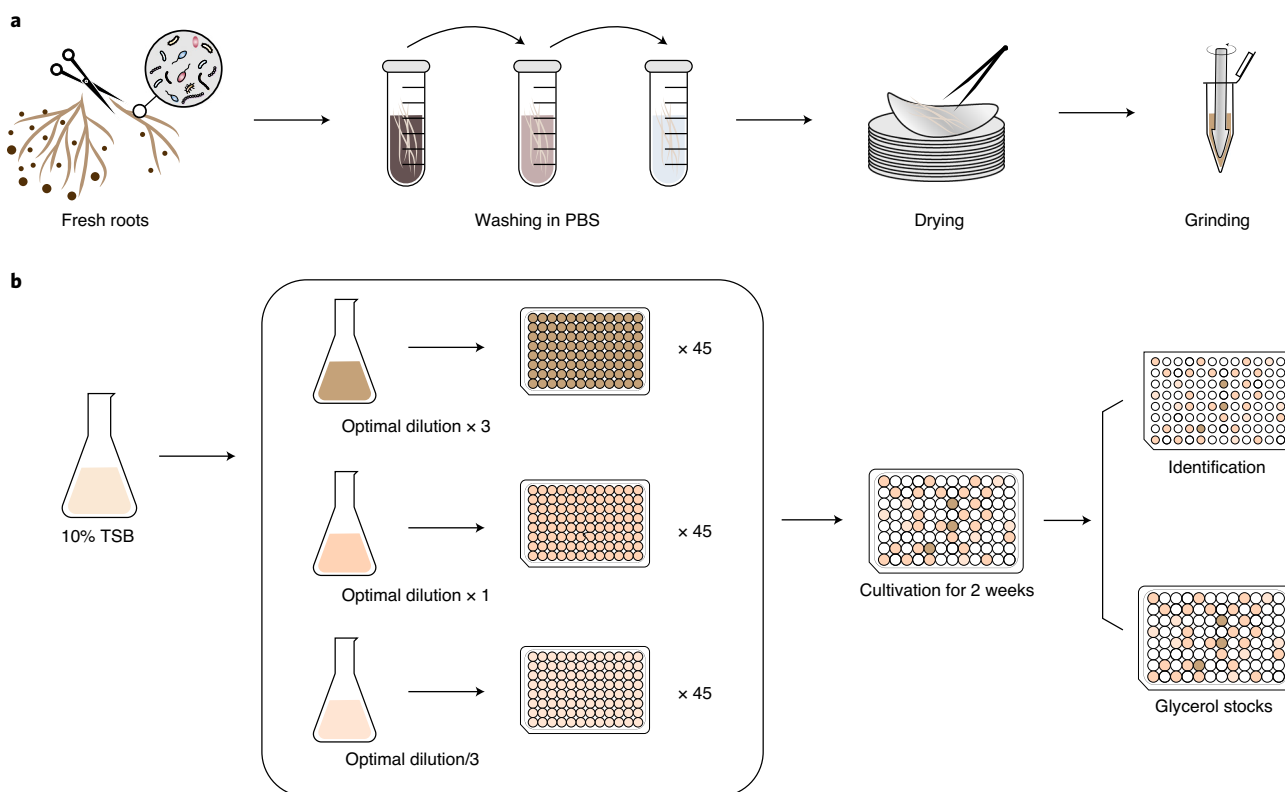
Fig. 2 | High-throughput bacterial isolation from root microbiota samples. a, Preparation of root samples for bacterial isolation. Fresh roots are washed three times in 1× PBS, dried with sterile filter paper, cut into 2-mm pieces and smashed. **b**, The limiting dilution method and incubation. The smashed root materials are diluted in 10% TSB medium to the optimal dilution centration based on the preliminary cultivation experiment. The dilutions should include the optimal dilution concentration (ODC) at 3× and 1/3× to account for experimental variation. The 96-well cell culture plates containing diluted bacteria are incubated at room temperature for 2 weeks for further identification and to create glycerol stocks.

complex insoluble polysaccharides that constitute plant cell walls render the root a carbon- and nutrient-rich niche for colonization by soil-dwelling bacteria. Tryptic soy broth (TSB) medium is sufficient to obtain most abundant root-associated bacteria (relative abundance greater than 0.1%)[9,12]. We used ten-fold diluted media to slow the growth of fast-growing bacteria. In addition, low-nutrient media, such as Reasoner's 2A, help capture bacteria that prefer to grow in oligotrophic enviroments[39]. If specific phylogenomic groups should be targeted, media designed to enrich for such taxa can be selected (e.g., TWYE for Actinobacteria and T3 for *Bacillus*)[40,41].

Cultivating root-associated bacteria with the limiting dilution method

Traditionally, bacteria are isolated by picking colonies from cultures growing in agar medium in plates[20,24,35]. Using this process, fast-growing bacteria tend to be over-represented because they inhibit the growth of slow-growing community members, which reduces the diversity of cultivated isolates. In addition, because different bacteria have different growth rates, researchers must continuously pick colonies depending on bacterial growth, which is tedious. To increase the isolation efficiency, we use the limiting dilution method to distribute root homogenates diluted using liquid media into 96-well cell culture plates, allowing fast-growing and slow-growing bacteria to be separated in different wells. The entire procedure can be completed in a single day. Simultaneous isolation of a large number of bacterial strains increases the likelihood of culturing community members found in low abundances in the original environment, also known as the rare biota. We use a concentration of diluted root homogenates that allows ~30% of the wells to show bacterial growth. In this case, most wells showing bacterial growth contain cultures originating from a single bacterial cell based on the assumption that the bacteria are distributed into different wells of 96-well cell culture plates following Poisson distribution[23]. In our experience, the dilution allowing 30~50% of wells showing bacterial

growth still generated a substantial proportion of pure cultures and is sufficient to obtain clonal cultures by further purification[12].

To determine the optimal dilution of fresh root samples to allow ~30% of wells to show bacterial growth, a preliminary limiting dilution experiment should be performed using a wide range of dilution rates of fresh root samples from plants grown under the same conditions used in high-throughput isolation. The optimal dilution should contain two culturable bacteria per milliliter (~33 bacteria in 96 wells, 160 μl of medium per well). A preliminary limiting dilution experiment should be performed if the protocol is optimized to isolate bacteria from other plant tissues or plants grown in different geographic locations or soil types.

The primer design of two-sided barcode PCR

We determined the *16S rRNA* gene sequences of the V5–V7 gene regions in cultivated bacteria with 799F and 1193R primers and cross-referenced these with *16S rRNA* gene sequences generated by *16S rRNA* gene amplicon sequencing of the same variable regions from root samples. The 799F primer is widely used to exclude amplification of plant plastids, as the primer does not amplify plant chloroplast DNA and produces larger PCR products from plant mitochondrial DNA compared to the bacterial *16S rRNA* gene[42]. For high-throughput identification of cultivated bacteria, we introduced well and plate barcodes to the forward and reverse primers, respectively (Supplementary Table 1). We added 96 sets of 10-base pair (bp) barcodes into the forward primer 799F to label individual wells in each 96-well plate. We introduced 48 sets of 6-bp barcodes into the reverse primer 1193R to label individual 96-well plates. The well and plate barcodes were designed with more than one bp difference in order to be tolerant to sequencing errors. We have tested the compatibility of *16S rRNA* gene primers and barcodes to ensure the PCR efficiency.

Identification of cultivated bacteria using a two-sided barcode PCR system

Bacteria isolated from natural samples tend to associate with other bacteria, resulting in problems during identification using Sanger sequencing, which requires pure templates. To identify all bacterial taxa in wells containing one or more bacteria in a high-throughput manner, we applied the two-sided barcode PCR system targeting bacterial *16S rRNA* gene sequences, followed by Illumina sequencing. First, the two-sided barcode system introduces the plate and well barcode sequences into PCR products in each well in a single step. Compared to protocols requiring two steps to introduce well and plate barcodes separately, our protocol is more convenient and avoids the formation of chimeric sequences from mislabeled barcodes when amplifying barcoded sequences. Second, different bacterial strains have different growth rates and DNA extraction efficiencies[43]. Bacterial DNAs with low concentration caused by slow growth or low efficiency of DNA extraction are not readily detected by Illumina sequencing, as substantially fewer PCR amplicons are generated from these bacteria compared to the others. To circumvent this problem, we first amplify the *16S rRNA* gene in DNAs extracted from each well using degenerate primers (799F and 1193R) without barcodes and adapters to increase amplification efficiency and obtain saturated PCR products. These saturated PCR products are then used as starting templates for the second round of PCR using primers connected to barcodes for labeling both plates and wells. In the end, different wells tend to generate similar amounts of PCR products, resulting in similar sequencing depths for bacteria in different wells.

Bioinformatic pipeline to analyze amplicon data from cultivated bacteria

The large amount of data generated from amplicon sequencing might hinder scientific progress owing to limited bioinformatic expertise and the need for expensive computational facilities[44–48]. To address this potential difficulty, we created an easy-to-use bioinformatic pipeline and a graphical user interface web server to help researchers analyze their sequencing data. The 'Culturome' pipeline (https://github.com/YongxinLiu/Culturome) and the graphical user interface web server (http://bailab.genetics.ac.cn/culturome/) were assembled using a combination of QIIME[49,50], USEARCH[51], VSEARCH[52] and in-house scripts[12,53] to assess sequence quality, split barcoded sequences, remove sequencing background and identify amplicon sequence variants (ASVs). We analyze sequences in each well at both the ASV and genus levels, as pure bacteria might contain *16S rRNA* gene variations, resulting in different ASVs. Thus, wells containing more than one ASV belonging to a single genus might contain pure cultures. Finally, we cross-reference the cultivated bacteria with the corresponding root microbiota profiling data using *16S rRNA* gene similarity, allowing for targeted

isolation of bacteria from relevant taxa (e.g., with high abundance or occupancy in natural communities).

Bacterial stocks and purification
To recover cultivated bacteria for functional and mechanistic studies, it is essential to generate bacterial stocks. We generate two types of stocks in 40% (vol/vol) glycerol for storage at −80 °C: (i) all bacterial strains in 96-well cell culture plates and (ii) high-quality stocks from selected bacteria after three rounds of continuous streaking. These two types of stocks serve different purposes. Bacterial stocks in 96-well cell culture plates are used as backups for purified high-quality stocks and as resources containing all isolates for successive generation of additional purified high-quality stocks. These bacteria must be purified by repeated streaking on agar plates containing the same medium before use because a substantial proportion of wells might still contain more than one bacterial strain, although most wells showing bacterial growth after limiting dilution are clonally pure based on the expectation from a Poisson distribution and further validated by the sequencing results. High-quality stocks can be subsequently used for whole-genome sequencing and functional studies. To maximize the genomic and phenotypic diversity of the bacterial stocks, we select and purify isolates with the same *16S rRNA* gene sequence but from different plants, which represent different colonization events.

Controls
We include two controls in the protocol. First, to ensure that the cultivated bacteria were exclusively derived from root microbiota samples, we distribute sterile medium in 96-well cell culture plates during the limiting dilution step as negative controls. Second, to test the PCR system and remove sequencing background when identifying cultivated bacteria, we amplify DNAs from a pure *Escherichia coli* culture and nuclease-free water as positive and negative controls, respectively.

Materials

Biological materials
• Roots were obtained from rice (*O. sativa* L. Nipponbare) or *A. thaliana* (Columbia-0). The plant materials are available from the authors upon request.

Reagents
• Tryptic soy broth (TSB) (Oxoid, cat. no. CM0129)
• Magnesium chloride hexahydrate (MgCl$_2$·6H$_2$O, Sigma-Aldrich, cat. no. 63068-250G)
• Sodium chloride (NaCl, Sigma-Aldrich, cat. no. S7653-250G)
• Sodium phosphate dibasic dihydrate (Na$_2$HPO$_4$·2H$_2$O, Sigma-Aldrich, cat. no. 04272-1KG)
• Sodium phosphate monobasic monohydrate (NaH$_2$PO$_4$·H$_2$O, Sigma-Aldrich, cat. no. 71504-250G)
• Sodium hydroxide (NaOH, Sigma-Aldrich, cat. no. S8045-500G)
• Na$_2$-EDTA·2H$_2$O (Amresco, cat. no. 0105)
• Tris-HCl (Genview, cat. no. BT504-500G)
• Glycerol (Sigma-Aldrich, cat. no. G7893)
• Agar (Sigma-Aldrich, cat. no. A7921-100G)
• Agarose (Biowest, cat. no. G-10)
• Ethanol (Sigma-Aldrich, cat. no. E7023-500ML) **! CAUTION** Ethanol is highly flammable. Keep it away from open flame.
• Nuclease-free water (Qiagen, cat. no. 129115) **▲ CRITICAL** Many brands of nuclease-free water exist, but their PCR performance highly vary because microbial DNAs are detected. We have had good experience with nuclease-free water from Qiagen, which is critical to avoid microbial contamination and negative control.
• PCR primers (Life Technologies; 799F and 1193R are purified by high-performance liquid chromatography; barcoded primers are purified by polyacrylamide gel electrophoresis) are listed in Supplementary Table 1.
• TaKaRa Taq Hot Start Version (Takara, cat. no. R007Z) **▲ CRITICAL** To avoid contamination of microbial DNAs in reagents, it is important to use a polymerase without microbial DNAs. We have had good experience with TaKaRa Taq Hot Start.
• 2,000-bp Plus DNA Ladder (DiNing, cat. no. DM1003)

- Midori Green Advance DNA Stain (Nippon Genetics Europe GmbH, cat. no. MG04) **! CAUTION** This stain can cause serious skin irritation. Handle it with gloves.
- Wizard SV Gel and PCR Clean-up System (Promega, cat. no. A9282)
- Agencourt AMPure XP beads (Beckman, cat. no. A63882) **! CAUTION** Beads contain sodium azide, which might contact heavy metals and form explosive compounds. Handle with appropriate safety equipment.
- Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, cat. no. P7589) including Quant-iT PicoGreen dsDNA reagent, 20× TE and Lambda DNA Standard.
- Tris (Vetec, cat. no. V900483-500G)
- Acetic acid (Rhawn, cat. no. R049946-500ML)
- *E. coli* (Tiangen Biotech, DH5α, cat. no. CB101-03)
- Loading buffer (Takara, cat. no. 9156)

Equipment
- Forceps (Jinzhong, cat. no. JD5020)
- Scissor (Jinzhong, cat. no. J21130)
- Scalpel (Jinzhong, cat. no. J11010)
- Filter papers (SEP, cat. no. DXLZ11F)
- Falcon tubes, polypropylene, 50 ml (BD Falcon, cat. no. 352070)
- Standard microcentrifuge tubes (1.5 and 5 ml; Eppendorf, cat. nos. 0030125.150, 30119401)
- Plastic grinding rods (Huaaobio, cat. no. SLXMB-1.5)
- 13-cm square Petri dish (Axygen, cat. no. ASJ-17-9142)
- 60-mm culture plate (Corning, cat. no. 430166)
- 96-well PCR plate (Jet Keen Biotechnology, cat. no. PC-0200-9B)
- 96-well cell culture plate (NEST, cat. no. 701001)
- 96-well EIA/PIA plate (Costar, cat. no. 3590)
- 2.0-ml cryogenic vial (NEST, cat. no. 607001)
- Microbank (Pro Lab, cat. no. PL.170/M)
- Multi 12-channel pipettes (10, 100, 300 μl; Eppendorf, cat. nos. 3122000027, 3122000043 and 3122000060)
- Single-channel pipettes (10, 20, 100, 200 and 1,000 μl; Eppendorf, cat. nos. 3120000020, 3120000038, 3120000046, 3120000054 and 3120000062)
- Pipette tips (nuclease-free; 10, 200 and 1,000 μl; Axygen, cat. nos. T-300, T-200-Y and T-1000-B)
- Parafilm (Bemis, cat. no. PM-996)
- Platemax Cycler Seal Sealing Film (Axygen, cat. no. PCR-TS)
- Class II biological safety cabinet (Thermo Fisher Scientific, cat. no. BSC-1000IIA2)
- TGyrate vortex basic (Tiagen Biotech, cat. no. OSE-VS-01)
- Electronic balance (Mettler Toledo, cat. no. AL104)
- Orbital shaker (Kylin-Bell Lab Instruments, model no. TS-2)
- Incubator-shaker (Shanghai Zhichu Instrument, cat. no. ZQZY-CF)
- Microplate centrifuge (Tiangen Biotech, cat. no. OSE-MP26)
- Thermal cycler (T100; Bio-Rad, model no. 1861096)
- Electrophoresis system (Junyi-Dongfang, cat. no. JY300HC)
- Gel Doc XR+ Imaging system (Bio-Rad, model no. Universal Hood II)
- SpectraMax Paradigm Multi-Mode Microplate Detection Platform (Molecular Devices, model no: Paradigm)
- UV-visible light spectrophotometer (Thermo Fisher Scientific, model no. NanoDrop ND-2000)
- High-speed freezing centrifuge (Eppendorf, model no. 5810R)
- DynaMag-2 magnet (Thermo Fisher Scientific, cat. no. 12321D)
- Tabletop high-speed micro centrifuge (Thermo Fisher Scientific, model no. Heraeus Pico 17)
- Ultra-low temperature freezer (Thermo Fisher Scientific, model no. 907)
- –20 °C medical cryogenic refrigerator (Haier, cat. no. DW-40L92)

Computer requirements
- Operating system requirements: 64-bit Linux, such as Ubuntu 16.04+ or CentOS 7.5+
- Hardware requirements: >4 GB of RAM and 30 GB of free disk space

Software
- Conda 4.8.3+, software management system (https://repo.anaconda.com/miniconda)
- R 3.4+ (https://cran.r-project.org). The analysis was performed using R 3.5.1 with ggplot2 (ref. [54]), pheatmap[55], dplyr and stringr packages installed
- QIIME v1.9.1 (ref. [49]) http://qiime.org for split barcodes
- USEARCH v10.0.240 (ref. [51]) http://www.drive5.com/usearch
- VSEARCH v2.7.1+[52] or later https://github.com/torognes/vsearch
- GraPhlAn 1.1.3-1+[56], for visualizing taxonomic tree visualization
- Culturome v1.0 analysis pipeline; related scripts and example sample results are available at https://github.com/YongxinLiu/Culturome

Input data files
- Sequencing data: the raw sequence data used in this paper have been deposited in the Genome Sequence Archive (GSA)[57] in the BIG Data Center[58] under accession number CRA002517
- RDP training set 16 (ref. [59]) as a taxonomic database. The USEARCH formatted taxonomy is available at http://www.drive5.com/sintax

Reagent setup
10× PBS stock, pH 7
10× PBS stock containing 1.3 M NaCl, 70 mM Na_2HPO_4 and 30 mM NaH_2PO_4 (pH 7). Autoclave and store at room temperature (22–25 °C) for up to 2 months.

1× PBS working solution
Prepare 1 L as follows. Add 100 ml of 10× PBS to 900 ml of deionized water. Store at room temperature for up to 1 month.

10% (vol/vol) liquid TSB medium
Prepare 1 L as follows. Dissolve 3 g of TSB in 1 L of deionized water and distribute into final containers. Sterilize by autoclaving at 121 °C for 15 min. The medium can be stored at 4 °C for up to 1 week.

1/2 liquid TSB medium
Prepare 1 L as follows. Dissolve 15 g of TSB in 1 L of deionized water and distribute into final containers. Sterilize by autoclaving at 121 °C for 15 min. The medium can be stored at 4 °C for up to 1 week.

1/2 solid TSB medium
Prepare 1 L as follows. Dissolve 15 g of TSB and 20 g of agar in 1 L of deionized water and distribute into final containers. Sterilize by autoclaving at 121 °C for 15 min. The medium can be stored at 4 °C for up to 1 week.

10 mM $MgCl_2$
Prepare 1 M $MgCl_2$ stock by adding 20.33 g of $MgCl_2·6H_2O$ to 100 ml of deionized water. Then, add 1 ml of 1 M $MgCl_2$ stock to 99 ml of deionized water. Sterilize by filtration and store at 4 °C for up to 1 month.

80% (vol/vol) glycerol
Prepare 1 L as follows. Add 800 ml of glycerin to 200 ml of deionized water. Autoclave twice and store at room temperature for up to 1 month.

Alkaline lysis buffer, pH 12
Alkaline lysis buffer containing 25 mM NaOH and 0.2 mM Na_2-EDTA at pH 12. Autoclave and store at 4 °C for up to 2 months. ▲ CRITICAL   We recommend preparing 100 mM NaOH stock and 100 mM Na_2-EDTA stock. Then, add 100 ml of NaOH stock and 0.8 ml of Na_2-EDTA stock to 299.2 ml of deionized water for working.

Neutralization buffer, pH 7.5
Neutralization buffer containing 40 mM Tris-HCl at pH 7.5. To prepare 400 ml, dissolve 1.936 g of Tris-HCl in 400 ml of deionized water. Autoclave and store at 4 °C for up to 2 months.

80% (vol/vol) ethanol

80% (vol/vol) ethanol is prepared by diluting absolute ethanol with molecular biology grade water.
▲ CRITICAL Always prepare fresh on the day of DNA purification with AMPure beads.

50× TAE solution

Prepare 1 L as follows. Dissolve 242 g of Tris and 37.2 g of Na_2-EDTA·$2H_2O$ in 800 ml of deionized water, add 57.1 ml of acetic acid until fully dissolved and fill water to 1 L. The solution can be stored at room temperature for up to 2 months.

Equipment setup

Downloading and installing software

1 To install Conda, download the latest version from https://docs.conda.io/en/latest/miniconda.html, install by bash and configure the Bioconda[60] channel for easy-install software in biology:

```
wget -c https://repo.continuum.io/miniconda/Miniconda2-latest-Linux-
x86_64.sh
bash Miniconda2-latest-Linux-x86_64.sh
conda config --add channels bioconda
```

2 To install the QIIME pipeline by Conda:

```
conda install qiime=1.9.1
```

3 To install the Culturome pipeline, download it from Github, unzip the zip file and add the directory to your PATH environment variable:

```
wget -c https://github.com/YongxinLiu/Culturome/archive/master.zip
unzip master.zip && cd Culturome-master
chmod +x script/*
export PATH=`pwd`/script:$PATH
```

4 To install USEARCH 10.0, download usearch10.0.240 Linux version to the script folder and rename it as 'usearch':

```
wget -c https://drive5.com/downloads/usearch10.0.240_i86linux32.gz
-O script/usearch.gz
gunzip script/usearch.gz
chmod +x script/usearch
```

5 To install VSEARCH, visit https://github.com/torognes/vsearch and follow the 'Download and install' section:

```
wget https://github.com/torognes/vsearch/archive·v2.14.2.tar.gz
tar xzf v2.14.2.tar.gz && cd vsearch-2.14.2
./autogen.sh
./configure
make && sudo make install
```

6 To install R language and related packages:

```
conda install r-base r-dplyr r-optparse r-stringr r-ggplot2 r-pheatmap
```

7 To install GraPhlAn 1.1.3-1 by Conda:

```
conda install graphlan=1.1.3-1
```

8   Prepare the input fastq file. If you want to use our web server, we recommend depositing data in GSA[58] (gsa.big.ac.cn), which greatly reduces the downloading time. Alternative public data centers such as the National Center for Biotechnology Information[61] or the European Bioinformatics Institute[62] can also be used to provide downloading links in online analysis mode. Here, we use CRR127980 from GSA as an example to test the pipeline and provide an overview of the results:

```
wget -c ftp://download.big.ac.cn/gsa3/CRA002517/CRR127980/CRR127980_f1.
fq.gz -O L1_1.fq.gz
wget -c ftp://download.big.ac.cn/gsa3/CRA002517/CRR127980/CRR127980_r2.
fq.gz -O L1_2.fq.gz
gunzip *.gz
```

9   Download the RDP taxonomy database; the USEARCH-compatible format is available at http://www.drive5.com/sintax:

```
wget -c http://www.drive5.com/sintax/rdp_16s_v16_sp.fa.gz
gunzip *.gz
```

(Optional) To simplify the installation procedure for all computational software, we provide a fully installed VirtualBox image: https://github.com/YongxinLiu/Culturome.

Procedure

Preliminary isolation to determine the optimal range of dilutions of fresh root samples ● Timing ~8 d (1 working day and a 7-d incubation period)

1   Harvest fresh roots from plants grown in natural soil. Using field-grown *O. sativa* L. or *A. thaliana* plants as examples, carefully collect roots from the plants (entire root system for *A. thaliana* and 10-cm upper rice root zone, corresponding to ~3 g of rice root tissue) and place the roots in a sterile 50-ml Falcon tube.

2   Remove soil particles and loosely attached microbes. To enrich bacteria closely associated with roots, wash the roots with sterile water until all visible soil particles have been removed. Next, transfer the roots to a new 50-ml Falcon tube containing 30 ml of sterile 1× PBS and shake at room temperature in a shaker at 180 r.p.m. for 15 min. Repeat this process three times. Finally, remove the roots from the Falcon tube and blot on sterile filter paper to remove the PBS from the root surface.

3   Grind the roots into homogeneous slurries (root homogenate). Cut the roots into 2-mm pieces and mix. Weigh 0.02 g of mixed root tissue and place in a 1.5-ml Eppendorf tube. Working in a laminar flow hood, resuspend the root tissue in 200 μl of 10 mM sterile $MgCl_2$ and grind the root tissue in the Eppendorf tube with a sterile pestle until it is homogenized.
▲ CRITICAL STEP   It is essential to grind roots with a sterile pestle in the Eppendorf tube. Do not use liquid nitrogen, because liquid nitrogen will likely introduce microbial contamination and kill bacteria. Avoid intensive smashing, such as bead beating, which might kill bacteria and influence isolation efficiency. For hard plant tissues, such as seeds, crush the sample first with sterile metal tools and then grind it with a sterile pestle.

4   Transfer the root slurry into a sterile Falcon tube containing 25 ml of 10 mM $MgCl_2$, mix well and incubate the samples for 15 min at room temperature to release the bacteria from the root tissue.

5   Prepare the dilution series. Mix the root slurry and transfer 4,500 μl, 1,500 μl, 500 μl, 167 μl, 56 μl and 19 μl of the root slurry to six bottles containing 1 L of 10% TSB to generate a 222×, 666×, 2,000×, 6,000×, 18,000× and 54,000× dilution series.

6   Distribute the diluted root slurries into 96-well cell culture plates.
   • For each dilution in the series, shake the bottle and transfer 50 ml of medium containing diluted root slurry to a 13-cm square Petri dish.
   • Place 160 μl of sample into each well of three 96-well cell culture plates using a multichannel pipette in a laminar flow hood. In total, 21 96-well cell culture plates are used, including three 96-well cell culture plates with sterile medium as the negative control.
   ▲ CRITICAL STEP   The laminar flow hood used to perform this step must be as clean as possible to avoid microbial contamination from the environment. Pipetting should be slow to avoid liquid splashing out of wells.

7   Incubate the samples. Seal the edge of each cell culture plate with Parafilm. Stack the plates and incubate them in the dark at room temperature for 1 week.
    ? TROUBLESHOOTING

8   Observe bacterial growth and determine the dilution at which ~30% of wells in 96-well cell culture plates show visible bacterial growth. This is the optimal dilution concentration (ODC) for high-throughput bacterial isolation.
    ▲ **CRITICAL STEP** Although most cultivated bacteria show visible growth after 1-week cultivation, some slow-growing bacteria might not be detected. Thus, the ODC determined in the preliminary isolation tends to be slightly lower than the actual ODC, which should be considered in the high-throughput isolation.

High-throughput bacterial isolation from root microbiota samples (Fig. 2) ● Timing ~16 d (2 working days and a 14-d incubation period)

9   Harvest fresh roots from plants grown in natural soil. Using field-grown *O. sativa* L. or *A. thaliana* plants as examples, carefully collect roots from the plants (full roots for *A. thaliana* and 3 g of 10-cm roots below the soil surface for rice) and place the root tissue in a sterile 50-ml Falcon tube.
    ▲ **CRITICAL STEP** It is important to obtain fresh root samples from plants grown under the same environmental conditions as the plants used in the preliminary isolation. We advise isolating bacteria from two or more independent plants to increase the chance of selecting and purifying isolates with the same *16S rRNA* gene sequence but from different colonization events, because this maximizes the functional diversity of the bacterial stocks.

10 Remove soil particles and loosely attached microbes. To enrich for bacteria closely associated with roots, wash the roots with sterile water until visible soil particles are removed. Transfer the roots to a new 50-ml Falcon tube containing 30 ml of sterile 1× PBS and shake at room temperature in a shaker at 180 r.p.m. for 15 min. Repeat this process three times. Finally, remove the roots from the Falcon tube and absorb any PBS on the root surface using sterile filter paper.

11 Grind the roots into homogeneous slurries. Cut the roots into 2-mm pieces and mix. Weigh 0.02 g of mixed root tissue and place in a 1.5-ml Eppendorf tube. Working in a laminar flow hood, resuspend the root tissue in 200 μl of 10 mM sterile $MgCl_2$. Grind the tissue in the tube with a sterile pestle until the root tissue is homogenized. Keep another 0.1 g of mixed root tissue at –20 °C for cross-referencing cultivated bacteria with root microbiota profiling using primers 799F and 1193R[7,12].

12 Release the bacteria from the root tissue. Transfer the root slurry into a sterile Falcon tube containing 25 ml of 10 mM $MgCl_2$, mix well and incubate the samples for 15 min at room temperature.

13 Prepare the dilution series. Mix and dilute the root slurries to the ODC determined in the preliminary isolation step by using 1 L of 10% TSB liquid medium as well as 1/3× ODC and 3× ODC to avoid experimental variation. If the ODC in the preliminary isolation was 6,000×, then add 500 μl, 167 μl and 56 μl of root slurry to three bottles containing 1 L of 10% TSB medium, respectively, obtaining 2,000×, 6,000× and 18,000× dilutions.
    ▲ **CRITICAL STEP** Owing to variation in plant materials and cultivation time in the preliminary and high-throughput isolation, we strongly suggest to also include 1/9× ODC and 9× ODC, if possible, to ensure the success of high-throughput isolation.

14 Distribute medium containing the diluted root slurries into 96-well cell culture plates. For each dilution, shake the bottle and transfer 50 ml of medium containing diluted root slurries to a 13-cm square Petri dish. Distribute 160 μl of medium from the square Petri dish to each well of 30–45 96-well cell culture plates using a multichannel pipette under a laminar flow hood and prepare three 96-well cell culture plates with sterile medium as the negative control.
    ▲ **CRITICAL STEP** Thirty to forty-five plates should be prepared to capture most abundant root-associated bacterial taxa. The laminar flow hood used to perform this step should be as clean as possible to avoid microbial contamination from the environment. Pipetting should be done slowly to avoid liquid splashing out of wells.

15 Incubate the bacterial cultures. Seal the edge of each cell culture plate with Parafilm. Incubate the 96-well cell culture plates at room temperature for 14 d.
    ▲ **CRITICAL STEP** A longer incubation time will help slow-growing bacteria reach saturation, but the cultivation time should be less than 4 weeks, because, at ~4 weeks of incubation, the liquid medium tends to dry, and bacteria might die due to nutrient depletion.
    ? TROUBLESHOOTING

16 Examine visible bacterial growth in the cell culture plate and select the dilution that allowed ~30% of wells in the 96-well cell culture plate to show visible bacterial growth.
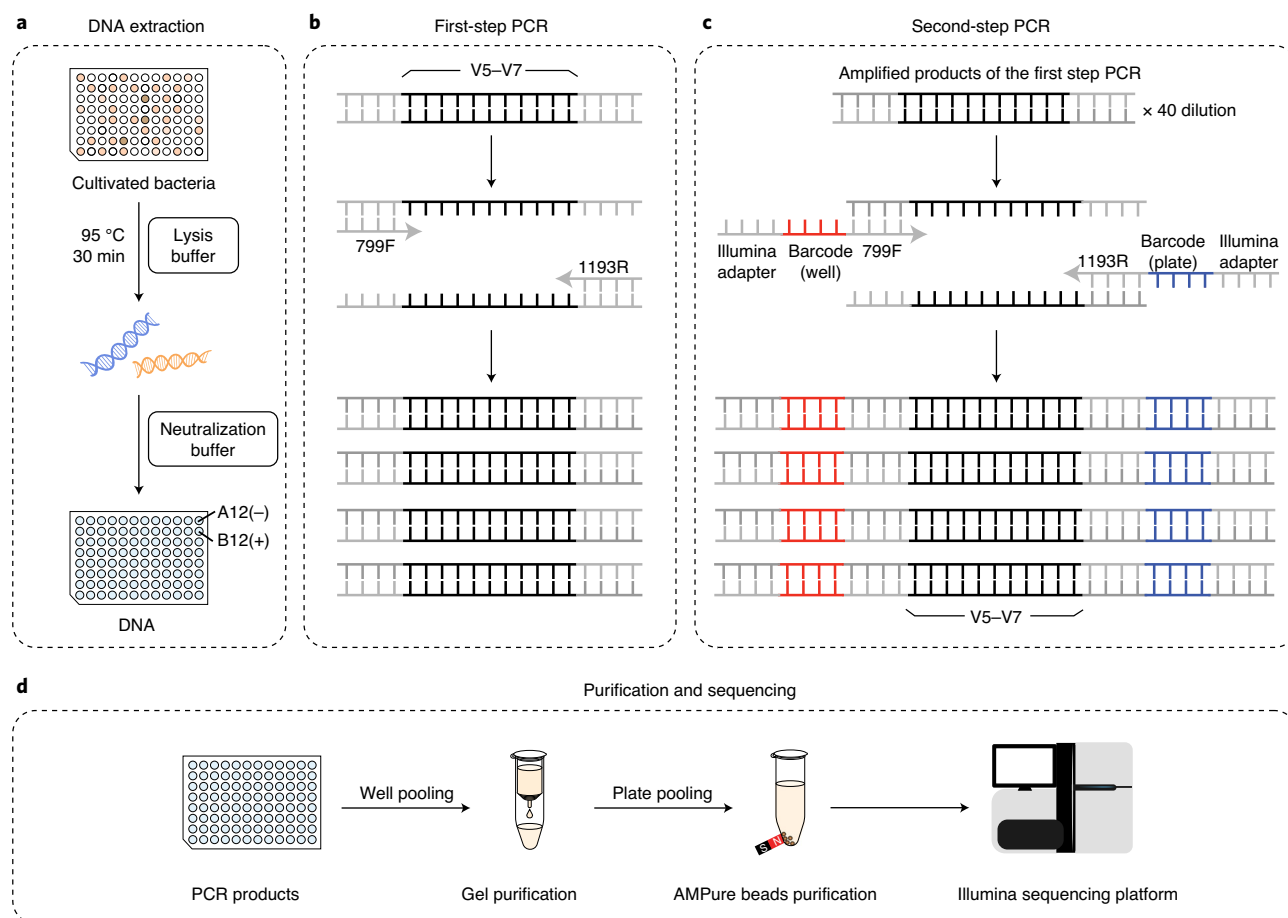
Fig. 3 | Bacterial identification using the two-sided barcode system. a, DNA extraction from cultivated bacteria in a 96-well cell culture plate. DNAs are extracted from the cultivated bacteria by the alkaline lysis method. **b**, Enrichment of the bacterial *16S rRNA* gene fragment. The V5–V7 regions of the bacterial *16S rRNA* gene are enriched using degenerate primers 799F and 1193R in the first-step PCR. **c**, Labeling of bacterial *16S rRNA* gene fragments with plate and well barcodes. The amplified *16S rRNA* gene fragments are labeled with plate and well barcodes as well as Illumina sequencing adapters. **d**, Pooling, purification and Illumina sequencing of PCR products. *16S rRNA* gene primers (799F and 1193R) and Illumina adapters are shown in gray. Well and plate barcodes are shown in red and blue, respectively.

17 Pipette 10 μl of the bacterial solution in each well of the 96-well cell culture plate into each well of a 96-well PCR plate for bacterial identification and store at −20 °C.
    ■ **PAUSE POINT** Bacteria for further identification can be stored at −20 °C for several months.
    ▲ **CRITICAL STEP** Before opening the seal on the 96-well PCR plate for any following steps, centrifuge the plate at 2,000*g* using a microplate centrifuge for 10 s making all liquid stay at the bottom of the 96-well plate to avoid cross-contamination between wells. Pipetting should be done slowly to avoid liquid splashing out of wells.

18 Add 140 μl of 80% (vol/vol) glycerol to the remaining bacterial culture in each well of the 96-well cell culture plate and store at −80 °C.

Amplify bacterial *16S rRNA* genes in the cultivated bacteria via two-sided barcode PCR ● Timing ~5 d
Extract DNA from the cultivated bacteria using the alkaline lysis method

19 Add 16.6 μl of lysis buffer to 10 μl of bacterial cultures in the 96-well PCR plates produced in Step 17 (Fig. 3a).

20 Mix the lysis buffer and bacterial culture with a multichannel pipette and incubate in a PCR machine at 95 °C for 30 min.

21 Cool the plates to room temperature, add 16.6 μl of neutralization buffer to each well, cover the PCR plates with polyester sealing film and store at −20 °C.

Amplify the bacterial *16S rRNA* gene in the first PCR using degenerate primers

22 Set up the PCR reaction mixture for each well on a laminar flow hood as shown below. Discard the original DNAs in the well positions of A12 and B12 and use nuclease-free water and *E. coli* DNA for the negative and positive controls in each plate of bacterial DNAs. (See also Fig. 3b.)

Component	Amount (µl)
10× Buffer	3
dNTPs (2.5 mM each)	2.4
799F (10 µM)	0.3
1193R (10 µM)	0.3
HS taq (5 U/µl)	0.15
Genomic DNA (gDNA)	3
Qiagen water	20.85
Total	30

▲ **CRITICAL STEP** It is important to use polymerase without bacterial DNA contamination. Other polymerases, such as DSF (Agilent) or HS primeSTAR (Kapa Biosystems), can be used instead of HS taq. The reaction mixtures should be set up on a laminar flow hood to avoid contamination from the environment. If a laminar flow hood is not available for this step, check the negative control carefully.

23 Perform PCR using the following cycling conditions:

Cycle no.	Denature	Anneal	Extend
1	94 °C, 2 min		
2–31	94 °C, 30 s	55 °C, 30 s	72 °C, 1 min
32			72 °C, 5 min

24 Examine the PCR products from the negative and positive controls by 1.0% (wt/vol) agarose gel electrophoresis using 3 µl of PCR product and 3 µl of loading buffer. Successful reactions should yield a single 400-bp product for the positive control and no visible product for the negative control.
    ? TROUBLESHOOTING

25 Dilute 2 µl of each first PCR product in each PCR plate (including the positive and negative controls) with 78 µl of nuclease-free water to reduce the amount of leftover PCR components from the first PCR.

Using the diluted first PCR products as templates, label each PCR product with barcoded primers in the second PCR (Supplementary Table 1)

26 Label each PCR product with well and plate barcodes. Set up the PCR reaction mixture for each well under a laminar flow hood as described below. (See also Fig. 3c.)

Component	Amount (µl)
10× Buffer	3
dNTPs (2.5 mM each)	2.4
Reverse-barcoded primer (10 µM)	0.6
Forward-barcoded primer (10 µM)	0.6
HS taq (5U/µl)	0.15
gDNA (diluted by 40× of PCR1)	3
Qiagen water	20.25
Total	30

▲ **CRITICAL STEP** It is important to use polymerase without bacterial DNA contamination. Other polymerases, such as DSF (Agilent) or HS primeSTAR (Kapa Biosystems), can be used instead of HS taq.

27 Perform PCR with the following cycling conditions:

Cycle number	Denature	Anneal	Extend
1	94 °C, 2 min		
2–26	94 °C, 30 s	55 °C, 30 s	72 °C, 1 min
27			72 °C, 5 min

28 Examine the PCR products of the negative and positive controls by 1.0% (wt/vol) agarose gel electrophoresis using 5 µl of PCR products and 5 µl of loading buffer. Successful reactions should yield a single 500-bp product in the positive control and no visible product in the negative control.
? TROUBLESHOOTING

Pooling, purification and Illumina sequencing of the PCR products ● Timing ~16 h

29 Take all the products from the second PCR from one plate and mix in a 5-ml Eppendorf tube. Do this for all individual plates (Fig. 3d).
30 Combine 40 µl of mixed product from each plate with 10 µl of loading buffer.
31 Run the samples, along with a 2,000-bp DNA ladder, in a 1.2% (wt/vol) agarose gel at 126 V for 40 min.
▲ CRITICAL STEP It is important to ensure that the 500-bp target sequences separate sufficiently from the other PCR byproducts, such as leftover primers, primer dimers and unspecified sequences.
32 Cut the gel to obtain 500-bp amplification fragments and extract pure DNA using a gel cleanup system according to the manufacturer's instructions. In our lab, we use Wizard SV Gel and PCR Clean-Up System from Promega. Store the DNA at –20 °C.
33 The DNA concentrations of the purified PCR products can be measured using a NanoDrop machine (Option A) or PicoGreen fluorescent dye (Option B). NanoDrop is more convenient, whereas PicoGreen-based measurement is more precise.
 (A) **Measure the DNA concentrations of purified PCR products using a NanoDrop device**.
   (i) Load 2 µl of nuclease-free water into the NanoDrop machine as a blank.
   (ii) Remove the nuclease-free water sample, load 2 µl of purified PCR product into the NanoDrop machine and record the result.
 (B) **Measure the DNA concentrations of purified PCR products using PicoGreen fluorescent dye**.
   (i) Prepare a standard curve with lambda DNA. Set up the reactions in an optical 96-well EIA/RIA plate as follows:

lambda DNA concentration (ng/µl)	1× TE volume (µl)	2 µg/ml lambda DNA volume (µl)
0.75	12.5	37.5
0.5	25	25
0.4	30	20
0.25	37.5	12.5
0.1	45	5
0.05	47.5	2.5
0.01	49.5	0.5
0	50	0

▲ CRITICAL STEP A standard curve must be set up for each optical 96-well EIA/RIA plate to reduce variation from plate to plate.
   (ii) In the other wells of the optical 96-well EIA/RIA plate, set up the following reactions:

Component	Amount (µl)
Purified PCR product	2
1× TE	48

(iii) Add 50 μl of PicoGreen fluorescent dye (1:200 diluted) to each well to measure the standard curve and purified PCR products. Mix by pipetting ten times. Wrap the 96-well EIA/RIA plate with aluminum foil.

▲ CRITICAL STEP PicoGreen fluorescent dye is easily degraded when exposed to light. When performing steps using PicoGreen fluorescent dye, strong light should be avoided.

(iv) Measure the fluorescence of each well in the 96-well EIA/RIA plate with an excitation light wavelength of 480 nm and absorbance wavelength of 520 nm.

(v) Calculate the DNA concentration of each purified PCR product based on the standard curve.

34 Combine 70 ng of each purified PCR product into a 1.5-ml Eppendorf tube and mix well.

35 Remove AMPure beads from a 4 °C refrigerator and allow them to reach room temperature.

36 Add the magnetic beads to each sample in an Eppendorf tube with a volume ratio of magnetic beads:sample volume = 0.9:1. Mix by pipetting 20 times and incubate at room temperature for 10 min.

▲ CRITICAL STEP The volume ratio of beads with PCR products works well from 0.8:1 to 1.8:1, in our experience. We use the ratio 0.9:1 instead of 1.8:1 that is suggested in the manufacturer's instruction to reduce experimental cost and keep purification efficiency.

37 Place the tubes in a magnetic rack and incubate the samples at room temperature for 5 min.

38 Carefully remove the supernatant without disturbing the beads.

39 Add 200 μl of 80% ethanol, incubate the samples for 30 s and remove the ethanol with a pipette. Repeat this step.

▲ CRITICAL STEP The ethanol needs to be freshly prepared, because the proportion of ethanol will decrease with time due to evaporation and absorbing water when contacting air. We suggest using 80% ethanol instead of 70% ethanol that is suggested in the manufacturer's instructions to reduce the likelihood that the ethanol concentration will decrease to below 70% during sample processing.

40 With the tubes still on the magnetic rack, air dry the beads for 4 min. Be careful not to overdry the beads, as it will significantly reduce DNA recovery rates.

41 Remove the tube from the magnetic rack, add 100 μl of 1× TE buffer and resuspend the beads by pipetting up and down. Close the tube lid and incubate at room temperature for 5 min.

▲ CRITICAL STEP A volume of TE buffer in the range from 40 μl to 100 μl is optimal, in our experience. We use 100 μl of TE buffer to improve the chance that we have eluted all the PCR products.

42 Place the tube back on the magnetic rack and incubate the sample for 5 min. Transfer 96 μl of the eluent to a new 1.5-ml tube.

43 To further purify and concentrate the purified PCR products, add 86.4 μl of magnetic beads to the tubes. Mix by pipetting 20 times and incubate the sample at room temperature for 10 min. Transfer the tubes to a magnetic rack and let stand for 5 min. Remove the supernatant, add 200 μl of 80% ethanol, incubate the samples for 30 s and remove the ethanol. Repeat this step once more for a total of two ethanol washes. With the tubes still on the magnetic rack, air dry the beads for 4 min.

▲ CRITICAL STEP This second round of purification using AMPure beads helps to completely remove primer dimmers and small DNA fragments in the PCR products, increasing the success of Illumina sequencing. If the quality of purified PCR products in Step 42 meets the requirement of the local sequencing facility, Steps 43–45 are optional.

44 Remove the tube from the magnetic rack, add 50 μl of 1× TE buffer and resuspend the beads by pipetting up and down. Incubate the samples at room temperature for 5 min.

45 Place the tube back on the magnetic rack and incubate the samples for 5 min. Transfer 48 μl of the eluent to a new 1.5-ml tube.

46 Quantify the purified samples using a Quant-iT PicoGreen dsDNA Assay Kit as described in Step 33.

47 Deliver 1,500 ng of purified PCR products for Illumina sequencing.

■ PAUSE POINT Purified PCR products can be stored at −20 °C for 3−6 months.

Bioinformatic analysis to identify cultivated bacteria (Fig. 4) ● Timing 1–3 d (~1 h for 1 GB of data). The exact time depends on the size of the dataset and computational resources and expertise.

▲ CRITICAL STEP The analysis can be performed either locally using the pipeline or in the web server http://bailab.genetics.ac.cn/culturome by providing the email, mapping file and links to download
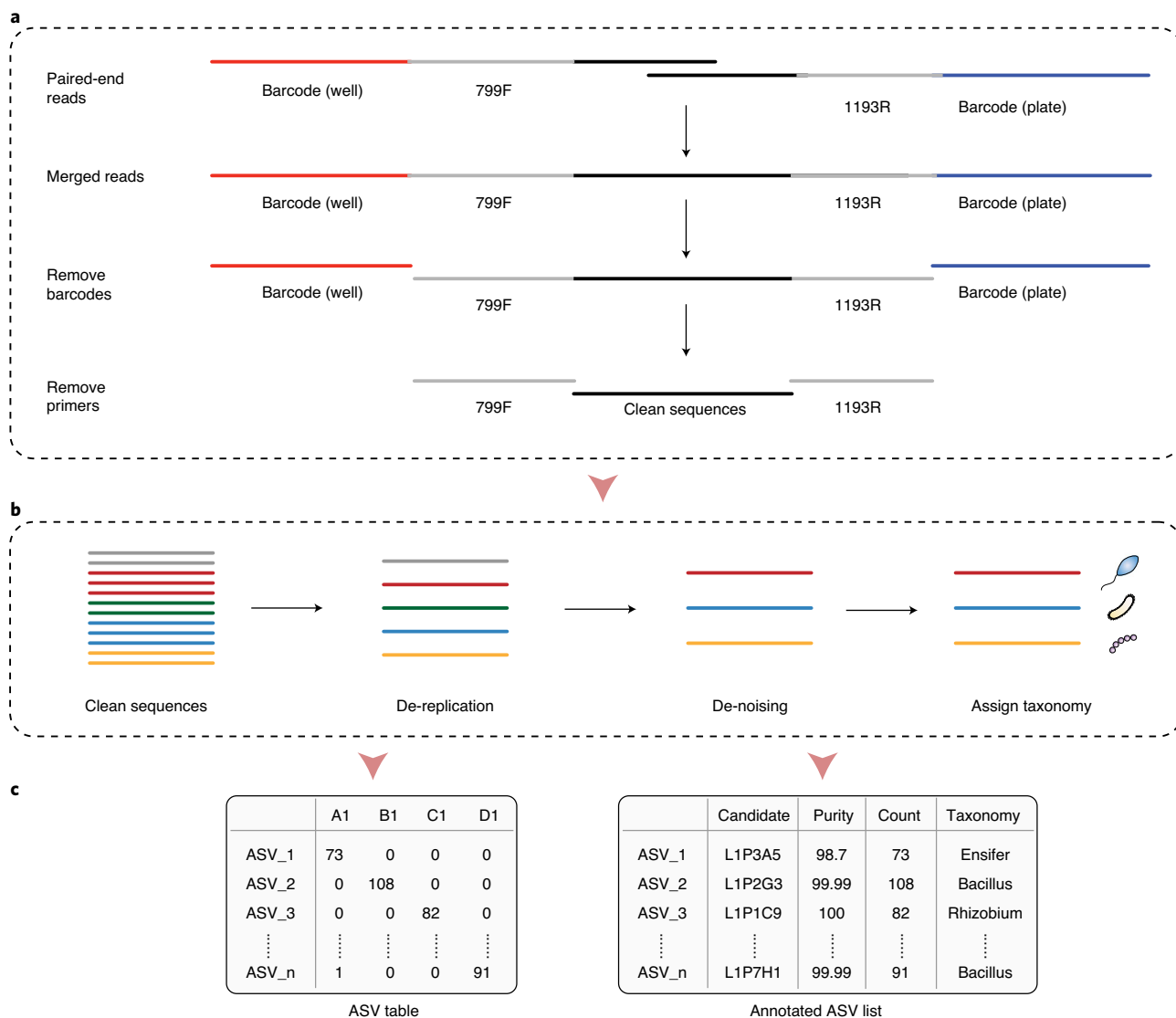
Fig. 4 | Workflow of bioinformatic analysis to identify cultivated bacteria. a, Summary of the steps from obtaining raw data to generating clean sequences of the amplified *16S rRNA* gene region. The amplified clean sequences are obtained by merging paired-end reads, demultiplexing plate and well barcodes and removing PCR primer sequences. **b**, Workflow to identify ASVs and taxonomy. Clean sequences are de-replicated with VSEARCH, de-noised with USEARCH to identify ASV sequences and taxonomically classified using the RDP database. **c**, Sequence and taxonomy information for cultivated bacteria in 96-well cell culture plates. The ASV table shows bacterial *16S rRNA* gene sequence information for each well in a 96-well cell culture plate. The annotated ASV list shows information for wells with bacteria belonging to the corresponding ASV, including purity, sequence count and taxonomy.

raw data. Detailed instructions for the use of the web server can be found on the website, http://bailab. genetics.ac.cn/culturome/manual.html. We suggest that researchers follow instructions for formatting fastq file names and sample IDs in the mapping file, which improves the performance of the web server bioinformatic pipeline.

48   For the local pipeline, install and double check software according to steps in 'Equipment setup' section.

49   Prepare the mapping file to analyze the amplicon data. The mapping file contains metadata such as the sample IDs, forward and reverse barcodes and forward and reverse primers. Although the mapping file can be manually generated according to the QIIME mapping file format, it is tedious and might generate errors when determining the positions of wells and each plate using combined well and plate barcodes, given the large number of rows (4,320 for 45 cell culture plates). Here we provide an in-house script (write_mapping_file.pl script) to generate a mapping file using an input including the well and plate barcodes (-i, -b), forward and reverse primers (-F, -R) and other

detailed information (-L, -p, -v, -c, -m, -s) (Supplementary Tables 1 and 2). For detailed descriptions of the parameters, type write_mapping_file.pl -h.

```
write_mapping_file.pl \
-i script/barcodeF96.txt \
-b script/barcodeR48.txt \
-F AACMGGATTAGATACCCKG -R ACGTCATCCCCACCTTCC \
-L L1 -p 45 -v Nipponbare -c Root -m TSB -s Rice \
-o L1.txt
```

50 Validate the mapping file. There are many formatting requirements for mapping files (Supplementary Table 2), and errors will affect subsequent analyses. Use validate_mapping_file. py from QIIME to determine whether the mapping file is formatted properly.

```
validate_mapping_file.py -m L1.txt -o temp/
```

If the message 'No errors or warnings were found in mapping file' appears, your file is correct. Otherwise, revise your mapping file based on the output report.
? TROUBLESHOOTING

51 Merge paired-end reads into single-end reads. The 250-bp paired reads from HiSeq2500/ NovaSeq6000 are merged according to the reverse complement of the end part of the sequences.

```
vsearch \
-fastq_mergepairs L1_1.fq \
-reverse L1_2.fq \
-fastqout temp/L1.fq
```

52 Demultiplex barcoded sequences into samples corresponding to each well (Supplementary Tables 1 and 2). According to the mapping file, remove plate and well barcodes (extract_barcodes.py), re-label sequences with the ID of plates and wells (split_libraries_fastq.py) and adjust the name of the sequences to the format compatible with USEARCH.

```
extract_barcodes.py \
-f temp/L1.fq -m L1.txt \
-c barcode_paired_stitched --bc1_len 10 --bc2_len 6 \
-a --rev_comp_bc2 \
-o temp/L1
split_libraries_fastq.py \
-i temp/L1/reads.fastq -b temp/L1/barcodes.fastq \
-m L1.txt -q 19 --max_barcode_errors 0 \
--barcode_type 16 --phred_offset 33 \
-o temp/L1
cut -f 1 -d ' ' temp/L1/seqs.fna | sed 's/_/./' > temp/qc.fa
```

53 Visualize sequence counts from each plate to check the success and evenness of amplicon sequencing. For more details, type stat_split_bar.R -h.

```
mkdir -p result/split
tail -n+16 temp/L1/split_library_log.txt| head -n-4 \
> result/split/L1.txt
stat_split_bar.R \
-i result/metadata.txt \
-d result/split/L1.txt \
-o result/split/
```

? TROUBLESHOOTING

54 Remove forward and reverse primer sequences.

```
usearch \
-fastx_truncate temp/qc.fa \
-stripleft 19 -stripright 18 \
-fastaout temp/filtered.fa
```

55 Identify ASVs by performing de-replication in VSEARCH and de-noising in USEARCH (Supplementary Table 3).

```
vsearch \
--derep_fulllength temp/filtered.fa \
--relabel Uni --minuniquesize 8 --sizeout \
--output temp/uniques.fa
usearch \
-unoise3 temp/uniques.fa \
-zotus temp/Zotus.fa
format_ASVID.sh -i temp/Zotus.fa -o result/ASV.fa
```

56 Identify bacteria in each well of a 96-well cell culture plate by quantifying the ASV abundance by mapping all clean sequences to ASVs using VSEARCH (Supplementary Table 4).

```
vsearch \
--usearch_global temp/filtered.fa \
--db result/ASV.fa --id 0.97 \
--otutabout temp/ASV_table.txt
```

57 Remove entries corresponding to wells containing only background noise. Amplicon sequencing from each well is easily contaminated by low-abundance DNAs from reagents, the environment or Illumina sequencing errors. Remove entries corresponding to wells containing fewer reads than the maximum number of reads in the negative controls. In each plate, A12 is the negative control using nuclease-free water and B12 is the positive control using *E. coli* DNA.

```
negative_threshold.R \
--input temp/ASV_table.txt --metadata L1.txt \
--threshold 1 --negative A12 --positive B12 \
--output result/fdr.txt
usearch -otutab_trim temp/ASV_table.txt \
-output result/ASV_table.txt \
-min_sample_size `cat result/fdr.txt`
```

58 Classify the taxonomy of ASVs from cultivated bacteria based on the RDP database and integrate into the tab-separated table (Supplementary Table 3).

```
usearch \
-sintax result/ASV.fa \
-db rdp_16s_v16_sp.fa \
-tabbedout temp/ASV.fa.tax \
-sintax_cutoff 0.6 -strand both
tax_sum.sh -i temp/ASV.fa.tax \
-d result/ASV_table.txt \
-o result/
```

59 Identify candidate isolates for each ASV. This step includes combing the ASV and taxonomy tables, evaluating the saturation of bacterial ASV diversity based on the number of wells containing bacteria, overviewing the distribution of wells containing different numbers of ASVs or genera and examining the purity of the cultivated bacteria in each well. The outputs include two tables: an ASV list (isolate_ASV.txt; Supplementary Table 5) including five best wells containing bacteria with the
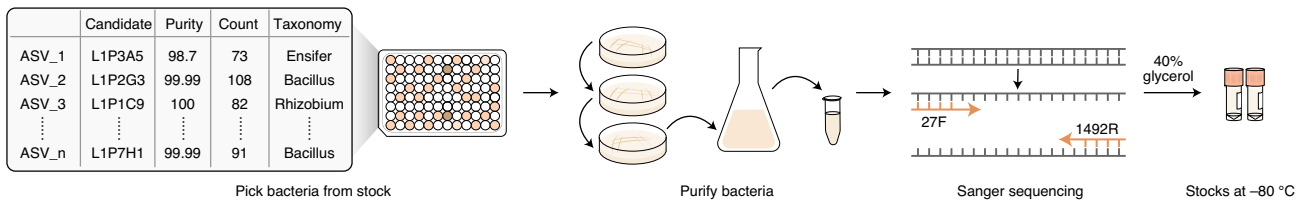
	Candidate	Purity	Count	Taxonomy
ASV_1	L1P3A5	98.7	73	Ensifer
ASV_2	L1P2G3	99.99	108	Bacillus
ASV_3	L1P1C9	100	82	Rhizobium
⋮	⋮	⋮	⋮	⋮
ASV_n	L1P7H1	99.99	91	Bacillus

Pick bacteria from stock   Purify bacteria   Sanger sequencing   Stocks at −80 °C

Fig. 5 | Preservation of cultivated bacteria. Bacteria in 96-well cell culture plates are selected based on the *16S rRNA* gene sequences, purified by continuous streaking and validated by Sanger sequencing. The validated bacteria are stored at -80 °C to study the functions and mechanisms of the root microbiota.

corresponding *16S rRNA* gene sequence and a well list (isolate_well.txt) including all detected ASV purity, read counts and taxonomy.

```
identify_isolate.R --input result/ASV_table.txt \
--taxonomy result/taxonomy_8.txt \
--output result/isolate
plot_purity.sh -i result/isolate_well.txt -l L1 \
-o result/purity
```

60 Summarize the taxonomic distribution and occurrence frequency of cultivated bacteria using GraPhlAn.

```
graphlan_prepare.R \
--input result/ASV_table.txt \
--output result/graphlan/ \
--taxonomy result/taxonomy_8.txt \
--abundance 0 --number 1000
graphlan_plot.sh -i ./ -o result/graphlan
```

61 Cross-reference the cultivated bacteria with the corresponding root microbiota profiling data using the similarity of the V5–V7 regions in the *16S rRNA* gene .

```
cross_reference.sh \
-i script/profiling_ASV.fa \
-d script/profiling_ASVtab.txt \
-r result/ASV.fa \
-o result/cross_reference.txt
```

Purification and preservation of cultivated bacteria (Fig. 5) ● Timing 16–24 d

62 For each ASV, select 2–3 wells containing the corresponding bacteria ideally derived from independent plants, scratch the surface of the frozen bacterial culture with a pipette tip and streak it onto a 1/2 TSB plate.

63 After 3–5 d of incubation at room temperature, pick a single colony and streak it onto a new 1/2 TSB plate. Repeat this process two more times.
▲ CRITICAL STEP Owing to the diverse growth rates of bacteria, the time for this step may vary from 9 to 15 d. It is important to pick single colonies to ensure that the bacterial culture is pure. If different colonies are growing on a single plate, pick the different colonies and purify them separately on plates containing bacterial medium.

64 Pick a single colony from the bacterial medium plate, incubate the colony in 1/2 TSB liquid medium and shake at 28 °C at 180 r.p.m. for 5–7 d.
? TROUBLESHOOTING

65 When the bacterial solution is turbid, transfer 10 μl of the bacterial culture into a PCR tube for DNA extraction following Steps 19–21.

66 Validate the bacterial *16S rRNA* gene sequence by Sanger sequencing. Detect the bacterial *16S rRNA* gene by PCR using primers 27F and 1492R (Supplementary Table 1) following Step 23. Sequence the PCR products using 1492R by Sanger sequencing.

67 Preserve the bacteria when the Sanger sequencing results are consistent with the high-throughput sequencing results. Bacterial stocks can be generated using either of the two methods described below (Options A and B) and are valuable materials for bacterial whole-genome sequencing and functional studies:

(A) **Storage in glycerol**

(i) Centrifuge 30 ml of bacterial solution for 10 min at 2,900*g* at room temperature.

(ii) Discard the excess medium to retain 1 ml of culture medium.

(iii) Vortex to completely suspend the bacteria in the remaining medium and combine 800 μl of bacterial solution and an equal volume of 80% sterile glycerol in a cryopreservation tube.

(iv) Mix the samples evenly by pipetting and store at –80 °C.

(B) **Microbank storage tube**

(i) Pick a single colony on a plate after three rounds of purification with a sterile pipette tip and transfer it to a Microbank storage tube.

(ii) Invert the samples four or five times to mix.

(iii) After 18 h of incubation at room temperature, discard the preservation solution in the tube and store at –80 °C.

▲ **CRITICAL STEP** We suggest using individual tubes to store pure bacteria; using 96-well plates might result in cross-contamination.

? **TROUBLESHOOTING**

Troubleshooting

Troubleshooting advice can be found in Table 1.

Table 1 | Troubleshooting table

Step	Problem	Possible reason	Solution
7,15	Drying of the liquid medium in border wells	Incomplete sealing with Parafilm	Tightly seal the edge of 96-well cell culture plates
24,28	Negative control showing visible amplification	Contamination in PCR system	Perform PCR reactions using a new set of reagents and nuclease-free water under laminar flow hood
50	Warning message	The format problem of the mapping file	Check carefully the mapping file with the provided example
53	Substantial variations in sequencing depth of plates	Unequal amount of PCR products from each plate in sequenced samples	Double check the concentration of purified PCR products in Step 33
64	Undetectable growth after 7 d	Low amount of inoculum; slow-growing characteristic in liquid medium	Inoculate with a bigger colony; proliferate bacteria on agar medium plates and perform colony PCRs
67	Low recovery rate	Sensitivity of thaw-frozen process for diverse natural isolates	Increase the bacterial concentration when preparing stocks; avoid thaw-frozen steps when using the glycerol stocks

Timing

Steps 1–6, harvesting fresh roots and diluting the root samples using several dilution series: 1 d

Steps 7 and 8, incubation of bacteria and determining the optimal range of dilutions: 7 d

Steps 9–14, harvesting fresh roots for high-throughput bacterial isolation and diluting the root samples based on the ODC: 1 d

Steps 15 and 16, incubation of isolated bacteria and selection of the 96-well cell culture plates containing ~30% wells showing visible bacterial growth: 14 d

Steps 17 and 18, generation of bacterial stocks in the 96-well cell culture plates and portion of the sample for bacterial identification: 1 d

Steps 19–21, DNA extraction from the cultivated bacteria: 1 d

Steps 22–25, amplification of the *16S rRNA* gene of the cultivated bacteria in the first PCR: 2 d

Steps 26–28, labeling the first PCR products with barcoded primers in the second PCR: 2 d

Fig. 6 | Anticipated results for bacterial cultivation from the *O. sativa* L. roots. a, Bar plot showing the distribution of ▶ read counts in each plate generated by the two-sided barcode system. Read counts of amplified *16S rRNA* gene sequences in each plate of cultivated bacteria reveal the depth and evenness of Illumina sequencing. This figure includes sequencing data from 45 cell culture plates of cultivated bacteria derived from Nippornbare rice roots. The original data have been used in *Nature Biotechnology*[12]. **b**, Read counts for the negative (nuclease-free water as the PCR template) and positive controls (*E. coli* DNA as the PCR template). **c**, Rarefaction curve of ASVs based on the number of wells containing bacteria. The curve reaches the plateau stage, indicating that a sufficient number of plates were used for high-throughput bacterial isolation. The iteration number for each well number is 30. The horizontal bars within the boxes represent medians. The tops and bottoms of the boxes represent the 75th and 25th percentiles, respectively. The upper and lower whiskers extend to data no more than 1.5× the interquartile range from the upper edge and lower edge of the box. **d**, Distribution of wells regarding the purity based on evaluation of ASVs of the bacterial *16S rRNA* gene. Consistent with the assumption that low-abundance sequences might come from multiple polymorphic copies of the *16S rRNA* gene and sequencing errors, we found that cultivated bacteria with a purity greater than 95% are likely to be pure when we performed streaking on agar plates. The ratio of wells with a purity greater than 95% to the rest of the wells containing multiple bacteria is 3.6:1. **e**, An example of the purity of wells containing cultivated bacteria in a 96-well cell culture plate. **f**, Cladogram showing the taxonomic distribution and occurrence frequency of cultivated bacteria. The inner ring represents the de-replicated ASVs from cultivated root bacteria. The heat map in the outer ring represents the $\log_2$-transformed number of cultivated bacterial isolates belonging to the corresponding ASV. All source data used in this figure are in Supplementary Table 6.

Step 29, pooling the PCR products from one plate in a 5-ml Eppendorf tube: 2 h

Steps 30 and 31, checking the PCR products by electrophoresis: 4 h

Step 32, cutting the amplified PCR fragments in the agarose gel and purification: 4 h

Step 33A, measurement of the DNA concentration of the purified PCR products using a NanoDrop machine: 1 h

Step 33B, measurement of the DNA concentrations of purified PCR products using PicoGreen fluorescent dye: 2 h

Step 34, pooling an equal amount of purified PCR products from different plates: 30 min

Steps 35–45, purification of the PCR products using magnetic beads twice: 2 h

Steps 46 and 47, quantification of the purified samples and delivery of the samples for Illumina sequencing: 1 h

Steps 48–61, bioinformatic analysis to identify cultivated bacteria: 1–3 d

Steps 62 and 63, purification of bacterial cultures in the 96-well cell culture plates using continuous streaking: 9–15 d

Step 64, incubation of the colony in the liquid medium: 5–7 d

Steps 65 and 66, DNA extraction of the bacterial solution and validation of the bacterial *16S rRNA* gene sequence by Sanger sequencing: 1 d

Step 67 (A), generation of high-quality stocks using a cryopreservation tube: 1 d

Step 67 (B), generation of high-quality stocks using a Microbank storage tube: 1 d

Anticipated results

This protocol provides detailed guidelines for systematically isolating bacteria from a given root microbiota sample. We isolated root bacteria from 8-week-old Nippornbare rice (*O. sativa* L.) grown under field conditions[12]. According to the preliminary isolation, we diluted the root slurry to three concentration: 666×, 2,000× and 6,000×. The 96-well cell culture plates at the 2,000× dilution contained 45% of wells showed visible bacterial growth and were suitable for further cultivation (Fig. 6 and Supplementary Table 6). The 45 cultivation plates showed relatively even sequencing depth with an average of 19,454 reads per plate (Fig. 6a), allowing efficient and unbiased identification of cultivated bacteria. The negative controls contained read counts below 20 (Fig. 6b), suggesting a low noise and high accuracy in strain identification. The diversity of cultivated bacteria reached the plateau stage, indicating that the number of isolates was sufficient under the tested condition (Fig. 6c). Of the wells, 44.8% contained 100% pure ASVs (a single ASV representing 100% of all sequences in the well), and 78.1% of wells showed purity greater than 95%. The ratio of wells with purity greater than 95% to the rest of the wells containing multiple bacteria was 3.6:1 (Fig. 6d,e). We found that cultivated bacteria with a purity greater than 95% (identified by Illumina) were likely to be pure when we performed streaking on agar plates, consistent with the assumption that low-abundance sequences might come from multiple polymorphic copies of the *16S rRNA* gene and sequencing errors. A total of 2,073 wells contained 278 distinct ASVs of bacterial isolates mainly belonging to Proteobacteria, Actinobacteria, Firmicutes and Bacteroidetes (Fig. 6f). These cultivated bacteria accounted for 57.6%
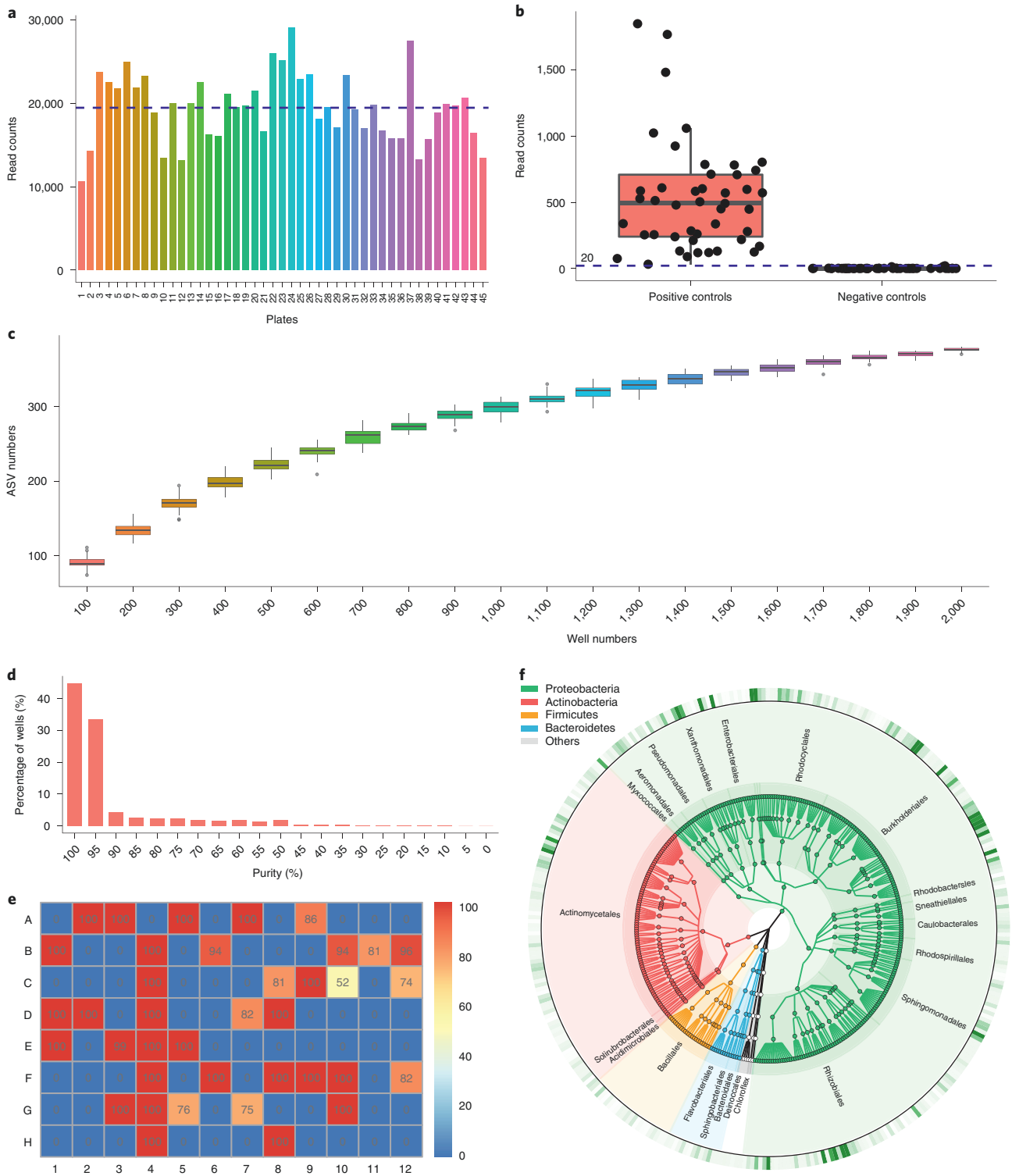
taxa and 78.4% cumulated relative abundance of rice root-associated ASVs (relative abundance > 0.1%, 89.8% of all sequencing reads). Note that the analysis described above can be performed with the bioinformatic pipeline (https://github.com/YongxinLiu/Culturome) as well as the graphical user interface web server (http://bailab.genetics.ac.cn/culturome). Based on our methods and cultivated bacterial resources from *A. thaliana*, we found that root-derived bacterial commensals protect plants against fungi and oomycetes in natural soil[7], and that *Arabidopsis* triterpene compounds selectively modulate root microbiota members[14]. In summary, this protocol provides a straightforward and

generic method to cultivate root bacteria at affordable costs and time. It can be adapted and used to cultivate members of the bacterial microbiota from other plant organs and environmental samples. The cultivated bacteria can be used as a key resource to study the functions, assembly process and molecular mechanisms of the root microbiota under normal or defined stressed conditions with a variety of gnotobiotic plant systems.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw sequence data used in this paper have been deposited in the Genome Sequence Archive[57] in the BIG Data Center[58] under accession number CRA002517. The authors declare that all data supporting the findings of this study are available in the original paper[12]. Because many labs might not frequently perform systematic bacterial isolation, to reduce the experimental cost we are happy to share the full set of barcoded primers with any scientists who want to use this protocol. Correspondence and requests for materials and computational scripts should be addressed to Y.B.

Code availability

The computational scripts of the data analysis pipeline (Culturome v1.0) are freely available at GitHub (https://github.com/YongxinLiu/Culturome) under a GNU General Public License v2 or later. The Culturome web server is freely available at http://bailab.genetics.ac.cn/culturome and continuously maintained by the authors.

References

1. Berendsen, R. L., Pieterse, C. M. J. & Bakker, P. A. H. M. The rhizosphere microbiome and plant health. *Trends Plant Sci.* **17**, 478–486 (2012).
2. Bulgarelli, D. et al. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* **488**, 91–95 (2012).
3. Lundberg, D. S. et al. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**, 86 (2012).
4. Peiffer, J. A. et al. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl Acad. Sci. USA* **110**, 6548–6553 (2013).
5. Edwards, J. et al. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl Acad. Sci. USA* **112**, E911–E920 (2015).
6. Wang, W. et al. An *Arabidopsis* secondary metabolite directly targets expression of the bacterial type III secretion system to inhibit bacterial virulence. *Cell Host Microbe* **27**, 601–613 (2020).
7. Durán, P. et al. Microbial interkingdom interactions in roots promote *Arabidopsis* survival. *Cell* **175**, 973–983 (2018).
8. Carrión, V. J. et al. Pathogen-induced activation of disease-suppressive functions in the endophytic root microbiome. *Science* **366**, 606–612 (2019).
9. Bai, Y. et al. Functional overlap of the *Arabidopsis* leaf and root microbiota. *Nature* **528**, 364–369 (2015).
10. Lebeis, S. L. et al. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349**, 860–864 (2015).
11. Zhang, J. et al. Root microbiota shift in rice correlates with resident time in the field and developmental stage. *Sci. China Life Sci.* **61**, 613–621 (2018).
12. Zhang, J. et al. *NRT1.1B* is associated with root microbiota composition and nitrogen use in field-grown rice. *Nat. Biotechnol.* **37**, 676–684 (2019).
13. Castrillo, G. et al. Root microbiota drive direct integration of phosphate stress and immunity. *Nature* **543**, 513–518 (2017).
14. Huang, A. C. et al. A specialized metabolic network selectively modulates *Arabidopsis* root microbiota. *Science* **364**, eaau6389 (2019).
15. Yuan, J. et al. Root exudates drive the soil-borne legacy of aboveground pathogen infection. *Microbiome* **6**, 156 (2018).
16. Lu, T. et al. Rhizosphere microorganisms can influence the timing of plant flowering. *Microbiome* **6**, 231 (2018).
17. Kwak, M.-J. et al. Rhizosphere microbiome structure alters to enable wilt resistance in tomato. *Nat. Biotechnol.* **36**, 1100 (2018).
18. Garrido-Oter, R. et al. Modular traits of the Rhizobiales root microbiota and their evolutionary relationship with symbiotic Rhizobia. *Cell Host Microbe* **24**, 155–167 (2018).
19. Liu, Y.-X., Qin, Y. & Bai, Y. Reductionist synthetic community approaches in root microbiome research. *Curr. Opin. Microbiol.* **49**, 97–102 (2019).

20. Lagier, J.-C. et al. Culturing the human microbiota and culturomics. *Nat. Rev. Microbiol.* **16**, 540–550 (2018).
21. Pham, V. H. T. & Kim, J. Cultivation of unculturable soil bacteria. *Trends Biotechnol.* **30**, 475–484 (2012).
22. Overmann, J., Abt, B. & Sikorski, J. Present and future of culturing bacteria. *Annu. Rev. Microbiol.* **71**, 711–730 (2017).
23. Goodman, A. L. et al. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc. Natl Acad. Sci. USA* **108**, 6252–6257 (2011).
24. Lagier, J.-C. et al. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol* **1**, 16203 (2016).
25. Forster, S. C. et al. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat. Biotechnol.* **37**, 186–192 (2019).
26. Zou, Y. et al. 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat. Biotechnol.* **37**, 179–185 (2019).
27. Ji, P., Zhang, Y., Wang, J. & Zhao, F. MetaSort untangles metagenome assembly by reducing microbial community complexity. *Nat. Commun.* **8**, 14306 (2017).
28. Wang, Y., Hammes, F., Boon, N., Chami, M. & Egli, T. Isolation and characterization of low nucleic acid (LNA)-content bacteria. *ISME J.* **3**, 889–902 (2009).
29. Nichols, D. et al. Use of Ichip for high-throughput in situ cultivation of "uncultivable" microbial species. *Appl. Environ. Microbiol.* **76**, 2445–2450 (2010).
30. Aoi, Y. et al. Hollow-fiber membrane chamber as a device for in situ environmental cultivation. *Appl. Environ. Microbiol.* **75**, 3826–3833 (2009).
31. Zhang, F. et al. Microbiota transplantation: concept, methodology and strategy for its modernization. *Protein Cell* **9**, 462–473 (2018).
32. Karasov, T. L. et al. *Arabidopsis thaliana* and *Pseudomonas* pathogens exhibit stable associations over evolutionary timescales. *Cell Host Microbe* **24**, 168–179 (2018).
33. Oberhardt, M. A. et al. Harnessing the landscape of microbial culture media to predict new organism–media pairings. *Nat. Commun.* **6**, 8493 (2015).
34. Atlas, R. M. *Handbook of Microbiological Media* (CRC Press, 2010).
35. Vartoukian, S. R., Palmer, R. M. & Wade, W. G. Strategies for culture of 'unculturable' bacteria. *FEMS Microbiol. Lett.* **309**, 1–7 (2010).
36. Davis, K. E. R., Joseph, S. J. & Janssen, P. H. Effects of growth medium, inoculum size, and incubation time on culturability and isolation of soil bacteria. *Appl. Environ. Microbiol.* **71**, 826–834 (2005).
37. Joseph, S. J., Hugenholtz, P., Sangwan, P., Osborne, C. A. & Janssen, P. H. Isolation and identification of Actinobacteria from surface-sterilized wheat roots. *Appl. Environ. Microbiol.* **69**, 7210–7215 (2003).
38. Hacquard, S. et al. Microbiota and host nutrition across plant and animal kingdoms. *Cell Host Microbe* **17**, 603–616 (2015).
39. Reasoner, D. J. & Geldreich, E. E. A new medium for the enumeration and subculture of bacteria from potable water. *Appl. Environ. Microbiol* **49**, 1–7 (1985).
40. Coombs, J. T. & Franco, C. M. M. Isolation and identification of Actinobacteria from surface-sterilized wheat roots. *Appl. Environ. Microbiol.* **69**, 5603–5608 (2003).
41. Travers, R. S., Martin, P. A. W. & Reichelderfer, C. F. Selective process for efficient isolation of soil *Bacillus* spp. *Appl. Environ. Microbiol* **53**, 1263–1266 (1987).
42. Chelius, M. K. & Triplett, E. W. The diversity of archaea and bacteria in association with the roots of *Zea mays* L. *Microb. Ecol.* **41**, 252–263 (2001).
43. Costea, P. I. et al. Towards standards for human fecal sample processing in metagenomic studies. *Nat. Biotechnol.* **35**, 1069–1076 (2017).
44. Knight, R. et al. Best practices for analysing microbiomes. *Nat. Rev. Microbiol.* **16**, 410–422 (2018).
45. Chong, J., Liu, P., Zhou, G. & Xia, J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nat. Protoc.* **15**, 799–821 (2020).
46. Liu, Y.-X. et al. A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* https://doi.org/10.1007/s13238-020-00724-8 (2020).
47. Liu, Y.-X., Qin, Y., Guo, X. & Bai, Y. [Methods and applications for microbiome data analysis]. *Yi Chuan* **41**, 845–862 (2019).
48. Qian, X.-B. et al. A guide to human microbiome research: study design, sample collection, and bioinformatics analysis. *Chin. Med. J.* **133**, 1844–1855 (2020).
49. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
50. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
51. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
52. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
53. Chen, Q. et al. Recently duplicated sesterterpene (C25) gene clusters in *Arabidopsis thaliana* modulate root microbiota. *Sci. China Life Sci.* **62**, 947–958 (2019).
54. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
55. Kolde, R. pheatmap: Pretty Heatmaps (2015).

56. Asnicar, F., Weingart, G., Tickle, T. L., Huttenhower, C. & Segata, N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **3**, e1029 (2015).
57. Wang, Y. et al. GSA: Genome Sequence Archive. *Genomics Proteom. Bioinforma.* **15**, 14–18 (2017).
58. Partners, N. G. D. C. M. Database resources of the National Genomics Data Center in 2020. *Nucleic Acids Res.* **48**, D24–D33 (2020).
59. Cole, J. R. et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642 (2014).
60. Grüning, B. et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat. Methods* **15**, 475–476 (2018).
61. Sayers, E. W. et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **48**, D9–D16 (2020).
62. Cook, C. E., Stroe, O., Cochrane, G., Birney, E. & Apweiler, R. The European Bioinformatics Institute in 2020: building a global infrastructure of interconnected data resources for the life sciences. *Nucleic Acids Res.* **48**, D17–D23 (2019).

Author contributions

J.Z. and X.G. performed the experiments. Y.-X.L. and Y.Q. performed the analysis. Y.Q., J.Z., Y.-X.L. and Y.B. designed the figures. R.G.-O., P.S.-L. and Y.B. supervised the project. J.Z., Y.-X.L., R.G.-O., P.S-L and Y.B. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-020-00444-7.

Correspondence and requests for materials should be addressed to R.G-O. or P.S.-L. or Y.B.

Peer review information *Nature Protocols* thanks Paul Dennis, Kiwamu Minamisawa and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Related links
Key references using this protocol
Zhang, J. et al. *Nat. Biotechnol.* **37**, 676–684 (2019): https://doi.org/10.1038/s41587-019-0104-4
Huang, A. C. et al. *Science* **364**, eaau6389 (2019): https://doi.org/10.1126/science.aau6389
Bai, Y. et al. *Nature* **528**, 364–369 (2015): https://doi.org/10.1038/nature16192

natureresearch

Corresponding author(s): Yang Bai, Paul Schulze-Lefert, Ruben Garrido-Oter

Last updated by author(s): Dec 20, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Culturome v1.0
Data analysis	R v3.5.1, QIIME v1.91., USEARCH v10.0.240, VSEARCH v.2.7.1, GraPhlAn v1.1.3-1, Culturome v1.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The raw sequence data used in this paper have been deposited in the Genome Sequence Archive (GSA)55 in the BIG Data Center56 under accession number CRA002517. The authors declare that all data supporting the findings of this study are available in the original paper. Since many labs may not frequently perform systematic bacterial isolation, to reduce the experimental cost, we are happy to share the full set of diluted barcoded primers with any scientists who wish to use this protocol. Correspondence and requests for materials and computational scripts should be addressed to Y.B. ybai@genetics.ac.cn.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	NA
Data exclusions	NA
Replication	NA
Randomization	NA
Blinding	NA

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology
☒ ☐	Animals and other organisms
☒ ☐	Human research participants
☒ ☐	Clinical data

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging