ARTICLES
https://doi.org/10.1038/s41587-019-0104-4
nature biotechnology

NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice

Jingying Zhang[1,2,10], Yong-Xin Liu[1,2,10], Na Zhang[1,2,3,10], Bin Hu[1,10], Tao Jin[4,5,6,10], Haoran Xu[1,2,3], Yuan Qin[1,2,3], Pengxu Yan[4,5,6], Xiaoning Zhang[1,2,3], Xiaoxuan Guo[1,2], Jing Hui[7], Shouyun Cao[1], Xin Wang[1,2], Chao Wang[1,2], Hui Wang[1,2,3], Baoyuan Qu[1,2], Guangyi Fan[4,5,6], Lixing Yuan[7], Ruben Garrido-Oter[8,9], Chengcai Chu[1,3*] and Yang Bai[1,2,3*]

Nitrogen-use efficiency of *indica* varieties of rice is superior to that of *japonica* varieties. We apply 16S ribosomal RNA gene profiling to characterize root microbiota of 68 *indica* and 27 *japonica* varieties grown in the field. We find that *indica* and *japonica* recruit distinct root microbiota. Notably, *indica*-enriched bacterial taxa are more diverse, and contain more genera with nitrogen metabolism functions, than *japonica*-enriched taxa. Using genetic approaches, we provide evidence that *NRT1.1B*, a rice nitrate transporter and sensor, is associated with the recruitment of a large proportion of *indica*-enriched bacteria. Metagenomic sequencing reveals that the ammonification process is less abundant in the root microbiome of the *nrt1.1b* mutant. We isolated 1,079 pure bacterial isolates from *indica* and *japonica* roots and derived synthetic communities (SynComs). Inoculation of IR24, an *indica* variety, with an *indica*-enriched SynCom improved rice growth in organic nitrogen conditions compared with a *japonica*-enriched SynCom. The links between plant genotype and root microbiota membership established in this study will inform breeding strategies to improve nitrogen use in crops.

The evolution of multiple plant traits is linked to the diverse microbe-rich soil environment[1,2]. Numerous microorganisms participate in the biogeochemical transformation of soil nutrients, including nitrogen and phosphorus cycling, which are essential for plant productivity[3]. Plant roots selectively recruit defined root microbiota from soil[4,5] in both model and crop plants[4–10]. It has been reported that root microbiota are determined by the rhizocompartment, soil type, geographical location, nutrient status, developmental stage and host genotype[11–17]. However, variation in the root microbiota within a plant population and the linked plant genetic basis are not yet clear.

Rice (*Oryza sativa* L.) is an important staple cereal that is grown worldwide. Two main types of Asian cultivated rice, *indica* and *japonica*, have been recognized[18,19]. Differences between *indica* and *japonica* rice are well documented at the genomic and phenotypic levels[20–22]. One striking phenotypic difference is the better nitrogen-use efficiency (NUE) of *indica* varieties compared with *japonica* varieties in the field[23,24]. Variation in NUE has been attributed, in part, to a natural variation in *NRT1.1B*, encoding a nitrate transporter and sensor (amino acid 327 is methionine in *indica* versus threonine in *japonica*)[25,26]. Nitrogen is mainly present as nitrate, ammonium and organic nitrogen in soil. Different forms of nitrogen can be metabolized by soil-borne bacteria[3,27] and this might affect the efficiency of nitrogen absorption by plant roots, as plants have a preference for inorganic nitrogen (nitrate (NO_3^-) and ammonium (NH_4^+)) rather than organic nitrogen[28], resulting in the hypothesis that root-associated bacteria might be involved in NUE variation between different plant subspecies, including *indica* and *japonica* rice. Moreover, rice roots in flooded soil recruit a specific root microbiota[29]. The roles of the rice genotype in shaping the root microbiota have been shown using a limited set of two to four rice varieties[7,30]. However, little is known about the pattern of variation and functional divergence in the root microbiota in rice populations. The availability of multiple *indica* and *japonica* varieties makes it possible to investigate the association of root microbiota variation with specific plant traits in rice subspecies and to identify key genes that contribute to the root microbiota variation during rice domestication.

In this study, we examined the composition and variation in the root microbiota of a landrace population of 68 *indica* and 27 *japonica* varieties grown under field conditions to identify links between the root microbiota, nutrient cycles in soil, host genetic regulation and NUE in the field.

Results

***Indica* and *japonica* have distinct root microbiota.** To explore variation in the root microbiota of cultivated rice, we grew 68 *indica* and 27 *japonica* varieties in two fields in separated locations on Lingshui Farm, Hainan Province, China (Supplementary Table 1). These varieties were collected from 44 countries as representatives of the rice mini-core collection from the US Department of Agriculture rice germplasm seed bank (Fig. 1a and Supplementary Table 2). Rarefaction analysis revealed that our population captured most root microbiota members from each rice subspecies (Supplementary Fig. 1). To avoid disturbance from seed-associated

[1]State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, The Innovative Academy of Seed Design, Chinese Academy of Sciences, Beijing, China. [2]CAS-JIC Centre of Excellence for Plant and Microbial Science, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [3]College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing, China. [4]BGI-Shenzhen, Shenzhen, China. [5]BGI-Qingdao, BGI-Shenzhen, Qingdao, China. [6]China National Genebank-Shenzhen, BGI-Shenzhen, Shenzhen, China. [7]Key Lab of Plant–Soil Interactions, MOE, College of Resources and Environmental Sciences, China Agricultural University, Beijing, China. [8]Department of Plant Microbe Interactions, Max Planck Institute for Plant Breeding Research, Cologne, Germany. [9]Cluster of Excellence on Plant Sciences, Dusseldorf, Germany. [10]These authors contributed equally: Jingying Zhang, Yong-Xin Liu, Na Zhang, Bin Hu, Tao Jin. *e-mail: ccchu@genetics.ac.cn; ybai@genetics.ac.cn

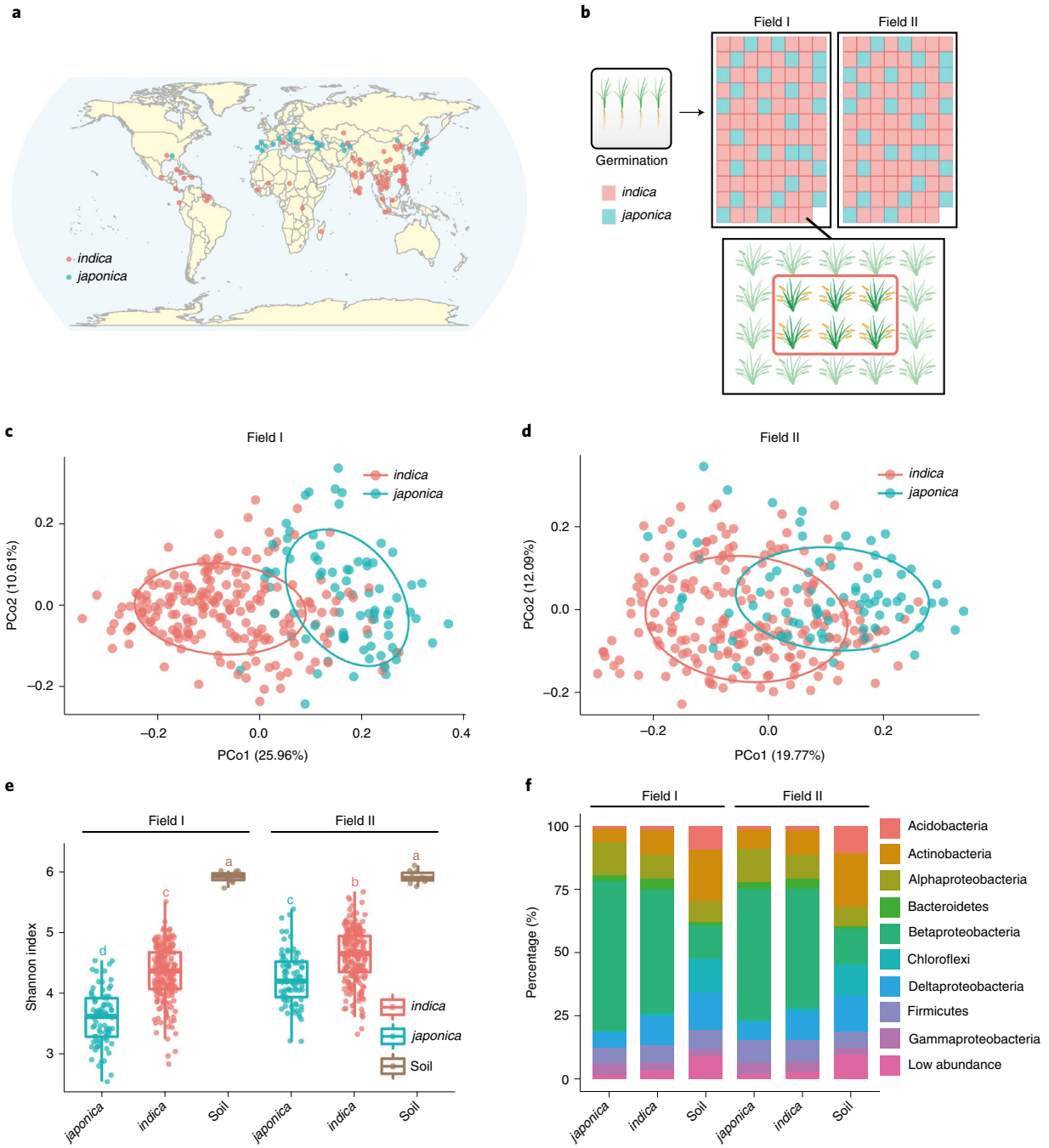

Fig. 1 | Root microbiota of *indica* and *japonica*. a, Diagram of original collection sites (44 countries) of *indica* (red) and *japonica* (blue) rice. **b**, Diagram of the experimental design for rice field trials. The *indica* and *japonica* varieties were arranged randomly. Harvested samples for each variety were surrounded by protection plants that separated the different varieties. **c**, Unconstrained PCoA (for principal coordinates PCo1 and PCo2) with Bray–Curtis distance showing that the root microbiota of *indica* separate from those of *japonica* in field I in the first axis ($P < 0.001$, permutational multivariate analysis of variance (PERMANOVA) by Adonis). Ellipses cover 68% of the data for each rice subspecies. **d**, Unconstrained PCoA with Bray–Curtis distance showing that the root microbiota of *indica* separate from those of *japonica* in field II in the first axis ($P < 0.001$, PERMANOVA by Adonis). **e**, Shannon index of the microbiota of roots from *indica*, *japonica* and the corresponding bulk soils in two fields. The horizontal bars within boxes represent medians. The tops and bottoms of 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, respectively. **f**, Phylum-level distribution of the *indica* and *japonica* root microbiota in two fields. Proteobacteria are shown at the class level. The numbers of replicated samples in this figure are as follows: in field I, *indica* ($n = 201$), *japonica* ($n = 80$), soil ($n = 12$); in field II, *indica* ($n = 201$), *japonica* ($n = 81$), soil ($n = 12$).

microbes, we germinated sterilized, dehulled rice seeds on MS agar medium and transferred 15-day-old seedlings to the fields in a random arrangement (Fig. 1b). The two fields had been used to grow rice for the past 10 years, with differential cultivation practices in 2014–2016 (Supplementary Table 1 and Methods), which enabled us to examine stable patterns of root microbiota membership between *indica* and *japonica* rice in representative agricultural soil. The root-associated microbiota were sampled 8 weeks after the

rice seedlings were transplanted, when the root microbiota were well established and stable[13,17]. For each variety, we harvested three root samples from three representative plants (Fig. 1b). We collected 12 unplanted soil samples from each field. We generated a bacterial community profile for each sample via PCR amplification of the 16S ribosomal RNA (rRNA) gene targeting regions V5–V7 using primers 799F and 1193R, followed by Illumina sequencing. We obtained 19,976,393 high-quality sequences from 563 samples (average, 35,482; range, 5,386–125,709 reads per sample). We analyzed high-quality reads with USEARCH, removing chimeric and organelle sequences, to produce 5,141 operational taxonomic units (OTUs; mean, 1,455 OTUs per sample; Supplementary Table 3; raw data are deposited in the Genome Sequence Archive in Beijing Institute of Genomics (BIG) Data Center with accession code CRA001372).

We found that the composition of the root bacterial microbiota differed in rice subspecies. Unconstrained principal coordinate analysis (PCoA) of Bray–Curtis distance revealed that the root microbiota of *indica* and *japonica* in field I formed two distinct clusters, which separated along the first coordinate axis (Fig. 1c and Supplementary Fig. 2), indicating that the largest source of variation in the rice root microbiota was proximity to the subspeciation between *indica* and *japonica*. As expected, the root microbiota in field II differed from that in field I due to soil differences (Supplementary Fig. 3 and Supplementary Table 1), but the separation of root microbiota between *indica* and *japonica* varieties was consistent in the two locations (Fig. 1d and Supplementary Figs. 2 and 3). Measurement of within-sample diversity (α-diversity) revealed a significant difference between *indica* and *japonica* varieties (Fig. 1e and Supplementary Fig. 4). The root microbiota of *indica* had higher diversity than those of *japonica* in both fields (Fig. 1e), indicating that *indica* roots recruited more bacterial species than *japonica* rice. The differences in the root microbiota between *indica* and *japonica* were significant and detectable at the phylum level (Fig. 1f). In both fields, Deltaproteobacteria, Actinobacteria, Acidobacteria, Bacteroidetes, Spirochaetes, Chloroflexi, Ignavibacteriae, Nitrospirae, and Verrucomicrobia were present at higher relative abundance in *indica* compared with *japonica*, whereas the relative abundance of Alphaproteobacteria was higher in *japonica* (false discovery rate (FDR) adjusted $P < 0.05$, Wilcoxon rank sum test; Supplementary Table 4). These data show that the root microbiota were reproducibly different in these rice subspecies.

Root microbiota as biomarkers for *indica* and *japonica*. Next, we analyzed whether root microbiota members can be used as biomarkers to differentiate *indica* and *japonica*, in common with other traits, such as grain shape, leaf color and genomic features[21,22]. We established a model using a random-forest machine-learning method[31] to correlate *indica* and *japonica* with root microbiota data in field II at the phylum, class, order, family, genus and OTU levels. In relation to rice subspecies, our model using bacterial families showed 83.7% accuracy of root microbiota classification, the highest within all taxonomic levels (Supplementary Table 5). We carried out ten-fold cross-validation with five repeats to evaluate the importance of indicator bacterial families. The cross-validation error curve stabilized when the 18 most relevant families were used. Thus, we defined these 18 families as biomarker taxa (Fig. 2a). Of these, 15 families showed higher relative abundance in *indica* than *japonica* rice; 3 families showed higher relative abundance in *japonica* (FDR adjusted $P < 0.05$, Wilcoxon rank sum test; Fig. 2b). The higher number of bacterial families enriched in *indica* was consistent with the observation that *indica* varieties showed higher microbial diversity.

Our model accurately predicted *indica* and *japonica* varieties in the tested fields. The model showed 86.0% average accuracy for the same set of *indica* and *japonica* varieties grown in field I, with 94.5% accuracy for *indica* and 77.5% for *japonica* rice (Fig. 2c and Supplementary Table 5). The model also successfully classified IR24 (*indica*) and ZH11 (*japonica*), which were not included within the training varieties and were grown on two farms in Beijing, 2,482 km away from Lingshui Farm (Fig. 2d and Supplementary Table 5). As the root microbiota are significantly affected by the geographical location[4,5] and our model was based on six fields over 2 years, data from more rice varieties grown in more different geographical locations are likely to further improve model accuracy. These results show that the root microbiota can serve as a biomarker to distinguish *indica* and *japonica* varieties.

Root microbiota and nitrogen use. Next we examined differences in the root microbiota of *indica* and *japonica* rice subspecies at the OTU level. First, we analyzed the enrichment of OTUs according to their taxonomy using Manhattan plots (Fig. 3a,b and Supplementary Table 6). In both fields, OTUs enriched in *indica* belonged to a wide range of bacterial phyla, including Acidobacteria, Proteobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Firmicutes and Verrucomicrobia (FDR adjusted $P < 0.05$, Wilcoxon rank sum test; Fig. 3a,b and Supplementary Table 6). *Japonica* roots retained the capacity to enrich OTUs belonging to Proteobacteria, Bacteroidetes and Firmicutes (FDR adjusted $P < 0.05$, Wilcoxon rank sum test; Fig. 3a,b and Supplementary Table 6). Second, we detected a notable overlap of *indica*-enriched OTUs or *japonica*-enriched OTUs in both fields (Fig. 3c,d): 141 OTUs (57.8% in field I; 95.9% in field II) were enriched in *indica* in both fields, with 51 OTUs enriched in *indica* roots compared with soil accounting for 76.4% of the cumulative abundance of *indica*-enriched OTUs (Fig. 3c, Supplementary Fig. 5 and Supplementary Table 6); 16 OTUs (47.1% in field I; 76.2% in field II) were enriched in *japonica*, with 14 OTUs enriched in *japonica* roots accounting for 99.0% of the cumulative abundance of *japonica*-enriched OTUs (Fig. 3d, Supplementary Fig. 5 and Supplementary Table 6). Most of the OTUs with differential relative abundance in *indica* and *japonica* were consistently present in both test fields.

We found that many OTUs enriched in *indica* varieties had functions related to the nitrogen cycle. We annotated the functions of OTUs that were specifically enriched in *indica* or *japonica* in both test fields using FAPROTAX, a database for converting microbial community profiles into putative functional profiles based on current literatures on cultivated strains[32]. These specific OTUs were related to 37 pathways (Supplementary Fig. 6). Notably, most *indica*-enriched OTUs were related to nutrient cycles, especially significant (25.5%, 36 out of 141 OTUs) in nitrogen metabolism, including nitrate ammonification, nitrate denitrification, nitrate reduction, nitrification and nitrite ammonification ($P = 0.025$, permutation test; Fig. 3e,g and Supplementary Table 7), reflecting the complexity of nitrogen forms used by rice in flooded soil. The cumulative abundance of nitrogen-related OTUs enriched in *indica* (36 OTUs, 11.7%) was superior to those enriched in *japonica* (5 OTUs, 3.4%; Supplementary Table 7). Further, six of the nine most abundant functions enriched in *indica* in both fields were related to the nitrogen cycle, accounting for 53.3% total cumulative abundance of *indica*-enriched functions (Fig. 3i,j and Supplementary Table 7), indicating that nitrogen transformation is more active in the root environment of *indica* than *japonica* varieties, which is consistent with previously reported higher NUE in *indica* varieties[25,33–35]. Notably, the relative abundance of nitrogen cycle-related OTUs increased during the later stage of rice growth in the field compared to plants at the seedling stage (Fig. 3f,h and Supplementary Table 7), suggesting that rice plants might actively coordinate the environmental microbial community to modulate soil nutrients for optimal plant growth.

***NRT1.1B* and root microbiota variation.** *NRT1.1B* functions as a nitrate transporter and sensor, and has been reported to contribute to the difference in NUE of *indica* and *japonica* varieties[25,26]. We

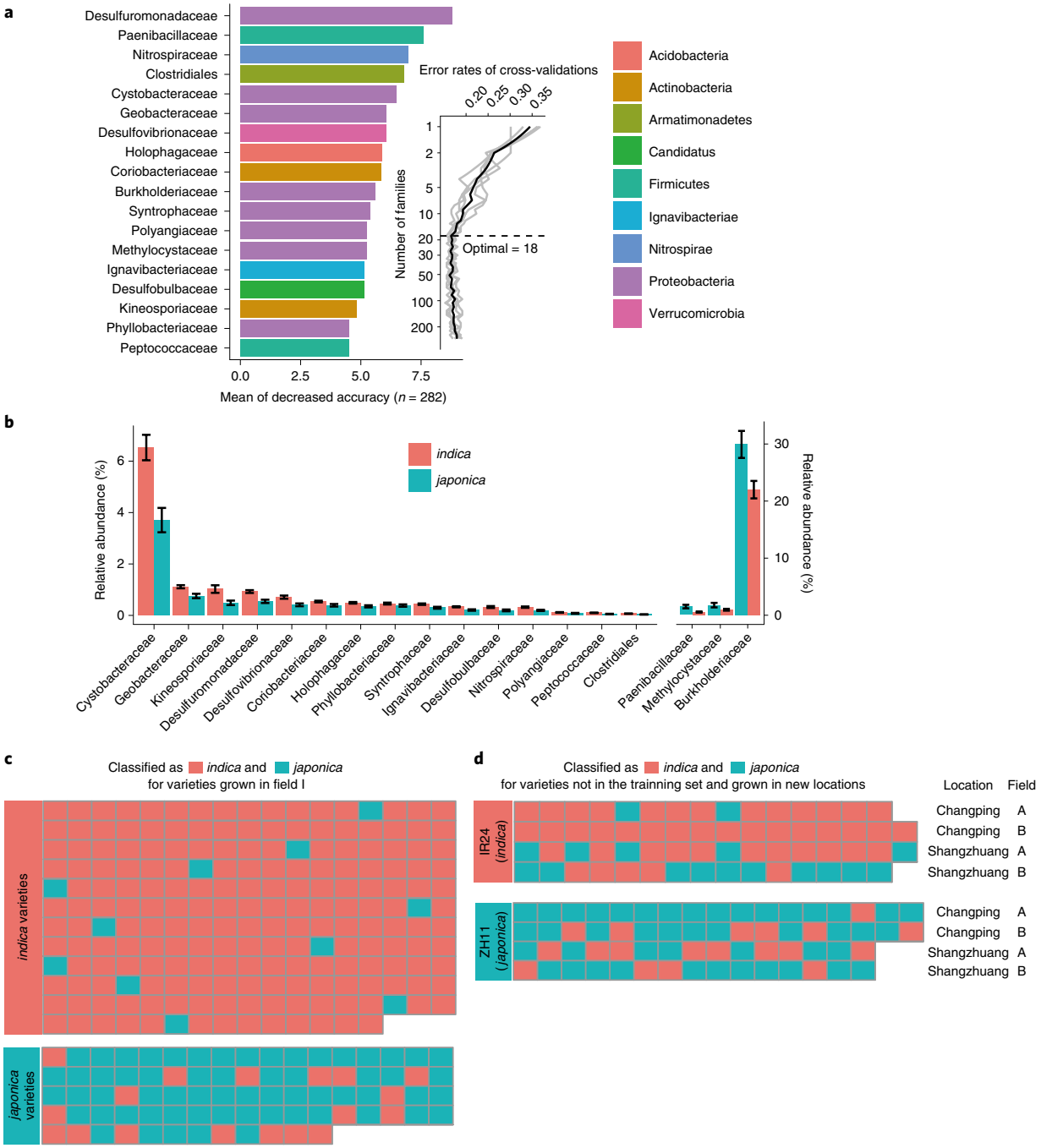

Fig. 2 | Random-forest model detects bacterial taxa that accurately predict *indica* and *japonica* subspeciation. a, The top 18 bacterial families were identified by applying random-forest classification of the relative abundance of the root microbiota in field II against *indica* and *japonica* rice. Biomarker taxa are ranked in descending order of importance to the accuracy of the model. The inset represents ten-fold cross-validation error as a function of the number of input families used to differentiate *indica* and *japonica* root microbiota in order of variable importance. **b**, Biomarker families with higher relative abundance in the root microbiota of *indica* and *japonica* rice. Data bars represent means. Error bars represent s.e.m. **c**, The results of the prediction of *indica* and *japonica* rice in field I according to the random-forest model. Each square represents an individual plant from 68 *indica* and 27 *japonica* varieties. The characteristics of the varieties are shown on the left. The predicted characteristics are shown on the right. *Indica* varieties are shown in red; *japonica* varieties are shown in blue. **d**, The results of the prediction of *indica* and *japonica* varieties that were not included in the training set and were grown in different geographical locations, according to the random-forest model. Each square represents an individual plant from tested varieties. The numbers of replicated samples in this figure are as follows: in field I, *indica* ($n = 201$), *japonica* ($n = 80$); in field II, *indica* ($n = 201$), *japonica* ($n = 81$).

hypothesize that the differences in the root microbiota of *indica* and *japonica* are regulated by *NRT1.1B*. We found that single nucleotide polymorphism (SNP) of *NRT1.1B* natural variation was correlated with the differential functions of the nitrogen metabolism in root microbiota between *indica* and *japonica* rice ($P = 2.2 \times 10^{-16}$ in field I; $P = 1.8 \times 10^{-12}$ in field II, two-sided *t*-test; Fig. 4a and Supplementary Fig. 7), indicating that *NRT1.1B* might be involved in establishment of the rice root microbiota.

To investigate the role of *NRT1.1B* in recruitment of the rice root microbiota, we carried out 16S rRNA gene amplicon sequencing

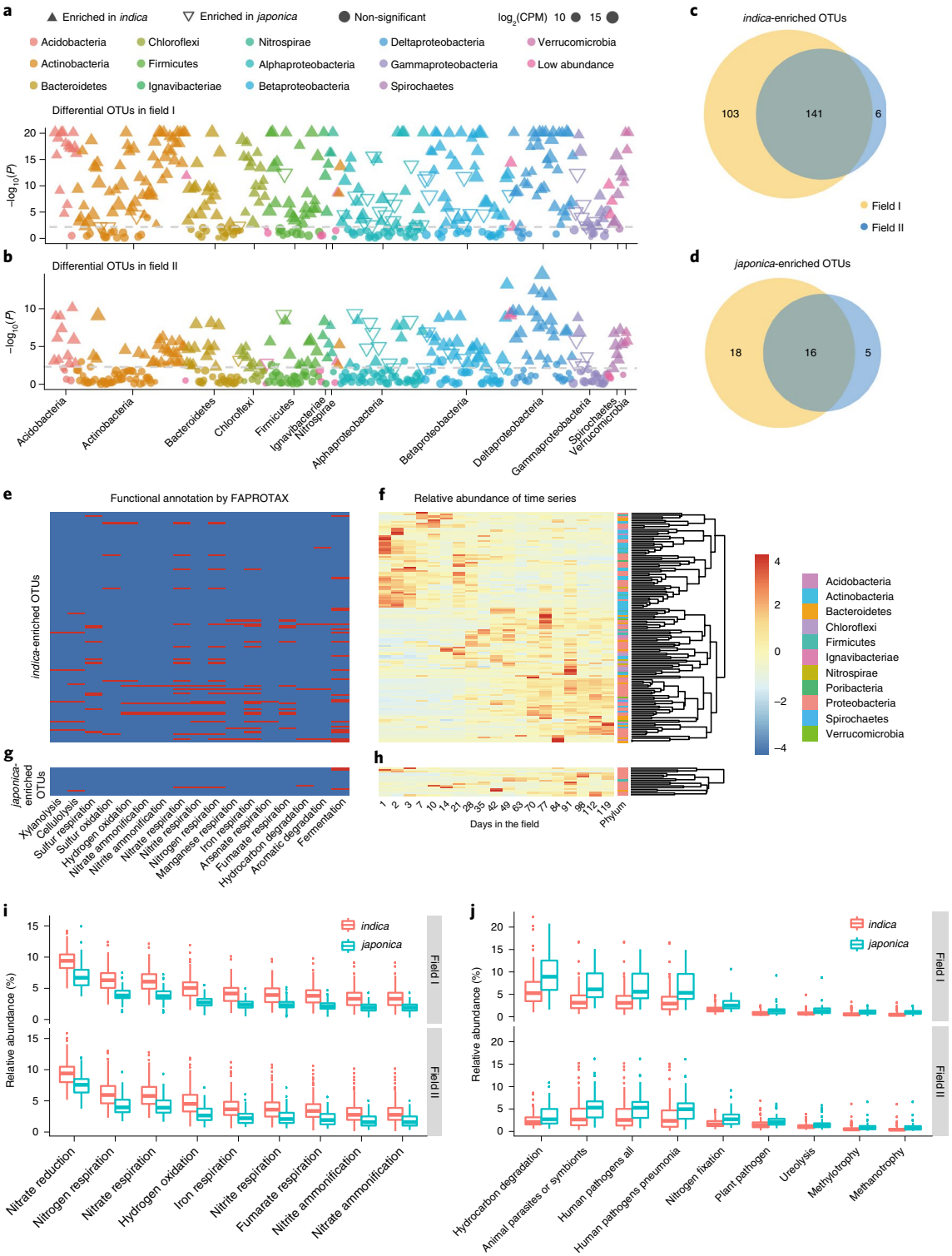

Fig. 3 | Taxonomic and functional characteristics of differential bacteria between the *indica* and *japonica* root microbiota. a,**b**, Manhattan plot showing OTUs enriched in *indica* or *japonica* in field I (**a**) and field II (**b**). Each dot or triangle represents a single OTU. OTUs enriched in *indica* or *japonica* are represented by filled or empty triangles, respectively (FDR adjusted $P < 0.05$, Wilcoxon rank sum test). OTUs are arranged in taxonomic order and colored according to the phylum or, for Proteobacteria, the class. CPM, counts per million. **c**,**d**, Overlapping OTUs enriched in *indica* (**c**) or *japonica* (**d**) in both fields. **e**,**g**, Metabolic and ecological functions of OTUs enriched in *indica* (**e**) and *japonica* (**g**) in two fields based on FAPROTAX. Each row represents an OTU. The presence of functions is shown in red. **f**,**h**, Pattern of relative abundance of the corresponding OTUs enriched in *indica* (**f**) or *japonica* (**h**) according to time-course data from the rice root microbiota in the field on Changping Farm[17]. Columns represent the days of rice growth in the field. **i**,**j**, Cumulative relative abundance of metabolic and ecological functions enriched in the root microbiota of *indica* (**i**) or *japonica* (**j**) varieties in two fields. The horizontal bars within 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, respectively. The numbers of replicated samples in this figure are as follows: in field I, *indica* ($n = 201$), *japonica* ($n = 80$); in field II, *indica* ($n = 201$), *japonica* ($n = 81$).

to profile the root microbiota of wild-type ZH11 and the loss-of-function mutant, *nrt1.1b*, in the field of Changping Farm, Beijing, China. We found that *NRT1.1B* was associated with the relative abundance of *indica*-enriched OTUs. Unconstrained PCoA of Bray–Curtis distance revealed that the *nrt1.1b* mutant and wild-type root microbiota formed separated clusters in the first two coordinate axes, which means that *NRT1.1B* has a role in the establishment of the rice root microbiota (Fig. 4b and Supplementary Fig. 8). Similar results were obtained in a full factorial replication (Supplementary Fig. 8). Notably, OTUs depleted in the *nrt1.1b* mutant showed significant overlap with *indica*-enriched OTUs in the two above-mentioned fields on Lingshui Farm (Fig. 4c,e, Supplementary Figs. 8 and 9, and Supplementary Table 8). For instance, 67.6% (119 out of 176) of OTUs depleted in the *nrt1.1b* mutant overlapped with *indica*-enriched OTUs in at least one field, and 49.6% (70 out of 141) of *indica*-enriched OTUs in both locations were depleted in the *nrt1.1b* mutant accounting for 86.5% total abundance of *indica*-enriched OTUs (Supplementary Table 8). Many bacteria involved in the nitrogen cycle, including *Anaeromyxobacter*, which has the highest relative abundance in *indica*-enriched bacteria and has a role in ammonification, are present in wild-type but are less abundant in the *nrt1.1b* mutant (Fig. 4c,e and Supplementary Table 8), suggesting that *NRT1.1B* contributes to the recruitment and abundance of a large proportion of *indica*-enriched bacteria in the rice root microbiota.

We also found a measurable difference between the root microbiota of the recipient parent Nipponbare *NRT1.1B*^{japonica} and near-isogenic line[25] Nipponbare *NRT1.1B*^{indica} on Changping Farm (Fig. 4b and Supplementary Fig. 8). Specifically, six out of seven OTUs enriched in *NRT1.1B*^{indica} were also present in *indica*-enriched OTUs in field I and II on Lingshui Farm (Fig. 4d,f, Supplementary Figs. 8 and 9 and Supplementary Table 8). The abundance of these six OTUs accounted for 29.0% of the total *indica*-enriched OTUs (Supplementary Table 8). These results provide further evidence that *NRT1.1B* is associated with establishment of the rice root microbiota and contributes to the divergence of the root microbiota between *indica* and *japonica* subspecies.

To further explore the role of *NRT1.1B* in root microbiome establishment, we carried out metagenomic sequencing of *nrt1.1b* and wild-type ZH11 root microbiomes. After removing plant-derived sequences, we obtained an average of 30.8 gigabases (Gb) of microbial sequences per sample (raw data are deposited in Genome Sequence Archive in the BIG Data Center with accession code CRA001362). We used a de novo assembly of reads to predict genes and annotate them by KEGG[36]. Genes related to the nitrogen cycle were selected from gene profiles. Consistent with our

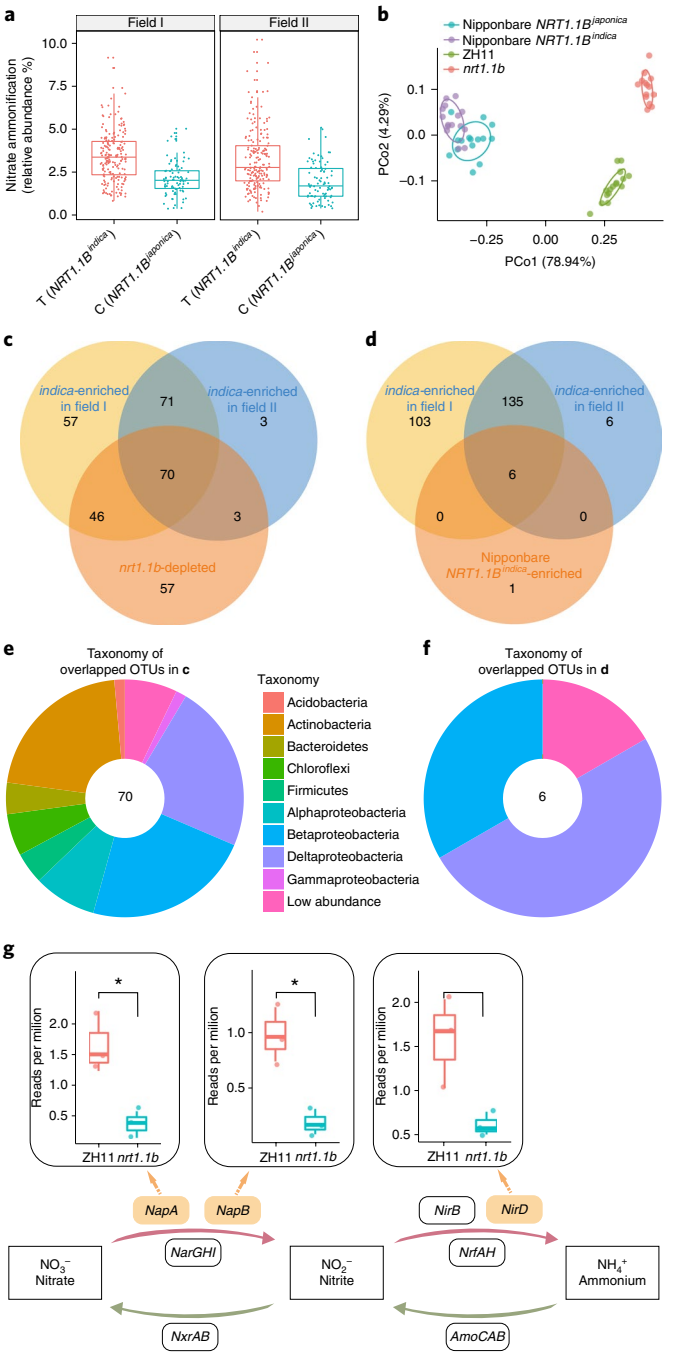

Fig. 4 | *NRT1.1B* contributes to the variation in the root microbiota of *indica* and *japonica*. **a**, Cumulative relative abundance of nitrate ammonification function correlates with the natural variation in *NRT1.1B* in *indica* and *japonica* populations ($P = 2.2 \times 10^{-16}$ in field I; $P = 1.8 \times 10^{-12}$ in field II, two-sided *t*-test). *NRT1.1B*^{indica} harbors a T 980 bp downstream of the ATG start codon and *NRT1.1B*^{japonica} harbors a C at the same position, resulting in an amino acid substitution (Met327Thr). The numbers of replicated samples are as follows: in field I, *indica* ($n = 192$), *japonica* ($n = 86$); in field II, *indica* ($n = 192$), *japonica* ($n = 87$). **b**, Unconstrained PCoA of Bray–Curtis distance showing that the root microbiota of ZH11 (wild-type), *nrt1.1b*, *NRT1.1B*^{indica} and *NRT1.1B*^{japonica} separate in the first two axes. Ellipses cover 68% of the data for each genotype. The numbers of replicated samples are as follows: ZH11 ($n = 16$), *nrt1.1b* ($n = 14$), *NRT1.1B*^{indica} ($n = 15$) and *NRT1.1B*^{japonica} ($n = 15$). **c**, The overlap of OTUs depleted in the *nrt1.1b* mutant and *indica*-enriched OTUs in two fields. **d**, The overlap of OTUs enriched in *NRT1.1B*^{indica} compared to *NRT1.1B*^{japonica} and *indica*-enriched OTUs in two fields. **e**, Taxonomic composition of the overlapped OTUs in **c** depleted in the *nrt1.1b* mutant and *indica*-enriched OTUs in two fields. **f**, Taxonomic composition of the overlapped OTUs in **d** enriched in *NRT1.1B*^{indica} compared to *NRT1.1B*^{japonica} and *indica*-enriched OTUs in two fields. **g**, *NRT1.1B* is associated with genes related to the formation of ammonium in the root microbiota. The diagram represents microbial pathways related to the soil nitrogen cycle. The relative abundances of *NapA* (FDR adjusted $P = 0.031$, Welch's *t*-test) and *NapB* (FDR adjusted $P = 0.027$, Welch's *t*-test) were significantly less abundant in the root microbiome of the *nrt1.1b* mutant. Although the change of *NirD* did not pass the statistical threshold after multiple test correction (FDR adjusted $P = 0.057$, Welch's *t*-test), it showed a clear trend of reduction in the *nrt1.1b* mutant. In this figure, the horizontal bars within 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, respectively. The asterisk indicates a significant difference at FDR adjusted $P < 0.05$. The numbers of replicated samples are as follows: ZH11 ($n = 3$), *nrt1.1b* ($n = 3$).

functional predictions, several nitrogen cycle-related genes showed a differential relative abundance between the root microbiome of ZH11 and *nrt1.1b* (Supplementary Table 9). The relative abundance of three key genes (*NapA*, *NapB* and *NirD*) in the ammonification process were reduced 2.8–5.9-fold in the root microbiome of *nrt1.1b* compared with the wild-type, whereas the abundance of other genes in the nitrogen cycle did not significantly differ (Fig. 4g and Supplementary Table 9). *NapA* and *NapB* encode nitrate reductases responsible for the reduction from nitrate to nitrite[37]; *NirD* is involved in the transformation of nitrite to ammonium[38]. These data show that *NRT1.1B* is associated with the relative abundance of root bacteria that harbor key genes in the ammonification process: these root microbes may catalyze the formation of ammonium in the root environment.

Cultivation of rice root microbiota. To investigate the function of specific OTUs enriched in *indica* or *japonica* root microbiota, we established taxonomically diverse bacterial culture collections of root microbiota from rice varieties IR24 (*indica*) and Nipponbare (*japonica*) grown on Changping Farm. To maximize the likelihood of isolating the bacteria recruited by IR24 and Nipponbare roots, each variety was grown in two different fields. Bacterial isolates were recovered from pooled roots of healthy plants using limiting dilution in liquid media in 96-well microtitre plates (Methods). Based on an identification protocol that involves 454 pyrosequencing[6], we established a two-step barcoded method to target the V5–V7 of the 16S rRNA gene for taxonomic classification of the cultivated bacteria by Illumina HiSeq (Supplementary Fig. 10). Compared with pooling of barcoded PCR products from each well in one plate in the 454 pyrosequencing protocol[6] (Supplementary Fig. 11), our method identified cultivated bacteria in each well separately, which avoided chimera formation between cultivated bacteria in different wells or plates (Supplementary Fig. 10). In parallel, parts of the roots were used for cultivation-independent 16S rRNA gene profiling to cross-reference OTU-defined taxa with cultivated colony-forming units (CFUs) in our culture collection.

In total, 13,512 CFUs were recovered from pooled root samples of IR24 or Nipponbare rice plants grown in agriculture soils (Supplementary Table 10). After removing low-quality sequences (quality < 20), we discarded chimera and low-abundance sequences ($n < 8$) using the UNOISE algorithm. These 13,512 CFUs comprised 1,041 unique bacteria with distinct 16S rRNA gene sequences and accounted for 68.6% of *indica* and 71.7% of *japonica* root-associated OTUs (relative abundance > 0.1% in at least one sample) that we identified in our culture-independent community profiles of the corresponding root samples (Fig. 5a,b and Supplementary Table 10). We were able to identify root-derived CFUs representing all four bacterial phyla and 27 bacterial families typically associated with rice roots. Thus, most bacterial taxa that were reproducibly associated with rice roots have culturable members.

We selected a taxonomically representative core set of bacterial strains from these 13,512 CFUs, and validated them by sequential steps of purification and Sanger sequencing of the 16S rRNA gene. To increase the intraspecies genetic diversity of our culture collection, we included multiple bacterial strains with the same 16S rRNA gene sequence that were isolated from different plant roots, which

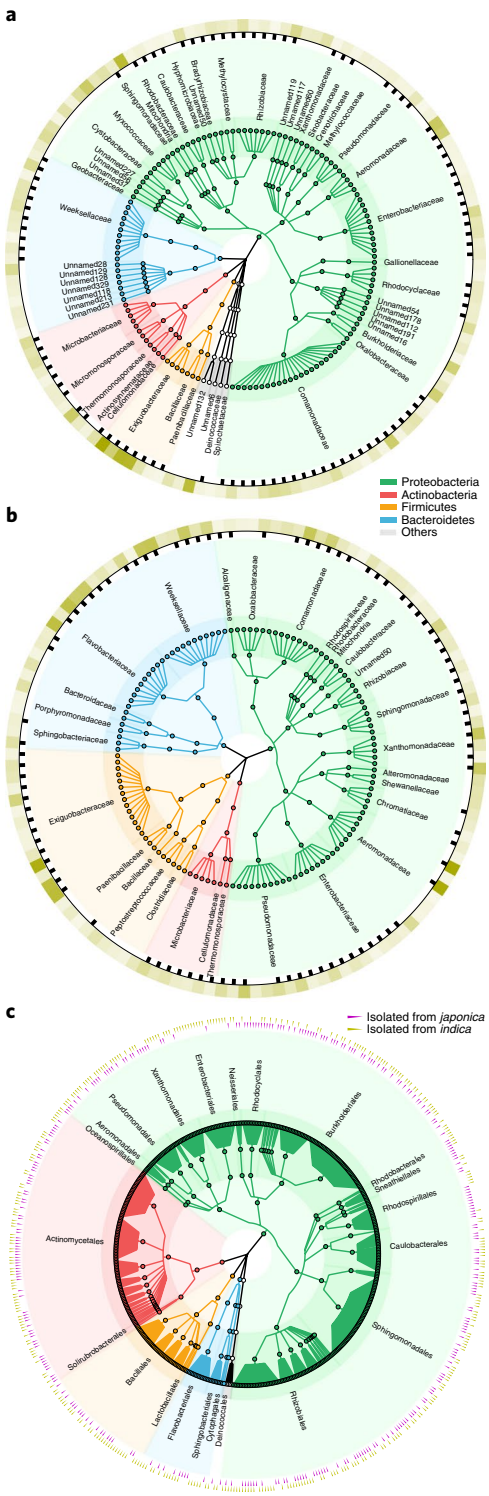

Fig. 5 | Rice root-associated bacterial culture collections capture the majority of bacterial species that are reproducibly detectable by culture-independent sequencing. a, Culture-dependent coverage (68.6%) of *indica* root-associated bacteria. The inner ring represents the root-associated OTUs reproducibly detected in the roots of IR24 (with relative abundance greater than 0.1%). The outer ring with black squares represents OTUs that are founded in the cultivated bacteria derived from IR24. The relative abundances of root-associated OTUs from IR24 are shown in the outermost ring of the white and olive heat map. **b**, Culture-dependent coverage (71.7%) of *japonica* root-associated bacteria. The inner ring represents the root-associated OTUs reproducibly detected in the roots of Nipponbare (with relative abundance greater than 0.1%). The outer ring with black squares represents OTUs that are founded in the cultivated bacteria derived from Nipponbare. The relative abundances of root-associated OTUs are shown in the outermost ring of the white and olive heat map. **c**, Taxonomic distribution of 515 unique bacterial 16S rRNA gene sequences from 1,079 purified *indica*- or *japonica*-derived bacteria (Supplementary Table 11). The origins of isolates are shown in the outer rings (*indica* as olive and *japonica* as magenta).

represented independent host colonization events (Fig. 5c). In total, we obtained a collection of 1,079 root-derived isolates that comprise 22 bacterial orders (Fig. 5c and Supplementary Table 11), which will enable future studies on the function of rice root-associated bacterial communities.

Synthetic communities improve nitrogen assimilation in rice. To probe nitrogen-related functions of *indica*- or *japonica*-enriched bacteria, we used a clay-based synthetic community (SynCom) system in which different nitrogen sources, such as nitrate, ammonium and soluble organic nitrogen (Supplementary Table 12), can be precisely controlled[6,28]. We designed SynComs based on *indica*- or *japonica*-enriched OTUs that were associated with *NRT1.1B* in our earlier comparative experiments. We selected 16 *indica*-enriched and 3 *japonica*-enriched OTUs that had representative cultivated bacteria in IR24- or Nipponbare-derived bacterial culture collections, respectively. One cultivated bacteria sharing > 97% 16S rRNA gene similarity with each OTU was used in the SynCom experiments. Under conditions of nitrate or ammonium, *indica*-enriched and *japonica*-enriched SynComs mainly reduced IR24 growth, which was probably because bacteria in the SynComs competed with rice plants for inorganic nitrogen (Supplementary Figs. 12 and 13 and Supplementary Table 13). Notably, we found that our *indica*-enriched SynCom significantly promoted IR24 growth when organic nitrogen was the sole nitrogen source (Fig. 6 and Supplementary Table 13). After 14-day co-cultivation with a 16-member *indica*-enriched SynCom, IR24 plants developed substantially longer roots ($P = 1.29 \times 10^{-12}$, ANOVA, Tukey honestly significant difference (HSD)) and larger shoots ($P = 5.74 \times 10^{-12}$, ANOVA, Tukey HSD), while a 3-member *japonica*-enriched SynCom promoted root and shoot growth but to a lesser extent compared with *indica*-enriched SynCom (Fig. 6 and Supplementary Table 13). Soil-enriched bacteria and heat-killed 16 *indica*-enriched bacteria or 3 *japonica*-enriched bacteria showed no growth promotion effect (Fig. 6 and Supplementary Table 13). These preliminary data suggest that the organic nitrogen-related function of *indica*-enriched bacteria may contribute to higher NUE in *indica* rice.

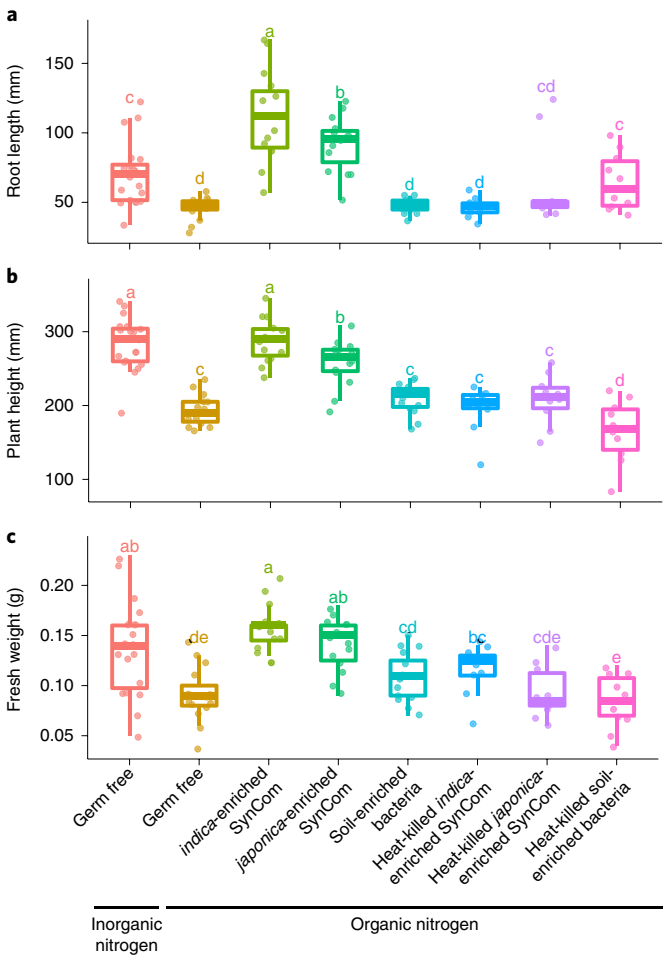

Fig. 6 | *Indica*-enriched SynCom have stronger ability to promote rice growth under the supply of organic nitrogen than *japonica*-enriched SynCom. a–c, IR24 rice plants were grown under inorganic nitrogen and organic nitrogen conditions (a mixture of five amino acids; Supplementary Table 12) with and without SynComs, including *indica*-enriched SynCom, *japonica*-enriched SynCom, soil-enriched bacteria and corresponding heat-killed bacteria as controls (Supplementary Table 12). After 2-week bacterial inoculation, rice plants were measured by root length (**a**), plant height (**b**) and shoot fresh weight (**c**). Box plots show combined data from three independent inoculation experiments with three to five technical replicates each (Supplementary Table 13). Different letters indicate significantly different groups ($P < 0.05$, ANOVA, Tukey HSD). The horizontal bars within boxes represent medians. The tops and bottoms of 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, respectively.

Discussion

During the long history of rice cultivation, two major types of Asian cultivated rice, *indica* and *japonica*, have emerged with distinct genotypes and phenotypes[20,22]. Here, we examined variations in the root microbiota of multiple representative varieties of *indica* and *japonica* rice grown under field conditions, and showed that *indica* and *japonica* varieties recruit distinct root microbiota (Fig. 1). We found that specific biomarker taxa can serve to differentiate *indica* and *japonica* rice subspecies (Fig. 2). Further, we report that *NRT1.1B* is associated with enrichment of subspecies-specific taxa in the root microbiota (Fig. 4), suggesting that the host genotype has a considerable impact on root microbiota establishment under fixed environmental conditions. As the natural variation of *NRT1.1B* explained 29.0% of the differential abundance in root microbiota between *indica* and *japonica*, it will be important to identify additional key genes and pathways involved in root microbiota variation in rice.

Coordination between the host plant and the root microbiota is essential for plant growth in natural environments[1]. We found that roots of *indica* varieties recruited a higher proportion of nitrogen cycle-related bacteria (Fig. 3), indicating that the nitrogen transformation process is probably more efficient in the root environment of *indica* than *japonica* varieties, which could explain the historical observations of the higher NUE in *indica* varieties than *japonica*[25,33–35]. Based on cultivated bacteria (Fig. 5) and the use of synthetic communities that we named SynComs, we found that an *indica*-enriched SynCom had a larger effect on rice growth than a *japonica*-enriched SynCom under conditions of organic nitrogen supply (Fig. 6). We hypothesize that SynCom bacteria transform organic nitrogen into nitrate and ammonium, so that it can be efficiently absorbed by rice roots. Further work with varied SynComs might help to identify the functions of *indica*- or *japonica*-enriched bacteria.

Intriguingly, the establishment and variation of root microbiota in rice subspecies were associated with the presence of *NRT1.1B*, which encodes a nitrate transporter and sensor. Therefore, rice roots might utilize a transporter responsible for nitrate uptake to monitor changes in nitrate concentrations in soil and in turn this could induce expression of nitrate-related genes or root metabolites, which might influence recruitment of root microbiota. Further studies

using a specific mutation in *NRT1.1B* (*chl1-9*) in *Arabidopsis*[39] will help to dissect nitrate transport and sensing activities of *NRT1.1B* in the regulation of root microbiota establishment.

Our data suggest that rice plants coordinate root microbial communities to acquire nitrogen from the soil environment and ensure optimal plant growth under agricultural conditions. The intersection between the root microbiota, soil nutrient cycles, gene regulation and host plant traits could pave the way for technologies to modulate the root microbiota that increase crop productivity and sustainability.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-019-0104-4.

Received: 19 November 2018; Accepted: 13 March 2019;
Published online: 29 April 2019

References

1. Bakker, P., Pieterse, C. M. J., De Jonge, R. & Berendsen, R. L. The soil-borne legacy. *Cell* **172**, 1178–1180 (2018).
2. Berendsen, R. L., Pieterse, C. M. & Bakker, P. A. The rhizosphere microbiome and plant health. *Trends Plant Sci.* **17**, 478–486 (2012).
3. Hacquard, S. et al. Microbiota and host nutrition across plant and animal kingdoms. *Cell Host Microbe* **17**, 603–616 (2015).
4. Bulgarelli, D. et al. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* **488**, 91–95 (2012).
5. Lundberg, D. S. et al. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**, 86–90 (2012).
6. Bai, Y. et al. Functional overlap of the *Arabidopsis* leaf and root microbiota. *Nature* **528**, 364–369 (2015).
7. Edwards, J. et al. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl Acad. Sci. USA* **112**, E911–E920 (2015).
8. 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).
9. Schlaeppi, K., Dombrowski, N., Oter, R. G., Ver Loren van Themaat, E. & Schulze-Lefert, P. Quantitative divergence of the bacterial root microbiota in *Arabidopsis thaliana* relatives. *Proc. Natl Acad. Sci. USA* **111**, 585–592 (2014).
10. Zarraonaindia, I. et al. The soil microbiome influences grapevine-associated microbiota. *mBio* **6**, e02527–14 (2015).
11. Breidenbach, B., Pump, J. & Dumont, M. G. Microbial community structure in the rhizosphere of rice plants. *Front. Microbiol.* **6**, 1537 (2016).
12. Castrillo, G. et al. Root microbiota drive direct integration of phosphate stress and immunity. *Nature* **543**, 513–518 (2017).
13. Edwards, J. A. et al. Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. *PLoS Biol.* **16**, e2003862 (2018).
14. Lebeis, S. L. et al. Salicylic acid modulates colonization of the root microbiome by specific bacterial taxa. *Science* **349**, 860–864 (2015).
15. Muller, D. B., Vogel, C., Bai, Y. & Vorholt, J. A. The plant microbiota: systems-level insights and perspectives. *Annu. Rev. Genet.* **50**, 211–234 (2016).
16. Walters, W. A. et al. Large-scale replicated field study of maize rhizosphere identifies heritable microbes. *Proc. Natl Acad. Sci. USA* **28**, 7368–7373 (2018).
17. Zhang, J. Y. 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).
18. Oka, H. I. Intervarietal variation and classification of cultivated rice. *Ind. J. Genet. Plant. Breed.* **18**, 10 (1958).
19. Zhang, Q. F., Maroof, M. A., Lu, T. Y. & Shen, B. Z. Genetic diversity and differentiation of *indica* and *japonica* rice detected by RFLP analysis. *Theor. Appl. Genet.* **83**, 495–499 (1992).
20. Huang, X. et al. A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490**, 497–501 (2012).
21. Morishima, H. & Oka, H. I. Phylogenetic differentiation of cultivated rice. XXVII. Numerical evaluations of the *indica-japonica* differentiation. *Jpn. J. Breed.* **31**, 402–413 (1981).
22. Wang, W. et al. Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* **557**, 43–49 (2018).
23. Koutroubas, S. D. & Ntanos, D. A. Genotypic differences for grain yield and nitrogen utilization in Indica and Japonica rice under mediterranean conditions. *Field Crop. Res.* **83**, 251–260 (2003).
24. Rakotoson, T. et al. Genetic variability of nitrogen use efficiency in rainfed upland rice. *Field Crop. Res.* **213**, 194–203 (2017).
25. Hu, B. et al. Variation in *NRT1.1B* contributes to nitrate-use divergence between rice subspecies. *Nat. Genet.* **47**, 834–838 (2015).
26. Wang, W. et al. Expression of the nitrate transporter *OsNRT1.1A/OsNPF6.3* confers high yield and early maturation in rice. *Plant Cell* **30**, 638–651 (2018).
27. Näsholm, T., Kielland, K. & Ganeteg, U. Uptake of organic nitrogen by plants. *New Phytol.* **182**, 31–48 (2009).
28. Bloom, A. J. The increasing importance of distinguishing among plant nitrogen sources. *Curr. Opin. Plant Biol.* **25**, 10–16 (2015).
29. Santos-Medellin, C., Edwards, J., Liechty, Z., Nguyen, B. & Sundaresan, V. Drought stress results in a compartment-specific restructuring of the rice root-associated microbiomes. *mBio* **8**, e00764–00717 (2017).
30. Shenton, M., Iwamoto, C., Kurata, N. & Ikeo, K. Effect of wild and cultivated rice genotypes on rhizosphere bacterial community composition. *Rice* **9**, 42 (2016).
31. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2**, 18–22 (2002).
32. Louca, S., Parfrey, L. W. & Doebeli, M. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**, 1272–1277 (2016).
33. Chanh, T. T., Tsutsumi, M. & Kurihara, K. Comparative study on the response of Indica and Japonica rice plants to ammonium and nitrate nitrogen. *Soil Sci. Plant Nutr.* **27**, 83–92 (1981).
34. Chao, D. Y. & Lin, H. X. Nitrogen-use efficiency: transport solution in rice variations. *Nat. Plants* **1**, 15096 (2015).
35. Ta, T. C. & Ohira, K. Effects of various environmental and medium conditions on the response of Indica and Japonica rice plants to ammonium and nitrate nitrogen. *Soil Sci. Plant Nutr.* **27**, 347–355 (1981).
36. Kanehisa, M. et al. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* **42**, 199–205 (2014).
37. Gao, H. et al. Reduction of nitrate in *Shewanella oneidensis* depends on atypical NAP and NRF systems with NapB as a preferred electron transport protein from CymA to NapA. *ISME J.* **3**, 966 (2009).
38. Malm, S. et al. The roles of the nitrate reductase NarGHJI, the nitrite reductase NirBD and the response regulator GlnR in nitrate assimilation of *mycobacterium tuberculosis*. *Microbiology* **155**, 1332–1339 (2009).
39. Ho, C. H., Lin, S. H., Hu, H. C. & Tsay, Y. F. CHL1 functions as a nitrate sensor in plants. *Cell* **138**, 1184–1194 (2009).

Acknowledgements

We thank P. Schulze-Lefert and S. Hacquard at the Max Planck Institute for Plant Breeding Research for their suggestions for improving the manuscript. This work was financially supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (grant nos. XDB11020700 to Y.B. and XDA08010104 to C.C.), the Key Research Program of Frontier Sciences of the Chinese Academy of Science (grant nos. QYZDB-SSW-SMC021 to Y.B. and QYZDJ-SSW-SMC014 to C.C.), the National Natural Science Foundation of China (grant nos. 31772400 to Y.B. and 31801945 to J.Z.), and the Key Research Program of the Chinese Academy of Sciences (grant no. KFZD-SW-219 to Y.B.). J.Z. is supported by the CPSF-CAS Joint Foundation for Excellent Postdoctoral Fellows (grant no. 2016LH00012). Y.B. is supported by the Thousand Youth Talents Plan (grant no. 2060299).

Author contributions

C.C. and Y.B. conceived the study and supervised the project. J.Z. and N.Z. performed the experiments. Y.-X.L. analyzed the data of 16S rRNA gene profiling. B.H. and L.Y. coordinated field experiments and revised the manuscript. T.J., P.Y. and R.G.-O. analyzed the metagenomic data. X.Z., Y.Q. and G.F. were involved in the informatics analysis. J.H. performed the soil properties analysis. S.C., H.X., X.W., C.W., H.W. and B.Q. participated in growing plants and harvesting samples. X.G. optimized the protocol of library preparation for the 16S rRNA gene profiling. J.Z., C.C. 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/s41587-019-0104-4.

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

Correspondence and requests for materials should be addressed to C.C. or Y.B.

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

Methods

Plant germination, transplantation and measurement of soil properties. Seeds from 95 cultivated varieties (68 *indica* and 27 *japonica*) were dehulled, surface-sterilized in 75% ethanol for 30 s and 2.5% sodium hypochlorite three times for 15 min, and sterilely germinated on MS agar media. The *indica* and *japonica* varieties were collected from 44 countries and were representative of the diversity of rice (Supplementary Table 2). After 15-day germination, the rice seedlings were transplanted into two flooded fields separated by a path 1.5 m wide on Lingshui Farm, Hainan Province, China (18.512° N, 110.044° E). The two fields had been used to cultivate rice for the past 10 years. For the past 3 years, the cultivation practice was different (Supplementary Table 1). Field I was used to grow Peiza1303, a hybrid rice variety with high taste quality, requiring relatively low nitrogen fertilizer input (1 kg N per 100 m^2). Field II was used to grow Teyou009, a hybrid rice variety with high yield requiring high nitrogen fertilizer input (2 kg N per 100 m^2). *Indica* and *japonica* varieties were grown in a random arrangement in both fields with nitrogen fertilizer input (1.5 kg N per 100 m^2) in January in 2017. Physical and chemical properties for soil were measured from three soil samples harvested at 5–6 cm depth in unplanted areas in each field. The sampling points were evenly distributed. The organic matter was determined by the oven heating method[40]. The available N, P and K in soil were determined according to the previous literature[41]. Soil micronutrient concentrations were measured using the diethylenetriaminepentaacetic acid (DTPA) extraction method[42]. A total of 10 g of air-dried soil was mixed with 20 ml of DTPA extraction solution (0.005 M DTPA, 0.01 M $CaCl_2$, 0.1 M triethanolamine, pH 7.3) by shaking for 2 h at room temperature. The liquid supernatants were filtered and subsequently analyzed by inductively coupled plasma-optical emission spectrometry (ICP-OES, Optima 7300 DV, PerkinElmer) for Fe, Mn, Cu and Zn concentrations. The soil analytes are listed in Supplementary Table 1.

Rice cultivation, sample collection and plant traits. Each rice variety was grown in a plot of 20 (4 × 5) rice plants with 20 cm spacing between individuals (Fig. 1b). The distance between plots was 30 cm. Plants on the borders of the plots were designed for protection and not harvested. Root samples were collected 8 weeks after the seedlings were transferred into the fields, at the late tillering stage. Three representative individuals from each rice variety were selected from a central position. Roots were shaken to remove the loosely adhering soil and then washed until no visible soil particles remained. Roots of length 10 cm from the ground were cut into 2 mm sections and placed into a 2-ml tube. In total, 563 rice root samples were used in this study. The details about rice varieties are provided in Supplementary Table 2. In each field, 12 of the corresponding bulk soil samples were collected from an unplanted site. All samples were immediately stored at –20 °C, transported to the laboratory with dry ice and stored at –80 °C. The *NRT1.1B*-related materials were grown on Changping farm (116.424° E, 40.109° N) in 2016 and 2017, and the planting mode and collection methods were the same as described above. Fifteen individuals of wild-type (ZH11), mutants, recipient parent (Nipponbare) and near-isogenic lines were collected for bacterial 16S rRNA gene profiling.

DNA extraction, PCR amplification and sequencing. Rice roots and corresponding unplanted soil samples were performed for bacterial 16S rRNA gene profiling by Illumina sequencing. The DNA for each sample was extracted with FastDNA SPIN Kit (MP Biomedicals). DNAs were measured by PicoGreen dsDNA Assay Kit (Life Technologies) and subsequently diluted to 3.5 ng μl^{-1}. The V5–V7 region of the bacterial 16S rRNA gene was amplified by degenerate PCR primers 799F and 1193R[4]. Each sample was amplified in triplicate (together with water control) in a 30 µl reaction system, which contained 3 µl of diluted DNA, 0.75 U PrimeSTAR HS DNA polymerase, 1x PrimSTAR buffer (Takara), 0.2 mM deoxyribonucleoside triphosphates (dNTPs) and 10 pM of barcoded forward and reverse primers. After an initial denaturation step at 98 °C for 30 s, the targeted region was amplified by 25 cycles of 98 °C for 10 s, 55 °C for 15 s and 72 °C for 60 s, followed by a final elongation step of 5 min at 72 °C. If there was no visible amplification from negative control (no template added), triplicate PCR products were mixed and purified using an AMPure XP Kit (Beckman Coulter). The purified PCR products were measured by Nanodrop (NanoDrop 2000C, Thermo Scientific), and diluted to 10 ng μl^{-1} as templates for the second step of the PCR. All samples were amplified in triplicate with second-step primers (Supplementary Table 14), using identical conditions to the first step of the PCR but with eight cycles. Technical replicates of each sample were combined and run on a 1.2% (w/v) agarose gel, and the bacterial 16S rRNA gene amplicons were extracted using a QIAquick Gel Extraction Kit (Qiagen). DNAs were subsequently measured with a PicoGreen dsDNA Assay Kit (Life Technologies) and 200 ng of each sample were mixed. Final amplicon libraries were purified twice using a Agencourt AMPure XP Kit (Beckman Coulter) and subjected to a single sequencing run on the HiSeq 2500 platform (Illumina Inc).

Bioinformatics analysis on 16S rRNA gene profiling. The 16S rRNA gene sequences were processed using QIIME v.1.9.1 (ref. [43]), USEARCH v.10.0 (ref. [44]) and in-house scripts[17]. Metadata are given in Supplementary Table 3. The quality of the paired-end Illumina reads was checked by FastQC v.0.11.5 (ref. [45]) and processed in the following steps by USEARCH: joining of paired-end reads and relabeling of sequencing names (-fastq_mergepairs); removal of barcodes and primers (-fastx_truncate); filtering of low-quality reads (-fastq_filter); and finding non-redundancy reads (-fastx_uniques). Unique reads were clustered into OTUs with 97% similarity. The representative sequences were picked by UPARSE[46]. OTUs were aligned to the SILVA 132 database to remove sequences from chimera and host plastids[47]. The OTU table was generated by USEARCH (-otutab). The taxonomy of the representative sequences was classified with the RDP classifier[48]. Diversity analysis was carried out using self-Rscript and QIIME2 (ref. [49]). Analysis of the differential OTU abundance and taxa was performed using Wilcoxon rank sum tests based on OTUs with median relative abundance from each genotype >0.2%, and corresponding *P* values were corrected for multiple tests using a FDR set at 0.05. Functional annotations of prokaryotic taxa were carried out using FAPROTAX v.1.1 (ref. [32]). Rice time-course data were used from previously published papers[17].

To acquire the best discriminant performance of taxa across *indica* and *japonica* varieties, we classified the relative abundances of bacterial taxa in the phylum, class, order, family, genus and OTU level against rice subspecies using the randomForest package v.4.6–14 in R using default parameters[31]. We used the data of field II as the training set and the randomForest (importance = TRUE, proximity = TRUE) function to generate the classification model for *indica* and *japonica*. Cross-validation was performed by the rfcv() function for selecting appropriate features. The varImpPlot function was used to show the importance of features in the classification. The importance of features and the cross-validation curve were visualized by using the ggplot2 v.2.2.1 package in R v.3.5.1 (ref. [50]). Data of field I and some varieties grown in Beijing were used as the test set for the model with default parameters.

Metagenome analysis. DNAs from three ZH11 root samples and three *nrt1.1b* mutant samples were sequenced on the BGISEQ500 platform using a pair-end 100 base pair (bp) sequencing strategy, with an average of 44 Gb per sample. All the raw data were trimmed by SOAPnuke v.1.5.2 (ref. [51]), then the clean reads were aligned against the *O. sativa* genome (gi: 996703420; accession: NW_015379174.1) to remove host contaminations using SOAP2.22 (ref. [52]). The rest of the reads (with an average of 30.8 Gb per sample) were assembled by MEGAHIT v.1.0 (ref. [53]). Assembled contigs with length less than 200 bp were discarded in the following analysis. Filtered contigs were aligned against the *O. sativa* genome to remove host contaminations by using blastn-2.2.31.

Genes were predicted over contigs by using MetaGeneMark_v1_mod[54]. Genes with length greater than or equal to 100 bp were retained. All predicted genes were combined together and then clustered using cd-hit[55] into a non-redundant gene set by setting the similarity threshold as 95%. Functional annotations were implemented by mapping non-redundant genes against KEGG v.81 (ref. [56]) by blastx-2.2.31, and results meeting the criterion identity > 30% and *E*-value score $< 1 \times 10^{-5}$ were retained. Next, clear reads were aligned against the unique gene set by using SOAP2.22 and the relative abundance of genes was calculated according to the amount of mapped reads. KEGG Orthology (KO) profiles were generated by summing up the abundance of genes affiliated to the same KO. The difference of the KO relative abundance was detected by Welch's *t*-test with FDR correction in STAMP[57].

Rice SNP data were downloaded from previously published paper[58]. The correlations between the functional SNP in *NRT1.1B* and nitrogen-related functions were evaluated by *t*-test and visualized by ggplot2 in R[50].

Isolation of root-derived bacteria. Rice plants were grown in different natural soils for bacterial isolations by limiting dilution as well as 16S rRNA gene-based community profiling. To obtain a library of representative root colonizing bacteria, we used two representative rice varieties, IR24 (*indica*) and Nipponbare (*japonica*). Each variety was grown in two fields on Changping Farm (116.424° E, 40.109° N, Beijing, China). Plants were harvested before flowering. Roots of three healthy plants of the same variety in each field were pooled to reduce the variation. Rice roots were washed three times in washing buffers (PBS) on a shaking platform for 15 min at 180 r.p.m. and then smashed. Homogenized roots from each root pool were sedimented for 15 min and the supernatants were empirically diluted, distributed and cultivated in 96-well microtiter plates in 1:10 (v/v) tryptic soy broth for 20 d at room temperature. In parallel to the isolation of root-derived bacteria, roots of rice plants grown in the aforementioned soil were harvested and used to assess bacterial diversity by culture-independent 16S rRNA gene sequencing.

We adopted a two-step barcoded PCR protocol in combination with Illumina HiSeq to define V5–V7 sequences of bacterial 16S rRNA genes of root-derived bacteria. DNA of isolates was extracted by lysis buffer with 6 µl of bacterial cultures in 10 µl of buffer I containing 25 mM NaOH and 0.2 mM EDTA, at pH 12 at 95 °C for 30 min, and then the pH value was lowered by addition of 10 µl of buffer II containing 40 mM Tris-HCl at pH 7.5. Positions of isolates in 96-well microtiter plates were indexed by a two-step PCR protocol using the degenerate primers 799F and 1193R containing well- and plate-specific barcodes to amplify the variable regions V5–V7. During the first step of PCR amplification, DNA from 3 µl of lysed cells was amplified using 0.75 U HS-Taq DNA polymerase, 10× buffer, 0.2 mM dNTPs (Takara), 0.1 µM forward primer (799F) and 0.1 µM

reverse primer (1193R) (Life Technologies) in a 30 µl reaction. PCR amplification was performed under the following conditions: DNA was initially denatured at 94 °C for 2 min, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min, and a final elongation step at 72 °C for 5 min. PCR products of each 96-well were diluted 40× for the second step of the PCR. During this second step, 3 µl of diluted PCR from the product from the first step was amplified by 0.75 U HS-Taq DNA polymerase, 10× buffer, 0.2 mM dNTPs (Takara), 0.1 µM one out of 96 barcoded forward primers (Supplementary Table 14) and 0.1 µM reverse barcoded primer in a 30 µl reaction. The PCR cycling conditions were as follows. First, denaturation at 94 °C for 2 min, followed by 25 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 1 min, and a final elongation at 72 °C for 5 min. PCR products were purified using the Agencourt AMPure XP Kit (Beckman Coulter) and Wizard SV Gel and PCR Clean-up System (Promega). DNA concentration was determined by PicoGreen dsDNA Assay Kit (Life Technologies) and samples were pooled in equal amounts (Supplementary Table 14). The final PCR product libraries were sequenced on an Illuminar HiSeq 2500. Each sequence contained a plate barcode, a well barcode and V5–V7 sequences. The sequences were quality-filtered and demultiplexed according to well and plate identifiers[43]. OTUs were defined after removing chimera and low-abundance sequences (read count < 8) by the UNOISE algorithm[59]. The taxonomy of the final stock was classified by the sintax algorithm in RDP training set 16 (ref. [60]). Cultivated bacteria were compared with OTUs in corresponding root microbiota members with greater than 97% 16S rRNA gene similarity to examine the culture-dependent coverage. Cladograms were visualized by GraPhlAn v.0.9.7 (ref. [61]).

Preparation of rice root-associated bacterial culture collections. Representative isolates for each unique OTUs identified from cultivated bacteria were purified by three consecutive platings on the respective solidified media before an individual colony was used to inoculate liquid cultures. These liquid cultures were used for validation by Sanger sequencing with both 27F and 1492R primers as well as for the preparation of glycerol stocks for the culture collections.

Plant growth promotion assay. Rice plants were grown in a clay-based SynCom containing defined nitrate, ammonium or soluble organic nitrogen, with or without bacteria. Calcined clay, an inert soil substitute, was washed with water, sterilized three times by autoclaving and heat-incubated until completely dehydrated. Rice seeds were surface-sterilized with 75% ethanol for 30 s, then in 2.5% sodium hypochlorite three times for 15 min, and germinated on MS agar media for 5 d. Root-derived bacteria of the culture collections were cultivated in 50 ml tubes in TSB medium at 28 °C for 5 d and subsequently pooled (in equal concentration) to prepare SynComs for inoculations below the carrying capacity of roots. To inoculate SynComs into the calcined clay matrix, the optical density OD_{600} of SynComs were adjusted to 0.5 and 2.0 ml SynComs were added to 250 ml 1× Kimura B nutrient solution (pH 5.8), and mixed with 200 g of calcined clay in Magenta boxes (~10^6 cells per g of calcined clay). Then 5-day-old sterile rice seedlings were transferred to the clay in the Magenta boxes. Plants were grown at 25 °C, 16 h light and 21% humidity. After 2 weeks, plant height, shoot fresh weight and root length were measured to evaluate plant growth promotion.

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

Data availability

Raw sequence data reported in this paper have been deposited (PRJCA001214) in the Genome Sequence Archive in the BIG Data Center[62], Chinese Academy of Sciences under accession codes CRA001372 for bacterial 16S rRNA gene sequencing data and CRA001362 for metagenomic sequencing data that are publicly accessible at http://bigd.big.ac.cn/gsa. All pure strains (Supplementary Table 11) are deposited in two national culture collection centers, the China Natural Gene Bank and the Agricultural Culture Collection of China. All information about these strains, such as the 16S rRNA gene sequences, taxonomy and isolation details, as well as any further updates are available at http://bailab.genetics.ac.cn/culture_collection/.

Code availability

Scripts employed in the computational analyses are available at https://github.com/microbiota/Zhang2019NBT.

References

40. Nuan, G. et al. Study on the determination of organic matter in soil by the oven heating method. *Acta Agricult. Boreali-Sinica* **32**, 343–346 (2017).
41. Zhang, L. H. et al. *NRT1.1B* improves selenium concentrations in rice grains by facilitating selenomethinone translocation. *Plant Biotechnol. J.* https://doi.org/10.1111/pbi.13037 (2019).
42. Lindsay, E. L. & Norvell, W. A. Development of a DTPA soil test for zinc, iron, manganese, and copper. *Soil Sci. Soc. Am. J.* **42**, 421–428 (1978).
43. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
44. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
45. Andrews, S. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc (2010).
46. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).
47. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, 590–596 (2012).
48. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microb.* **73**, 5261–5267 (2007).
49. Bolyen, E. et al. QIIME 2: reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ Preprints* **6**, e27295v27291 (2018).
50. Wickham, H. *ggplot2: Elegant Graphics For Data Analysis* (Springer, 2016).
51. Chen, Y. et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *GigaScience* **7**, 1–6 (2017).
52. Li, R. et al. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966–1967 (2009).
53. Li, D. et al. MEGAHITv1. 0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**, 3–11 (2016).
54. Zhu, W., Lomsadze, A. & Borodovsky, M. *Ab initio* gene identification in metagenomic sequences. *Nucleic Acids Res.* **38**, e132 (2010).
55. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
56. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
57. Parks, D. H., Tyson, G. W., Hugenholtz, P. & Beiko, R. G. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123–3124 (2014).
58. Wang, H. et al. The power of inbreeding: NGS-based GWAS of rice reveals convergent evolution during rice domestication. *Mol. Plant* **9**, 975–985 (2016).
59. Edgar, R. C. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. Preprint at https://doi.org/10.1101/081257 (2016).
60. Cole, J. R. et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642 (2014).
61. 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).
62. Members, B. I. G. D. C. Database resources of the BIG Data Center in 2018. *Nucleic Acids Res.* **46**, D14–D20 (2018).

nature research

Corresponding author(s): Chengcai Chu & Yang Bai

Last updated by author(s): Feb 22, 2019

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	No software was used.
Data analysis	Software packages for 16S amplicon data included FastQC 0.11.5, QIIME 1.9.1, QIIME 2, and USEARCH10. Databases for 16S amplicon data included RDP train set 16, SILVA 132, and FAPROTAX 1.1. Software packages for metagenomics data included SOAPnuke v1.5.2, SOAP v2.22, Megahit 1.0, MetaGeneMark v1 and blast 2.2.31. Databases for metagenomics data included Oryza sativa genome (gi: 996703420), and KEGG v81. We used R 3.5.1 for statistical analyses and plotting, including ggplot2 2.2.1 and randomForest 4.6-14. Scripts employed in the computation analyses are available under the following link: https://github.com/microbiota/Zhang2019NBT.

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

Raw sequence data reported in this paper have been deposited (PRJCA001214) in the Genome Sequence Archive in Beijing Institute of Genomics (BIG) Data Center, Chinese Academy of Sciences under accession numbers, CRA001372 for Bacterial 16S rRNA gene sequencing data and CRA001362 for metagenomic sequencing data that are public accessible at http://bigd.big.ac.cn/gsa.

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	Sample sizes were determined based on the Authors' experience of what is necessary to generate a convincing and compelling result.
Data exclusions	No data points were excluded from analysis in any experiment depicted in this manuscript.
Replication	The findings in this paper were remarkably reproducible. Every experiment was performed multiple times.
Randomization	Indica and japonica varieties were grown in the fields in a random arrangement (Figure 1B).
Blinding	For the indica and japonica root microbiome data, the investigator was blinded to genotypes and allowing them to choose each subject blindly.

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