

Article

Rice Plant–Soil Microbiome Interactions Driven by Root and Shoot Biomass

Cristina P. Fernández-Baca¹, Adam R. Rivers² , Jude E. Maul³ , Woojae Kim^{1,4}, Ravin Poudel², Anna M. McClung¹, Daniel P. Roberts³, Vangimalla R. Reddy⁵ and Jinyoung Y. Barnaby^{1,*} 

¹ Dale Bumpers National Rice Research Center, USDA Agricultural Research Service, Stuttgart, AR 72160, USA; cristina.fernandez@usda.gov (C.P.F.-B.); suwonman@korea.kr (W.K.); anna.mcclung@usda.gov (A.M.M.)

² Genomics and Bioinformatics Research Unit, USDA Agricultural Research Service, Gainesville, FL 32608, USA; adam.rivers@usda.gov (A.R.R.); ravin.poudel@usda.gov (R.P.)

³ Beltsville Agricultural Research Center, Sustainable Agricultural Systems Laboratory, USDA Agricultural Research Service, Beltsville, MD 20705, USA; jude.maul@usda.gov (J.E.M.); dan.roberts@usda.gov (D.P.R.)

⁴ Rural Development Administration, National Institute of Crop Science, Wanju 55365, Korea

⁵ Beltsville Agricultural Research Center, Adaptive Cropping Systems Laboratory, USDA Agricultural Research Service, Beltsville, MD 20705, USA; vr.reddy@usda.gov

* Correspondence: jinyoung.barnaby@usda.gov; Tel.: 1-301-504-8436

Abstract: Plant–soil microbe interactions are complex and affected by many factors including soil type, edaphic conditions, plant genotype and phenotype, and developmental stage. The rice rhizosphere microbial community composition of nine recombinant inbred lines (RILs) and their parents, Francis and Rondo, segregating for root and shoot biomass, was determined using metagenomic sequencing as a means to examine how biomass phenotype influences the rhizosphere community. Two plant developmental stages were studied, heading and physiological maturity, based on root and shoot biomass growth patterns across the selected genotypes. We used partial least squares (PLS) regression analysis to examine plant trait-driven microbial populations and identify microbial species, functions, and genes corresponding to root and shoot biomass as well as developmental stage patterns. Species identified correlated with increases in either root or shoot biomass were widely present in soil and included species involved in nitrogen cycling (*Anaeromyxobacter* spp.) and methane production (*Methanocella avoryzae*), as well as known endophytes (*Bradyrhizobium* spp.). Additionally, PLS analysis allowed us to explore the relationship of developmental stage with species, microbial functions, and genes. Many of the community functions and genes observed during the heading stage were representative of cell growth (e.g., carbohydrate and nitrogen metabolism), while functions correlated with physiological maturity were indicative of cell decay. These results are consistent with the hypothesis that microbial communities exist whose metabolic and gene functions correspond to plant biomass traits.

Keywords: rhizosphere microbiome; shotgun metagenomic sequencing; partial least squares (PLS) regression analysis; rice; recombinant inbred lines; root and shoot biomass; heading and physiological maturity stage



Citation: Fernández-Baca, C.P.; Rivers, A.R.; Maul, J.E.; Kim, W.; Poudel, R.; McClung, A.M.; Roberts, D.P.; Reddy, V.R.; Barnaby, J.Y. Rice Plant–Soil Microbiome Interactions Driven by Root and Shoot Biomass. *Diversity* **2021**, *13*, 125. <https://doi.org/10.3390/d13030125>

Academic Editors: Ipek Kurtboke, Milko A. Jorquera and Jacqueline Acuña

Received: 22 January 2021

Accepted: 10 March 2021

Published: 15 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Soil microbial communities can increase nutrient availability to plants and influence plant growth and overall health [1]. In turn, rhizosphere soil microorganisms rely on root exudates, such as carbon metabolites and other nutrients, as growth substrates. Plant species directly influence soil microbial communities through these root exudates which change as the plant matures [2,3]. Even within a plant species, the rhizosphere soil microbial community can be altered by plant genotypic differences [4,5]. Plant breeding efforts have the potential to make use of beneficial plant–soil microbiome interactions to increase the health and productivity of a crop [6]. However, more needs to be learned regarding

how plant developmental stages and their physiological traits influence soil microbial communities before this plant breeding potential can be realized. Rice is a staple crop for half the world's population; thus understanding rice-soil microbial community interactions is important [7–15].

Previous studies exploring plant–soil microbiome interactions report plant developmental stage, soil type, and genotype influence the rhizosphere soil microbial community [2,16,17]. It is established that plant developmental stage influences rhizosphere microbial community structure [4,5,8,18] in part, as root exudation patterns change with plant age [3]. Plant developmental impacts on the rhizosphere soil microbiome have been studied for rice [8,13,14,19] and other crops including potato [4,5,20] and maize [3]. Developmental stages from vegetative to maturity result in changes in numerous plant physiological traits. For example, roots and young leaves are major carbon and nutrient sinks during the vegetative stage whereas, after anthesis (i.e., heading) and during the subsequent grain fill stages, fruits and seeds become the dominant metabolic sinks as plants reallocate carbon to these reproductive tissues [21]. The transition that occurs between heading and physiological maturity is, therefore, a critical period to study in order to capture the interactions between root and shoot biomass growth and soil microbial community composition.

Likewise, plant genotypic variation has been shown to modulate the soil microbial community structure in rice [7–9,22] and other crops such as maize [23–25]. Many studies examined the plant genotype effect on soil microbe interactions using one or a few genetically diverse cultivars, whereas other studies used plants possessing a single gene mutation that modifies a trait of interest. A study by Zhang et al. (2019) examined the effect of mutating *NRT1.1B*, a nitrate transporter in rice, on the soil microbial community composition compared to that of the wild type. They found the single gene mutation in *nrt1.1b* decreased the relative abundance of nitrogen-cycling microbial populations compared to the wild-type [26].

Although it is well known that plants can shape the rhizosphere soil microbial community structure through qualitative traits such as root exudates, it remains largely elusive whether quantifiable traits, such as shoot biomass, play a similar role. However, a study by van der Heijden (1990) showed that both root and shoot biomass were positively correlated with mycorrhizal fungal community diversity in grassland species [27]. More recent studies have examined the connection between aboveground plant quantitative traits and belowground soil microbial communities [28,29].

We performed this study with the globally important rice crop, to understand how changes in growth patterns of plant traits, such as root and shoot biomass, during the developmental transition from heading to ripening stages impact plant trait-driven shifts in microbial community structure. This study used nine rice recombinant inbred lines (RILs), and their parents, differing for root and shoot biomass to examine the impact of these traits, measured at heading and physiological maturity, on the rice rhizosphere microbial community structure. RILs represent a genetic composition of the genomes from two parents generally having contrasting phenotypes. They are valuable for studying the impact of differences in plant quantitative traits on the soil microbial community because of their shared genetic background.

Many soil microbiome studies use exploratory methods, such as principal coordinate analyses (PCoA) among others, to characterize the observed shifts in the overall microbial communities between treatments [30]. However, these methods may obscure specific taxa or genes that significantly affect growth but are present as a small proportion of the overall microbial community. For example, it has been observed that relatively high abundance taxa are found across all rice growth stages but low abundance taxa can be characteristic of specific developmental stages [19] and this principle may apply to other plant traits. Partial least squares (PLS) analysis is a statistical tool that can discriminate between treatment groups by a trait of interest. Several soil microbiome studies have used PLS methods to differentiate and group microbial taxa by treatment [31–33]. Here we used nine rice

recombinant inbred lines, and their parents, segregating for root and shoot biomass at two important developmental stages, i.e., heading and ripening, to understand how changes in microbial taxa, functions, or genes are related to plant biomass traits and developmental changes using both PCoA and PLS regression analysis.

2. Materials and Methods

2.1. Rice Genotype Selection and Experimental Design

All studies were conducted in fields or greenhouses at the Dale Bumpers National Rice Research Center (DBNRRRC) in Stuttgart, AR in 2017. From a Francis (FRCS) and Rondo (ROND) (FR) mapping population of 217 recombinant inbred lines (RILs) grown to maturity in the field, 62 RILs were chosen based on similarity in developmental and physiological stages, i.e., heading date similarity (within 10 days) and plant height (within 30 cm) for a preliminary greenhouse study. The 62 selected RILs were grown for six weeks to evaluate their shoot and root biomass traits (Figure S1). Seeds were planted (roughly 5 seeds per pot) in small pots (7.6 cm square and 10.2 cm high). Pots used for transplanting were filled with soil collected from a field with a history of rice cultivation at the DBNRRRC in Stuttgart, AR. The soil is characterized as a Dewitt silt loam soil (fine, smectite, thermic, Typic Albaqualf), slightly acidic (pH of 5.6), with total C and N contents of 0.66% and 0.085%, respectively. Prior to soil collection for the greenhouse study, basal fertilizer of P (29 kg P ha⁻¹) and K (84 kg K ha⁻¹) was applied to the field and incorporated. Flooded conditions were maintained in the pots beginning at transplanting and continued for 6-weeks with a water depth of 5–7 cm above the soil surface.

Based on this preliminary study, nine RILs and their parents (Francis and Rondo) were selected which displayed a range in shoot and root biomass. Three replicates of the nine selected FR-RILs (except for RIL 9 which only had 2 replicates), and the two parents were grown in a greenhouse (plant date: 4 April 2017) using a completely randomized design, as described above, and after emergence, seedlings were thinned to one per pot. Four weeks after sowing, seedlings were transplanted into larger pots (27 l volume) with one seedling per pot and pots were placed at a uniform distance from each other. Additionally, two pots, containing soil but no plant, were used as replicated soil controls. Flooded conditions were maintained in the pots beginning at transplanting and continuing to maturity with a water depth of 5–7 cm above the soil surface. Each pot received nitrogen fertilizer in the form of urea (80 kg N ha⁻¹) applied in a three-way split during the two months following transplanting. Total photosynthetic active radiation (PAR) fluxes in the greenhouse during the study were 2398.3 mmol m⁻² and the average air temperature in the greenhouse was 24.9 ± 0.14 °C.

2.2. Plant and Soil Sampling and DNA Extraction

Plants were destructively sampled for biomass traits at the reproductive stage (heading, R4-R5), and ripening stage (maturity, R8-9) [34]. Plant samples and their associated rhizosphere soil samples were taken based on plant developmental stages as the selected FR-RILs matured at different rates. The entire plant was removed from the pot, and soil samples were taken from the center of the soil-root aggregate at a depth between 13–18 cm below the crown, where roots were most abundant. The soil-root aggregate was shaken to remove loose soil, leaving roughly a 1–2 cm soil layer on the roots. Approximately 5 g of soil covered roots were placed in a sterile flask with 50 mL of sterilized phosphate-buffered saline (PBS) solution and shaken gently to wash soil from the roots. Sterile forceps were used to remove roots from the soil slurry solution which was subsequently used for DNA extraction (see below). From the remaining soil-root aggregate, the soil was washed off. Plant tissue was divided into root biomass (RB) (below the crown) and shoot biomass (SB) (crown and above). Plant tissues were weighed after drying at 60 °C for 4–5 days to determine root and shoot dry weight.

Soil DNA was extracted from samples according to Kepler et al. (2018). The DNeasy PowerSoil HTP 96 Kit (Qiagen, Germantown, MD, USA) was used to extract gDNA from

800 μL of soil slurry according to manufacturer's instructions [35]. DNA was quantified using a Qubit dsDNA high sensitivity assay (Invitrogen, Waltham, MA, USA). All samples were diluted to $2.5 \text{ ng } \mu\text{L}^{-1}$ using deionized, autoclaved water for sequencing preparation. A subsample of soil was taken for total C and N analysis using a TruSpec CN analyzer (LECO Corp., Saint Joseph, MI, USA).

2.3. Shotgun Metagenomic Library Construction and Illumina Sequencing

Shotgun metagenomic sequencing was conducted in order to capture both the taxonomy and functions of the microbial community associated with SB, RB, and developmental stages heading and maturity. Diluted DNA samples were used for the shotgun metagenomic library construction using the Nextera XT DNA Library Preparation Kit (96 samples) and the Nextera XT Index Kit v2 Set A (96 indexes) (Illumina Inc., www.illumina.com, accessed on 14 March 2021) following the Illumina protocol. Briefly, DNA was fragmented and tagged with adapter sequences. Tagmented DNA was then amplified to add indexed adapters. After amplification, libraries were cleaned up using AMPure XP beads and quality checked on the Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Libraries were normalized to 0.5 nM using 10 nM TRIS-HCl and 0.1% Tween 20 (pH 8.5) and pooled. Rhizosphere soil samples were sequenced on the Illumina NextSeq 500 platform at the USDA Agricultural Research Service, Beltsville Area Research Center, Beltsville, MD, USA.

2.4. Sequence Processing

Reads were trimmed of adaptors and contaminant reads were removed using BBDuk version 37.66 (<https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbdduk-guide/>, accessed on 1 February 2018). BBDuk was then used to filter reads with kmers matching to the reference rice genome. DIAMOND (v0.9.17.118) [36] was used to perform searches of the reads against the NCBI NR database retrieved (February 2018) in protein space. Taxonomic and functional gene assignments to the NCBI taxonomy database (November 2018) and SEED DB, a functional gene database, (May 2015) [37,38], respectively, were made using the Megan 6 (version 6.15.2) command line tool "daa-meganizer" with the weighted lowest common ancestor (LCA) algorithm using a top percent setting of 3.0. Read counts from the analysis were exported from Megan Community Edition [39]. The data produced by sequencing is compositional count data [40]. For that reason, samples were closed, multiplicative replacement was run, and a centered log-ratio (CLR) transformation was performed using Skbio 0.5.5 (<https://scikit-bio.org>, accessed on 1 November 2018).

2.5. Community and Multivariate Statistical Analyses (PCoA and PLS)

Principal coordinate analysis (PCoA), using Bray-Curtis distance, was used to examine the soil microbial community structure as a whole using Megan Community Edition. Alpha diversity was measured by the Shannon index and calculations were performed using R (version 3.5.1) and Welch's *t*-tests were conducted to test significant differences at $p = 0.05$ between developmental stages by genotype [41]. Partial least squares (PLS) regression analysis was performed on the whole soil microbial community to differentiate species, metabolic pathways, and genes as a function of SB or RB dry weight. All PLS regression analysis steps were carried out in JMP v. 14.2.0 (2018 SAS Institute Inc., Cary, NC, USA).

As both PCoA and PLS regression analyses were performed on the entire microbial community, specific species, functions, or gene responses may have been masked by more abundant taxa, functions, or genes. For this reason, we performed a linear least squares fit analysis for each, individual dependent variable (species, function, or gene) as a function of SB or RB. Candidate lists were created for further analyses based on analysis of variance (ANOVA) significance tests ($p < 0.05$) of the linear least squares fit. These *p*-values were used only for creating candidate lists, and not used for multiple comparisons between species.

PLS analysis was performed on the selected candidate list of variables (species, function, or gene level), to differentiate their patterns with traits, i.e., SB, RB, and developmental

stages (heading and physiological maturity), and to further identify which variable(s) significantly contribute to trait differences. PLS regression analyses can be used to discriminate between treatment groups by a trait of interest and are particularly useful when the number of independent variables (e.g., number of microbial taxa or genes) are significantly greater than the number of response variables (e.g., traits of interest). For the PLS analysis, SB and RB were used as continuous variables and developmental stage (heading or physiological maturity) was used as pseudo-continuous variable by coding samples as either members (1) or non-members (0) for the specific developmental stage [42]. Regression beta coefficients were calculated and used to present the correlation patterns of traits with the selected variables. Developmental stages coefficient values are presented alongside SB and RB coefficients, to visualize their responses, however because they are compared to each other in the model (i.e., members vs. non-members), the coefficient values display equal and opposite patterns. Variable importance in projection (VIP) values were used to approximate trait variation. VIP scores greater than 0.8 are generally considered significant [43,44] and this cut-off was used in our study. Hierarchical cluster analyses (using the Ward method) based on beta coefficient data were used only for ease of visualization to identify groups of variables showing similar trends with traits of interest. These clusters were then used to create coefficient plots which grouped variables by their similarity in response to our traits of interest.

Following selection of metabolic functions related to SB and RB through PLS analyses, a regression analysis of selected dependent variables as a function of plant biomass traits (i.e., RB or SB) meeting a $p < 0.05$ were examined using CLR transformed count data. This was done to confirm PLS results and examine the correlation between microbial functions and SB and/or RB.

3. Results

3.1. Selection of Recombinant Inbred Lines Segregating for Root and Shoot Biomass

From a Francis and Rondo (FR) mapping population of 217 recombinant inbred lines (RILs), 62 were chosen based on similar developmental and physiological stages, i.e., heading date similarity (within 10 days, Figure S1A) and plant height (within 30 cm, Figure S1B). From this candidate list, nine Francis (FRCS) and Rondo (ROND) recombinant inbred lines (FR-RILs), along with their two parents, were selected for the soil microbiome study. The selected genotypes resulted in a range of low, intermediate, and high shoot biomass (SB) and root biomass (RB) phenotypes (range of 29.3–188.5 g dry weight and 4.6–46.3 g dry weight for shoot and root biomass, respectively, Figure S2) at heading and maturity. In general, root biomass decreased between heading and harvest maturity (harvest maturity was considered to be R7-9) while shoot biomass continued to increase at this stage (Figure S2).

3.2. Whole Microbial Community Structure and Impact of Shoot and Root Biomass Traits

Principal coordinate analysis (PCoA) plots demonstrated that microbial community composition varies by plant genotype and developmental stage. Plant genotype influences microbial community structure at both heading (Figure 1A) and at maturity (Figure 1B). Francis and Rondo have distinct microbial communities at heading and maturity differentiated in PC1, with the exception of one Rondo replicate that groups close to Francis at maturity. Taxa identified in both developmental stages include the genera *Sphingomonas*, *Bradyrhizobium*, and *Anaeromyxobacter* and the phyla *Acidobacteria* and *Chloroflexi*. Both shoot biomass and root biomass influenced the community structure during heading and at maturity. In contrast, soil percent total C and N did not influence the community structure strongly likely because it did not vary much between samples (mean total C and N $0.73\% \pm 0.02\%$ and $0.083\% \pm 0.002\%$, respectively). Alpha diversity, measured by the Shannon index, showed sample variance by genotype and developmental stage. The mean alpha diversity of each genotype tended to decrease from heading to maturity for RIL1, RIL5, RIL6, and Francis, increase for RIL2, RIL4, and RIL7 but did not change

for RIL3, RIL8, RIL9, Rondo, and the soil control samples (Figure S3). These differences by developmental stage were significant for Francis and RIL4 ($p = 0.128$ and $p = 0.0322$, respectively). Additionally, the top 1% of phyla (Figure S4) and top 2% of species (Figure S5) by genotype were plotted to examine their overlap with those species influencing the microbial community structure in the PCoA. The relative abundance plots revealed that all taxa influencing the community structure were highly abundant at both heading and maturity stages across all genotypes.

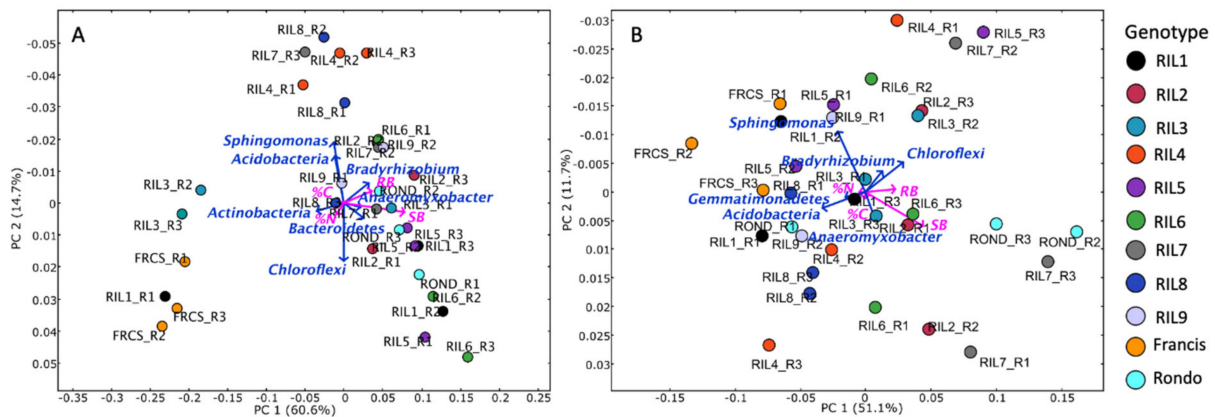


Figure 1. Principal coordinate analysis (PCoA) of whole soil microbial community for two parents Francis (FRCS) and Rondo (ROND) and 9 RILs at two developmental stages heading (A) and physiological maturity (B). Microbial populations influencing community structure are shown in blue. Shoot biomass (SB) and root biomass (RB) as well as soil percent total nitrogen (%N) and soil percent total carbon (%C) influences on community structure are shown in pink.

3.3. Shoot and Root Biomass Driven Species Level Analysis

We used PLS to identify taxa, functions and genes that are related to changes in SB and RB at heading and physiological maturity. To focus our analysis on those species, functions, or genes most affected by SB and RB, we used a linear least squares regression analysis to identify those variables from within the larger microbial population. The species, functions, or genes identified as correlated to either biomass trait and passing an ANOVA significance test of $p < 0.05$ were then used in the PLS regression analyses to further distinguish which displayed differential responses by SB and/or RB as well as their patterns by developmental stages. Regression beta coefficients and variable importance in projection (VIP) scores were plotted to visualize microbial species, function, or gene patterns with SB, RB, and/or the two developmental stages. Hierarchical clustering was used to visualize groups of species, functions, or genes based on their response patterns for SB, RB and the two developmental stages.

Approximately 100 candidate species were identified through ANOVA significance testing as a function of SB and RB (Table 1). PLS beta coefficient values were clustered for visualization, based on their SB, RB, and two developmental stage patterns, resulting in 7 species clusters (SC) ordered by increasing magnitude from SC1-SC7 (Table 1 and Figure 2). The greater the difference between coefficient values of a dependent variable, i.e., species, the greater the contribution of that variable to the model differentiating traits of interest. Additionally, variable importance in projection (VIP) values are presented and are used to approximate trait variation. VIP scores greater than 0.8 were considered significant [43,44].

Species in SC1 through SC4 had a higher abundance at heading when SB was lower as compared to maturity when SB was greater. In contrast, taxa in SC5 to SC7 showed the opposite pattern and were more abundant at maturity. The response due to changes in RB were the same direction as SB but were comparatively small (Table 1 and Figure 2). Species clusters SC4 and SC7 contain the species most significantly differentiated by biomass traits and developmental stages as demonstrated by the comparatively large

coefficient values and the VIP scores greater than 1. Species which increase as both shoot and root biomass increase, include several bacterial species from the *Planctomycetes*, *Proteobacteria* (*Alphaproteobacteria*), and *Verrucomicrobia* phyla, and two archaea from the phyla *Euryarchaeota* and *Thaumarchaeota* (SC7 in Figure 2). In SC4, species displaying a negative correlation with both SB and RB include mostly those in the *Actinobacteria*, *Chloroflexi*, and *Proteobacteria* (*Alpha-* and *Gamma-* *proteobacteria*) phyla.

Table 1. Species clusters (SC1-SC7) significantly correlated with shoot biomass (SB) or root biomass (RB) across two developmental stages. Species names, as well as numerical species ID used for coefficient plots, are shown. Beta coefficient values are reported for SB and RB. Variable importance in projection (VIP) scores are shown, VIP scores greater than 0.8 are considered significantly different by responses, i.e., biomass traits and developmental stages, in the model. All species in clusters 4 and 7 have VIP scores greater than 1.

Species Cluster	Species Name	ID	SB Coefficient	RB Coefficient	VIP
SC1	<i>Acidobacteria</i> bacterium 13_2_20CM_2_66_4	25	-4.5×10^{-3}	-6.8×10^{-5}	0.491
	<i>Acidobacteria</i> bacterium RIFCSPLOWO2_12_FULLL_66_10	34	-4.0×10^{-3}	-6.0×10^{-5}	0.433
	<i>Niastella korensis</i>	42	-3.6×10^{-3}	-5.4×10^{-5}	0.393
	<i>Bacterioidetes</i> bacterium 13_1_20CM_4_60_6	44	-3.9×10^{-3}	-5.8×10^{-5}	0.422
	<i>Gemmatimonadetes</i> bacterium	59	-4.7×10^{-3}	-7.0×10^{-5}	0.508
	<i>Gemmatimonadetes</i> bacterium SCN 70–22	64	-4.5×10^{-3}	-6.7×10^{-5}	0.484
	<i>Cyanobacteria</i> bacterium 13_1_20CM_4_61_6	269	-3.6×10^{-3}	-5.5×10^{-5}	0.395
	<i>Betaproteobacteria</i> bacterium RIFCSPLOWO2_12_FULLL_62_13	116	-4.0×10^{-3}	-6.0×10^{-5}	0.431
	<i>Deltaproteobacteria</i> bacterium RIFCSPLOWO2_12_FULLL_60_19	135	-4.4×10^{-3}	-6.5×10^{-5}	0.471
SC2	<i>Rhizobacter</i> sp. Root404	110	-5.7×10^{-3}	-8.5×10^{-5}	0.614
	<i>Hyalangium minutum</i>	124	-5.4×10^{-3}	-8.1×10^{-5}	0.583
	<i>Singulisphaera</i> sp. GP187	152	-7.6×10^{-3}	-1.1×10^{-4}	0.824
	<i>Gemmatirosa kalamazoonesis</i>	58	-6.1×10^{-3}	-9.1×10^{-5}	0.660
	<i>Gemmatimonadetes</i> bacterium 21-71-4	61	-6.0×10^{-3}	-9.0×10^{-5}	0.652
	<i>Alphaproteobacteria</i> bacterium 64-11	106	-7.0×10^{-3}	-1.0×10^{-4}	0.756
	<i>Dactylosporangium aurantiacum</i>	190	-7.7×10^{-3}	-1.2×10^{-4}	0.832
	<i>Gaiella</i> sp. SCGC AG-212-M14	196	-7.6×10^{-3}	-1.1×10^{-4}	0.819
	<i>Solirubrobacter soli</i>	199	-6.3×10^{-3}	-9.5×10^{-5}	0.686
	<i>Solirubrobacteriales</i> bacterium URHD0059	203	-6.8×10^{-3}	-1.0×10^{-4}	0.733
	<i>Actinobacteria</i> bacterium 13_1_20CM_4_68_12	208	-7.6×10^{-3}	-1.1×10^{-4}	0.826
	<i>Actinobacteria</i> bacterium 13_1_20CM_4_69_9	209	-7.8×10^{-3}	-1.2×10^{-4}	0.848
<i>Fimbriimonas ginsengisoli</i>	218	-7.2×10^{-3}	-1.1×10^{-4}	0.778	
SC3	<i>Betaproteobacteria</i> bacterium GR16-43	114	-8.7×10^{-3}	-1.3×10^{-4}	0.938
	<i>Rudaea cellulositytica</i>	141	-9.7×10^{-3}	-1.5×10^{-4}	1.048
	<i>Solirubrobacter</i> sp. URHD0082	200	-8.6×10^{-3}	-1.3×10^{-4}	0.927
	<i>Solirubrobacteriales</i> bacterium 67–14	201	-9.3×10^{-3}	-1.4×10^{-4}	1.012
	<i>Actinobacteria</i> bacterium 13_2_20CM_68_14	212	-9.1×10^{-3}	-1.4×10^{-4}	0.988
	<i>Actinobacteria</i> bacterium RBG_16_68_12	215	-8.5×10^{-3}	-1.3×10^{-4}	0.924
<i>Kouleothrix aurantiaca</i>	241	-8.2×10^{-3}	-1.2×10^{-4}	0.883	
SC4	<i>Phycoccus cremeus</i>	184	-1.1×10^{-2}	-1.6×10^{-4}	1.178
	<i>Phenylobacterium</i> sp. RIFCSPHIGHO2_01_FULLL_69_31	66	-1.1×10^{-2}	-1.6×10^{-4}	1.192
	<i>Gammaproteobacteria</i> bacterium RIFCSPHIGHO2_12_FULLL_63_22	138	-1.1×10^{-2}	-1.6×10^{-4}	1.159
	<i>Dokdonella immobilis</i>	139	-1.2×10^{-2}	-1.8×10^{-4}	1.319
	<i>Jatrophihabitans endophyticus</i>	181	-1.2×10^{-2}	-1.8×10^{-4}	1.316
	<i>Kineosporia</i> sp. A_224	182	-1.2×10^{-2}	-1.8×10^{-4}	1.321
	<i>Actinoplanes awajinensis</i>	186	-1.1×10^{-2}	-1.7×10^{-4}	1.242
	<i>Nocardioideus halotolerans</i>	192	-1.1×10^{-2}	-1.6×10^{-4}	1.144
	<i>Actinobacteria</i> bacterium IMCC26256	195	-1.1×10^{-2}	-1.7×10^{-4}	1.220
	<i>Thermoleophilum album</i>	204	-1.3×10^{-2}	-1.9×10^{-4}	1.364
	<i>Actinobacteria</i> bacterium 13_1_20CM_3_71_11	207	-1.2×10^{-2}	-1.8×10^{-4}	1.275
	<i>Actinobacteria</i> bacterium 13_2_20CM_2_66_6	210	-1.1×10^{-2}	-1.7×10^{-4}	1.233
	<i>Actinobacteria</i> bacterium RBG_16_67_10	214	-1.2×10^{-2}	-1.8×10^{-4}	1.309
	<i>Chloroflexi</i> bacterium 13_1_40CM_4_68_4	246	-1.1×10^{-2}	-1.6×10^{-4}	1.137
	<i>Chloroflexi</i> bacterium GWC2_73_18	250	-1.1×10^{-2}	-1.7×10^{-4}	1.231
	bacterium JGI 053	278	-1.2×10^{-2}	-1.8×10^{-4}	1.272

Table 1. Cont.

Species Cluster	Species Name	ID	SB Coefficient	RB Coefficient	VIP
SC5	<i>Bacteroidetes</i> bacterium RBG_13_42_15	50	5.4×10^{-3}	8.1×10^{-5}	0.583
	<i>Bacteroidetes</i> bacterium RBG_13_43_22	51	7.2×10^{-3}	1.1×10^{-4}	0.781
	<i>Rhodopseudomonas palustris</i>	73	7.2×10^{-3}	1.1×10^{-4}	0.778
	<i>Pseudorhodoplanes sinuspersici</i>	84	3.4×10^{-3}	5.1×10^{-5}	0.372
	<i>Planctomycetes</i> bacterium RBG_16_55_9	165	6.1×10^{-3}	9.2×10^{-5}	0.665
	<i>Anaerolinea thermophila</i>	223	6.3×10^{-3}	9.5×10^{-5}	0.687
	<i>Bellilinea caldifistulae</i>	224	4.8×10^{-3}	7.2×10^{-5}	0.521
	<i>Leptolinea tardivitalis</i>	225	5.5×10^{-3}	8.2×10^{-5}	0.596
	<i>Levilinea saccharolytica</i>	226	7.5×10^{-3}	1.1×10^{-4}	0.809
	<i>Longilinea arvoryzae</i>	227	5.7×10^{-3}	8.5×10^{-5}	0.614
	<i>Chloroflexi</i> bacterium HGW-Chloroflexi-10	252	5.4×10^{-3}	8.1×10^{-5}	0.583
	<i>Chloroflexi</i> bacterium RBG_16_54_18	258	6.5×10^{-3}	9.8×10^{-5}	0.706
	<i>Chloroflexi</i> bacterium RBG_16_57_11	259	7.2×10^{-3}	1.1×10^{-4}	0.779
	<i>Candidatus Nitrososphaera evergladensis</i>	283	5.1×10^{-3}	7.6×10^{-5}	0.550
	<i>Oxytricha trifallax</i>	285	7.7×10^{-3}	1.1×10^{-4}	0.828
<i>Acidobacteria</i> bacterium RBG_13_68_16	28	4.6×10^{-3}	6.9×10^{-5}	0.500	
SC6	<i>Pseudolabrys</i> sp. Root1462	86	1.0×10^{-2}	1.5×10^{-4}	1.097
	<i>Bradyrhizobium erythrophilei</i>	70	9.1×10^{-3}	1.4×10^{-4}	0.983
	<i>Rhodospirillales</i> bacterium 69-11	94	1.0×10^{-2}	1.5×10^{-4}	1.094
	<i>Rhodospirillales</i> bacterium URHD0088	96	8.8×10^{-3}	1.3×10^{-4}	0.956
	<i>Anaeromyxobacter dehalogenans</i>	118	9.3×10^{-3}	1.4×10^{-4}	1.008
	<i>Anaeromyxobacter</i> sp. RBG_16_69_14	120	8.4×10^{-3}	1.3×10^{-4}	0.913
	<i>Labilithrix luteola</i>	128	8.2×10^{-3}	1.2×10^{-4}	0.886
	<i>Myxococcales</i> bacterium 68-20	134	9.0×10^{-3}	1.3×10^{-4}	0.973
	<i>Planctomycetes</i> bacterium GWF2_41_51	158	8.7×10^{-3}	1.3×10^{-4}	0.945
	<i>Planctomycetes</i> bacterium RBG_13_50_24	161	9.7×10^{-3}	1.4×10^{-4}	1.047
	<i>Anaerolineae</i> bacterium CG2_30_64_16	231	9.1×10^{-3}	1.4×10^{-4}	0.983
	<i>Candidatus Nitrososphaera gargensis</i>	284	9.0×10^{-3}	1.3×10^{-4}	0.970
<i>Verrucomicrobia</i> bacterium 13_2_20CM_54_12	175	8.1×10^{-3}	1.2×10^{-4}	0.872	
SC7	<i>Alphaproteobacteria</i> bacterium 13_2_20CM_2_64_7	105	1.3×10^{-2}	1.9×10^{-4}	1.382
	<i>Phycisphaerae</i> bacterium SG8_4	143	1.3×10^{-2}	2.0×10^{-4}	1.452
	<i>Verrucomicrobia</i> bacterium 13_1_20CM_4_54_11	170	1.1×10^{-2}	1.7×10^{-4}	1.207
	<i>Verrucomicrobia</i> bacterium 13_2_20CM_55_10	176	1.4×10^{-2}	2.1×10^{-4}	1.528
	<i>Methanocella arvoryzae</i>	279	1.3×10^{-2}	2.0×10^{-4}	1.427
	<i>Candidatus Nitrosocosmicus oleophilus</i>	282	1.2×10^{-2}	1.8×10^{-4}	1.276
	<i>Acidobacteria</i> bacterium RBG_16_70_10	29	1.3×10^{-2}	1.9×10^{-4}	1.409
	<i>Bradyrhizobium elkanii</i>	69	1.2×10^{-2}	1.8×10^{-4}	1.306
	<i>Bradyrhizobium japonicum</i>	71	1.2×10^{-2}	1.8×10^{-4}	1.290
	<i>Rhodospirillales</i> bacterium 20-64-7	93	1.3×10^{-2}	1.9×10^{-4}	1.381
	<i>Anaeromyxobacter</i> sp. Fw109-5	119	1.1×10^{-2}	1.6×10^{-4}	1.182
	<i>Phycisphaerae</i> bacterium SM23_33	145	1.1×10^{-2}	1.7×10^{-4}	1.214
	<i>Planctomycetes</i> bacterium RBG_13_60_9	162	1.3×10^{-2}	1.9×10^{-4}	1.400
	<i>Planctomycetes</i> bacterium RBG_13_62_9	163	1.3×10^{-2}	1.9×10^{-4}	1.391
	<i>Planctomycetes</i> bacterium RBG_16_64_12	166	1.4×10^{-2}	2.1×10^{-4}	1.518
	<i>Verrucomicrobia</i> bacterium 13_1_20CM_3_54_17	169	1.1×10^{-2}	1.6×10^{-4}	1.164
	<i>Pedospaera paroula</i>	179	1.3×10^{-2}	2.0×10^{-4}	1.421

3.4. Rhizosphere Soil Microbial Community Function Analysis

At the functional level, three distinct clusters were associated with SB and RB at the two developmental stages (Figure 3A). To confirm observations made with PLS analyses, CLR transformed read counts for one to two representative microbial functions from each of the three clusters are shown that were significantly correlated to SB, RB, or both (Figure 3B–F). All other metabolic functions were plotted to confirm results and are presented in Supplementary Materials (Figure S6). All microbial functions from each of the three clusters exhibited the same pattern as the representative function shown in Figure 3B–F. Microbial functional cluster 1 (FC1) was divided into those functions (depicted by hashed bars in Figure 3A) displaying significant negative correlations with shoot biomass and significant positive correlations with root biomass ($p < 0.05$) (Figure 3B), and those functions with a negative correlation to shoot biomass ($p < 0.05$) and no significant

relationship to root biomass ($p > 0.05$) (Figure 3C). CLR transformed functional count data was plotted as a function of SB and RB to verify the correlations found by PLS analyses for fatty acid synthesis (Figure 3B), metabolism of aromatic compounds, amino acids and derivatives, metabolite damage and repair, cofactor, vitamins, prosthetics groups, and pigment (Figure S6A). The remaining four functions identified in FC1, had VIP scores greater than 0.8 and showed significant negative correlations with shoot biomass, but weak or no correlation to root biomass including carbohydrate metabolism (Figure 3C and Figure S6B). FC2 contains metabolic functions considered of lesser importance to the model (VIP < 0.8 for all) compared to all functions in FC1 and three functions in FC3. Among these FC2 functions are stress response and cell division and cell cycle functions (Figure 3D). In FC3, dormancy and sporulation are significantly positively correlated with shoot biomass but negatively correlated with root biomass (Figure 3E). All other functions in FC3 display the same significant positive correlation with SB. However, this was not the case for RB as the negative correlation pattern predicted from the PLS model was not present when CLR-transformed count data was plotted as a function of RB, indicating no correlation with RB. These functions include virulence, disease, and defense (Figure 3F) and motility and chemotaxis (Figure S6C).

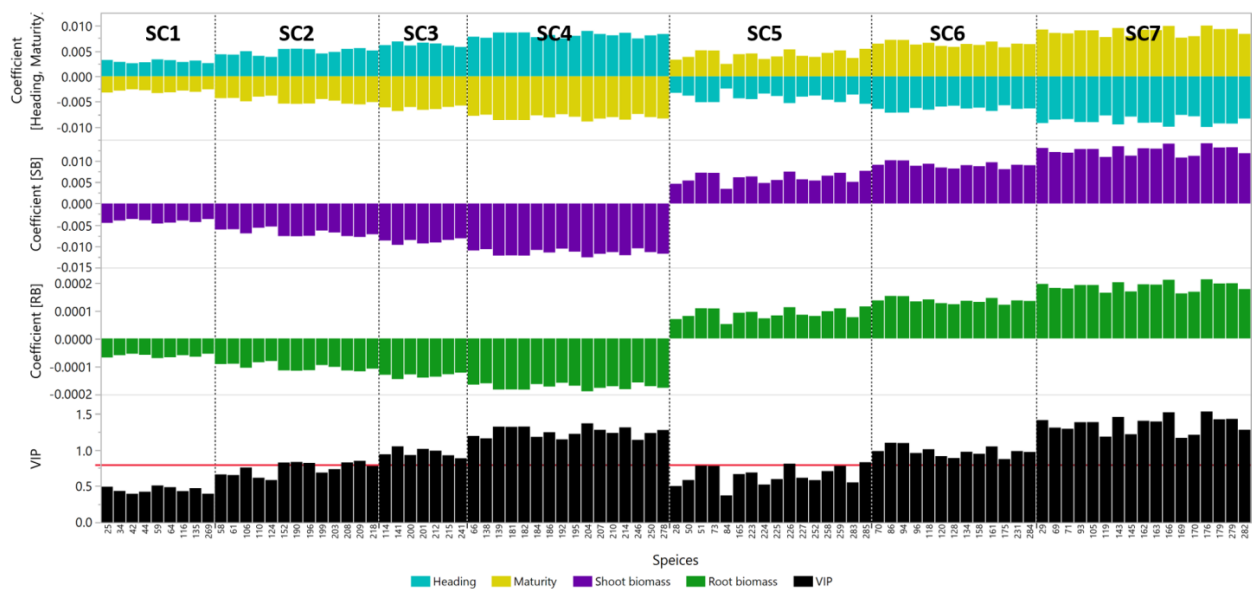


Figure 2. Partial least squares (PLS) regression coefficients of species differentiated by shoot biomass (SB), root biomass (RB), as well as the developmental stages, heading vs. maturity. The larger the difference between coefficient values, the greater that variable contributes to the model. Developmental stage was visualized as either a member or non-member, and thus heading and maturity show equal and opposite trends. Variable importance in projection (VIP) scores are shown, VIP scores greater than 0.8 (red line) are considered significantly different by variable responses, i.e., biomass traits and developmental stages, in the model. Species clusters (SC1–SC7) show species based on their coefficient variance patterns of responses, i.e., SB, RB, and two developmental stages, in the PLS model.

Interestingly, microbial community functions correlated with root biomass were also related to heading stage, while shoot biomass and maturity stage displayed similar functional trends. However, functions correlated with RB and heading showed opposite trends to those correlated with SB and maturity. This opposing trend of RB and SB associated genes was observed for all functions in the PLS-selected candidate list.

3.5. Metagenomic Gene Level Analysis With PLS

The PLS analysis used in this study was particularly powerful when examining the more than 7000 microbial gene annotations, to discover genes differentially correlated with plant biomass traits and developmental stage. From the thousands of annotated genes,

PLS analysis revealed 249 genes that were significantly associated with both SB and RB (Figure 4 and Table S1). Genes displaying particularly contrasting patterns in SB and RB, and VIP scores greater than 0.8, are summarized in Table 2. Genes grouped into 11 clusters that displayed similar gene coefficient patterns with RB or SB (Figure 4).

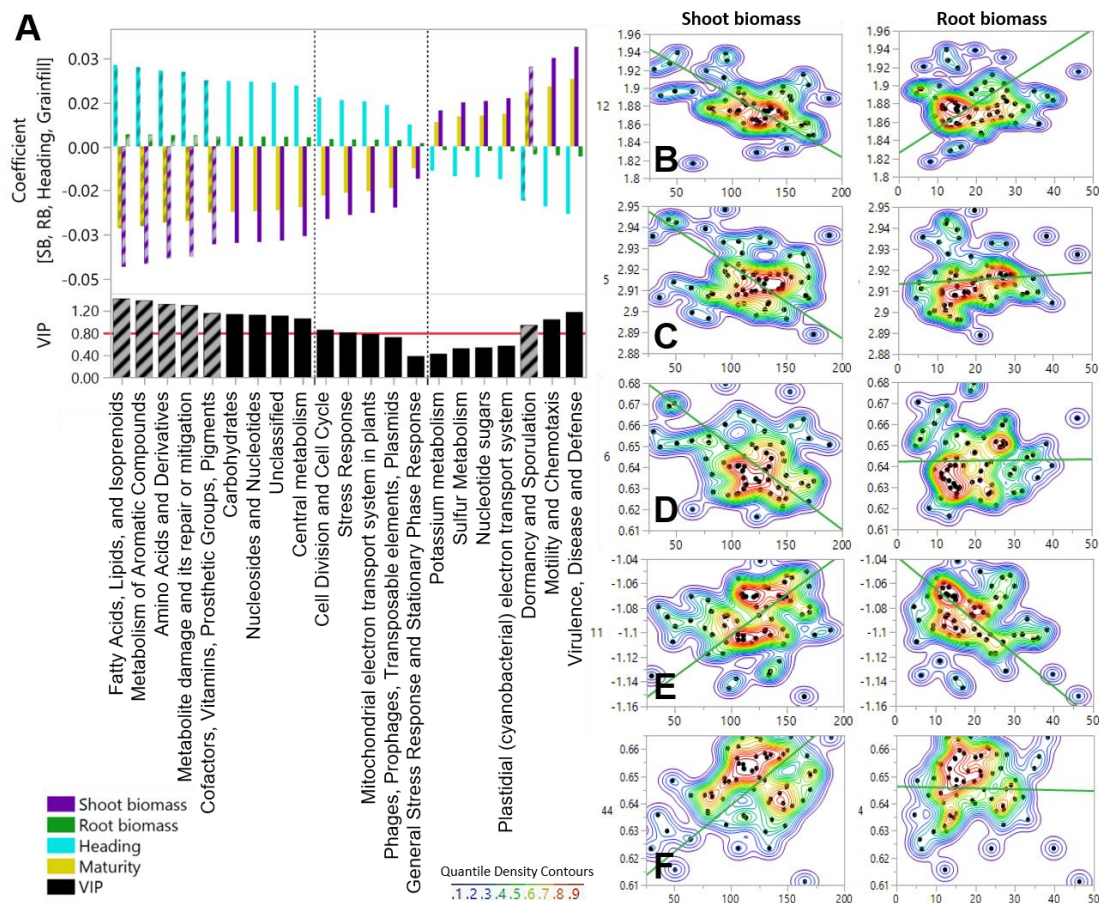


Figure 3. Microbial community functions identified with PLS analysis for shoot biomass (SB) and root biomass (RB) across two developmental stages (heading and maturity). (A) Coefficient values from PLS depicting directionality of trait relationships with functions and variable importance in projection (VIP) scores indicate importance in the model (VIP > 0.8 are considered significantly different by responses, i.e., biomass traits and developmental stages, in the model). Hashed bars in functional cluster 1 (FC1) indicate functions that are negatively correlated to SB and positively correlated to RB, while non-hashed bars in FC1 are negatively correlated to SB but show no correlation with RB. SB and RB patterns with center-log ratio transformed count data are shown for selected functions: fatty acid synthesis (B), carbohydrate metabolism (C), cell division and cell cycle (D), dormancy and sporulation (E), and virulence, defense, and disease (F). Contour lines indicate quantile density at 5% intervals. This means that approximately 5% of points generated from the estimated non-parametric distribution are below the lowest contour, 10% are below the next contour, etc. The highest contour has about 95% of the points below it.

Genes significantly positively correlated with RB and heading (and negatively correlated with SB and maturity) are largely found in gene clusters 1 through 5 (GC1–GC5). The most frequent genes identified in GC1 and GC2 were involved in amino acid metabolism, carbohydrate metabolism of mannose and maltose, fatty acid synthesis, and several stress response genes. Additionally, genes involved in iron acquisition by siderophore biosynthesis, polysaccharide synthesis, nitrogen metabolism genes, cell signaling genes, and metabolism of aromatic compounds, were all found to be negatively correlated with SB/maturity and positively correlated with RB/heading.

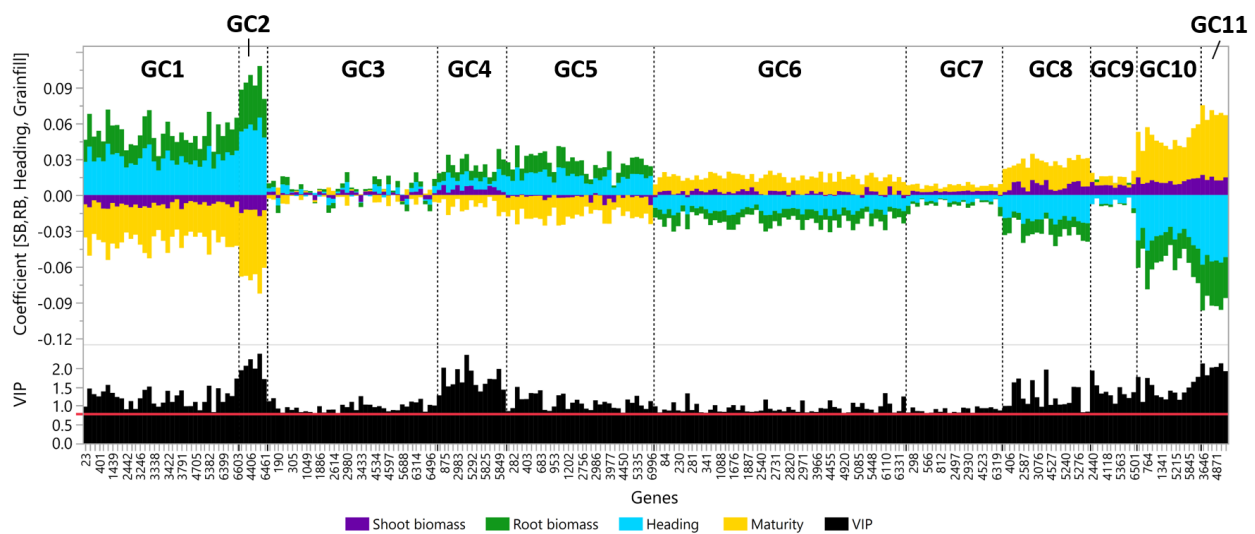


Figure 4. Partial least squares (PLS) regression coefficients for genes associated with shoot biomass (SB), root biomass (RB) as well as heading and maturity stages, presented in a stacked bar plot. Variable importance in projection (VIP) scores are shown, VIP scores greater than 0.8 (red line) are considered significantly different by responses, i.e., biomass traits and developmental stages, in the model. Gene clusters (GC1–GC11) are indicated on the plot and reveal genes with differential coefficient patterns based on the SB and RB as well as developmental stage in the PLS model.

Table 2. Genes identified as significantly correlated with shoot biomass (SB) and root biomass (RB) and showing contrasting trends with biomass traits (gene clusters 1, 2, 8, 10, 11). Gene names as well as numerical IDs (gene ID) used for coefficient plots are shown. Beta coefficient values are reported for SB and RB. Variable importance in projection (VIP) scores greater than 0.8 are considered significantly different by responses, i.e., biomass traits and developmental stages, in the model.

Gene Cluster	Gene	Gene ID	SB Coefficient	RB Coefficient	VIP
GC1	2-methylcitrate dehydratase FeS dependent (EC 4.2.1.79)	23	−0.008	0.018	0.98
	Arginine ABC transporter, permease protein ArtM	36	−0.010	0.027	1.46
	Ornithine aminotransferase (EC 2.6.1.13)	63	−0.004	0.021	1.31
	3,5-diaminohexanoate dehydrogenase (EC 1.4.1.11)	401	−0.006	0.022	1.25
	L-threonine transporter, anaerobically inducible	555	−0.011	0.017	1.38
	2-ketogluconate kinase (EC 2.7.1.13)	626	−0.012	0.029	1.55
	Maltose operon transcriptional repressor MalR, LacI family	1439	−0.007	0.024	1.34
	Putative regulator of the mannose operon, ManO	1509	−0.010	0.023	1.23
	poly(beta-D-mannuronate) lyase (EC 4.2.2.3)	1899	−0.007	0.023	1.19
	Substrate-specific component YkoE of thiamin-regulated ECF transporter for HydroxyMethylPyrimidine	2442	−0.008	0.014	0.91
	DNA polymerase-like protein MT3142	2904	−0.010	0.016	1.12
	Polyketide beta-ketoacyl synthase WhiE-KS paralog	3131	−0.008	0.016	0.91
	Phytoene desaturase, neurosporene or lycopene producing (EC 1.3.--)	3246	−0.011	0.019	1.20
	Fatty acyl-coenzyme A elongase	3263	−0.009	0.027	1.42
	Acyl carrier protein (ACP1)	3331	−0.011	0.029	1.51
	FIG027190: Putative transmembrane protein	3338	−0.007	0.020	1.06
	Triacylglycerol lipase precursor (EC 3.1.1.3)	3372	−0.009	0.012	0.97
	UPF0225 protein YchJ	3385	−0.008	0.021	1.08
	Haemin uptake system permease protein	3422	−0.011	0.024	1.40
	Probable Lysine n(6)-hydroxylase associated with siderophore S biosynthesis (EC 1.14.13.59)	3502	−0.006	0.020	1.10
	Dipeptide transport system permease protein DppC (TC 3.A.1.5.2)	3605	−0.004	0.019	1.27
	Transcriptional regulator of fimbriae expression FimZ (LuxR/UhpA family)	3791	−0.012	0.020	1.41
	Phenylacetaldehyde dehydrogenase (EC 1.2.1.39)	3929	−0.007	0.019	0.99
	Vanillate O-demethylase oxygenase subunit (EC 1.14.13.82)	4059	−0.005	0.018	0.99
	Protein gp47, recombination-related [Bacteriophage A118]	4705	−0.006	0.020	1.10
	Uncharacterized transporter, similarity to citrate transporter	4969	−0.008	0.015	0.88
	SSU ribosomal protein S10p (S20e), chloroplast	5339	−0.006	0.020	1.10
	SSU ribosomal protein S13p (S18e), mitochondrial	5382	−0.009	0.029	1.53
	Putative succinate dehydrogenase cytochrome b subunit	6059	−0.007	0.014	0.83
	Sigma factor RpoE negative regulatory protein RseB precursor	6373	−0.006	0.025	1.46
	tRNA methylase YGL050w homolog Wyeosine biosynthesis	6399	−0.006	0.020	1.10
	Diaminobutyrate-pyruvate aminotransferase (EC 2.6.1.46)	6476	−0.008	0.025	1.33
	Glutaredoxin 1	6486	−0.010	0.022	1.26
	RsbS, negative regulator of sigma-B	6603	−0.005	0.028	1.72

Table 2. Cont.

Gene Cluster	Gene	Gene ID	SB Coefficient	RB Coefficient	VIP
GC2	(GlcNAc) ₂ ABC transporter, permease component 2	621	−0.015	0.035	1.94
	Alpha-N-acetylglucosaminidase (EC 3.2.1.50)	811	−0.012	0.039	2.06
	Cyanate ABC transporter, ATP-binding protein	4406	−0.012	0.041	2.23
	Phage capsid and scaffold	4639	−0.012	0.038	1.99
	Autoinducer 2 (AI-2) ABC transport system, membrane channel protein LsrC	5486	−0.018	0.043	2.38
	High-affinity choline uptake protein BetT	6461	−0.013	0.032	1.71
GC8	Arginine pathway regulatory protein ArgR, repressor of arg regulon	44	0.003	−0.014	0.99
	Acetyl-CoA acetyltransferase (EC 2.3.1.9)	406	0.004	−0.013	1.00
	Methionyl-tRNA formyltransferase (EC 2.1.2.9)	2587	0.006	−0.016	1.19
	O-succinylbenzoate synthase (EC 4.2.1.113)	2616	0.003	−0.014	1.05
	Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)	4527	0.004	−0.014	1.02
	Pyridine nucleotide-disulphide oxidoreductase associated with reductive pyrimidine catabolism	4616	0.004	−0.018	1.19
	Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7)	4993	0.004	−0.014	1.03
	LSU ribosomal protein L5p (L11e)	5240	0.005	−0.011	1.04
	UbiD family decarboxylase, MJ1133 type	412	0.011	−0.005	1.62
	Lacto-N-biose phosphorylase (EC 2.4.1.211)	1381	0.011	−0.006	1.72
	Putative DNA-binding protein in cluster with Type I restriction-modification system	3047	0.013	−0.011	1.68
	Dipicolinate synthase subunit B	3076	0.010	−0.011	1.22
	Spore germination protein GerKB	3109	0.008	−0.012	0.97
	Stage IV sporulation protein A	3177	0.013	−0.005	1.96
	LSU ribosomal protein L18e	5267	0.009	−0.012	1.07
	LSU ribosomal protein L23Ae (L23p)	5271	0.011	−0.009	1.50
	LSU ribosomal protein L30e	5276	0.012	−0.011	1.49
SSU ribosomal protein S27e	5362	0.007	−0.015	0.83	
DNA-directed RNA polymerase II second largest subunit (EC 2.7.7.6)	6261	0.007	−0.015	0.84	
GC10	Meso-diaminopimelate D-dehydrogenase (EC 1.4.1.16)	374	0.015	−0.023	1.76
	S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1A	502	0.009	−0.017	1.10
	Predicted cellobiose ABC transport system, ATP-binding protein 1	764	0.010	−0.032	1.74
	Multiple sugar ABC transporter, substrate-binding protein	1123	0.012	−0.024	1.55
	Predicted regulator of fructose utilization, DeoR family	1138	0.011	−0.022	1.27
	Predicted L-rhamnose permease RhaY	1341	0.010	−0.022	1.23
	Formylmethanofuran dehydrogenase (tungsten) operon gene G	1526	0.010	−0.020	1.17
	Potassium uptake protein, integral membrane component, KtrB	4938	0.012	−0.015	1.40
	Similar to ribosomal large subunit pseudouridine synthase D, CAC1266-type	5215	0.009	−0.026	1.35
	SSU ribosomal protein S4p (S9e), mitochondrial	5397	0.009	−0.021	1.16
	Signal peptidase, type IV – prepilin/preflagellin	5413	0.012	−0.017	1.35
	Coenzyme F ₄₂₀ H ₂ dehydrogenase (methanophenazine) subunit FpoM	5845	0.013	−0.022	1.48
	Sulphydrogenase II subunit g	5959	0.014	−0.027	1.64
	Conjugative transfer protein TrbG	6889	0.014	−0.029	1.77
	GC11	IcmB (DotO) protein	3646	0.017	−0.038
Possible alpha/beta hydrolase superfamily, slr1917 homolog		4225	0.013	−0.034	1.81
photosystem I subunit XI (PsaL)		4833	0.015	−0.037	2.01
Phycobilisome rod-core linker polypeptide, phycocyanin-associated		4871	0.012	−0.038	2.03
Conjugative signal peptidase TrhF		6881	0.012	−0.040	2.13
Inclusion membrane protein-52		6958	0.015	−0.034	1.92

GC6 to GC11 represent genes positively correlated with SB and maturity and negatively correlated with RB and heading. Clusters GC8, GC10 and GC11 include virulence related genes involved in conjugative signaling and transfer and genes involved in potassium uptake, photosynthesis, and nucleotide synthesis. Additionally, several genes related to spore formulation are strongly related to increasing SB and decreasing RB. Many genes related to protein metabolism, specifically large subunit ribosomal proteins, and amino acid synthesis and degradation functions (including Acetyl-CoA acetyltransferase involved in lysine degradation) were also identified. Interestingly, the gene level patterns observed in GC8, GC10, and GC11 mirror the patterns observed in the functional level cluster 3 (FC3). PLS analyses were likewise carried out for genes specifically associated with SB or RB, individually and can be found in Supplementary Materials (Figures S7 and S8, respectively).

4. Discussion

A number of studies have shown that microbial community structure changes with plant genotype and developmental stage and this trend was observed in our study in which microbial communities differed across the FR-RILs and shifted between heading and maturity stages within genotypes [3,8,17]. Although we did observe genotypic differences, they were not based on the observed SB and RB trend that we selected for (i.e., two

parents, Francis and Rondo, displaying contrasting SB and RB traits and their nine FR–RILs, encompassing a range of SB and RB) when the soil microbial community was examined as a whole.

4.1. Soil Microbial Populations Associated With Rice Shoot and Root Biomass

To identify specific taxa, microbial functions, and genes related to biomass traits of interest, we used two multivariate analysis tools: PCoA and PLS regression analysis. The PCoA was conducted first to examine the whole microbial community shifts in the heading and physiological maturity stages. RB and SB explained some of the rhizosphere community structure in both stages, with SB being dominant over RB in both stages. Several taxa explained the microbial community structure in both stages and were also found to be highly abundant (Figures S5 and S6) including the genera *Bradyrhizobium* and *Anaeromyxobacter*. However, PCoA did not allow us to interrogate which specific taxa, functions, and genes were correlated with SB and RB, to do this we used PLS regression analysis.

PLS regression analyses were performed on the entire microbial community to identify microbial species correlated with SB and RB. However, when linear regressions of species CLR–transformed count data were plotted against SB or RB to confirm the model, many of the coefficient patterns identified through the PLS model showed a very weak, non-significant correlation ($p > 0.05$). We speculated that the presence of abundant microbial taxa that do not necessarily respond to SB and RB may have obscured changes in community structure of SB or RB associated species. For this reason, we used linear least squares regression analyses to identify species, functions, and genes that are associated with SB or RB for use in the PLS analysis. In this discussion we consider only species showing the most differentiation in PLS results, i.e., species in SC4, SC6, and SC7, with VIP scores greater than 0.8 as they are considered significant effectors in the PLS models [43,44]. Microbial taxa identified through the PLS analysis are all widely present in soil, including species from the *Actinobacteria*, *Chloroflexi*, *Planctomycetes*, *Proteobacteria*, and *Verrucomicrobia* phyla. Within the *Proteobacteria*, several *Alphaproteobacteria* species were identified as correlated with shoot and root biomass. Within this phylum, a *Phenyllobacterium* sp. (ID = 66, VIP = 1.192), aerobic or facultatively anaerobic non-spore forming bacteria [45,46], was negatively correlated with RB, SB, and maturity. Perhaps unsurprisingly, several *Bradyrhizobium* spp., which are known rice root endophytes capable of nitrogen–fixation [47], were identified as positively associated with root biomass, but, interestingly, also with shoot biomass and physiological maturity (Table 1 IDs 69, 70, 71 and VIP scores of 1.306, 1.209, 0.983, respectively). A previous study by Hirano et al. (2001), found that rice grown without nitrogen fertilizer, in soils with a high abundance of nitrogen–fixing bacteria in late growth stages, had similar yields to rice grown under conventional fertilization practices [48,49]. This underscores the importance of rice endophytes, including *Bradyrhizobium* spp., in rice cultivation. The fact that the abundance of these bacteria was found to correlate with both root and shoot biomass patterns with 11 genotypes suggests these traits could be used for breeding to select for nitrogen–fixing microbial populations.

Several species showed a significant positive correlation with rice shoot biomass and this pattern was well–presented at maturity. Especially interesting, *Anaeromyxobacter* spp. (Table 1 IDs 118, 119, 120 and VIP scores of 1.008, 1.182, 0.913, respectively) are Gram negative facultative anaerobes that are known to perform dissimilatory iron reduction [50] and have been shown to use nitrate [50] as an electron acceptor making it an important species for biogeochemical cycling of nutrients. A methanogen, *Methanocella avoryzae* (ID 279, VIP = 1.427), was also positively correlated with increased shoot biomass and physiological maturity. Peak methane emissions are typically observed during heading stage [51], however, higher shoot biomass likely creates a larger conduit for methane to diffuse from the soil to the atmosphere through aerenchyma tissue [9,52]. Additionally, *Candidatus Nitrosocosmicus oleophilus* (ID 282, VIP = 1.276), an ammonia oxidizing archaea (AOA) was found to correlate positively with biomass traits and physiological maturity.

AOA are autotrophs that fix inorganic carbon and are thought to be the dominant microbial population performing ammonia oxidation in soil [53,54]. AOA and ammonia oxidizing bacteria (AOB) are involved in nitrate formation on rice root surfaces which have been shown to be an important nitrogen source for rice grown in submerged conditions [55,56]. They are key players in the nitrification pathway, which oxidize ammonia to nitrite, a nitrogen form less amenable for assimilation by rice. Interestingly, this archaea is negatively correlated during the heading stage, when rice is transitioning from its reproductive phase to the ripening phase and requires more nitrogen in the easily assimilated form of ammonia [22].

There were a number of overlapping taxa between the predominant taxa present in the soil samples (Figures S4 and S5) and the taxa identified by PCoA (Figure 1) and PLS (Figure 2) analysis. *Methanocella avoryzae* and several *Anaeromyxobacter* species were among those identified as having relatively high abundance (Figure S5) and positively correlated with shoot and root biomass (Figure 2 and Table 1). However, other species, including *Bradyrhizobium* spp. and *Candidatus Nitrosocosmicus oleophilus*, were not in the top 2% of taxa but were significantly positively correlated with higher biomass. Species within *Bradyrhizobium* and *Anaeromyxobacter* influenced the community structure at both heading and maturity (Figure 1) with the latter species having relatively high abundance (in the top 2%) across genotypes and developmental stages (Figure S5). Other high abundance taxa including *Sphingomonas* spp., and unclassified taxa within the *Actinobacteria*, *Chloroflexi*, *Planctomycetes*, *Proteobacteria*, and *Verrucomicrobia* phyla, likewise influenced the PCoA structure (Figure 1 and Figure S4). Thus, while the PCoA did identify some of the same species as PLS, these results underscore the importance of the PLS analysis which specifically identified those species positively correlated with biomass traits regardless of abundance.

PLS analysis was able to differentiate microbial species based on root biomass; however, coefficient values for root biomass were one to two orders of magnitude smaller than those of shoot biomass, indicating RB had less impact in the species level model than SB. The nine selected RILs and their parents exhibited a lower and wider range in root biomass (roughly 10-fold difference) as compared to shoot biomass (roughly 6.5-fold difference). This likely allowed for stronger trends to arise from PLS analysis with SB data, which surpasses any relationships seen with RB. Additionally, processing soil samples for root quantification is difficult and time-consuming which may lead to underestimation of plant root biomass [57]. Despite the potential for underestimation in RB quantification, PLS analyses were still able to identify functions related to RB. This demonstrates that PLS analyses, used in conjunction with ANOVA significance selection from the microbial community by traits of interest, have the potential to be used to identify trait-associated microbial populations and to understand the relationship between traits of interest.

Overall, through PLS analysis, we were able to identify species that were abundant with lines that had a lower range of root and shoot biomass at heading, and other species that were more abundant with lines that had a higher range of root and shoot biomass at maturity. Our results also indicate that microbial populations are influenced by RB and SB in the same manner. This suggests that shoot biomass has potential to be used as a proxy for root biomass in breeding to select for beneficial soil microbial species.

4.2. Microbial Community Functions Correlated to Biomass Traits

Microbial community functions identified in our analysis were distinct between the two developmental stages, heading and physiological maturity. Heading stage represents a dynamic phase of plant growth as rice transitions from reproductive to ripening stages. This growth phase was reflected in the microbial functions related to cell growth (synthesis of amino acids and fatty acids) and metabolism (carbohydrate and aromatic compound metabolism) in our study. In contrast, many of the microbial functions indicative of cell growth plateauing or decaying, such as sporulation, virulence defense, and motility and chemotaxis, were represented at maturity. Chemotaxis functions may indicate bacterial

populations are searching for nutrients as root exudates decrease with plant maturation and they are required to search further out to find growth substrates [3,58].

Shoot and root biomass related functions followed the same trends as developmental stages maturity and heading, respectively. Root biomass related functions showed an opposing pattern to SB related functions. Although differences in coefficient values for RB were smaller than those of SB at the species level, they were greater at the functional level, indicating the PLS model was able to discern microbial functions differentially correlated with RB. This suggests that RB may more strongly influence overall community function than it does the abundance of specific taxa.

4.3. Gene Trends Related to Shoot and Root Biomass

Many of the genes correlated with RB or SB belonged to the pathways identified at the functional level. However, in contrast to the functional level analysis, RB showed a stronger pattern with gene abundance than SB. This suggests that PLS analyses can make use of relatively small variation in observed traits to differentiate changes in abundance among thousands of genes and identify those related to traits of interest. Even though only a few RB-associated species and functions were identified with PLS as compared to SB, it appears RB may be a more dominant driver of soil microbial community gene abundance than SB. This implies there may be functional redundancy across species; thus, although individual species or functions do not correlate with the RB phenotype, specific genes do.

As expected, the genes most frequently identified were involved in amino acid, carbohydrate, and protein metabolism, all processes central to microbial life. Genes identified as specifically positively associated with RB and heading include those involved in amino acid degradation including an aminotransferase. Additionally, two genes for mannose and maltose metabolism as well as a 2-ketogluconate kinase, part of the pentose phosphate metabolic pathway, were identified as positively associated with RB. A gene related to nitrogen metabolism by nitrite transport showed a similar positive correlation with RB and heading which indicates bacterial nitrogen cycling is prominent during the heading stage. Fatty acid synthesis genes and several stress response genes were likewise positively correlated with RB. Genes involved in iron acquisition by siderophore biosynthesis, polysaccharide synthesis, cell signaling, and metabolism of the phenolic compound vanillic acid, found in plant root exudates [59], are all representative of cell growth, and were found to be positively correlated with RB/heading and negatively correlated with SB/maturity.

Genes that were identified as being positively correlated with SB and maturity and strongly negatively correlated with RB and heading belong to many of the cell decaying functions observed at the functional level. Virulence related genes, for example, were identified as uniquely related to RB and maturity, particularly genes predicted to be involved in conjugative signaling and transfer, indicate cells are undergoing attack. Several genes involved in spore formulation are strongly related to increasing SB and decreasing RB, suggesting microbial populations are preparing for dormancy as plants mature from heading to maturity and RB begins to decrease. Two photosynthesis-related genes were identified as positively correlated with SB and maturity and negatively correlated with RB and heading. As roots decay and root exudates decrease with age, a decrease in available carbon metabolites that bacteria rely on may cause the soil microbial community to favor autotrophs capable of photosynthesis resulting in a negative correlation between photosynthesis-related genes and root biomass.

4.4. Relationship of Developmental Stage to Microbial Community Structure and Functions

In general, beta coefficient patterns during heading stage followed those of SB while coefficient patterns for maturity were more closely associated with RB patterns in our study. During heading stage, plants experience rapid growth whereas during maturation stages rice is remobilizing resources to grain development, and as a consequence, belowground carbon allocation decreases [60] as do overall root exudates [61]. As plants age, root exudates change and consequently modulate the microbial community structure and

function. In our study, FR–RIL SB increased from heading through the grain fill stages while RB decreased between these stages. Thus, up until maturity, SB is in a growth phase, influencing, and likewise being influenced by, the microbial community. At maturity, FR–RIL root biomass decreased relative to heading. Since RB is in direct contact with the soil microbial community, it follows that during RB growth (i.e., heading stage), when root exudates are high, it is driving the microbial community structure and function.

It is difficult to disentangle the effect of biomass traits from developmental stage, thus focusing on two developmental stages, heading and physiological maturity, that show contrasting biomass traits was a key component of this study. However, further studies validating the results herein are needed. For example, inoculation studies using PLS–identified species or species harboring PLS–identified genes could be used to verify if they positively impact RB and SB growth during the critical heading and maturing stages. A study by Mayer et al. (2019), did just this, using identified endophyte isolates to inoculate several plant species and evaluating their impact on root and shoot growth [62]. Future work in this vein could prove useful for validation, following identification of specific taxa or genes related to RB and SB growth.

5. Conclusions

Overall, these results suggest species, functions, and gene abundance patterns vary by shoot and root biomass, and across developmental stages of heading and maturity. This study indicates changes in root and shoot biomass may play an equally important role in driving the microbial community as does developmental stage. However, our results suggest that SB may be used as a surrogate for root biomass in future plant trait-based breeding, as we found that rhizosphere microbial species vary with SB and RB in the same manner. Using SB as a breeding trait would be advantageous over RB, as RB quantification is time–consuming and prone to error, while SB quantification is more efficient and accurate. This study underscores the potential of exploiting rice phenotypic variation in plant breeding to promote beneficial plant–soil microbiome interactions.

Supplementary Materials: The following are available online at <https://www.mdpi.com/1424-2818/13/3/125/s1>, Figure S1: Rice genotype selection based on shoot and root biomass, Figure S2: Biomass trends of nine rice genotypes, Figure S3: Shannon alpha diversity, Figure S4: Relative abundance of the top 1% of phyla, Figure S5: Relative abundance of the top 2% of species, Figure S6: Linear regressions of biomass traits and metabolic function abundance, Figure S7: Partial least squares results for genes related to shoot biomass, Figure S8: Partial least squares results for genes related to root biomass, Table S1: Genes correlated to biomass traits.

Author Contributions: Conceptualization, J.Y.B.; methodology, J.Y.B., C.P.F.-B., A.R.R., W.K., R.P.; software, C.P.F.-B., J.Y.B., A.R.R., R.P.; formal analysis, C.P.F.-B., J.Y.B., A.R.R.; investigation, J.Y.B., C.P.F.-B., A.R.R., W.K., J.E.M.; resources, J.Y.B., A.M.M., J.E.M., D.P.R., A.R.R., V.R.R.; data curation, C.P.F.-B., J.Y.B., A.R.R., R.P.; writing—original draft preparation, C.P.F.-B.; writing—review and editing, C.P.F.-B., J.Y.B., A.M.M., D.P.R., A.R.R., J.E.M., R.P., W.K., V.R.R.; visualization, J.Y.B., C.P.F.-B.; validation, J.Y.B., C.P.F.-B., A.R.R., J.E.M., W.K., R.P., A.M.M., D.P.R., V.R.R.; supervision, J.Y.B., D.P.R., A.M.M.; project administration, J.Y.B.; funding acquisition, J.Y.B., A.M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Cooperative Research Program of Rural Development Administration, Republic of Korea and United State Department of Agriculture, Agricultural Research Service, USA (Project title: Development of a screening methodology for assisting in rice cultivar selection for reduced methane emissions, Project No.: PJ012431).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All sequencing data and associated metadata have been deposited as an NCBI BioProject PRJN663614 under can be found under accession numbers SAMN16226752—SAMN16226799.

Acknowledgments: This research utilized facilities and assistance provided by USDA Agricultural Research Service Dale Bumpers National Rice Research Center, Stuttgart, Arkansas and Beltsville Agricultural Research Center, Beltsville, Maryland. We would like to thank Jonathan Moser, and Sarah Emche and Seonwoo Kim for their technical assistance in sample collection and sequencing preparation. USDA is an equal opportunity provider and employer.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Friesen, M.L.; Porter, S.S.; Stark, S.C.; Von Wettberg, E.J.; Sachs, J.L.; Martinez-Romero, E. Microbially Mediated Plant Functional Traits. *Annu. Rev. Ecol. Syst.* **2011**, *42*, 23–46. [\[CrossRef\]](#)
2. Berg, G.; Smalla, K. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol. Ecol.* **2009**, *68*, 1–13. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Baudoin, E.; Benizri, E.; Guckert, A. Impact of growth stage on the bacterial community structure along maize roots, as determined by metabolic and genetic fingerprinting. *Appl. Soil Ecol.* **2002**, *19*, 135–145. [\[CrossRef\]](#)
4. Inceoğlu, O.; Salles, J.F.; van Overbeek, L.; van Elsas, J.D. Effects of plant genotype and growth stage on the betaproteobacterial communities associated with different potato cultivars in two fields. *Appl. Environ. Microbiol.* **2010**, *76*, 3675–3684. [\[CrossRef\]](#)
5. Van Overbeek, L.; Van Elsas, J.D. Effects of plant genotype and growth stage on the structure of bacterial communities associated with potato (*Solanum tuberosum* L.). *FEMS Microbiol. Ecol.* **2008**, *64*, 283–296. [\[CrossRef\]](#)
6. Philippot, L.; Raaijmakers, J.M.; Lemanceau, P.; Van Der Putten, W.H. Going back to the roots: The microbial ecology of the rhizosphere. *Nat. Rev. Genet.* **2013**, *11*, 789–799. [\[CrossRef\]](#)
7. Santos-Medellín, C.; Edwards, J.; Liechty, Z.; Nguyen, B.; Sundaresan, V. Drought Stress Results in a Compartment-Specific Restructuring of the Rice Root-Associated Microbiomes. *mBio* **2017**, *8*, e00764-17. [\[CrossRef\]](#)
8. Edwards, J.A.; Santos-Medellín, C.M.; Liechty, Z.S.; Nguyen, B.; Lurie, E.; Eason, S.; Phillips, G.; Sundaresan, V. Compositional shifts in root-associated bacterial and archaeal microbiota track the plant life cycle in field-grown rice. *PLoS Biol.* **2018**, *16*, e2003862. [\[CrossRef\]](#)
9. Edwards, J.; Johnson, C.; Santos-Medellín, C.; Lurie, E.; Podishetty, N.K.; Bhatnagar, S.; Eisen, J.A.; Sundaresan, V. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E911–E920. [\[CrossRef\]](#)
10. Okubo, T.; Tokida, T.; Ikeda, S.; Bao, Z.; Tago, K.; Hayatsu, M.; Nakamura, H.; Sakai, H.; Usui, Y.; Hayashi, K.; et al. Effects of Elevated Carbon Dioxide, Elevated Temperature, and Rice Growth Stage on the Community Structure of Rice Root-Associated Bacteria. *Microbes Environ.* **2014**, *29*, 184–190. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Knief, C.; Delmotte, N.; Chaffron, S.; Stark, M.; Innerebner, G.; Wassmann, R.; Von Mering, C.; Vorholt, J. Metaproteomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *ISME J.* **2012**, *6*, 1378–1390. [\[CrossRef\]](#)
12. Hussain, Q.; Pan, G.; Liu, Y.; Zhang, A.; Li, L.; Zhang, X.; Jin, Z. Microbial community dynamics and function associated with rhizosphere over periods of rice growth. *Plant Soil Environ.* **2012**, *58*, 55–61. [\[CrossRef\]](#)
13. Wu, Z.; Liu, Q.; Li, Z.; Cheng, W.; Sun, J.; Guo, Z.; Li, Y.; Zhou, J.; Meng, D.; Li, H.; et al. Environmental factors shaping the diversity of bacterial communities that promote rice production. *BMC Microbiol.* **2018**, *18*, 51. [\[CrossRef\]](#)
14. Breidenbach, B.; Pump, J.; Dumont, M.G. Microbial Community Structure in the Rhizosphere of Rice Plants. *Front. Microbiol.* **2016**, *6*, 1537. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Knief, C.; Delmotte, N.; Vorholt, J.A. Bacterial adaptation to life in association with plants A proteomic perspective from culture to in situ conditions. *Proteomics* **2011**, *11*, 3086–3105. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Qiao, Q.; Wang, F.; Zhang, J.; Chen, Y.; Zhang, C.; Liu, G.; Zhang, H.; Ma, C.; Zhang, J. The Variation in the Rhizosphere Microbiome of Cotton with Soil Type, Genotype and Developmental Stage. *Sci. Rep.* **2017**, *7*, 1–10. [\[CrossRef\]](#)
17. Lundberg, D.S.; Lebeis, S.L.; Paredes, S.H.; Yourstone, S.; Gehring, J.; Malfatti, S.; Tremblay, J.; Engelbrektson, A.; Kunin, V.; Del Rio, T.G.; et al. Defining the core Arabidopsis thaliana root microbiome. *Nat. Cell Biol.* **2012**, *488*, 86–90. [\[CrossRef\]](#)
18. Veach, A.M.; Morris, R.; Yip, D.Z. Rhizosphere microbiomes diverge among Populus trichocarpa plant–host genotypes and chemotypes, but it depends on soil origin. *Microbiome* **2019**, *7*, 76. [\[CrossRef\]](#)
19. Breidenbach, B.; Econrad, R. Seasonal dynamics of bacterial and archaeal methanogenic communities in flooded rice fields and effect of drainage. *Front. Microbiol.* **2015**, *5*, 752. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Pfeiffer, S.; Mitter, B.; Oswald, A.; Schloter-Hai, B.; Schloter, M.; Declerck, S.; Sessitsch, A. Rhizosphere microbiomes of potato cultivated in the High Andes show stable and dynamic core microbiomes with different responses to plant development. *FEMS Microbiol. Ecol.* **2017**, *93*. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Wardlaw, I.F. Tansley Review No. 27 The control of carbon partitioning in plants. *New Phytol.* **1990**, *116*, 341–381. [\[CrossRef\]](#)
22. Briones, A.M.; Okabe, S.; Umehiya, Y.; Ramsing, N.-B.; Reichardt, W.; Okuyama, H. Influence of Different Cultivars on Populations of Ammonia-Oxidizing Bacteria in the Root Environment of Rice. *Appl. Environ. Microbiol.* **2002**, *68*, 3067–3075. [\[CrossRef\]](#)
23. Pfeiffer, J.A.; Spor, A.; Koren, O.; Jin, Z.; Tringe, S.G.; Dangl, J.L.; Buckler, E.S.; Ley, R.E. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6548–6553. [\[CrossRef\]](#)
24. Aira, M.; Gómez-Brandón, M.; Lazcano, C.; Bååth, E.; Domínguez, J. Plant genotype strongly modifies the structure and growth of maize rhizosphere microbial communities. *Soil Biol. Biochem.* **2010**, *42*, 2276–2281. [\[CrossRef\]](#)

25. Bouffaud, M.; Kyselková, M.; Gouesnard, B.; Grundmann, G.; Muller, D.; Moëgne-Loccoz, Y. Is diversification history of maize influencing selection of soil bacteria by roots? *Mol. Ecol.* **2011**, *21*, 195–206. [[CrossRef](#)] [[PubMed](#)]
26. Zhang, J.; Liu, Y.-X.; Zhang, N.; Hu, B.; Jin, T.; Xu, H.; Qin, Y.; Yan, P.; Zhang, X.; Guo, X.; et al. NRT1.1B is associated with root microbiota composition and nitrogen use in field-grown rice. *Nat. Biotechnol.* **2019**, *37*, 676–684. [[CrossRef](#)] [[PubMed](#)]
27. Van Der Heijden, M.G.A.; Klironomos, J.N.; Ursic, M.; Moutoglis, P.; Streitwolf-Engel, R.; Boller, T.; Wiemken, A.; Sanders, I.R. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nat. Cell Biol.* **1998**, *396*, 69–72. [[CrossRef](#)]
28. Wardle, D.A.; Bardgett, R.D.; Klironomos, J.N.; Setälä, H.; Van Der Putten, W.H.; Wall, D.H. Ecological Linkages between Aboveground and Belowground Biota. *Science* **2004**, *304*, 1629–1633. [[CrossRef](#)] [[PubMed](#)]
29. Leff, J.W.; Bardgett, R.D.; Wilkinson, A. Predicting the structure of soil communities from plant community tax-onomy, phylogeny, and traits. *ISME J.* **2018**, *12*, 1794–1805. [[CrossRef](#)]
30. Paliy, O.; Shankar, V. Application of multivariate statistical techniques in microbial ecology. *Mol. Ecol.* **2016**, *25*, 1032–1057. [[CrossRef](#)]
31. Ullah, A.; Akbar, A.; Luo, Q. Microbiome Diversity in Cotton Rhizosphere under Normal and Drought Condi-tions. *Microb. Ecol.* **2019**, *77*, 429–439. [[CrossRef](#)]
32. Ma, J.; Tang, J.Y.; Wang, S. Illumina sequencing of bacterial 16S rDNA and 16S rRNA reveals seasonal and spe-cies-specific variation in bacterial communities in four moss species. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 6739–6753. [[CrossRef](#)]
33. Dang, P.; Gao, Y.; Liu, J.; Yu, S.; Zhao, Z. Effects of thinning intensity on understory vegetation and soil microbial communities of a mature Chinese pine plantation in the Loess Plateau. *Sci. Total. Environ.* **2018**, *630*, 171–180. [[CrossRef](#)] [[PubMed](#)]
34. Counce, P.A.; Keisling, T.C.; Mitchell, A.J. A Uniform, Objective, and Adaptive System for Expressing Rice Development. *Crop Sci.* **2000**, *40*, 436–443. [[CrossRef](#)]
35. Kepler, R.M.; Schmidt, D.J.E.; Yarwood, S.A.; Cavigelli, M.A.; Reddy, K.N.; Duke, S.O.; Bradley, C.A.; Williams, M.M.; Buyer, J.S.; Maul, J.E. Soil Microbial Communities in Diverse Agroecosystems Exposed to the Herbicide Glyphosate. *Appl. Environ. Microbiol.* **2020**, *86*, e01744-19. [[CrossRef](#)] [[PubMed](#)]
36. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60. [[CrossRef](#)]
37. Huerta-Cepas, J.; Szklarczyk, D.; Heller, D.; Hernández-Plaza, A.; Forslund, S.K.; Cook, H.; Mende, D.R.; Letunic, I.; Rattei, T.; Jensen, L.J.; et al. eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **2019**, *47*, D309–D314. [[CrossRef](#)]
38. Overbeek, R.; Olson, R.; Pusch, G.D.; Olsen, G.J.; Davis, J.J.; Disz, T.; Edwards, R.A.; Gerdes, S.; Parrello, B.; Shukla, M.; et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* **2014**, *42*, D206–D214. [[CrossRef](#)] [[PubMed](#)]
39. Huson, D.H.; Beier, S.; Flade, I.; Górska, A.; El-Hadidi, M.; Mitra, S.; Ruscheweyh, H.-J.; Tappu, R. MEGAN Community Edition—Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Comput. Biol.* **2016**, *12*, e1004957. [[CrossRef](#)]
40. Gloor, G.B.; Macklaim, J.M.; Pawlowsky-Glahn, V.; Egozcue, J.J. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* **2017**, *8*, 2224. [[CrossRef](#)]
41. Oksanen, A.J.; Blanchet, F.G.; Kindt, R. Vegan: Community Ecology Package Version 2.5-7 2020. Available online: <http://CRAN.R-project.org/package=vegan> (accessed on 1 February 2021).
42. Lee, L.C.; Liong, C.-Y.; Jemain, A.A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps. *Analysis* **2018**, *143*, 3526–3539. [[CrossRef](#)] [[PubMed](#)]
43. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom Intell Lab Syst* **2001**, *58*, 109–130. [[CrossRef](#)]
44. Wold, S.; Geladi, P.; Esbensen, K.; Öhman, J. Multi-way principal components-and PLS-analysis. *J. Chemom.* **1987**, *1*, 41–56. [[CrossRef](#)]
45. Lingens, F.; Blecher, R.; Blecher, H.; Blobel, F.; Eberspächer, J.; Fröhner, C.; Gorisch, H.; Layh, G. *Phenylobacterium immobile* gen. nov., sp. nov., a Gram-Negative Bacterium That Degrades the Herbicide Chloridazon. *Int. J. Syst. Bacteriol.* **1985**, *35*, 26–39. [[CrossRef](#)]
46. Abraham, W.-R.; Macedo, A.J.; Lunsdorf, H. Phylogeny by a polyphasic approach of the order Caulobacterales, proposal of *Caulobacter mirabilis* sp. nov., *Phenylobacterium haematophilum* sp. nov. and *Phenylobacterium conjunc-tum* sp. nov., and emendation of the genus *Phenylobacterium*. *Int. J. Syst. Evol. Microbiol.* **2008**, *58*, 1939–1949. [[CrossRef](#)]
47. Ding, L.-J.; Cui, H.-L.; Nie, S.-A.; Long, X.-E.; Duan, G.-L.; Zhu, Y.-G. Microbiomes inhabiting rice roots and rhizosphere. *FEMS Microbiol. Ecol.* **2019**, *95*, 1–13. [[CrossRef](#)]
48. Ishii, S.; Ikeda, S.; Minamisawa, K.; Senoo, K. Nitrogen Cycling in Rice Paddy Environments: Past Achievements and Future Challenges. *Microbes Environ.* **2011**, *26*, 282–292. [[CrossRef](#)]
49. Hirano, K.; Sugiyama, T.; Kosugi, A.; Nioh, I.; Asai, T.; Nakai, H. Relationship between number of nitrogen-fixing rihzobacteria and growth pattern of rice varieties in the nature farming. *Breed Res.* **2001**, *3*, 3–12. [[CrossRef](#)]
50. Sanford, R.A.; Cole, J.R.; Tiedje, J.M.; Tiedje, J.M. Characterization and description of *Anaeromyxobacter dehalogenans* gen. nov., sp. nov., an aryl-halorespiring facultative anaerobic myxobacterium. *Appl. Environ. Microbiol.* **2002**, *68*, 893–900. [[CrossRef](#)] [[PubMed](#)]

51. Lee, H.J.; Kim, S.Y.; Kim, P.J.; Madsen, E.L.; Jeon, C.O. Methane emission and dynamics of methanotrophic and methanogenic communities in a flooded rice field ecosystem. *FEMS Microbiol. Ecol.* **2014**, *88*, 195–212. [[CrossRef](#)] [[PubMed](#)]
52. Neue, H.-U. Methane Emission from Rice Fields. *Bioscience* **1993**, *43*, 466–474. [[CrossRef](#)]
53. Sauder, L.A.; Albertsen, M.; Engel, K.; Schwarz, J.; Nielsen, P.H.; Wagner, M.; Neufeld, J.D. Cultivation and characterization of *Candidatus Nitrosocosmicus exaquare*, an ammonia-oxidizing archaeon from a municipal wastewater treatment system. *ISME J.* **2017**, *11*, 1142–1157. [[CrossRef](#)] [[PubMed](#)]
54. Jung, M.-Y.; Kim, J.-G.; Sinnighe Damsté, J.S. A hydrophobic ammonia-oxidizing archaeon of the Nitrosocosmicus clade isolated from coal tar-contaminated sediment. *Environ. Microbiol. Rep.* **2016**, *8*, 983–992. [[CrossRef](#)]
55. Kirk, G.J.D.; Kronzucker, H.J. The Potential for Nitrification and Nitrate Uptake in the Rhizosphere of Wetland Plants: A Modelling Study. *Ann. Bot.* **2005**, *96*, 639–646. [[CrossRef](#)]
56. Kirk, G. Plant-mediated processes to acquire nutrients: Nitrogen uptake by rice plants. *Plant Soil* **2001**, *232*, 129–134. [[CrossRef](#)]
57. Park, B.B.; Yanai, R.D.; Vadeboncoeur, M.A.; Hamburg, S.P. Estimating Root Biomass in Rocky Soils using Pits, Cores, and Allometric Equations. *Soil Sci. Soc. Am. J.* **2007**, *71*, 206–213. [[CrossRef](#)]
58. Pandey, G.; Jain, R.K. Bacterial Chemotaxis toward Environmental Pollutants: Role in Bioremediation. *Appl. Environ. Microbiol.* **2002**, *68*, 5789–5795. [[CrossRef](#)]
59. Philippot, L. Denitrifying genes in bacterial and Archaeal genomes. *Biochim. Biophys. Acta (BBA) Gene Struct. Expr.* **2002**, *1577*, 355–376. [[CrossRef](#)]
60. Kuzyakov, Y.; Domanski, G. Carbon input by plants into the soil. Review. *J. Plant Nutr. Soil Sci.* **2000**, *163*, 421–431. [[CrossRef](#)]
61. Mougel, C.; Offre, P.; Ranjard, L. Dynamic of the genetic structure of bacterial and fungal communities at different developmental stages of *Medicago truncatula* Gaertn. cv. Jemalong line J5. *New Phytol.* **2006**, *170*, 165–175. [[CrossRef](#)]
62. Mayer, E.; Dörr de Quadros, P.; Fulthorpe, R. *Plantibacter flavus*, *Curtobacterium herbarum*, *Paenibacillus tai-chungensis*, and *Rhizobium selenitireducens* Endophytes Provide Host-Specific Growth Promotion of Arabidopsis thaliana, Basil, Lettuce, and Bok Choy Plants. *Appl. Environ. Microbiol.* **2019**, *85*, e00383-19. [[CrossRef](#)] [[PubMed](#)]