microorganisms

MDPI

Article

Mapping genetic variation of Arabidopsis in response to Plant Growth Promoting Bacterium *Azoarcus olearius* DQS-4T

Fernanda Plucani do Amaral[1,6], Juexin Wang[7], Jacob Williams[3], Thalita R. Tuleski[4], Trupti Joshi[2,5], Marco A. R. Ferreira[3] and Gary Stacey[1]*

[1] Division of Plant Sciences and Technology, C. S. Bond Life Science Center, University of Missouri-Columbia, Missouri, USA.
[2] Department of Electrical Engineering and Computer Science, C. S. Bond Life Science Center, University of Missouri-Columbia, Missouri, USA.
[3] Department of Statistics, Virginia Tech.
[4] Department of Biochemistry and Molecular Biology, Federal University of Parana´, Curitiba-PR 19046, Brazil.
[5] Department of Health Management and Informatics, MU Institute for Data Science and Informatics, C. S. Bond Life Science Center, University of Missouri-Columbia, Missouri, USA.
[6] Ginkgo Bioworks, Ag Biologicals, 890 Embarcadero Dr., West Sacramento, California – 95605, USA.
[7] Department of BioHealth Informatics, School of Informatics and Computing, Indiana University Purdue University Indianapolis, Indianapolis, IN 46202, USA.
* Correspondence: **author**: Gary Stacey – staceyg@missouri.edu

Keywords: Plant growth promoting bacteria1 (PGPB), *Arabidopsis thaliana*; Natural genetic variation; Genome wide association study (GWAs), Agronomic traits

Abstract: Plant growth promoting bacteria (PGPB) can enhance plant health by facilitating nutrient uptake, nitrogen fixation, protection from pathogens, stress tolerance and/or boosting plant productivity. The genetic determinants that drive the plant – bacteria association remain understudied. To identify genetic loci highly correlated with traits responsive to PGPB, we performed a genome wide association study (GWAS) using an *Arabidopsis thaliana* population treated with *Azoarcus olearius* DQS-4T. Phenotypically, the 305 Arabidopsis accessions tested responded differently to bacterial treatment by improving, inhibiting, or not affecting root system or shoot traits. GWA mapping analysis identified several predicted loci associated with primary root length or root fresh weight. Two statistical analyses were performed to narrow down potential gene candidates followed by haplotype block analysis, resulting in identification of 11 loci associated with the responsiveness of Arabidopsis root fresh weight to bacterial inoculation. Our results showed considerable variation in the ability of plants to respond to inoculation by *A. olearius* DQS-4T while revealing considerable complexity regarding statistically associated loci with the growth traits measured. This investigation is a promising starting point for sustainable breeding strategies for future cropping practices that may employ beneficial microbes and/or modifications of the root microbiome.

Introduction

Plant growth promoting bacteria (PGPB) benefit plant growth notably by enhancing nutrient uptake, nitrogen fixation, protection from pathogens, stress tolerance and/or boosting plant productivity. Changes in root system architecture and shoot biomass are common plant responses to PGPB as evidenced on a variety of crops, including maize, rice, wheat, and various bioenergy grasses [1-4]. Despite an extensive literature documenting the beneficial effect of PGPB's on plant growth, we still know relatively little about the molecular details of their mode of action. Genetic variation in the plant host has been reported to modulate the composition of the root associated microbial popula-

tion and has been suggested to have important adaptative consequences for plant health [5, 6]. The process of domestication has profound consequences on crops, where the domesticate has moderately reduced genetic diversity relative to the wild ancestor across the genome, and severely reduced diversity for genes targeted by domestication [7], which can also impact plant-microbe interaction [8]. Hence, the identification and characterization of genes associated with the capacity of plants to maximize profitable responses to their associated beneficial bacteria could form the basis for breeding approaches to enhance yield, as well as increasing the sustainability of cropping systems. Taking advantage of genetic variation within a natural population, genome wide association analysis (GWAS) offers the opportunity to analyze associations between single nucleotide polymorphisms (SNPs) and phenotypic variance, a powerful tool for identification of genetic loci associated with agronomic traits. For example, GWA mapping analyses have revealed novel and unknown genes impacting a variety of agronomic traits (e.g. flowering time, defense, drought, root cell development, plant architecture, disease resistance) in numerous plant species, including Arabidopsis [8-10], rice [11, 12], maize [13, 14], and soybean [15-17].

Previous studies identified a number of bacterial genes involved in PGPB-plant association, including genes involved in nitrogen fixation [18], siderophore production and iron uptake, phosphate solubilization, production of volatile organic compounds, as well as phytohormones[19-21]. There is also a growing realization that plant symbionts can suppress the plant immune system in order to promote their colonization and infection of their plant host [22].

Azoarcus olearius DQS-4^T is a nitrogen-fixing, PGPB with the ability of colonize plant roots and enhance plant growth in monocots, such as rice and *Setaria viridis* [23, 24]. In a recent study, we utilized transposon mutagenesis to identify essential bacterial genes that modulated colonization of *Setaria viridis* roots by *A. olearius* DQS-4T and another PGPB, *Herbaspirillum seropedicae* SmR1[25]. Surprisingly, this study identified very few genes that were critical for both bacteria to colonize *S. viridis* roots. Instead, the data suggest that each bacterium requires a unique set of genes required for root colonization. This genetic diversity in the bacterial partner suggests that a similar level of complexity may exist with regard to how various plant hosts respond to specific PGPB strains. Although only two strains were used in our earlier study, the results appear to be quite distinct from other well-studied, plant-microbial associations. For example, in the rhizobial-legume symbiosis, a core set of both microbial and plant genes appear to be critical for establishment of the symbiosis [26-28]. For example, at this point there appears to be no evidence for the involvement of a common symbiosis pathway (CSP), as defined for the rhizobial and mycorrhizal-plant symbioses, in the intimate association of diverse plant hosts with PGPB [27]. It should be noted that a significant contributor to the identification of the CSP in legumes was the adoption by the research community of model plant species (i.e., *Medicago truncatula* and *Lotus japonicus*) that greatly aided the identification of plant genes essential for establishment of the symbiosis. In a recent review, we argued that the research community interesting in PGPB-plant associations would also greatly benefit by adopting model organisms to speed the molecular investigation of these interactions [29].

In the current study, we investigated the phenotypic responses of a natural population of 305 accessions [30] of the model plant *Arabidospsis thaliana* to inoculation by *A. olearius* DQS-4^T, a PGPB, and performed GWA mapping of four growth traits to identify genetic regions that contribute to bacterial plant growth promotion. The basis of GWAS is the ability to statistically associate specific genetic loci to measured phenotypic diversity within the population and depends on the [31] large linkage disequilibrium (LD) in plants [11, 32]. To explore further the output of our analysis we used two independent, statistical methods to analyze our dataset on the four root growth traits measured. Overlapping SNPs were identified associated with changes in root fresh weight and confirmed by haplotype block analysis. Here, we present the predicted candidate genetic

loci from the statistical analysis and discuss their possible relevance to the ability of *A. olearius* DQS-4^T to promote plant growth.

GWA mapping resulted in the identification of genetic loci associated with PGPB-induced changes in primary root length or root fresh weight.

Material and methods

Bacteria cultivation

Azoarcus olearius DQS-4^T [33] was grown overnight at 30 °C on liquid NFbHP-malate modified medium (DL malic acid 20 g L^{-1}) supplemented with potassium phosphate (K_2HPO_4 17.8 g L^{-1}, KH_2PO_4 159.5g L^{-1}) and ammonium chloride (NH_4Cl 20 mM) [34, 35]. Antibiotics were added to the culture at the following concentrations 100 µg mL^{-1} streptomycin and 10 µg mL^{-1} nalidixic acid. Subsequently, the DQS-4^T culture (OD_{600} = 0.9) was centrifuged at 3000g for 1 min, the pellet was washed three times by resuspension in 0.9 % (w/v) NaCl. After that, the optical density was adjusted by dilution to 0.005 cells.ml^{-1} (2.3 x10^5 CFU.ml^{-1}) in 50 ml of NaCl solution.

Plant growth conditions and bacterial treatment

A collection of 305 natural accessions of *Arabidopsis thaliana* [30] (Supplementary Material S1) was used to investigate the response to inoculation by *Azoarcus olearius* DQS-4^T. Arabidopsis seeds were sterilized by vortexing once with 70% ethanol for 1 min, twice with 70% ethanol plus 0.01% Triton X-100 for 2 min followed by once with 100% ethanol for 2 min. After the ethanol removal, seeds were allowed to dry in a sterile hood and then 1 ml sterile water was added prior to vernalization at 4 °C for 3 days. Sterilized seeds were sown on square Petri dishes containing agar-solidified ½ Murashige and Skoog (MS) medium supplemented with 0.5 % sucrose and incubated in a vertical position in a plant growth chamber at 21 °C with 16 h light / 8 h dark cycle. After 5 days of germination, seedlings of similar size were transferred to circular Petri dishes containing ½ MS agar medium. Plants were inoculated by applying 150 µl of DQS-4^T culture containing 2 x 10^5 cell ml^{-1} onto the agar medium, approximately 5 cm below the root tip, which allowed the roots to grow into the inoculant. The same procedure was done for control treatments using 150 µl of saline solution (0.9 % NaCl) without bacteria. The plates were briefly dried in a sterile hood, sealed with Parafilm and placed vertically in a growth chamber until phenotypic analysis.

Phenotypic response of Arabidopsis accessions to Azoarcus olearius DQS-4^T

For each accession, 5 seedlings were grown on a ½ MS agar plate. The reference accessions Col-0 and WS were used in each experiment since they represent non-responding and responding ecotypes, respectively. A total of 3 replicate plates (15 seedlings) of control and DQS-4T treated seedlings were analyzed for each of the 305 accessions tested. Growth parameters were analyzed 7 days upon treatment by counting lateral root number and measuring primary root length (cm), then average primary root length and lateral root number per seedling were determined. Root and shoot fresh weight were analyzed 8 days after treatment. Data were acquired simultaneously for inoculated and control samples from the three replicates. To determine statistical significance between control and inoculated plants, the one-way analysis (ANOVA) was used with Tukey (p-value < 0.05). Accessions with significant differences between control and inoculated were categorized as: A) Positive – genotypes that showed trait enhancement due inoculation; B) Negative - genotypes where inoculation inhibited growth and, C) Non-responsive – genotypes that showed no growth response to bacterial inoculation.

Data analysis

In this study, a natural population of 305 accessions of *Arabidopsis thaliana* [30] was used to investigate the genetic basis of the growth response to bacterial inoculation. Mapping analysis was performed on the following root parameters, primary root length (**Δ**PRL), lateral root number (**Δ**LRN), root fresh weight (**Δ**RFW), and shoot fresh weight (**Δ**SFW). For all traits, means per seedling (n = 5) per biological replicate (n = 3) were used to calculate the mean per treatment per accession. The mean value of the control treatment was subtracted from the value of the inoculated plants to generate the data sets used for GWAS analysis. Genome wide association analysis employed a Mixed Linear Model (MLM) using Tassel 5.0 software (http://www.maizegenetics.net/tassel) [36] incorporated with population structure by principal component analysis (PCA) and kinship matrix acquired from 1001 Genomes (https://1001genomes.org/) with minor allele frequency (MAF)= as 0.05, and the data was inferred as a normal distribution by Kolmogorov-Smirnov test. All accessions were genotyped against the Col-0 reference genome with ~214 k single nucleotide polymorphism (SNPs) markers [37]. Significant SNPs were identified with a strict threshold of significance by Bonferroni correction with *p-value* = 2.34 x 10^{-7}. Annotations of candidate genes were retrieved from TAIR10 (http://www.arabidopsis.org). In order to narrow the list of candidate SNPs to a more focused set of SNPs, the data were also analyzed using the two-steps method implemented in the R package GWAS.BAYES [38]. The two steps are called screening and model selection. The screening step of the GWAS.BAYES method performs a usual GWAS analysis with a linear mixed effects models with a SNP fixed effect and kinship random effects. The screening step provides the usual list of significant SNPs. The model selection step of the GWAS.BAYES method performs a genetic algorithm search through model space, where candidate models are linear mixed effects models with kinship random effects and may contain multiple SNPs. The genetic algorithm in the model selection step forces the SNPs to compete to appear in the highest ranked models. As shown in [39], combining a screening step and a model selection step provides a much shorter list of significant SNPs and leads to a much higher true discovery rate.

Manhattan plots and Linkage Disequilibrium (LD) plots were generated using the R statistical software [40].

Validation using quantitative reverse transcription PCR (qRT-PCR)

Gene expression was evaluated by qRT-PCR. RNA was isolated from roots 7 days after mock or bacterial treatment using Direct-zol RNA kit treated with DNase (Zymo Research) following the manufacturer's instructions. cDNA synthesis was performed with M-MLV reverse transcriptase (Promega). qRT-PCR was carried out as follow: 10 ng cDNA, 3 pM of each primer and PowerUp™ SYBR™ Green master mix (Applied Biosystems) were mixed and amplified in Applied Biosystems ABI PRISM™ 7500 detection system (Applied Biosystems). Three biological replicates and 3 technical replicates for each transcript were analyzed using LinReg PCR 11.1 [41]. Quantitative amplifications were performed for different genes and ubiquitin 10 (At4G05320) was used as an internal reference. Primers used are listed in Supplementary Material S2.

Results

Arabidopsis thaliana response to Azoarcus olearius DQS-4^T inoculation.

Previous, published research, including from our own laboratory [23] documented that *A. olearius* DQS-4^T can produce strong, positive effects on root growth in both rice and *Setaria viridis* [24]. As a prelude to our larger GWAS study, we initially tested only a few Arabidopsis ecotypes regarding their response to DQS-4^T inoculation. These initial experiments revealed that *A. thaliana* ecotype Columbia (Col-0) had no measurable response to bacterial inoculation, while ecotype Wassilewskija (Ws) showed a robust and significant increase in all the traits measured (lateral root number, root and shoot fresh weight) (Supplementary Material S3). Although limited, these initial experiments were

important in showing phenotypic diversity in the plant response to inoculation, as well as providing both a negative (Col-O) and positive (Ws) control for future experiments.

Natural variation in the response of Arabidopsis accessions to A. olearius inoculation

In order to investigate the natural variation in the responsiveness of Arabidopsis to *A. olearius* DQS-4^T, a total of 305 Arabidopsis accessions were analyzed for changes in root and shoot traits upon bacterial treatment. Analysis of correlation to measure the direction and strength between control and DQS-4^T treated samples showed a low correlation coefficient between control and treated samples for shoot fresh weight, primary root length and root fresh weight (R^2 = 0.452, R^2 = 0.349 and R^2 = 0.106, respectively). However, for lateral root number the correlation was only slightly stronger (R^2 = 0.501) between control and treated samples (Fig. 1A-H), suggesting that the magnitude of these DQS-4^T induced growth responses were weakly related to the intrinsic growth capacity for these parameters under the tested conditions. Hence, faster growing accessions or accessions that form more lateral roots in the experimental setup are not necessarily stronger responders to bacterial treatment.

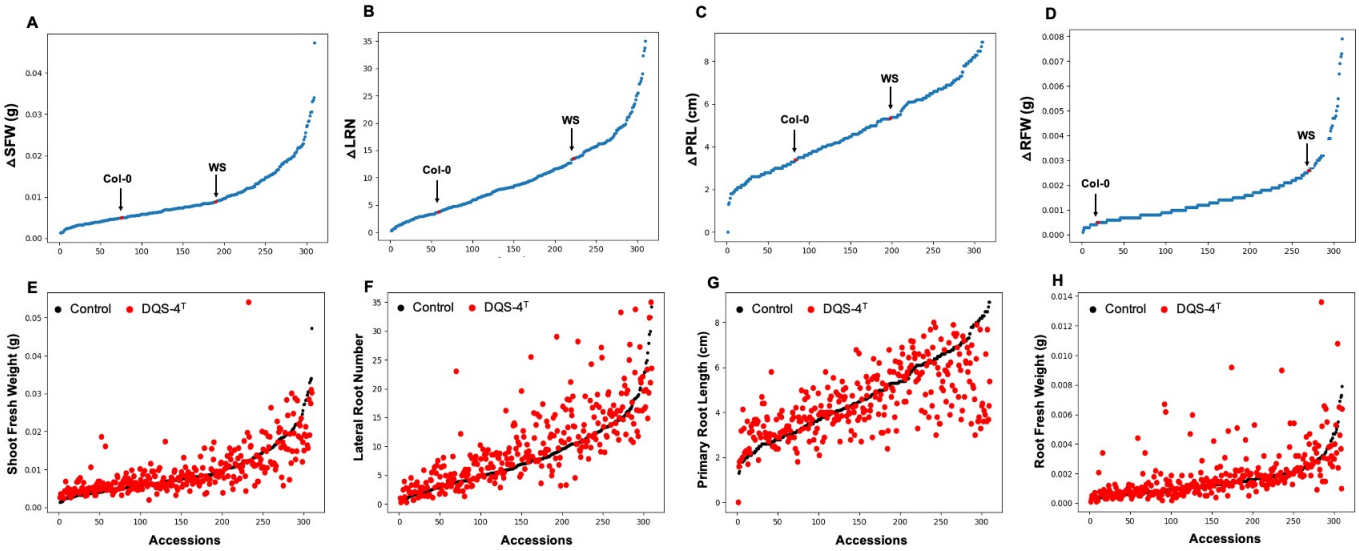


Figure 1. Natural variation of 305 *A. thaliana* accessions in response to the plant growth-promoting *Azoarcus olearius* DQS-4^T. A. Accessions sorted for increase in shoot fresh weight (ΔSFW) in response to DQS-4T (Col-0 and Ws are indicated with black or red arrow dot). B. Accessions sorted for increase in lateral root number (ΔLRN) in response to DQS-4T. C. Accessions sorted for increase in primary root length (ΔPRL) in response to DQS-4T. D. Accessions sorted for increase in root fresh weight (ΔRFW) in response to DQS-4T. E**.** Average shoot fresh weight (ΔSFW) of control (black dots) and DQS-4T (red dots) plants. F. Number of lateral roots (ΔLRN) formed in control (black dots) and DQS-4T-treated (red dots) plants. G. Primary root length (ΔPRL) of control (black dots) and DQS-4T (red dots) plants. H. Average root fresh weight (ΔRFW) of control (black dots) and DQS-4T (red dots) plants. Each dot represents the average of 3 biological replicates.

Response categories of Arabidopsis thaliana growth traits to treatments

The measured, phenotypic variability, relative to lateral root number (LRN), primary root length (PRL), root and shoot fresh weight (RFW, SFW), among the 305 Arabidopsis accessions classified into 3 categories: 1. Positive responsive: accessions that demonstrate a significant, positive change upon inoculation compared to mock samples; 2. Non-responsive: accessions where bacterial treatment did not affect growth and, 3. Negative responsive: accessions where treatment inhibited growth (Fig. 2). Within the positive responsive category, eleven genotypes showed growth enhancement in all four parameters analyzed, where most of the changes were statistically significant in lateral root number followed by shoot fresh weight and root fresh weight (Fig. 2 A and C). The

growth of 13 ecotypes clearly benefited from DQS-4^T inoculation (Fig. 2B-C). For instance, the genotype Ws showed an increase in biomass and root growth when inoculated (Fig. 2C).

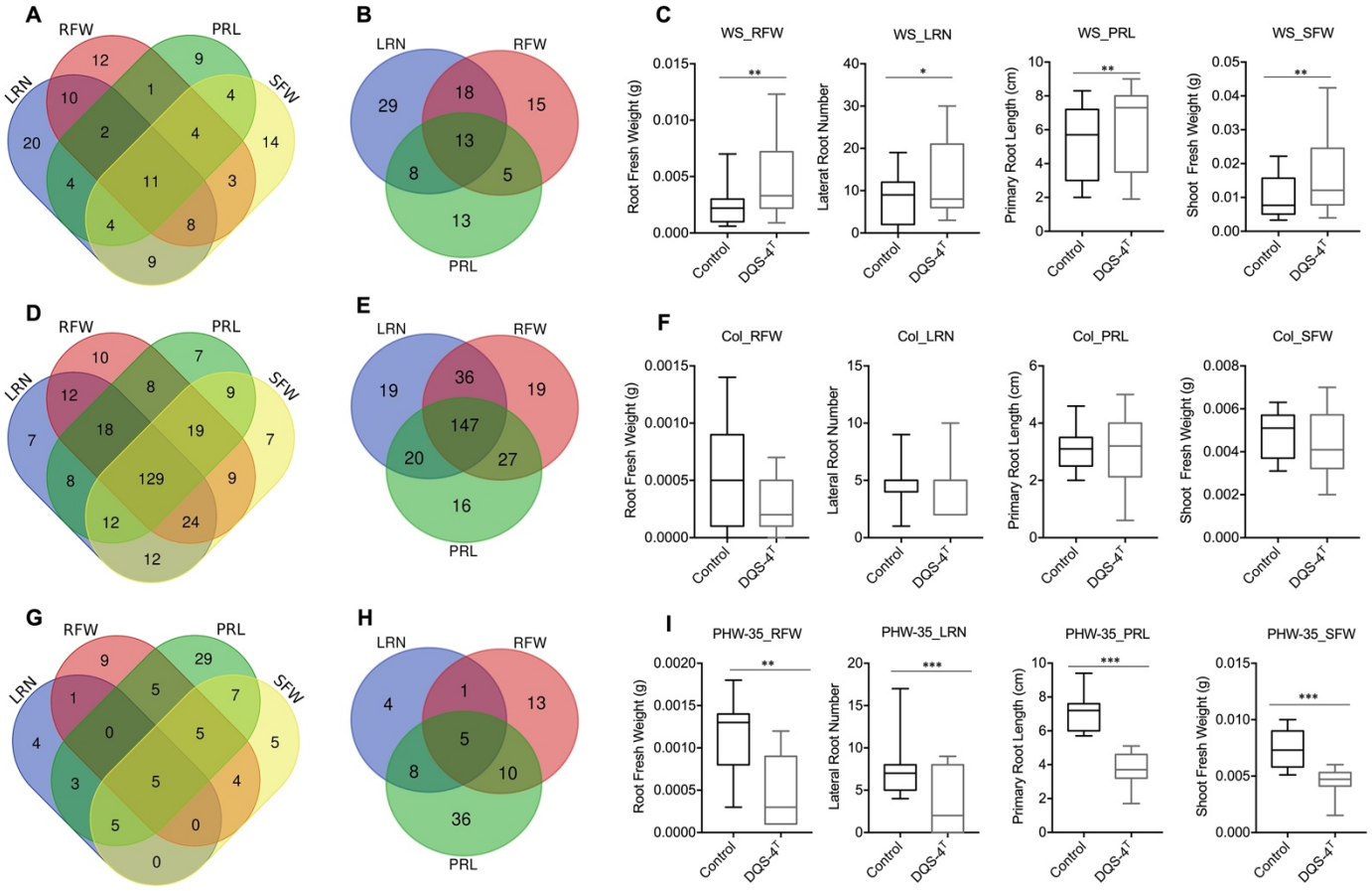


Figure 2. Growth responses of control and inoculated Arabidopsis accessions. Venn diagram showing the total number of common accessions across traits in each response category. A. Responsive accessions to DQS-4T inoculation for root and shoot traits. B. Responsive accessions in root growth parameters. C. Significant response of Wassilewskija (WS) genotype upon inoculation with DQS-4T in all four traits. D. Venn diagram of non-responsive accessions to DQS-4T inoculation for each root and shoot traits. E. Significant Non-responsive accessions in root growth parameters. F. No significant response of Col-0 upon inoculation with DQS-4T in all four traits. G. Accessions that responded negatively to DQS-4T inoculation in each trait analyzed. H. Negative response accessions in root growth parameters. I. Significant response of PHW-35 control upon inoculation with DQS-4T in all four traits analyzed. Bars are an average of 3 biological replicates (n=15). Statistical analysis was carried out using One-way Anova with *Tukey p-value* = 0.05.

However, the largest number of ecotypes (129) fell within the non-responsive group, which showed no statistically significant positive or negative response to bacterial inoculation for the parameters analyzed (Fig. 2D-F). As mentioned above, Col-0 is a good example of an ecotype within this non-responsive category (Fig. 2F). Interestingly, 82 genotypes showed a negative growth response to bacterial inoculation (Fig. 2G). An example is ecotype PHW-35 (Fig. 2H-I), which showed a negative response in each of the four parameters measured. However, when present, the positive effects on root architecture can be dramatic (Fig. 3); for example, as shown in Figure 3A for ecotype Kr-0 (Fig. 3A).

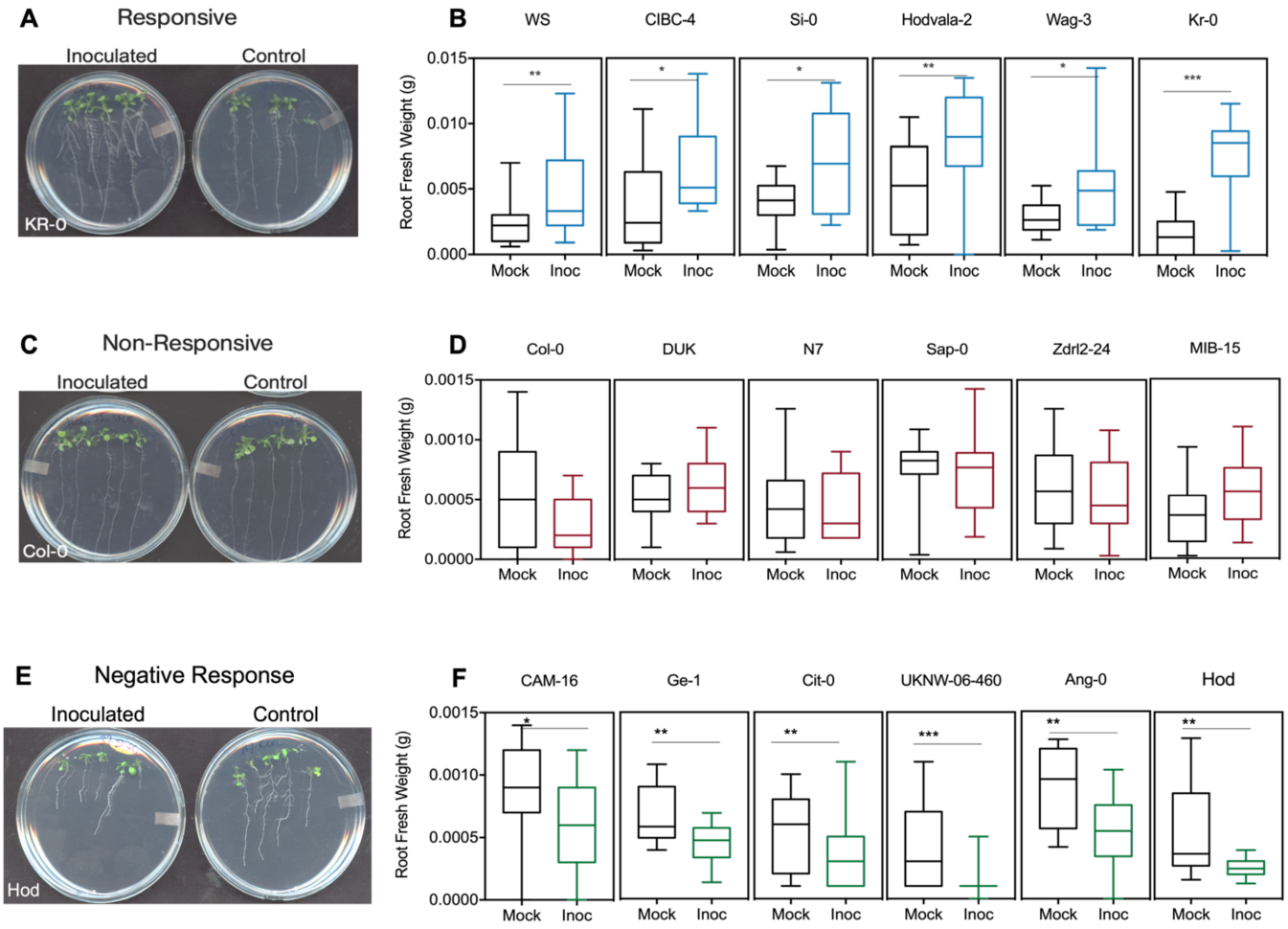


Figure 3. Phenotyping of Arabidopsis accessions for responsiveness to DQS-4T mediated effect in root fresh weight (RFW). Agar plates with 5 seedlings showing the different response of Arabidopsis accessions to *Azoarcus olearius*. A-B. Accessions that increased RFW upon DQS-4T treatment. C-D. Non-responsive accessions to bacteria treatment. E-F. Accessions where RFW was inhibited by DQS-4T treatment. Bars are an average of 3 biological replicates (n=15). Statistical analysis was carried out using one-way Anova with *Tukey p-value* < 0.05. RFW = root fresh weight. Photographs were taken by scanning the plates using a photo scanner with resolution 640 x 480.

Other ecotypes, such as Ws, CIBC-4, Si-0, Hodvala-2 and wag-3, demonstrated similar increases in root fresh weight upon bacterial treatment (Fig. 3B) and could be easily identified from non-responsive ecotypes, such as Col-0, DUK, N7, Sap-0, Zdrl2-24 and MIB-15 (Fig. 3C and D). However, it is important to note that it was common to find ecotypes that were not consistently responsive for all parameters measured (Fig. 3E and F). For instance, ecotype Aa-0 showed a significant increase in LRN upon inoculation but was non-responsive for the other root and shoot parameters. Another example of trait-related variation is exemplified by ecotype Hod which showed no significant changes to LRN and PRL; however, RFW was significantly reduced by inoculation while increasing shoot biomass. This variability within overall response, responses in individual parameters, and opposing responses (i.e., negative, and positive) suggests significant underlying complexity in the genetics of the plant response, as well as the molecular mechanisms involved. However, unlike Aa-0 and Hod, some ecotypes gave robust responses for all four traits analyzed, including ecotypes Ws, Bla-1, KI-5, Kr-0 or showed a consistent, negative response, such as ecotypes UKNW-06-460, UKSE 06-349, PHW-35 and PHW-37 (Table 1 and Supplementary Material S1 for a complete dataset).

Table 1. Genotypes with statistical significance for the four plant growth parameters measured with regard to DQS-4^T treatment.

		LRN	RFW	PRL	SFW
Response	**Genotype**	*p-value*			
Positive	WS	2.22E-03	9.60E-03	1.34E-02	4.67E-03
	Bla-1	3.27E-05	3.56E-04	3.40E-04	4.81E-06
	KI-5	8.21E-03	5.13E-02	1.83E-02	3.94E-02
	Ka-0	7.67E-04	2.30E-02	3.51E-04	9.94E-03
	LDV-25	5.90E-13	4.02E-08	8.31E-11	2.17E-08
	HS-0	3.92E-03	1.59E-03	6.93E-03	1.72E-03
	DralV-15	2.65E-03	9.26E-05	1.12E-17	1.52E-05
	In-0	3.00E-03	3.63E-02	2.21E-03	1.34E-02
	Hodvala-2	2.22E-02	5.20E-03	7.74E-03	1.66E-03
	Kr-0	2.36E-04	5.51E-06	6.12E-04	1.10E-04
	JEA	9.02E-04	1.87E-02	3.63E-03	9.04E-05
Negative	PHW-35	2.87E-03	9.19E-05	5.88E-10	1.01E-05
	PHW-37	4.90E-05	1.55E-03	8.45E-09	8.30E-06
	UKSE 06-349	1.65E-04	1.56E-04	1.76E-04	4.12E-04
	UKNW-06-460	7.31E-05	8.33E-03	3.26E-05	9.82E-06
	Lis-1	3.33E-02	3.59E-02	3.12E-05	1.50E-04

Genome Wide Association loci mapping in the Arabidopsis population

The dataset collected for DQS-4^T induced changes in lateral root number (ΔLRN), primary root length (ΔPRL), root fresh weight (ΔRFW) and shoot fresh weight (ΔSFW) were averaged and analyzed against the Col-0 reference genome using a mixed linear model (MLM) algorithm in Tassel 5.0 Software. To determine the distribution and quality of the data within the population a quantile-quantile plot (Q-Q plot) was generated for each growth trait (Supplementary Material S4). The GWA mapping results showed highly significant SNPs for two traits ΔPRL and ΔRFW (Fig. 4). No significant SNPs were significantly associated to ΔLRN and ΔSFW (Supplementary Material S4). With a threshold of $-\log_{10}(P) > 7$ adjusted by Bonferroni correction, a total of 63 loci were detected for root fresh weight and 55 loci correlated to primary root length (Supplemental Material S5). We observed only one SNP associated with both traits ΔRFW and ΔPRL, mapping close to the gene encoding AtFKGP, bifunctional fucokinase/fucose pyrophosphorylase (At1G01220, *p-value* = 2.19 x 10^{-08} and 3.32 x 10^{-07}, adjusted by Bonferroni correction respectively). Given this close association, we measured the expression of *AtFKGP* by qRT-PCR using mRNA extracted from roots of 7-day old seedlings of ecotypes representing non-responsive (Col-0, N7), responsive (Ws, Kr-0 and CIBC-4), and negative re-

sponding lines (PHW-37 and Lis-1) either mock non-inoculated or inoculated with DQS-4^T. However, this experiment failed to find either down or upregulated gene expression of *AtFKGP* in response to bacterial inoculation in any of the Arabidopsis accessions tested (Supplementary Material S6).

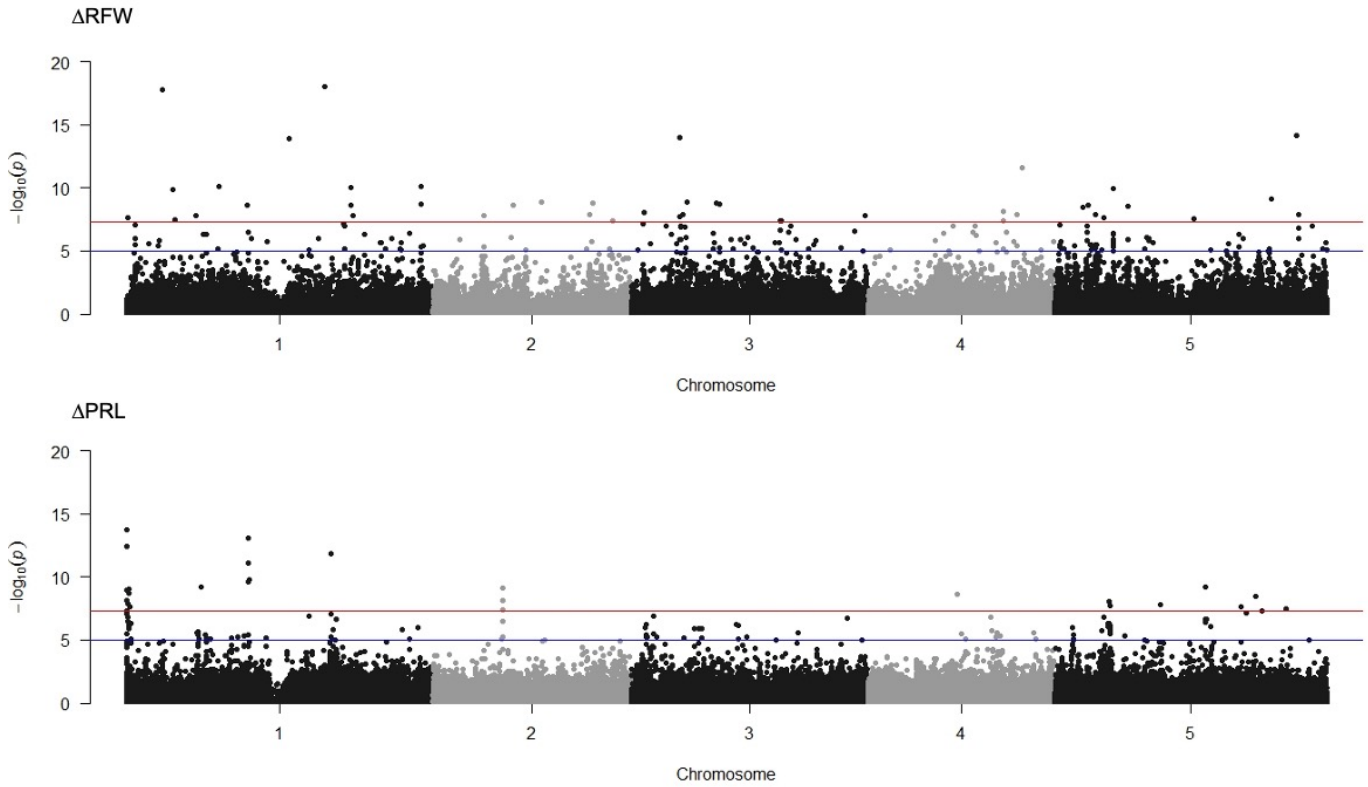


Figure 4. Manhattan plot of GWA mapping of *A. olearius* DQS-4^T effects on Root Fresh Weight (ΔRFW) and Primary Root Length (ΔPRL) in $-\log_{10}(P)$. SNP marker trait associations are shown as black and grey dots for each chromosome. The red and blue lines indicate the arbitrary thresholds of -log10(P) = 5 and -log10(P) = 7.

Primary Root Length highly correlated SNPs and candidate genes

We identified 55 SNPs highly correlated to measured changes in PRL (Supplemental Material S5). An example are SNPs mapping close to gene At1G33410, encoding the suppressor of auxin resistance 1 (*sar1*) gene. SAR1 and SAR3 are proteins similar to vertebrate nucleoporins that are part of the nuclear pore complex (NPC). Plants deficient in either protein exhibit pleiotropic growth defects partially affecting the translocation of proteins involved in hormonal signaling and plant development [42, 43]. A large LD resulted in inclusion of many polymorphisms in this candidate region; for example, loci within a gene predicted to encode a tetratricopeptide repeat 9 (TPR9) protein involved in gibberellic acid regulation [44], fascinated stem 4 (Atfas4) protein and a Ring/Ubox super family (At1g01660) protein. Moreover, a SNP highly correlated to primary root length corresponded to the gene encoding a late elongate hypocotyl LHY1, a MYB-related putative transcription factor implicated in circadian regulation of flowering time [45].

Root Fresh weight highly correlated SNPs and candidate genes

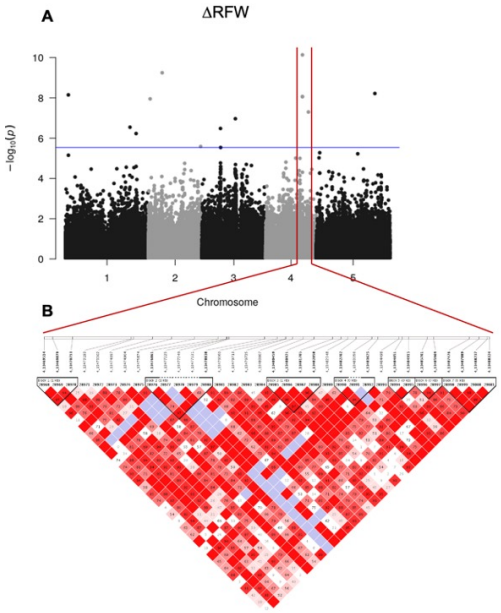
We identified 65 SNPs highly correlated to changes in RFW (Supplemental Material S5). In order to narrow the list of candidate SNPs to a more focused set of SNPs, the data were also analyzed using the two-stages procedure implemented in the R package

GWAS.BAYES [38,39]. This analysis reduced the number of SNPs highly associated with ΔRFW to 11 (Table 2); however, no SNPs were highly associated with ΔPRL in this analysis. In most GWA studies, the highly trait-correlated SNPs can present alleles in linkage disequilibrium (LD) at two or more loci in a population. However, because the peak of selection signals is relatively large in the GWA peak region, it is difficult to conclude whether the target of selection is the causative SNP or other alleles are significantly associated with this genome region. To examine the relationship between SNPs and regions, a haplotype block was generated for 10 kb upstream or downstream of the most significantly associated SNP at position 13459922 on chromosome 4 observed for ΔRFW (Fig. 5A-B). The selected SNP 13459922 (p = 7.55 x 10^{-09}; At4G26690) mapped close to a glycerophosphodiester phosphodiesterase like 3 GDPDL3 gene, also known as Shaven 3 (SHV3) that is involved in cell wall organization and root hair growth [46]. This SNP was identified as significant in both statistical methods used. SNP 13459922 is located at an intronic 5′ untranslated region (5′UTR) with a modifier predicted effect [47]. Next, we determined which gene in proximity to this highly associated genomic region underlies the variation. As shown in Fig. 5, six haplotype blocks were significantly associated with SNP 13459922, At4G26690. Based on haplotype analysis, seven SNP regions were identified associated, respectively, with genes encoding a member of the vacuolar-type ATPase family (At4G26710) and a protein phosphatase x-1(PP4) (At4G26720), extensin like protein (Lip5, At4G26750), microtubule-associated protein 65-2 (MAP65-2, At4G26760), phosphatidate cytidylyltransferase (CDS3, At4G26770), mitochondrial GRPE 2 (At4G26780) and Lipase acyl hydrolase superfamily (GDSL-like, At4G26790) (Fig. 5C). Because the polymorphisms are difficult to identify, we also carried out a matrix analysis, that showed SNPs at position 13477249 (gene Lip5) and 13483356 (CDS3 gene) highly correlated to SNP 13459922, GDPDL3 corroborating the haplotype block (Supplementary Material S7). Next, we determined the expression level of the genes correlated to the GWA peak by quantitative RT-PCR (qRT-PCR) using mRNA isolated from root tissue. We assumed that the expression of a candidate gene might be altered in accessions of different responsive categories. Hence, we extracted mRNA from the roots of selected accessions responsive (Ws and Kr-0), non-responsive (Col-0 and N7) and negative response (PHW-37 and Lis-1) regarding the RFW trait. The data show that none of the accessions tested showed a significant change in expression level upon inoculation for the various genes tested (i.e., those encoding GDPDL3, Lip5 and CDS3) (Supplementary Material S8). Of course, the lack of a transcriptional response to bacterial inoculation does not rule out the possibility that a specific gene could be playing an important role in the response to *A. olearius* DQS-4^T treatment.

Table 2. List of candidate genes from GWAs analysis of the *A. olearius* DQS-4^T -mediated plant effects on root fresh weight (ΔRFW) of 305 *Arabidopsis thaliana* population.

Chr	Candidate gene	Loci position	*p-value*	Gene annotation
1	At1G03530	882791	8.32E-08	Nuclear assembly factor 1 (ATNAF1)
1	At1G10660	3534853	1.75E-18	Transmembrane protein
1	At1G14040	4812798	3.23E-08	PHO1 homolog 3
1	At1G22550	7967378	4.68E-07	NPF5.16
1	At1G52710	19638846	9.13E-19	Rubredoxin-like superfamily protein
2	At2G18245	7939481	2.18E-09	alpha/beta-Hydrolases superfamily protein
3	At3G14400	4812265	1.84E-08	Ubiquitin-specific protease 25

4	At4G14820	8507871	1.15E-07	Pentatricopeptide repeat (PPR) superfamily protein
4	At4G26690	13459922	7.55E-09	Glycerophosphodiester phosphodiesterase like 3 (GDPDL3)
5	At5G08640	2804242	3.19E-09	Flavonol synthase 1
5	At5G35630	13833427	2.63E-08	Glutamine synthetase 2
5	At5G60070	24191284	1.57E-07	Ankyrin repeat family protein



C

SNP Position 13459922	p-value 7.55E-09	ID At4G26690	Gene name GDPDL3 / SHV3	Gene annotation Glycerophosphodiester phosphodiesterase like3	Uniprot function Root hair growth defect
Haplotype	SNP Position	Chr	ID	Gene annotation	
H1	13469514	4	At4G26710	VHA-E2 (Member of V-ATPase family. Vacuolar-type H + -ATPase)	
	13469879	4	At4G26711		
	13470713	4	At4G26720	Protein phosphatase x-1 (PP4)	
H2	13476861	4	At4G26750	Lip 5 - Extensin Like protein 5	
	13478838	4	At4G26760	Microtubule-associated protein 65-2 - MAP65-2	
H3	13480459	4	At4G26760	Microtubule-associated protein 65-2 - MAP65-2	
	13480921	4	At4G26760		
	13481705	4	At4G26760		
H4	13482702	4	At4G26770	Phosphatidate cytidylyltransferase – CDS3	
	13483625	4	At4G26770	Mitochondrial GRPE 2	
H5	13484651	4	At4G26770		
	13484915	4	At4G26780		
H6	13487476	4	At4G26790	GDSL-like Lipase/Acyl hydrolase superfamily	

Figure 5. A. Manhattan plot (-$\log_{10}(P)$) of a genome-wide association study (GWAS) of 305 Arabidopsis accessions treated with *Azoarcus olearius* DQS-4T. The GWAS significance level was set at 2.34 x 10^{-7} and plotted as a red line. B. Haplotype physical location around the highly correlated SNP At4G26690. C. Haplotype analysis of alleles in linkage disequilibrium (LD) with a highly correlated SNP of root fresh weight (ΔRFW).

Discussion

The benefits of PGPB in promoting plant growth, improving nutrient uptake and plant resilience to biotic and abiotic stress, and boosting crop production are documented by an expansive literature [3, 48-50]. While the molecular mechanisms and specific pathways that underlie the growth promoting responses in plants by PGPB have been investigated to some extent regarding the bacterial functions, left largely unexplored are the plant functions involved. Better defining these functions is important since they may help address the problems of consistency and efficiency that are found commonly when PGPB are used under field conditions to enhance crop yield and sustainability [51-53]. GWAS is a now popular method to harness natural genetic variation in a population to identify genetic loci critical for specific agronomic traits and in support of breeding improvement programs [13, 54-56]. Although used with great success in many studies, classical GWAS relying on SNPs has its limitations due to 'missing heritability' [57]. Failure to capture rare variants, allelic heterogeneity, epistasis, and/or epigenetic variation often decreases the detecting capacity of GWAS [58-62] [63]. To test the feasibility of this approach to investigate PGPG-plant interactions, we applied GWAS to map loci within the model plant Arabidopsis crucial for the beneficial response to the PGPB *Azoarcus olearius* DQS-4^T. Such an approach has been used previously; for example, to ex-

amine the response of an *Arabidopsis* natural population to inoculation with the PGPB *Pseudomonas fluorescens*. This study identified 10 potential genes candidates involved in changes root architecture and shoot biomass but found none strongly correlated to growth responses to bacterial inoculation with no common gene that could be correlated to a PGPB mediated effect [64].

A few general conclusions can be made from our study. Consistent with published reports, including some from our own studies [2, 23, 65] plant genotype largely determines whether a given PGPB strain will or will not enhance or inhibit plant growth. This could be a major factor in field-to-field variation in published PGPB studies [51-53]. However, perhaps most impactful, is that the data point to considerable complexity in the mechanisms that underlie a beneficial plant response to PGPB inoculation. For example, within the Arabidopsis population, 27% of ecotypes showed no response to inoculation, while others showed either a negative (12% PRL, 4% LRN, 6% RFW and 6%SFW) or positive (8% PRL, 13% LRN, 10% RFW and 11% SFW) response to a specific trait. Considering the four growth parameters tested, 11 ecotypes showed consistently positive response whereas 6 accessions responded negatively. Indeed, some ecotypes showed different responses regarding a specific phenotypic parameter. For instance, ecotype Hod showed increased shoot fresh weight while root biomass was negatively affected. This complexity correlates well with our recent, mutational analysis of two PGPB strains that suggested that the gene functions necessary for plant root colonization are unique to a given strain, with only a few genes appearing essential for both strains tested. This large variation, coupled with normal issues found when applying biological inoculation to cropping systems, could, in large part, explain why it is not uncommon to find very variable, inconsistent results when PGPB are used under field conditions [51-53].

The power of the GWAS method is the ability to statistically correlate specific genetic loci with the phenotypic variation measured across the entire population [66]. Hence, we used two, orthogonal methods for data analysis with increased statistical stringency. For example, our initial analysis using a mixed effect model identified several candidate SNPs associated with changes in primary root length and root fresh weight. However, in root length none of those SNPs were significant when submitted to a stricter statistical analysis. We encountered two major challenges to identify statistically significant SNP associations with the phenotypes measured. First, although several statistically robust models have been developed, false positives can still arise from population structure. Second is the large extended, linkage disequilibrium (LD) which results in the inclusion of many candidate genes within a single LD block, making it if very hard to ultimately identify the underlying, causative gene. For example, this might explain why none of the candidate genes associated with ΔPRL were found to respond transcriptionally to PGPB inoculation.

Among the candidate genes identified were a phosphate transporter essential for arbuscular mycorrhizal symbiosis [67] and glutamine synthetase (GS) known for its role in assimilation of nitrogen and biosynthesis of glutamine in plants. Transcriptional levels of the GS2 gene showed no changes in sugarcane leaves inoculated with several PGPB bacteria strains. However, these authors did detect biochemical differences relative to concentrations of glutamine and glutamate [68]. The ankyrin repeat family protein, was previously found in a study that characterized genes that contribute to bacterial adaptation isolated from Arabidopsis, maize, and poplar trees [69]. In rice, ankyrin-repeat protein was shown to be upregulated in response to rice leaf bright pathogen *Xanthomonas oryzae* (*xoo*) suggesting its role as a positive regulator in basal defense pathways [70].

Unfortunately, ecotype Col-0 did not respond to inoculation for any of the growth parameters measured. Hence, all the genetic resources available for this ecotype were of little use for follow-up genetic studies. For example, the gene encoding Lyst-interacting protein 5, (LIP5) was within the haplotype associated with ΔRFW. Arabidopsis LIP5 is

part of the endosomal sorting complexes required for transporters (ESCORTs) for sustained protein trafficking. Lyst-interacting protein 5 (LIP5), interacts with MAP kinase 3 (MPK3) and MPK6 in response to pathogen infection playing a critical role in plant basal resistance [71, 72]. Mutational disruption of LIP5 expression had little effect on pathogen associated molecular pattern (i.e., flagellin) or salicylic acid-induced defense responses but compromised basal resistance in response to the bacterial pathogen *Pseudomonas syringae* [71]. We obtained the available T-DNA insertion line of AtLIP5 within the Col-O background and found, not surprisingly, that this mutation had no effect on the response of seedlings to *A. olearius* DQS-4^T inoculation (data not shown). Hence, confirmation of those candidate genes implicated by our GWAS analysis in the PGPB response awaits the ability to use gene-editing to create mutations in those specific ecotype backgrounds that do respond to inoculation.

In summary, our study of the natural genetic variation within an Arabidopsis population showed considerable variation in the ability of plants to respond to inoculation by *A. olearius* DQS-4^T while revealing considerable complexity regarding statistically associated loci with the growth traits measured and, in the patterns (positive, neutral, negative) of those responses. Considering all candidate genes with SNP-trait associations in the GWA analysis, several have known or predicted functions that hold promise for being functional in mediating the PGPB growth effects.

Supplementary Materials:

Author Contributions: F.P.A designed and performed the experiments, analyzed results and wrote the manuscript. J.W conducted the GWAs mapping and reviewed the manuscript. J.W developed the GWAS.Bayes R package, conducted GWA mapping and reviewed the manuscript. T.R.T helped setup samples and phenotypic analysis. T.J GWA mapping discussion and manuscript review. M.A.R.F developed the GWAS.Bayes R package, conducted GWA mapping and reviewed the manuscript. G.S designed and discussion of experiment results, wrote and reviewed the manuscript.

Funding:

Institutional Review Board Statement:

Informed Consent Statement:

Data Availability Statement:

Acknowledgments: We would like to thank Dr. Sung-Hwan Cho for helping with results analysis and plant phenotyping discussions. The work of Ferreira and Williams was supported in part by National Science Foundation Grants DMS 1853549 and DMS 2054173. This work was supported by award NSF DMS 1853556. This version of the article has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.3390/microorganisms11020331.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. de Brito Ferreira, E.P., A.M. Knupp, and C.C. Garcia Martin-Didonet, *GROWTH OF RICE CULTIVARS (Oryza sativa L.) AS AFFECTED BY INOCULATION WITH PLANT GROWTH-PROMOTING BACTERIA.* Bioscience Journal, 2014. **30**(3): p. 655-665.
2. do Amaral, F.P., et al., Differential growth responses of Brachypodium distachyon genotypes to inoculation with plant growth promoting rhizobacteria. Plant Mol Biol, 2016. **90**(6): p. 689-97.
3. Pedrosa, F.O., et al., The ammonium excreting Azospirillum brasilense strain HM053: a new alternative inoculant for maize. 2019: Plant Soil. p. 1-12.

4. Zeffa, D.M., et al., Azospirillum brasilense promotes increases in growth and nitrogen use efficiency of maize genotypes. Plos One, 2019. **14**(4).
5. Micallef, S.A., M.P. Shiaris, and A. Colon-Carmona, Influence of Arabidopsis thaliana accessions on rhizobacterial communities and natural variation in root exudates. Journal of Experimental Botany, 2009. **60**(6): p. 1729-1742.
6. Haney, C., et al., Associations with rhizosphere bacteria can confer an adaptive advantage to plants. 2015: Nature Plants. p. 1-9.
7. Flint-Garcia, S.A., *Genetics and Consequences of Crop Domestication.* Journal of Agricultural and Food Chemistry, 2013. **61**(35): p. 8267-8276.
8. Perez-Jaramillo, J., Mendes R, and J. Raaijmakers, Impactof plant domestication on rhizosphere microbiome assembly and functions. 2016: Plant Mol Biol.
9. Francisco, M., et al., Genome Wide Association Mapping in Arabidopsis thaliana Identifies Novel Genes Involved in Linking Allyl Glucosinolate to Altered Biomass and Defense. Front Plant Sci, 2016. **7**: p. 1010.
10. Angelovici, R., et al., Network-Guided GWAS Improves Identification of Genes Affecting Free Amino Acids. Plant Physiology, 2017. **173**(1): p. 872-886.
11. Yano, K., et al., Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. Nature Genetics, 2016. **48**(8): p. 927-+.
12. Guo, Z., et al., Genome-Wide Association Studies of Image Traits Reveal Genetic Architecture of Drought Resistance in Rice. Molecular Plant, 2018. **11**(6): p. 789-805.
13. Gyawali, A., et al., Single-plant GWAS coupled with bulk segregant analysis allows rapid identification and corroboration of plant-height candidate SNPs. Bmc Plant Biology, 2019. **19**(1).
14. Jiang, S., et al., Genome-Wide Association Study Dissects the Genetic Architecture of Maize Husk Tightness. Frontiers in Plant Science, 2020. **11**.
15. Fang, C., et al., Genome-wide association studies dissect the genetic networks underlying agronomical traits in soybean. Genome Biology, 2017. **18**.
16. Qin, J., et al., Genome Wide Association Study and Genomic Selection of Amino Acid Concentrations in Soybean Seeds. Frontiers in Plant Science, 2019. **10**.
17. Assefa, T., et al., Deconstructing the genetic architecture of iron deficiency chlorosis in soybean using genome-wide approaches. Bmc Plant Biology, 2020. **20**(1).
18. Bashan, Y. and L.E. De-Bashan, How the Plant Growth - Promoting Bacterium Azospirillum Promotes PlantGrowth — ACriticalAssessment Vol. 108. 2010, Advances in Agronomy.
19. Spaepen, S. and J. Vanderleyden, *Auxin and Plant-Microbe Interactions.* Cold Spring Harbor Perspectives in Biology, 2011. **3**(4).
20. Spaepen, S., J. Vanderleyden, and R. Remans, *Indole-3-acetic acid in microbial and microorganism-plant signaling.* Fems Microbiology Reviews, 2007. **31**(4): p. 425-448.
21. Wintermans, P.C., P.A. Bakker, and C.M. Pieterse, Natural genetic variation in Arabidopsis for responsiveness to plant growth-promoting rhizobacteria. Plant Mol Biol, 2016. **90**(6): p. 623-34.
22. Cao Y , et al., *The Role of Plant Innate Immunity in the Legume-Rhizobium Symbiosis*. 2017: Annual Review of Plant Biology. p. 535-561
23. Pankievicz, V.C.S., et al., Robust biological nitrogen fixation in a model grass-bacterial association. Plant Journal, 2015. **81**(6): p. 907-919.
24. Faoro, H., et al., The oil-contaminated soil diazotroph Azoarcus olearius DQS-4(T) is genetically and phenotypically similar to the model grass endophyte Azoarcus sp. BH72. Environ Microbiol Rep, 2017. **9**(3): p. 223-238.
25. Amaral, F.P., et al., Diverse bacterial genes modulate plant root association by beneficial bacteria. 2020: mBio.
26. Gourion, B., et al., *Rhizobium-legume symbioses: the crucial role of plant immunity.* Trends in Plant Science, 2015. **20**(3): p. 186-194.
27. Roy, S., et al., Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation(OPEN). Plant Cell, 2020. **32**(1): p. 15-41.
28. Nakagawa, T. and H. Imaizumi-Anraku, Rice arbuscular mycorrhiza as a tool to study the molecular mechanisms of fungal symbiosis and a potential target to increase productivity. Rice, 2015. **8**.
29. Pankievicz, V., et al., Diazotrophic Bacteria and their Mechanisms to Interact and Benefit Cereals. 2021: Molecular Plant-Microbe Interactions.
30. Nordborg, M., et al., *The pattern of polymorphism in Arabidopsis thaliana.* PLoS Biol, 2005. **3**(7): p. e196.
31. Curtin, S.J., et al., Validating Genome-Wide Association Candidates Controlling Quantitative Variation in Nodulation. Plant Physiol, 2017. **173**(2): p. 921-931.
32. Flint-Garcia, S.A., J.M. Thornsberry, and E.S. Buckler, *Structure of linkage disequilibrium in plants.* Annual Review of Plant Biology, 2003. **54**: p. 357-374.
33. Chen, M.H., et al., *Azoarcus olearius sp nov., a nitrogen-fixing bacterium isolated from oil-contaminated soil.* International Journal of Systematic and Evolutionary Microbiology, 2013. **63**: p. 3755-3761.
34. Okon, Y. and C.A. Labandera-Gonzalez, *Agronomic applications of azospirillum: An evaluation of 20 years worldwide field inoculation.* Soil Biology and Biochemistry, 1994. **26**(12): p. 1591-1601.
35. Klassen, G., et al., Effect of nitrogen compounds on nitrogenase activity in Herbaspirillum seropedicae SMR1. Canadian Journal of Microbiology, 1997. **43**(9): p. 887-891.

36. Bradbury, P.J., et al., TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics, 2007. **23**(19): p. 2633-5.
37. Kim, S., et al., Recombination and linkage disequilibrium in Arabidopsis thaliana. Nat Genet, 2007. **39**(9): p. 1151-5.
38. Williams J, Ferreira M.A.R, and Ji. T, GWAS.BAYES: An R package for Bayesian analysis of GWAS data. R package version 1.0.0. 2020, Bioconductor.
39. Williams J, Ferreira M.A.R, and Ji T., BICOSS: Bayesian iterative conditional stochastic search for GWAS. BMC Bioinformatics, 2022, 23, 475.
40. Team, R.C., *R: A language and environment for statistical computing*. 2022: R Foundation for Statistical Computing.
41. Ruijter, J.M., et al., Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res, 2009. **37**(6): p. e45.
42. Parry, G., et al., The Arabidopsis SUPPRESSOR OF AUXIN RESISTANCE proteins are nucleoporins with an important role in hormone signaling and development. 2006: The Plant cell. p. 1590-1603.
43. Li, C., et al., Nucleoporin 160 Regulates Flowering through Anchoring HOS1 for Destabilizing CO in Arabidopsis. 2020: Plant communications. p. 100033.
44. Silverstone, A.L., et al., *Functional analysis of SPINDLY in gibberellin signaling in Arabidopsis.* Plant Physiology, 2007. **143**(2): p. 987-1000.
45. Dong, M.A., E.M. Farre, and M.F. Thomashow, CIRCADIAN CLOCK-ASSOCIATED 1 and LATE ELONGATED HYPOCOTYL regulate expression of the C-REPEAT BINDING FACTOR (CBF) pathway in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(17): p. 7241-7246.
46. Hayashi, S., et al., The Glycerophosphoryl Diester Phosphodiesterase-Like Proteins SHV3 and its Homologs Play Important Roles in Cell Wall Organization. Plant and Cell Physiology, 2008. **49**(10): p. 1522-1535.
47. Cingolani P, et al., A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. 2012: Fly (Austin). p. 80-92.
48. Compant, S., et al., Use of plant growth-promoting bacteria for biocontrol of plant diseases: Principles, mechanisms of action, and future prospects. Applied and Environmental Microbiology, 2005. **71**(9): p. 4951-4959.
49. Fukami, J., et al., Accessing inoculation methods of maize and wheat with Azospirillum brasilense. Amb Express, 2016. **6**.
50. Di Benedetto, N.A., et al., The role of Plant Growth Promoting Bacteria in improving nitrogen use efficiency for sustainable crop production: a focus on wheat. Aims Microbiology, 2017. **3**(3): p. 413-434.
51. Timmusk, S., et al., Perspectives and Challenges of Microbial Application for Crop Improvement. Frontiers in Plant Science, 2017. **8**.
52. Backer, R., et al., Plant Growth-Promoting Rhizobacteria: Context, Mechanisms of Action, and Roadmap to Commercialization of Biostimulants for Sustainable Agriculture. Frontiers in Plant Science, 2018. **9**.
53. Gange, A.C. and K.R. Gadhave, Plant growth-promoting rhizobacteria promote plant size inequality. Scientific Reports, 2018. **8**.
54. Battenfield, S.D., et al., Breeding-assisted genomics: Applying meta-GWAS for milling and baking quality in CIMMYT wheat breeding program. Plos One, 2018. **13**(11).
55. Gupta, P.K., P.L. Kulwal, and V. Jaiswal, *Association mapping in plants in the post-GWAS genomics era.* Advances in Genetics, Vol 104, 2019. **104**: p. 75-154.
56. Tsai, H.Y., et al., Genomic prediction and GWAS of yield, quality and disease-related traits in spring barley and winter wheat (vol 10, 3347, 2020). Scientific Reports, 2020. **10**(1).
57. Eichler, E.E., et al., Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet, 2010. **11**(6): p. 446-50.
58. Johannes, F., et al., Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genet, 2009. **5**(6): p. e1000530.
59. Manolio, T.A., et al., *Finding the missing heritability of complex diseases.* Nature, 2009. **461**(7265): p. 747-53.
60. Bergelson, J. and F. Roux, Towards identifying genes underlying ecologically relevant traits in Arabidopsis thaliana. Nat Rev Genet, 2010. **11**(12): p. 867-79.
61. Bodenhausen, N., M.W. Horton, and J. Bergelson, Bacterial Communities Associated with the Leaves and the Roots of Arabidopsis thaliana. Plos One, 2013. **8**(2).
62. Brachi, B., G.P. Morris, and J.O. Borevitz, Genome-wide association studies in plants: the missing heritability is in the field. Genome Biol, 2011. **12**(10): p. 232.
63. Wang, J., et al., A Bayesian model for detection of high-order interactions among genetic variants in genome-wide association studies. BMC Genomics, 2015. **16**: p. 1011.
64. Wintermans, P.C.A., P.A.H.M. Bakker, and C.M.J. Pieterse, *Natural genetic variation in Arabidopsis for responsiveness to plant growth-promoting rhizobacteria.* Plant Molecular Biology, 2016. **90**(6): p. 623-634.
65. Beneduzi, A., A. Ambrosini, and L.M.P. Passaglia, *Plant growth-promoting rhizobacteria (PGPR): Their potential as antagonists and biocontrol agents.* Genetics and Molecular Biology, 2012. **35**(4): p. 1044-1051.
66. Zhang, Z.W., et al., Mixed linear model approach adapted for genome-wide association studies. Nature Genetics, 2010. **42**(4): p. 355-U118.

67. Javot, H., et al., A Medicago truncatula phosphate transporter indispensable for the arbuscular mycorrhizal symbiosis. Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1720-5.
68. Cipriano, M.A.P., et al., Plant-Growth Endophytic Bacteria Improve Nutrient Use Efficiency and Modulate Foliar N-Metabolites in Sugarcane Seedling. Microorganisms, 2021. **9**(3).
69. Levy, A., Salas Gonzalez, I., Mittelviefhaus, M., *Genomic features of bacterial adaptation to plants*. 2018: *Nat Genet*. p. 138-150.
70. Mou, S., et al., Functional analysis and expressional characterization of rice ankyrin repeat-containing protein, OsPIANK1, in basal defense against Magnaporthe oryzae attack. PLoS One, 2013. **8**(3): p. e59699.
71. Wang F, et al., Arabidopsis LIP5, a Positive Regulator of Multivesicular Body Biogenesis, Is a Critical Target of Pathogen- Responsive MAPK Cascade in Plant Basal Defense. 2014: PLoS Pathog.
72. Buono, R.A., et al., Role of SKD1 Regulators LIP5 and IST1-LIKE1 in Endosomal Sorting and Plant Development. Plant Physiology, 2016. **171**(1): p. 251-264.