The genetic architecture of *Arabidopsis thaliana* in response to native non-pathogenic leaf bacterial species revealed by GWA mapping in field conditions

Daniela Ramírez-Sánchez[¶], Rémi Duflos[¶], Chrystel Gibelin-Viala, Rémy Zamar, Fabienne Vailleau[*] and Fabrice Roux[*]

LIPME, INRAE, CNRS, Université de Toulouse, Castanet-Tolosan, France

[¶] These authors have contributed equally to this work and share first authorship.

[*] The authors share senior authorship.

[*] Corresponding author: fabrice.roux@inrae.fr (F. Roux)

Competing Interests: The authors declare that there are no competing financial interests in relation to the work described

ABSTRACT

Non-pathogenic bacteria can largely contribute to plant health by mobilizing and supplying nutrients and by providing protection against pathogens and resistance to abiotic stresses. Yet, the number of GWAS reporting the genetic architecture of the response to individual members of the beneficial microbiota remains limited. In this study, we established a GWAS under field conditions to estimate the level of genetic variation and the underlying genetic architecture, among 162 accessions of *Arabidopsis thaliana* originating from 54 natural populations located south-west of France, in response to 13 strains of seven of the most abundant and prevalent non-pathogenic bacterial species isolated from the leaf compartment of *A. thaliana* in the same geographical region. Using a high-throughput phenotyping methodology to score vegetative growth-related traits, extensive genetic variation was detected within our local set of *A. thaliana* accessions in response to these leaf bacteria, both at the species and strain levels. The presence of crossing reaction norms among strains indicates that declaring a strain as a plant-growth promoting bacterium is highly dependent on the host genotype tested. In line with the strong genotype-by-genotype interactions, we detected a complex and highly flexible genetic architecture between the 13 strains. Finally, the candidate genes underlying the QTLs revealed a significant enrichment in several biological pathways, including cell, secondary metabolism, signalling and transport. Altogether, plant innate immunity appears as a significant source of natural genetic variation in plant-microbiota interactions and opens new avenues for better understanding the ecologically relevant molecular dialog during plant-microbiota interactions.

INTRODUCTION

Both wild plant species and crops are consistently challenged by pathogens, making infectious disease often the major selective agent in nature [1–5]. In wild species, pathogen attacks can significantly decrease the number of offspring, which in turn affects host population growth rate [6–8]. Yield losses resulting from pathogen attacks can reach several tens of percent in crops [9–12], thereby threatening global food security [10, 13]. A major challenge in plant breeding and in ecological genomics is therefore to characterize the genetic architecture of response to pathogen attacks [14, 15]. Identifying the genetic and molecular bases for natural variation in response to pathogen attacks might lead to fundamental insights in the prediction of evolutionary trajectories of natural populations [16–19] and have enormous practical implications by increasing crop yield and quality [20–22].

Over the last decade, whole-genome sequencing made possible through the development of cutting-edge next-generation sequencing (NGS) technologies, combined with the development of increasingly sophisticated statistical methods in quantitative genetics [23, 24], led to a burst in the number of genome-wide association studies (GWAS) that were successful in both wild and cultivated plants. This allowed detecting genomic regions associated with natural variation of response to experimental inoculation with, in most cases, individual pathogenic strains [15, 25–27]. GWAS in plants revealed that the genetic architecture of response to pathogen attacks was highly polygenic [15], highly dependent on the abiotic environment [28, 29] and dynamic along the infection stages [28, 30]. In addition, the functional validation of few quantitative trait loci (QTLs) combined with transcriptomic analyses revealed both the involvement of a broad range of rarely considered molecular functions in plant immunity [17, 31, 32] as well as a new biomolecular network of the signaling machineries underlying disease resistance [33–35].

However, the entire set of microbial pathogens - also called pathobiota - represent only a small fraction of the entire set of microbes inhabiting plants, the so-called plant microbiota [14, 36,

37]. For instance, in the leaf compartment of 163 natural populations of *Arabidopsis thaliana* located in the south-west of France and characterized for bacterial communities using a metabarcoding approach allowing distinguishing pathogenic bacteria from other bacterial species [38], the relative abundance of pathobiota in microbiota was on average 1.6% in asymptomatic plants and 4.5% in plants with visible disease symptoms [38]. Furthermore, microbiota can largely contribute to plant health by (i) providing direct (production of antimicrobial components, niche competition) or indirect (triggering immune defense) protection against pathogens, (ii) mobilizing and provisioning nutrients, and (iii) providing resistance to abiotic stresses (such as drought) [37, 39–46]. Yet, the number of GWAS reporting the genetic architecture of the response to experimental inoculation with individual members of the beneficial microbiota remains limited in comparison to the number of GWAS on response to pathogens. In addition, despite the fact that the phyllosphere represents 60% of the total biomass on Earth and concentrating 10^{26} bacteria (Vorholt, 2012), most GWAS conducted in response to non-pathogenic bacteria focused on symbiotic bacteria or non-symbiotic plant-growth promoting bacteria (PGPB) at the below-ground level and in laboratory controlled conditions [47–52].

In this study, we established a GWAS under field conditions to estimate the level of genetic variation and the underlying genetic architecture, among 162 whole-genome sequenced accessions of *A. thaliana* originating from 54 natural populations located south-west of France, in response to 13 strains of seven of the most abundant and prevalent bacterial species isolated from the leaf compartment of *A. thaliana* in the same geographical region [53]. To do so, we first developed a high-throughput phenotyping methodology to score vegetative growth-related traits on tens of thousands of plants. We then combined GWA mapping derived from a Bayesian hierarchical model (BHM) [54], with a local score (LS) approach [55] to fine map QTLs down to the gene level, a combination that was successfully applied to detect and/or functionally

validate QTLs involved in biotic interactions in *A. thaliana* [28, 30, 55, 56]. We finally identified the main biological pathways associated with all the candidate genes and discussed the function of the main candidate genes.

MATERIAL AND METHODS

Plant material

A total of 54 populations (each represented by three accessions) were chosen to represent both the genomic and ecological diversity identified among a set of 168 natural populations of *A. thaliana* located southwest of France [57, 58] (Supplementary Table S1). Seeds from maternal plants sampled in natural populations were collected in June 2016. Differences in the maternal effects among the 162 seed lots were reduced by growing one plant of each accession for one generation (Supplementary Text).

Bacterial material

We considered two strains of seven (*i.e.* OTU2, OTU3, OTU4, OTU5, OTU6, OTU13 and OTU29) out of the 12 most abundant and prevalent non-pathogenic leaf bacterial OTUs identified across the 168 natural populations of *A. thaliana* [38], with the exception of OTU4 for which only one strain was available [53]. Based on whole-genome sequencing, the closest taxonomic classification for OTU2, OTU3, OTU4, OTU5, OTU6, OTU13 and OTU 29 was *Paraburkholderia fungorum*, *Oxalobacteraceae* bacterium, *Comamonadaceae* bacterium, *Pseudomonas moraviensis*, *Pseudomonas siliginis*, *Methylobacterium* sp. and *Sphingomonadaceae* bacterium, respectively [53]. For the purpose of another study, we also included the strain JACO-CL of the bacterial pathogen *Pseudomonas viridiflava* (OTU8), which is with *Xanthomonas campestris* the most abundant and prevalent bacterial pathogen across the 168 natural populations of *A. thaliana* [58].

Experimental design and growth conditions

A field experiment of 15,552 plants was set up at the INRAE center of Auzeville-Tolosane using a split-plot design arranged as a randomized complete block design (RCBD) with 16 treatments nested within six experimental blocks (Figure S1). The 16 treatments correspond to two mock treatments and the individual inoculation of 14 bacterial strains, namely OTU2_*Pfu*_1, OTU2_*Pfu*_2, OTU3a_*Oxa*_1, OTU3a_*Oxa*_2, OTU4_*Com*_1, OTU5_*Pmo*_1, OTU5_*Pmo*_2, OTU6_*Psi*_1, OTU6_*Psi*_2, OTU13_*M*sp_1, OTU13_*M*sp_2, OTU29_*Sph*_1, OTU29_*Sph*_2 and OTU8_JACO-CL. Each block was represented by 48 trays of 54 individual bottom-pierced wells (Ø4.7 cm, vol. ~70 cm²) (SOPARCO, reference 4920) filled with PROVEEN® Semi-Bouturage 2. In each block, each treatment corresponded to three trays stuck to each other and containing 162 plants, with one replicate per accession (54 populations * 3 accessions). Randomization of accessions was kept identical among treatments within a block, but differed among the six blocks. Randomization of the 16 treatments differed between the six blocks, with the exception of the two mock treatments that were kept at the same position (Supplementary Figure S1).

All seeds were sown on March 18th 2021, with several seeds sown in each well. Two weeks after sowing, seedlings were thinned to one per well, keeping the seedling the closest to the center of the well. During the entire growing period, the plants were watered as needed, *i.e.* manual watering morning and evening on hot and dry days and no watering on rainy days. A molluscicide (Algoflash® Naturasol) was regularly applied around the trays.

Inoculation procedure

Bacterial strains were grown on solid medium in Petri dishes (TSA for OTUs 5, 6 and 8, TSB for OTU2, R2A for OTUs 3, 4 and 29, R2A for OTU13). The day of inoculation, bacterial colonies were resuspended in sterile deionized water and bacterial solutions were diluted to

reach an $OD_{600\ nm}$ of 0.1. To facilitate the penetration of bacteria cells into plant organs, the Tween® 20 surfactant was added to each bacterial solution at a final concentration of 0.01%. Inoculation was performed 27 days after sowing (April 14, 2021), when most plants reached a 5-6 leaf stage. Using a Multipette® with a Combitips advanced® 50 mL, a volume of 1 mL of inoculum was dispensed on each rosette. A volume of 1mL of sterile water with a Tween® concentration of 0.01% was dispensed on each rosette of the plants of the two mock treatments. In order to increase relative humidity, plants were watered with a water mist spray system the seven days following the inoculation.

Phenotyping

Following [59], a non-destructive imaging approach (Supplementary Figure S2) was used to measure each plant for nine traits related to vegetative growth (Supplementary Data Set 1): projected rosette surface area measured at 1day before inoculation (dbi) (**area-1dbi**), 5 days after inoculation (dai) (**area-5dai**) and 9 dai (**area-9dai**); rosette perimeter measured at 1 dbi (**perimeter-1dbi**), 5 dai (**perimeter-5dai**) and 9 dai (**perimeter-9dai**); maximal rosette diameter measured at 1 dbi (**diameter-1dbi**), 5 dai (**diameter-5dai**) and 9 dai (**diameter-9dai**). To estimate plant growth relative to size, three relative growth rates (RGR) were estimated based on the rosette surface area: RGR between 5 dai and 1 dbi (**RGR-5dai-1dbi**), RGR between 9 dai and 5 dai (**RGR-9dai-5dai**) and RGR between 9 dai and 1 dbi (**RGR-9dai-1dbi**). The procedure and methodologies are detailed in Supplementary Text.

Data analyses

For the purpose of this study, the strain JACO-CL (OTU8, *P. viridiflava*) was not considered in any data analysis.

Investigation of the extent of natural genetic variation

To test the homogeneity of plant growth across the field trial and the presence of genetic variation for the three vegetative growth related traits measured before inoculation, data from the two mock treatments were pooled and the following mixed model (PROC MIXED procedure in SAS v. 9.4, SAS Institute Inc., Cary, NC, USA) was then used:

$$Y_{ijklmn} = \mu_{trait} + Block_i + Treatment_j + Block_i * Treatment_j + Population_k + Population_k * Treatment_j + Line_l(Tray_n) + Column_m(Tray_n) + \varepsilon_{ijklmn} \text{ (Model 1)}$$

where Y is one of the three phenotypic traits measured before inoculation (*i.e.* area-1dbi, perimeter-1dbi and diameter-1dbi), μ is the overall mean of the phenotypic data, 'Block' accounts for differences in micro-environmental conditions among blocks, 'Line(Tray)' and 'Column(Tray)' accounts for difference in micro-environmental conditions within 54-well trays, 'Treatment' tests for difference among the 14 treatments (*i.e.* mock treatment and 13 treatments with non-pathogenic bacterial strains), 'Population' corresponds to the genetic differences among the 54 populations, 'Population*Treatment' tests whether the rank among the 54 populations differs among the 14 treatments, and 'ε' is the residual term.

While the terms 'Treatment' and 'Population*Treatment' were not significant, we detected a highly significant 'Population' effect (Supplementary Table S2), thereby indicating that the level of significant genetic variation observed among the 54 populations was homogeneous across the field trial before inoculation.

To estimate the natural genetic variation of the response of the 162 accessions nested within 54 populations to the 13 non-pathogenic bacterial strains, the following mixed model (PROC MIXED procedure in SAS v. 9.4, SAS Institute Inc., Cary, NC, USA) was used for each of the 15 treatments:

$$Y_{ijklm} = \mu_{trait} + Block_i + Population_j + Accession_k(Population_j) + Line_l(Tray_i) + Column_m(Tray_i) + \varepsilon_{ijklm} \text{ (Model 2)}$$

where Y corresponds to one of the nine traits (area-5dai, area-9dai, perimeter-5dai, perimeter-9dai, diameter-5dai, diameter-9dai, RGR-5dai-1dbi, RGR-9dai-5dai and RGR-9dai-1dbi). All the terms are identical to the ones described in Model (1), with the exception of 'Accession' that accounts for mean genetic differences among the three accession within populations.

For each of the 126 'phenotypic trait * treatment' combinations (*i.e.* nine traits *14 treatments), genotypic values of the 54 populations were estimated by calculating least-squares (LS) mean values of the term 'Population' by the following linear model (PROC MIXED procedure in SAS v. 9.4, SAS Institute Inc., Cary, NC, USA) was used:

$$Y_{ijkl} = \mu_{trait} + Block_i + Population_j + Line_k(Tray_l) + Column_l(Tray_l) + \varepsilon_{ijkl} \text{ (Model 3)}$$

For each of the nine phenotypic traits, the estimated genotypic values (Supplementary Data Set 2) were then used to (i) compare phenotypic variation among the 14 treatments, (ii) estimate the level of 'Population*Treatment' interactions by calculating pairwise non-linear correlation coefficients (Spearman's *rho*) among the 14 treatments, and (iii) run GWA analyses (see below).

To estimate broad-sense heritability values (H^2) for each of the 126 'phenotypic trait * treatment' combinations, the following linear model (PROC MIXED procedure in SAS v. 9.4, SAS Institute Inc., Cary, NC, USA) was used:

$$Y_{ijkl} = \mu_{trait} + Block_i + Accession_j + Line_k(Tray_l) + Column_l(Tray_l) + \varepsilon_{ijkl} \text{ (Model 4)}$$

After considering the effects of the terms 'Line(Tray)' and 'Column(Tray)', the percentage of phenotypic variance explained by each other term of Model 4 was estimated by the PROC VARCOMP procedure (REML method, SAS v. 9.4, SAS Institute Inc., Cary, NC, USA). Following [33], H^2 values were estimated using the following formula:

$$H^2_{Trait} = \frac{VF}{VF + \frac{VR}{N}}$$

where 'VF' corresponds to the genetic variance among the 162 accessions, "VR" is the residual variance, and '*N*' is the mean number of biological replicates per accession (N = 6 in this study).

196 In Models 1, 2, 3 and 4, all factors were treated as fixed effects. For calculating *F*-values, terms
197 were tested over their appropriate denominators. A correction for the number of tests was
198 performed to control the False Discover Rate (FDR) at a nominal level of 5%.
199
200 *Combining GWA mapping with a local score approach (GW-LS)*
201 Based on a Pool-Seq approach, a representative picture of within-population genetic variation
202 was previously obtained for 168 natural populations of *A. thaliana* located southwest of France
203 [58], leading to the estimation of standardized allele frequencies corrected for the effect of
204 population structure within each population for 1,638,649 SNPs across the genome [57, 58].
205 For the purpose of this study, standardized population allele frequencies were retrieved for the
206 54 populations. Then, for each of the 126 'phenotypic trait * treatment' combinations, a genome
207 scan was first launched by estimating for each SNP Spearman's *rho* and associated *p* values
208 between standardized allele frequencies and population genotypic values. Thereafter, to
209 increase (i) the resolution in fine mapping genomic regions associated with genetic variation in
210 response to bacterial strains, and (ii) the identification of QTLs with small effects, we followed
211 [28, 30, 55, 56] by implementing a local score approach (with tuning parameter $\xi = 2$) on these
212 *p* values. Finally, significant SNP-phenotype associations were identified by estimating a
213 chromosome-wide significance threshold for each chromosome [55].
214
215 *Enrichment in biological processes*
216 A custom script written under the *R* environment [56] was used to retrieve the candidate genes
217 underlying detected QTLs for each of the 126 'phenotypic trait * treatment' combinations. For
218 each of the 14 treatments, we merged the lists of candidate genes of the nine phenotypic traits
219 and removed duplicates. For each of the 13 treatments with a non-pathogenic bacterial strain,
220 only candidate genes not found in the mock treatment were kept. To identify biological

pathways significantly over-represented ($P < 0.01$), each of the 14 resulting lists of unique candidate genes were submitted to the classification SuperViewer tool on the university of Toronto website (http://bar.utoronto.ca/ntools/cgibin/ntools_classification_superviewer.cgi) using the MAPMAN classification.

RESULTS

Genetic variation of *Arabidopsis thaliana* in response to non-pathogenic bacterial strains in field conditions

In agreement with previous experiments conducted in *in vitro* conditions [53], no disease symptoms were observed in our field conditions. For each of the 14 treatments (mock treatment and 13 treatments with a non-pathogenic bacterial strain), highly significant genetic variation was detected both between the 54 populations (Figure 1, Supplementary Figure S3, Supplementary Table S3) and between the 162 accessions (Supplementary Table S4) for each of the nine phenotypic traits, with the exception of (i) the rosette perimeter at 9 dai in presence of OTU3a_*Oxa*_1 at the population level (Supplementary Table S3), and (ii) the rosette area at 9 dai in presence of OTU5_*Pmo*_1 and OTU6_*Psi*_1 at the accession level (Supplementary Table S4). Across the 126 'phenotypic trait * treatment' combinations, the mean broad-sense heritability (H^2) estimate was 0.49 (median = 0.56, quantile 5% = 0.16, quantile 95% = 0.70). These results indicate that a non-negligible fraction of phenotypic variance was explained by genetic variation among populations and accessions (Supplementary Table S4).

A significant variation was observed among the 14 treatments for each phenotypic trait (Figure 1, Supplementary Figure S3). However, significant differences between the response to any bacterial strain and the mock treatment were only observed for three traits (*i.e.* area-9dai, diameter-9dai and RGR-5dai-1dbi) (Figure 1, Supplementary Figure S3). For instance, the rosette area at 9 dai was on average bigger and smaller in response to

246 OTU2_*Pfu*_1/OTU3a_*Oxa*_2 and OTU29_*Sph*_2 than in the mock treatment, respectively
247 (Figure 1a). The relative growth rate between 5 dai and 1 dbi was significantly higher in
248 response to OTU5_*Pmo*_2 than in seven treatments, including the mock treatment (Figure 1b).
249 More importantly, for each phenotypic trait, we observed a strong genetic variation among the
250 54 populations in their response to each of the 13 non-pathogenic bacterial strains (Figure 2,
251 Supplementary Figure 4). Indeed, values of genetic correlations between the mock treatment
252 and each treatment with a bacterial strain were largely deviating from 1, in particular at 9 dai
253 (Figure 2a), with the exception of relative growth rate for with lower values of genetic
254 correlations were observed within 5 dai than within 9 dai (Figure 2b, Supplementary Figure
255 S4). In addition, the response of the 54 populations greatly varied among the 13 bacterial strains,
256 even between two strains belonging to the same bacterial species (Figure 2, Supplementary
257 Figure 4). For instance, while most populations present either a positive, neutral or negative
258 response to either OTU13_*M*sp strain (*i.e*. presence of crossing reaction norms), the direction
259 and/or the strength of response of given population can largely differ between the two
260 OTU13_*M*sp strains, as illustrated by the populations FERR-A and LUZE-B (Figure 3).
261 Altogether, the presence of (i) genetic variation at the population and accession levels for most
262 'phenotypic trait * treatment' combinations, (ii) crossing reaction norms between the mock
263 treatment and each treatment with a bacterial strain, and (ii) crossing reaction norms among the
264 13 treatments with a bacterial strain, suggests a genetic architecture that largely differs among
265 the 14 treatments, whatever the phenotypic trait considered.
266
267 **A genomic map of local adaptation to prevalent and/or abundant leaf bacterial species**
268 Based on the allele frequencies of 1,638,649 SNPs obtained by a Pool-Seq approach for each
269 of the 54 populations (Frachon et al., 2018), a GWA mapping analysis combining a Bayesian
270 hierarchical model with a local score approach (BMH-LS) was conducted to characterize the

genetic architecture of response to the 13 non-pathogenic bacterial strains. Across the 126 'phenotypic trait * treatment' combinations, we detected 2,064 QTLs with a mean length of QTL interval equal to ~837bp (quantile 5% ~ 38bp, quantile 95% = 3.12kb) (Supplementary Data Set 3). The number of QTLs per 'phenotypic trait * treatment' combination ranged from 6 to 34 (mean =16.4), suggesting a polygenic architecture for the response to members of the most prevalent and/or abundant non-pathogenic bacterial species of the leaf compartment of *A. thaliana* located south-west of France (Figure 4a).

In agreement with the level of genetic correlations observed among the 14 treatments (mock treatment and 13 treatments with a non-pathogenic bacterial strain) and the presence of crossing reaction norms (Figures 2 and3, Supplementary Figures S3 and S4), the genetic architecture was highly flexible between the mock treatment and treatments with bacterial strains, as well as among treatments with bacterial strains at the interspecific and intraspecific levels, as illustrated for the rosette area at 9 dai (Figure 5). For instance, most candidate genes underlying detected QTLs and not shared with the mock treatment were specific to a given treatment with a bacterial strain (Figure 4b, Supplementary Figure 5, Supplementary Data Set 4), in particular at 9 dai (Figure 4b, Supplementary Figure 5). For instance, for the maximal rosette diameter, while the percentage of candidate genes specific to a given treatment with a bacterial strain ranged from 57.7% to 86.1% (mean = 75.2%) at 9 dai, it ranged from 26.9% to 81.1% (mean = 46.0%) at 5 dai (Supplementary Figure 5).

Identification of enriched biological processes and candidate genes associated with the response to prevalent and/or abundant leaf bacterial species

The first approach to identify relevant candidate genes involved in the response to the 13 non-pathogenic bacterial species was to focus on candidate genes underlying the most pleiotropic QTLs. Here, we focused on QTLs detected for the response in more than six bacterial strains,

but not detected for the mock treatment. We identified seven such pleiotropic QTLs encompassing 17 candidate genes (Table 1, Supplementary Data Set 5). In agreement with the highly flexible genetic architecture observed between strains within a bacterial species (Figure 4b, Supplementary Figure 5), the high level of pleiotropy observed for these QTLs was more dependent on the identity of the bacterial strains than the identity of the bacterial species (Table 1). Among the 17 candidate genes, eight genes have functions in relation with plant development and organ growth, *i.e. At2g40650* [60], *At2g40670* [61, 62], *At2g44710* [63], *At2g47190* [64–66], *At4g14713* [67–70], *At4g14716* [71], *At4g14720* [72] and *At5g42360* [73]. Interestingly, three genes have a link with plant immunity, *i.e.* the genes *MEMB12* (*At5g50440*) that is silenced by a microRNA during *Pseudomonas syringae* bacterial infection [74], *ARR16* (*At2g40670*) that is repressed by *Botrytis cinerea* fungal infection [75], and *TIFY4B/ PEAPOD 2* (*At4g14720*) that interacts with the begomovirus AL2 transcriptional activator protein, an inhibitor of plant basal defense [76].

Based on the lists of unique candidate genes identified for each treatment and the list of unique candidate genes identified across the 13 treatments with a bacterial strain, the second approach was to identify biological processes significantly over-represented in frequency compared to the overall class frequency in the *A. thaliana* MapMan annotation. When considering both the 14 treatments individually and the 13 treatments with a bacterial strain altogether, we identified 19 significantly enriched classes, among which five were also enriched in the mock treatment, i.e. ‘development’, ‘hormone metabolism’, ‘lipid metabolism’, ‘protein’ and ‘RNA’ (Figure 6a, Supplementary Data Set 5). Amongst the 14 over-represented classes not detected in the mock treatment, most of them were highly dependent on the identity of the bacterial strain, suggesting the involvement of diverse pathways in response to representative members of the non-pathogenic microbiota down to the intraspecific level (Figure 6a). We nonetheless identified four classes that were significantly over-represented for at least three treatments with a bacterial

strain and when considering the 13 treatments with a bacterial strain altogether, *i.e*. 'cell', 'secondary metabolism', 'signalling' and 'transport' (Figure 6a). Interestingly, amongst the 99 'signalling' genes, we identified (i) 54 kinase-related genes including 24 leucine-rich repeat (LRR) kinases, 8 cysteine-rich receptor-like kinases (CRK) and 6 MAP kinases, and (ii) 23 genes associated with calcium signalling, in particular for the two strains of *P. fungorum* (OTU2), the two strains of *Oxalobacteraceae* bacterium (OTU3) and one strain of *P. siliginis* (OTU6) (Figure 6b).

DISCUSSION

Extensive genetic variation within a local set of *A. thaliana* accessions in response to non-pathogenic leaf bacteria at the species and strain levels

Extensive genetic variation was previously observed in two worldwide collections of *A. thaliana*, each challenged at the root level in *in vitro* conditions with a single PGPB strain isolated on another plant species than *A. thaliana*, *i.e*. the strain *Pseudomonas simiae* WCS417r isolated from the rhizosphere of wheat [47] and the strain *Bacillus pumilus* TUAT-1 isolated from rice roots [49]. In this study, in line with the need to bring evolutionary and ecological functional genomics from the lab to the wild [23, 77–79], the ecological realism of plant-microbiota interactions was increased by phenotyping in field conditions, the rosette growth response of *A. thaliana* accessions collected south-west of France to non-pathogenic bacterial strains isolated from the leaf compartment of *A. thaliana* in the same geographical region.

In agreement with previous observations with bacterial pathogens [15, 31, 80], the extent of genetic variation of response to non-pathogenic bacterial strains was more dependent on the identity of the bacterial strain than the identity of the bacterial species. In addition, the presence of crossing reaction norms indicates that declaring a strain as a PGPB is highly dependent on the host genotype tested. Whether the genotype-dependent plant-growth promoting effect of a

346 particular strain on aboveground vegetative growth is also observed at the below-ground level
347 would deserve investigation, for instance by estimating root growth and root/shoot biomass
348 ratios [81, 82].
349 Interestingly, while genetic variation in response to bacterial strains was observed within few
350 days after inoculation in field conditions, such a genetic variation was mainly observed after
351 several weeks in *in vitro* conditions [53]. Since the bacterial strains used in this study have been
352 isolated from complex microbiota they used to interact and/or coevolve with in the native
353 habitats of *A. thaliana*. Hence, the effect of bacterial strains may require the presence of
354 additional microbiota members in the plant, a prerequisite not achieved in germ-free plants in
355 *in vitro* conditions [53].
356
357 **A complex and highly flexible genetic architecture underlies adaptive plant-microbiota**
358 **interactions**
359 So far, the five GWAS [83–86] and the single genome-environment association study (GEAS)
360 (Roux et al. 2022) conducted on the leaf compartment and using bacterial community
361 descriptors as phenotypic traits, revealed a polygenic architecture controlling microbiota
362 assembly, which is in line with the small percentage of variance explained by the phenotyping
363 of individual mutant lines [87]. In agreement with those association genetic studies and the two
364 GWAS conducted on *A. thaliana* in response to a PGPB strain [47, 49], we identified a complex
365 genetic architecture for the response of *A. thaliana* to 13 non-pathogenic bacterial strains. In
366 addition, this polygenic architecture was highly flexible among the 13 bacterial strains, with the
367 detection of a few number of highly pleiotropic QTLs. Similar results were observed in recent
368 GWAS conducted both in crops and wild species in response to experimental inoculation with
369 individual pathogenic bacterial strains [88]. For instance, challenging 130 natural accessions of
370 *A. thaliana* with 22 strains of the bacterial pathogen *Xanthomonas arboricola* revealed a clear

host-strain specificity in quantitative disease resistance [80]. The complex genetic interactions observed between *A. thaliana* and the main members of its leaf microbiota should maintain high levels of diversity at the candidate genes, which in turn should result in complex co-evolutionary dynamics [16].

Beyond the question of the effects of strong genotype-by-genotype (GxG) interactions on the nature and strength of footprints of natural selection on the genome of *A. thaliana*, whether the genetic architecture underlying the response of *A. thaliana* to co-inoculation corresponds to the sum of QTLs that are specific to the response to mono-inoculations and/or to the emergence of new QTLs, remains on an open question in the research area of plant-microbe interactions. Experimental studies on plant-plant interactions demonstrated that the genetic architecture of the response of *A. thaliana* in a plurispecific neighborhood was not predictable from the genetic architecture of the response of *A. thaliana* in the corresponding bispecific neighborhoods [56, 57]

Plant innate immunity is a significant source of natural genetic variation in plant-microbiota interactions

The candidate genes underlying the most pleiotropic QTLs have functions mainly related to plant development and/or stresses (biotic or abiotic stresses). Nevertheless, a more global approach identified four biological classes that were significantly and specifically over-represented for at least three bacterial strains but not with the mock, *i.e.* 'cell', 'secondary metabolism', 'signalling' and 'transport'. These four classes were also over-represented in a GEAS performed on 163 natural populations of *A. thaliana* located south-west of France (including the 54 populations considered in this study) (Roux et al. 2022) and characterized *in situ* for bacterial communities in the leaf and root compartments using a metabarcoding

395 approach [38], thereby strengthening the importance of these four classes in mediating host
396 response to the 13 bacterial strains tested here.
397 Strikingly, we found a clear enrichment for signalling genes underlying QTLs in response to
398 the 13 bacterial strains tested in this study. Signalling genes have been extensively described as
399 being involved in plant-microbe interactions. Of particular note, we identified 8 genes
400 belonging to the CRK family, which represents one of the largest group of RLKs with 44
401 members in *A. thaliana* [89, 90]. Some CRKs are involved in the regulation of plant
402 developmental processes, while others are involved in stress and pathogen response [89].
403 Interestingly, by assessing host transcriptional and metabolic adaptations to 39 bacterial strains
404 in the leaf compartment of *A. thaliana*, a core set of 24 genes consistently induced by the
405 presence of most strains was identified and thereby referred as a molecular process called
406 general non-self-response (GNSR) [91]. Importantly, one gene of this core set (*CRK6*) was also
407 identified as a candidate genes in our GWAs, reinforcing the importance of CRKs in plant-
408 microbiota interactions.
409 Another interesting result is that while few classical *R* genes involved in specific recognition of
410 microbial effectors have been identified in this study, we highlighted many candidate genes
411 related to pattern-triggered immunity (PTI), including receptor-like kinases (RLKs) and
412 receptor-like proteins (RLPs). PTI relies on the perception of specific molecular patterns such
413 as microbe- or pathogen-associated molecular patterns (MAMPs/PAMPs), or self-molecules
414 (damage-associated molecular patterns, DAMPs) [92]. In particular, we identified a main actor
415 of PTI as a candidate gene, the *FLS2* gene in response to the two strains of OTU6 and one strain
416 of OUT 13 (Supplementary Data Set 6), the best-characterized pattern-recognition receptor
417 (PRR) gene, encoding an LRR-RLK protein that acts as a receptor for flg22 bacterial PAMP
418 [92, 93]. Moreover, it was previously shown that CRK6 and CRK36 are part of the PRR FLS2
419 protein complex, modulating PTI response through an association with FLS2 [94]. Two recent

works dissected the interplay between FLS2 and numerous flg22 variants, studying how *A. thaliana* association with different evolved flg22 variants from bacterial microbiota differentially fine-tune the balance between bacterial motility and defense activation [95, 96]. PTI response is also characterized by the production of reactive oxygen species (ROS) and by the activation of the mitogen-activated protein kinases (MAPKs) cascade [97]. In our study, we identified four NADPH oxidase *RBOH* genes, among them *RBOHD* that is required for microbiota homeostasis in leaves [98]. Another candidate gene is *MPK4*, a main actor of PTI signalling (Bazin et al., 2020). Even if few mutant lines related to signalling and PTI have been tested for their effect on microbiota assembly [87, 98, 99], our results strengthen the need for a deeper investigation of some of our most promising candidate genes in relationship with the 13 strains used in this study. Importantly, the *de-novo* whole-genome sequence of the 13 strains tested in this study have been recently obtained with long-read sequencing technology [53]. Comparative genomics, and notably for their PAMP sequences (*i.e.* flagelline, EF-TU), may bring very informative data on their potential recognition by the plant, thereby making a link between plant-microbiota recognition and plant innate immunity. This is directly in line with a recent study that nicely shows how root commensal bacteria modulate host susceptibility to pathogens by either eliciting or dampening PTI responses [100].

ACKNOWLEDGMENTS

We are grateful to the members of the ECOGEN tem for their assistance during the sowing and thinning. D.R.S. was funded by a PhD fellowship from CONACYT. R.D. was funded by a grant from the French Ministry of National Education and Research. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 951444 –PATHOCOM). This study was performed at the LIPME belonging to the Laboratoire d'Excellence (LABEX) entitled TULIP (ANR-10-LABX-41).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Raw phenotypic data are available in Supplementary Data Set 1.

REFERENCES

1. Stukenbrock EH, McDonald BA. The origins of plant pathogens in agro-ecosystems. *Annu Rev Phytopathol* 2008; **46**: 75–100.
2. Alexander HM. Disease in natural plant populations, communities, and ecosystems: Insights into ecological and evolutionary processes. *Plant Dis* 2010; **94**: 492–503.
3. Burdon JJ, Zhan J. Climate change and disease in plant communities. *PLoS Biol* 2020; **18**: e3000949.
4. Jeger MJ. The impact of climate change on disease in wild plant populations and communities. *Plant Pathol* 2022; **71**: 111–130.
5. Desaint H, Aoun N, Deslandes L, Vailleau F, Roux F, Berthomé R. Fight hard or die trying: when plants face pathogens under heat stress. *New Phytol* 2021; **229**: 712–734.
6. Jarosz AM, Davelos AL. Effects of disease in wild plant populations and the evolution of pathogen aggressiveness. *New Phytol* 1995; **129**: 371–387.
7. Burdon JJ. Fungal pathogens as selective forces in plant populations and communities. *Austral Ecol* 1991; **16**: 423–432.
8. Zhan J, Ericson L, González-Jiménez J, Burdon JJ. Disease influences host population growth rates in a natural wild plant–pathogen association over a 30-year period. *J Ecol* 2022; **110**: 173–184.
9. Oerke E. Crop losses to pests. *J Agric Sci* 2006; **144**: 31–43.
10. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. The global burden of pathogens and pests on major food crops. *Nat Ecol Evol* 2019; **3**: 430–439.
11. McDonald BA, Stukenbrock EH. Rapid emergence of pathogens in agro-ecosystems: global threats to agricultural sustainability and food security. *Philos Trans R Soc Lond B Biol Sci* 2016; **371**.

12. Brown JKM. Yield penalties of disease resistance in crops. *Curr Opin Plant Biol* 2002; **5**: 339–344.

13. Ristaino JB, Anderson PK, Bebber DP, Brauman KA, Cunniffe NJ, Fedoroff NV, et al. The persistent threat of emerging plant disease pandemics to global food security. *Proc Natl Acad Sci U S A* 2021; **118**: e2022239118.

14. Roux F, Bergelson J. The genetics underlying natural variation in the biotic interactions of *Arabidopsis thaliana*: the challenges of linking evolutionary genetics and community ecology. *Curr Top Dev Biol* 2016; **119**: 111–156.

15. Bartoli C, Roux F. Genome-Wide Association Studies in plant pathosystems: toward an ecological genomics approach. *Front Plant Sci* 2017; **8**: 763.

16. Karasov TL, Kniskern JM, Gao L, DeYoung BJ, Ding J, Dubiella U, et al. The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature* 2014; **512**: 436–440.

17. Roux F, Voisin D, Badet T, Balagué C, Barlet X, Huard-Chauveau C, et al. Resistance to phytopathogens *e tutti quanti*: placing plant quantitative disease resistance on the map. *Mol Plant Pathol* 2014; **15**: 427–432.

18. Tack AJM, Laine A-L. Spatial eco-evolutionary feedback in plant-pathogen interactions. *Eur J Plant Pathol* 2014; **138**: 667–677.

19. Burdon JJ, Thrall PH, Ericson AL. The current and future dynamics of disease in plant communities. *Annu Rev Phytopathol* 2006; **44**: 19–39.

20. Li W, Deng Y, Ning Y, He Z, Wang G-L. Exploiting broad-spectrum disease resistance in crops: From molecular dissection to breeding. *Annu Rev Plant Biol* 2020; **71**: 575–603.

21. Deng Y, Ning Y, Yang D-L, Zhai K, Wang G-L, He Z. Molecular basis of disease resistance and perspectives on breeding strategies for resistance improvement in crops. *Mol Plant* 2020; **13**: 1402–1419.

500 22. Karasov TL, Shirsekar G, Schwab R, Weigel D. What natural variation can teach us about
501 resistance durability. *Curr Opin Plant Biol* 2020; **56**: 89–98.
502 23. Bergelson J, Roux F. Towards identifying genes underlying ecologically relevant traits in
503 *Arabidopsis thaliana*. *Nat Rev Genet* 2010; **11**: 867–879.
504 24. Monnot S, Desaint H, Mary-Huard T, Moreau L, Schurdi-Levraud V, Boissot N.
505 Deciphering the genetic architecture of plant virus resistance by GWAS, state of the art
506 and potential advances. *Cells* 2021; **10**: 3080.
507 25. Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, et al. Genome-wide
508 association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 2010;
509 **465**: 627–631.
510 26. Xiao Y, Liu H, Wu L, Warburton M, Yan J. Genome-wide association studies in maize:
511 Praise and stargaze. *Mol Plant* 2017; **10**: 359–374.
512 27. Gupta PK, Kulwal PL, Jaiswal V. Association mapping in plants in the post-GWAS
513 genomics era. *Adv Genet* 2019; **104**: 75–154.
514 28. Aoun N, Desaint H, Boyrie L, Bonhomme M, Deslandes L, Berthomé R, et al. A complex
515 network of additive and epistatic quantitative trait loci underlies natural variation of
516 *Arabidopsis thaliana* quantitative disease resistance to *Ralstonia solanacearum* under heat
517 stress. *Mol Plant Pathol* 2020; **21**: 1405–1420.
518 29. Aoun N, Tauleigne L, Lonjon F, Deslandes L, Vailleau F, Roux F, et al. Quantitative
519 disease resistance under elevated temperature: genetic basis of new resistance mechanisms
520 to *Ralstonia solanacearum*. *Front Plant Sci* 2017; **8**: 1387.
521 30. Demirjian C, Razavi N, Desaint H, Lonjon F, Genin S, Roux F, et al. Study of natural
522 diversity in response to a key pathogenicity regulator of *Ralstonia solanacearum* reveals
523 new susceptibility genes in *Arabidopsis thaliana*. *Mol Plant Pathol* 2022; **23**: 321–338.

524 31. Debieu M, Huard-Chauveau C, Genissel A, Roux F, Roby D. Quantitative disease
525 resistance to the bacterial pathogen *Xanthomonas campestris* involves an Arabidopsis
526 immune receptor pair and a gene of unknown function. *Mol Plant Pathol* 2016; **17**: 510–
527 520.
528 32. Badet T, Voisin D, Mbengue M, Barascud M, Sucher J, Sadon P, et al. Parallel evolution
529 of the POQR prolyl oligo peptidase gene conferring plant quantitative disease resistance.
530 *PLoS Genet* 2017; **13**: e1007143.
531 33. Huard-Chauveau C, Perchepied L, Debieu M, Rivas S, Kroj T, Kars I, et al. An atypical
532 kinase under balancing selection confers broad-spectrum disease resistance in
533 Arabidopsis. *PLoS Genet* 2013; **9**: e1003766.
534 34. Delplace F, Huard-Chauveau C, Dubiella U, Khafif M, Alvarez E, Langin G, et al.
535 Robustness of plant quantitative disease resistance is provided by a decentralized immune
536 network. *Proc Natl Acad Sci U S A* 2020; **117**: 18099–18109.
537 35. Delplace F, Huard-Chauveau C, Berthomé R, Roby D. Network organization of the plant
538 immune system: from pathogen perception to robust defense induction. *Plant J* 2021.
539 36. Müller DB, Vogel C, Bai Y, Vorholt JA. The plant microbiota: systems-level insights and
540 perspectives. *Annual Review of Genetics* . 2016. , **50**: 211–234
541 37. Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant
542 health. *Trends Plant Sci* 2012; **17**: 478–486.
543 38. Bartoli C, Frachon L, Barret M, Rigal M, Huard-Chauveau C, Mayjonade B, et al. In situ
544 relationships between microbiota and potential pathobiota in *Arabidopsis thaliana*. *ISME*
545 *J* 2018; **12**: 2024–2038.
546 39. Bulgarelli D, Schlaeppi K, Spaepen S, Ver Loren van Themaat E, Schulze-Lefert P.
547 Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* 2013;
548 **64**: 807–838.

549 40. Pieterse CMJ, Zamioudis C, Berendsen RL, Weller DM, Van Wees SCM, Bakker PAHM.
550 Induced systemic resistance by beneficial microbes. *Annu Rev Phytopathol* 2014; **52**: 347–
551 375.
552 41. Escudero-Martinez C, Bulgarelli D. Tracing the evolutionary routes of plant-microbiota
553 interactions. *Curr Opin Microbiol* 2019; **49**: 34–40.
554 42. Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. Plant-microbiome interactions: from
555 community assembly to plant health. *Nat Rev Microbiol* 2020; **18**: 607–621.
556 43. Glick BR, Gamalero E. Recent developments in the study of plant microbiomes.
557 *Microorganisms* 2021; **9**: 1533.
558 44. Bai B, Liu W, Qiu X, Zhang J, Zhang J, Bai Y. The root microbiome: Community
559 assembly and its contributions to plant fitness. *J Integr Plant Biol* 2022; **64**: 230–243.
560 45. Kroll S, Agler MT, Kemen E. Genomic dissection of host–microbe and microbe–microbe
561 interactions for advanced plant breeding. *Curr Opin Plant Biol* 2017; **36**: 71–78.
562 46. Naitam M, Kaushik R, Sharma A. Crop microbiome engineering and relevance in abiotic
563 stress tolerance. *Soil Biology*. 2021. Springer International Publishing, Cham, pp 253–277.
564 47. Wintermans PCA, Bakker PAHM, Pieterse CMJ. Natural genetic variation in Arabidopsis
565 for responsiveness to plant growth-promoting rhizobacteria. *Plant Mol Biol* 2016; **90**:
566 623–634.
567 48. Vidotti MS, Lyra DH, Morosini JS, Granato ÍSC, Quecine MC, Azevedo JL de, et al.
568 Additive and heterozygous (dis)advantage GWAS models reveal candidate genes involved
569 in the genotypic variation of maize hybrids to *Azospirillum brasilense*. *PLoS One* 2019;
570 **14**: e0222788.
571 49. Cotta MS, do Amaral FP, Cruz LM, Wassem R, de Oliveira Pedrosa F, Yokoyama T, et
572 al. Genome-Wide Association Studies reveal important candidate genes for the *Bacillus*
573 *pumilus* TUAT-1-*Arabidopsis thaliana* interaction. *bioRxiv* . 2020. bioRxiv.

574 50. Stanton-Geddes J, Paape T, Epstein B, Briskine R, Yoder J, Mudge J, et al. Candidate
575 genes and genetic architecture of symbiotic and agronomic traits revealed by whole-
576 genome, sequence-based association genetics in *Medicago truncatula*. *PLoS One* 2013; **8**:
577 e65688.
578 51. Torkamaneh D, Chalifour F-P, Beauchamp CJ, Agrama H, Boahen S, Maaroufi H, et al.
579 Genome-wide association analyses reveal the genetic basis of biomass accumulation under
580 symbiotic nitrogen fixation in African soybean. *Züchter Genet Breed Res* 2020; **133**: 665–
581 676.
582 52. Curtin SJ, Tiffin P, Guhlin J, Trujillo DI, Burghart LT, Atkins P, et al. Validating genome-
583 wide association candidates controlling quantitative variation in nodulation. *Plant Physiol*
584 2017; **173**: 921–931.
585 53. Ramirez-Sanchez D, Gibelin-Viala C, Mayjonade B, Duflos R, Belmonte E, Pailler V, et
586 al. Investigating genetic diversity within the most abundant and prevalent non-pathogenic
587 leaf-associated bacteria interacting with *Arabidopsis thaliana* in natural habitats. *bioRxiv*
588 . 2022.
589 54. Gautier M. Genome-Wide Scan for adaptive divergence and association with population-
590 specific covariates. *Genetics* 2015; **201**: 1555–1579.
591 55. Bonhomme M, Fariello MI, Navier H, Hajri A, Badis Y, Miteul H, et al. A local score
592 approach improves GWAS resolution and detects minor QTL: application to *Medicago*
593 *truncatula* quantitative disease resistance to multiple *Aphanomyces euteiches* isolates.
594 *Heredity* 2019; **123**: 517–531.
595 56. Libourel C, Baron E, Lenglet J, Amsellem L, Roby D, Roux F. The genomic architecture
596 of competitive response of *Arabidopsis thaliana* is highly flexible among plurispecific
597 neighborhoods. *Front Plant Sci* 2021; **12**: 741122.

598 57. Frachon L, Mayjonade B, Bartoli C, Hautekèete N-C, Roux F. Adaptation to plant
599 communities across the genome of *Arabidopsis thaliana*. *Mol Biol Evol* 2019; **36**: 1442–
600 1456.
601 58. Frachon L, Bartoli C, Carrère S, Bouchez O, Chaubet A, Gautier M, et al. A genomic map
602 of climate adaptation in *Arabidopsis thaliana* at a micro-geographic scale. *Front Plant Sci*
603 2018; **9**: 967.
604 59. Roux F, Mary-Huard T, Barillot E, Wenes E, Botran L, Durand S, et al. Cytonuclear
605 interactions affect adaptive traits of the annual plant *Arabidopsis thaliana* in the field. *Proc*
606 *Natl Acad Sci U S A* 2016; **113**: 3687–3692.
607 60. Watanabe E, Mano S, Nishimura M, Yamada K. *AtUBL5* regulates growth and
608 development through pre-mRNA splicing in *Arabidopsis thaliana*. *PLoS One* 2019; **14**:
609 e0224795.
610 61. Srivastava AK, Dutta S, Chattopadhyay S. MYC2 regulates ARR16, a component of
611 cytokinin signaling pathways, in Arabidopsis seedling development. *Plant Direct* 2019; **3**:
612 e00177.
613 62. Yang W, Choi M-H, Noh B, Noh Y-S. De Novo shoot regeneration controlled by HEN1
614 and TCP3/4 in Arabidopsis. *Plant Cell Physiol* 2020; **61**: 1600–1613.
615 63. Zhang Y, Fan S, Hua C, Teo ZWN, Kiang JX, Shen L, et al. Phase separation of HRLP
616 regulates flowering time in Arabidopsis. *Sci Adv* 2022; **8**: eabn5488.
617 64. Guo X, Li L, Liu X, Zhang C, Yao X, Xun Z, et al. MYB2 is important for tapetal PCD
618 and pollen development by directly activating protease expression in Arabidopsis. *Int J*
619 *Mol Sci* 2022; **23**: 3563.
620 65. Jia T, Zhang K, Li F, Huang Y, Fan M, Huang T. The AtMYB2 inhibits the formation of
621 axillary meristem in Arabidopsis by repressing RAX1 gene under environmental stresses.
622 *Plant Cell Rep* 2020; **39**: 1755–1765.

66. Nguyen HTK, Hyoung S, Kim HJ, Cho KM, Shin JS. The transcription factor γMYB2 acts as a negative regulator of secondary cell wall thickening in anther and stem. *Gene* 2019; **702**: 158–165.

67. White DWR. PEAPOD repressors modulate and coordinate developmental responses to light intensity in Arabidopsis. *New Phytol* 2022; **235**: 1470–1485.

68. González R, Butković A, Rivarez MPS, Elena SF. Natural variation in *Arabidopsis thaliana* rosette area unveils new genes involved in plant development. *Sci Rep* 2020; **10**: 17600.

69. White DWR. PEAPOD regulates lamina size and curvature in Arabidopsis. *Proc Natl Acad Sci U S A* 2006; **103**: 13238–13243.

70. Zhu Y, Luo X, Liu X, Wu W, Cui X, He Y, et al. Arabidopsis PEAPODs function with LIKE HETEROCHROMATIN PROTEIN1 to regulate lateral organ growth. *J Integr Plant Biol* 2020; **62**: 812–831.

71. Friedman EJ, Wang HX, Jiang K, Perovic I, Deshpande A, Pochapsky TC, et al. Acireductone dioxygenase 1 (ARD1) is an effector of the heterotrimeric G protein beta subunit in Arabidopsis. *J Biol Chem* 2011; **286**: 30107–30118.

72. Baekelandt A, Pauwels L, Wang Z, Li N, De Milde L, Natran A, et al. Arabidopsis leaf flatness is regulated by PPD2 and NINJA through repression of CYCLIN D3 genes. *Plant Physiol* 2018; **178**: 217–232.

73. Franciosini A, Lombardi B, Iafrate S, Pecce V, Mele G, Lupacchini L, et al. The Arabidopsis COP9 SIGNALOSOME INTERACTING F-BOX KELCH 1 protein forms an SCF ubiquitin ligase and regulates hypocotyl elongation. *Mol Plant* 2013; **6**: 1616–1629.

74. Chung KP, Zeng Y, Li Y, Ji C, Xia Y, Jiang L. Signal motifs-dependent ER export of Qc-SNARE BET12 interacts with MEMB12 and affects PR1 trafficking in Arabidopsis. *J Cell Sci* 2017; **131**: jcs202838.

75. Li B, Wang R, Wang S, Zhang J, Chang L. Diversified regulation of cytokinin levels and signaling during Botrytis cinerea infection in Arabidopsis. *Front Plant Sci* 2021; **12**: 584042.

76. Chung HY, Sunter G. Interaction between the transcription factor AtTIFY4B and begomovirus AL2 protein impacts pathogenicity. *Plant Mol Biol* 2014; **86**: 185–200.

77. Zaidem ML, Groen SC, Purugganan MD. Evolutionary and ecological functional genomics, from lab to the wild. *Plant J* 2019; **97**: 40–55.

78. Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, et al. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet* 2010; **6**: e1000940.

79. Pujol B, Blanchet S, Charmantier A, Danchin E, Facon B, Marrot P, et al. The Missing Response to Selection in the Wild. *Trends Ecol Evol* 2018; **33**: 337–346.

80. Wang M, Roux F, Bartoli C, Huard-Chauveau C, Meyer C, Lee H, et al. Two-way mixed-effects methods for joint association analysis using both host and pathogen genomes. *Proc Natl Acad Sci U S A* 2018; **115**: E5440–E5449.

81. Mantelin S, Touraine B. Plant growth-promoting bacteria and nitrate availability: impacts on root development and nitrate uptake. *J Exp Bot* 2004; **55**: 27–34.

82. Brock AK, Berger B, Mewis I, Ruppel S. Impact of the PGPB *Enterobacter radicincitans* DSM 16656 on growth, glucosinolate profile, and immune responses of *Arabidopsis thaliana*. *Microb Ecol* 2013; **65**: 661–670.

669 83. Horton MW, Bodenhausen N, Beilsmith K, Meng D, Muegge BD, Subramanian S, et al.
670 Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. *Nat*
671 *Commun* 2014; **5**: 5320.
672 84. Brachi B, Filiault D, Whitehurst H, Darme P, Le Gars P, Le Mentec M, et al. Plant genetic
673 effects on microbial hubs impact host fitness in repeated field trials. *Proc Natl Acad Sci U*
674 *S A* 2022; **119**: e2201285119.
675 85. Walters WA, Jin Z, Youngblut N, Wallace JG, Sutter J, Zhang W, et al. Large-scale
676 replicated field study of maize rhizosphere identifies heritable microbes. *Proc Natl Acad*
677 *Sci U S A* 2018; **115**: 7368–7373.
678 86. Roman-Reyna V, Pinili D, Borja FN, Quibod IL, Groen SC, Alexandrov N, et al.
679 Characterization of the leaf microbiome from whole-genome sequencing data of the 3000
680 rice genomes project. *Rice* 2020; **13**: 72.
681 87. Bergelson J, Brachi B, Roux F, Vailleau F. Assessing the potential to harness the
682 microbiome through plant genetics. *Curr Opin Biotechnol* 2021; **70**: 167–173.
683 88. Zhang F, Hu Z, Wu Z, Lu J, Shi Y, Xu J, et al. Reciprocal adaptation of rice and
684 *Xanthomonas oryzae* pv. *oryzae*: cross-species 2D GWAS reveals the underlying genetics.
685 *Plant Cell* 2021; **33**: 2538–2561.
686 89. Bourdais G, Burdiak P, Gauthier A, Nitsch L, Salojärvi J, Rayapuram C, et al. Large-scale
687 phenomics identifies primary and fine-tuning roles for CRKs in responses related to
688 oxidative stress. *PLoS Genet* 2015; **11**: e1005373.
689 90. Wrzaczek M, Brosché M, Salojärvi J, Kangasjärvi S, Idänheimo N, Mersmann S, et al.
690 Transcriptional regulation of the CRK/DUF26 group of receptor-like protein kinases by
691 ozone and plant hormones in Arabidopsis. *BMC Plant Biol* 2010; **10**: 95.
692 91. Maier BA, Kiefer P, Field CM, Hemmerle L, Bortfeld-Miller M, Emmenegger B, et al. A
693 general non-self response as part of plant immunity. *Nat Plants* 2021; **7**: 696–705.

694 92. Macho AP, Zipfel C. Plant PRRs and the activation of innate immune signaling. *Mol Cell*
695 2014; **54**: 263–272.
696 93. Gómez-Gómez L, Boller T. FLS2: an LRR receptor-like kinase involved in the perception
697 of the bacterial elicitor flagellin in Arabidopsis. *Mol Cell* 2000; **5**: 1003–1011.
698 94. Yeh Y-H, Chang Y-H, Huang P-Y, Huang J-B, Zimmerli L. Enhanced Arabidopsis
699 pattern-triggered immunity by overexpression of cysteine-rich receptor-like kinases. *Front*
700 *Plant Sci* 2015; **6**: 322.
701 95. Colaianni NR, Parys K, Lee H-S, Conway JM, Kim NH, Edelbacher N, et al. A complex
702 immune response to flagellin epitope variation in commensal communities. *Cell Host*
703 *Microbe* 2021; **29**: 635-649.e9.
704 96. Parys K, Colaianni NR, Lee H-S, Hohmann U, Edelbacher N, Trgovcevic A, et al.
705 Signatures of antagonistic pleiotropy in a bacterial flagellin epitope. *Cell Host Microbe* .
706 2021. Elsevier BV. , **29**: 620-634.e9
707 97. Monaghan J, Zipfel C. Plant pattern recognition receptor complexes at the plasma
708 membrane. *Curr Opin Plant Biol* 2012; **15**: 349–357.
709 98. Pfeilmeier S, Petti GC, Bortfeld-Miller M, Daniel B, Field CM, Sunagawa S, et al. The
710 plant NADPH oxidase RBOHD is required for microbiota homeostasis in leaves. *Nat*
711 *Microbiol* 2021; **6**: 852–864.
712 99. Chen T, Nomura K, Wang X, Sohrabi R, Xu J, Yao L, et al. A plant genetic network for
713 preventing dysbiosis in the phyllosphere. *Nature* 2020; **580**: 653–657.
714 100. Ma K-W, Niu Y, Jia Y, Ordon J, Copeland C, Emonet A, et al. Coordination of microbe-
715 host homeostasis by crosstalk with plant innate immunity. *Nat Plants* 2021; **7**: 814–825.
716
717 Roux F, Frachon L, Bartoli C. The genetic architecture of adaptation to leaf and root bacterial
718 communities in *Arabidopsis thaliana*. In preparation 2022.

719 **FIGURE LEGENDS**

720 **Figure 1.** Phenotypic variation of the response to the mock treatment and the 13 bacterial strains
721 in field conditions. **a** Box-plots illustrating the variation among the 14 treatments for the trait
722 'area-9dai'. **b** Box-plots illustrating the variation among the 14 treatments for the trait 'RGR-
723 5dai-1dbi'. For each treatment, each dot corresponds to the genotypic value of one of the 54
724 populations of *A. thaliana*. For each trait, different letters indicate different groups according
725 to the treatments after a Ryan-Einot-Gabriel-Welsh (REGWQ) multiple-range test at $P = 0.05$.
726 dai: days after inoculation, dbi: day before inoculation.

727

728 **Figure 2.** Genetic variation of 54 natural populations of *A. thaliana* in response to the 13
729 bacterial strains in field conditions. **a** Box-plots illustrating the range of genetic correlations
730 between each treatment with a bacterial strain and the remaining 13 treatments for the traits
731 'area-5dai' and 'area-9dai'. **b** Box-plots illustrating the range of genetic correlations between
732 each treatment with a bacterial strain and the remaining 13 treatments for the traits 'RGR-5dai-
733 1dbi' and 'RGR-9dai-1dbi'. Red triangle: genetic correlation with the mock treatment, black
734 dots: genetic correlations with other treatments with a bacterial strain (*ggplot2* library
735 implemented in the R environment). dai: days after inoculation, dbi: day before inoculation.
736 Treatments are ranked according to their mean genetic correlation with other treatments.

737

738 **Figure 3.** Interaction plots illustrating the reaction norms observed at the population level
739 between the mock treatment and the treatment with OTU13_*M*sp_1 (left panel) and
740 OTU13_*M*sp_2 (right panel). Each dot corresponds to the genotypic value of one of the 54
741 populations of *A. thaliana*. Each line corresponds to the response of one of the 54 populations
742 to the inoculation with either OTU13_*M*sp strain. The blue and red lines correspond to two
743 populations FERR-A and LUZE-B, respectively, with an opposite response to the strain

744 OTU13_*M*sp_2. Pictures illustrate representative plants of the two populations highlighted in
745 blue and red for the mock treatment and the treatment with either OTU13_*M*sp strain.
746
747 **Figure 4.** Genetic architecture of the response of 54 natural populations of *A. thaliana* to the 13
748 bacterial strains in field conditions. **a** Number of QTLs per treatment for each of the nine
749 phenotypic traits. **b** An UpSet plot illustrating the flexibility of genetic architecture among the
750 13 treatments with bacterial strains for the trait 'area-9dai' (*upset* library implemented in the *R*
751 environment). 'Number of genes': Total number of candidate genes underlying detected QTLs
752 and not shared with the mock treatment. A single dot indicates the number of candidate genes
753 specific to a given treatment. Candidate genes shared between two or more treatments are
754 represented by a line connecting two or more dots.
755
756 **Figure 5.** Manhattan plots of the Lindley process for the trait 'area_9dai' for the mock treatment
757 and the treatments with the bacterial strains OTU13_*M*sp_2, OTU5_*Pmo*_1 and OTU5_*Pmo*_2.
758 The *x*-axis corresponds to the physical position of 1,638,649 SNPs on the five chromosomes.
759 The dashed lines indicate the minimum and maximum of the five chomosome-wide significance
760 thresholds.
761
762 **Figure 6.** Enriched biological processes in response to the 13 bacterial strains in field
763 conditions. **a** Enriched biological processes for the list of unique candidate genes for each of
764 the 14 treatments and for the list of unique candidate genes from the combined 13 bacterial
765 strains ('All strains'), obtained with the MapMan classification superviewer tool. The color of
766 the dots corresponds to the level of significance. **b** Number of candidate genes belonging to the
767 different sub-categories of the enriched 'signalling' biological process for each treatment with

768 a bacterial strain. LRR: leucine rich repeat, CRK: cystein-rich receptor-like kinase, MAP:

769 mitogen-activated protein.

SUPPLEMENTARY FIGURES AND TABLES

Supplementary Figure S1. Experimental design of the field experiment.

Supplementary Figure S2. Phenotyping of three traits related to vegetative growth by imaging (AREA, PERIMETER and DIAMETER).

Supplementary Figure S3. Phenotypic variation of the response to the 13 bacterial strains in field conditions.

Supplementary Figure S4. Box-plots illustrating the range of genetic correlations between each treatment with a bacterial strain and the remaining 13 treatments for the traits 'perimeter-5dai', 'perimeter-9dai', 'diameter-5dai', 'diameter-9dai' and 'RGR-9dai-5dai'.

Supplementary Figure S5. UpSet plots illustrating the flexibility of genetic architecture among the 13 treatments with bacteria strains for the traits 'area-5dai', 'perimeter-5dai', 'perimeter-9dai', 'diameter-5dai', 'diameter-9dai', 'RGR-5dai-1dbi', 'RGR-9dai-5dai' and 'RGR-9dai-1dai'.

Supplementary Table S1. Names and GPS coordinates (expressed in degrees) of the 54 populations used in this study.

Supplementary Table S2. Homogeneity of plant growth across the field trial and presence of genetic variation for the three resource acquisition traits measured on the plants before inoculation.

Supplementary Table S3. Genetic variation of nine traits related to resource acquisition among the 162 accessions nested within 54 populations of *A. thaliana* for each of the 14 treatments.

SUPPLEMENTARY DATA SETS

Supplementary Data Set 1. Raw data for the 12 phenotypic traits scored on 14,580 plants in a field experiment conducted at INRAE Toulouse (France).

Supplementary Data Set 2. Genotypic values of the 54 natural populations of *A. thaliana* for the nine traits 'area-5dai', 'area-9dai', 'perimeter-5dai', 'perimeter-9dai', 'diameter-5dai', 'diameter-9dai', 'RGR-5dai-1dbi', 'RGR-9dai-5dai' and 'RGR-9dai-1dai' for each of the 14 treatments.

Supplementary Data Set 3. Genetic architecture of the 126 'phenotypic trait * treatment' combinations.

Supplementary Data Set 4. List of all candidate genes identified for each of the 126 'phenotypic trait * treatment' combinations.

Supplementary Data Set 5. List of unique candidate genes identified for each trait of each treatment with a bacterial strain.

Supplementary Data Set 6. List of the 1,962 candidate genes unique to the treatments with a bacterial strain. The pleiotropic level corresponds to the number of treatments with a bacterial strain for which the candidate gene was detected by GWA mapping.

Table 1. List of pleiotropic candidate genes associated with more than six bacterial strains but not detected in the mock treatment. Colored squares indicate the strains for which the candidate genes were identified. The different colors correspond to the seven QTLs in which the pleiotropic QTLs are located.

ATG number	Annotation	OTU2_*Pfu*_1	OTU2_*Pfu*_2	OTU3a_*Oxa*_1	OTU3a_*Oxa*_2	OTU4_*Com*_1	OTU5_*Pmo*_1	OTU5_*Pmo*_2	OTU6_*Psi*_1	OTU6_*Psi*_2	OTU13_*Msp*_1	OTU13_*Msp*_2	OTU29_*Sph*_1	OTU29_*Sph*_2
At2g40640	PUB62 (Plant U-box type E3 ubiquitin ligase)													
At2g40650	PRP38													
At2g40660	Nucleic acid-binding, OB-fold-like protein													
At2g40670	ARR16 (ARABIDOPSIS THALIANA RESPONSE REGULATOR 16)													
At2g44710	HNRNP R-LIKE PROTEIN, HRLP													
At2g44730	Alcohol dehydrogenase transcription factor Myb/SANT-like family protein													
At2g44735	transmembrane protein													
At2g47180	ATGOLS1, GALACTINOL SYNTHASE 1													
At2g47190	ATMYB2, MYB DOMAIN PROTEIN 2													
At2g47250	RNA helicase family protein													
At4g14713	PEAPOD 1, PPD1, TIFY4A													
At4g14716	ACIREDUCTONE DIOXYGENASE 1, ARD1, ATARD1, SGB3, SUPPRESSOR OF G BETA 3													
At4g14720	PEAPOD 2, PPD2, TIFY4B													
At5g42360	CFK2, COP9 SIGNALOSOME INTERACTING F-BOX KELCH 2													
At5g42370	Calcineurin-like metallo-phosphoesterase superfamily protein													
At5g50440	ATMEMB12, MEMB12, MEMBRIN 12													
At5g50450	HCP-like superfamily protein with MYND-type zinc finger													

Figure 1

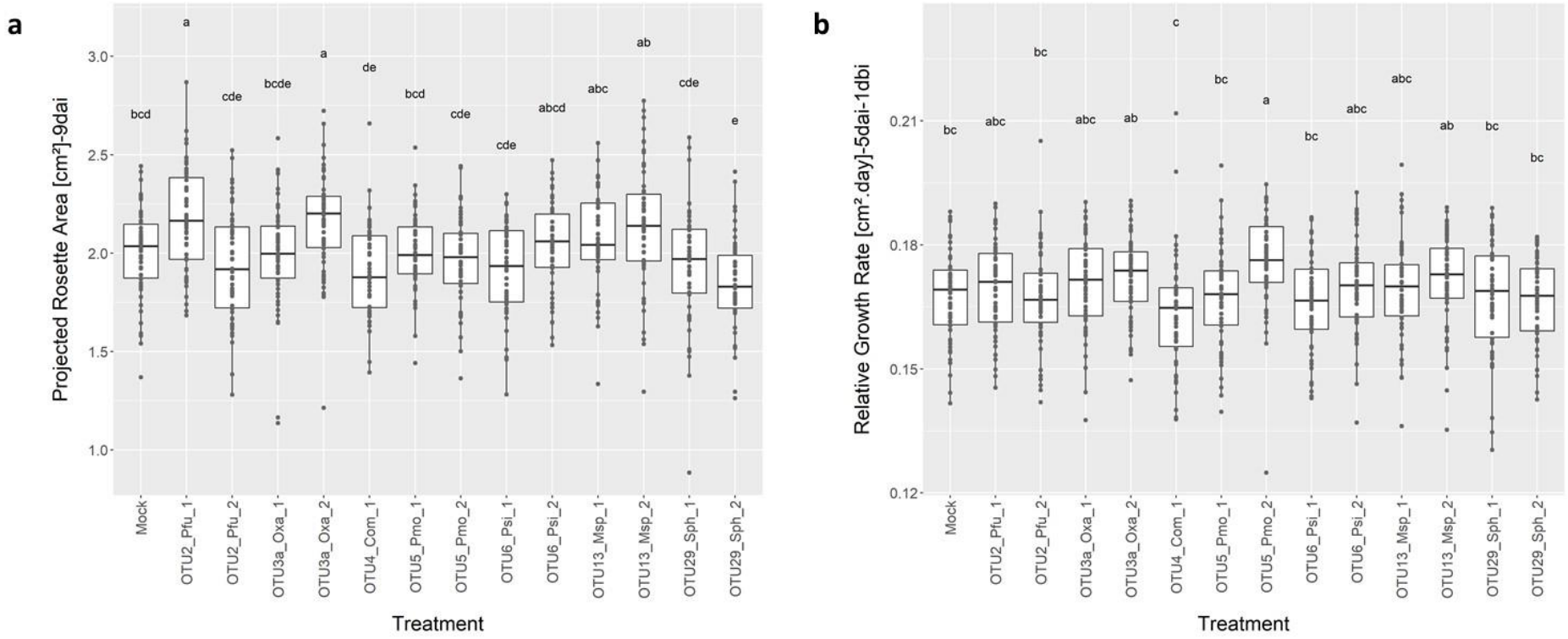
a
Projected Rosette Area [cm²]-9dai
3.0
2.5
2.0
1.5
1.0
bcd
a
cde
bcde
a
de
bcd
cde
cde
abcd
abc
ab
cde
e
Mock
OTU2_Pfu_1
OTU2_Pfu_2
OTU3a_Oxa_1
OTU3a_Oxa_2
OTU4_Com_1
OTU5_Pmo_1
OTU5_Pmo_2
OTU6_Psi_1
OTU6_Psi_2
OTU13_Msp_1
OTU13_Msp_2
OTU29_Sph_1
OTU29_Sph_2
Treatment
b
Relative Growth Rate [cm².day]-5dai-1dbi
0.21
0.18
0.15
0.12
bc
abc
bc
abc
ab
c
bc
a
bc
abc
abc
ab
bc
bc
Mock
OTU2_Pfu_1
OTU2_Pfu_2
OTU3a_Oxa_1
OTU3a_Oxa_2
OTU4_Com_1
OTU5_Pmo_1
OTU5_Pmo_2
OTU6_Psi_1
OTU6_Psi_2
OTU13_Msp_1
OTU13_Msp_2
OTU29_Sph_1
OTU29_Sph_2
Treatment

Figure 2

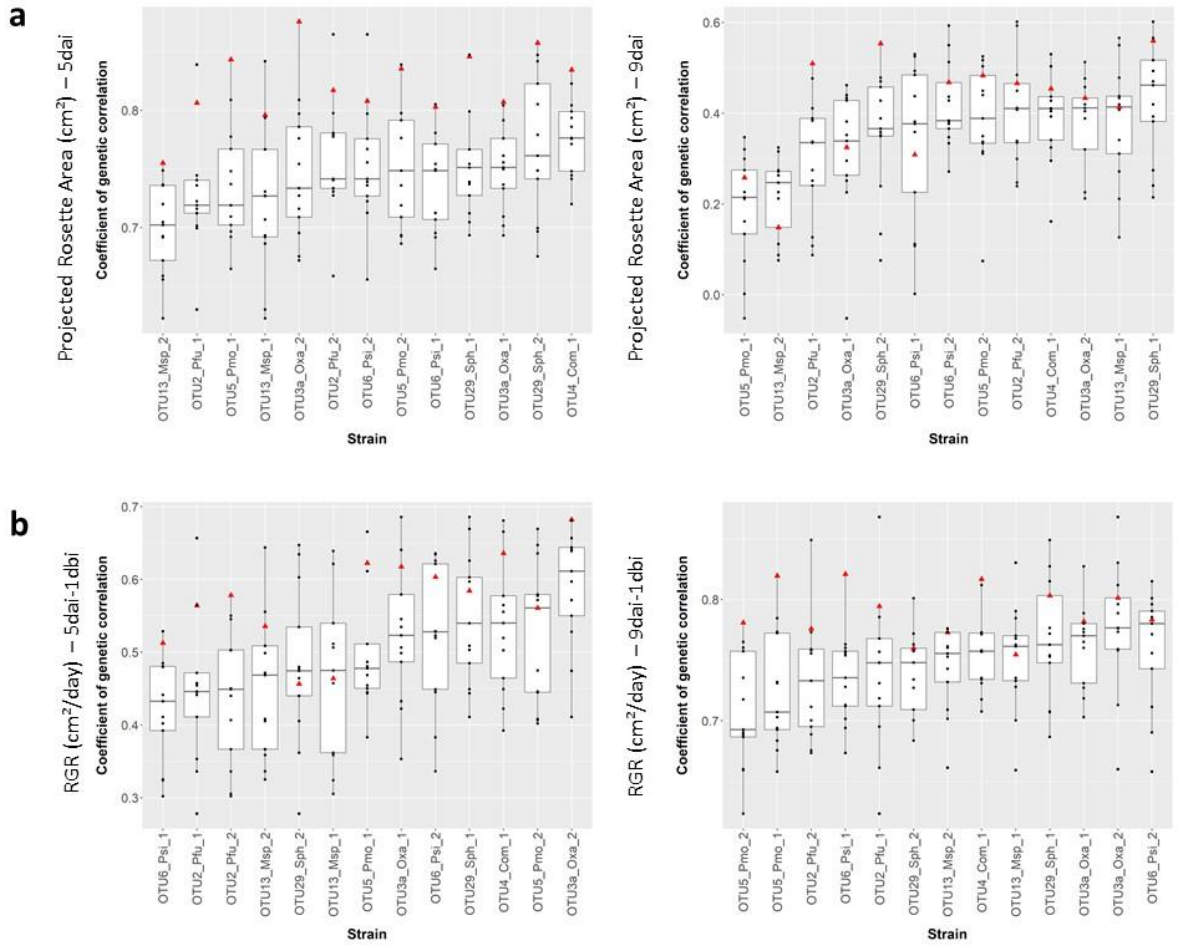

Figure 3

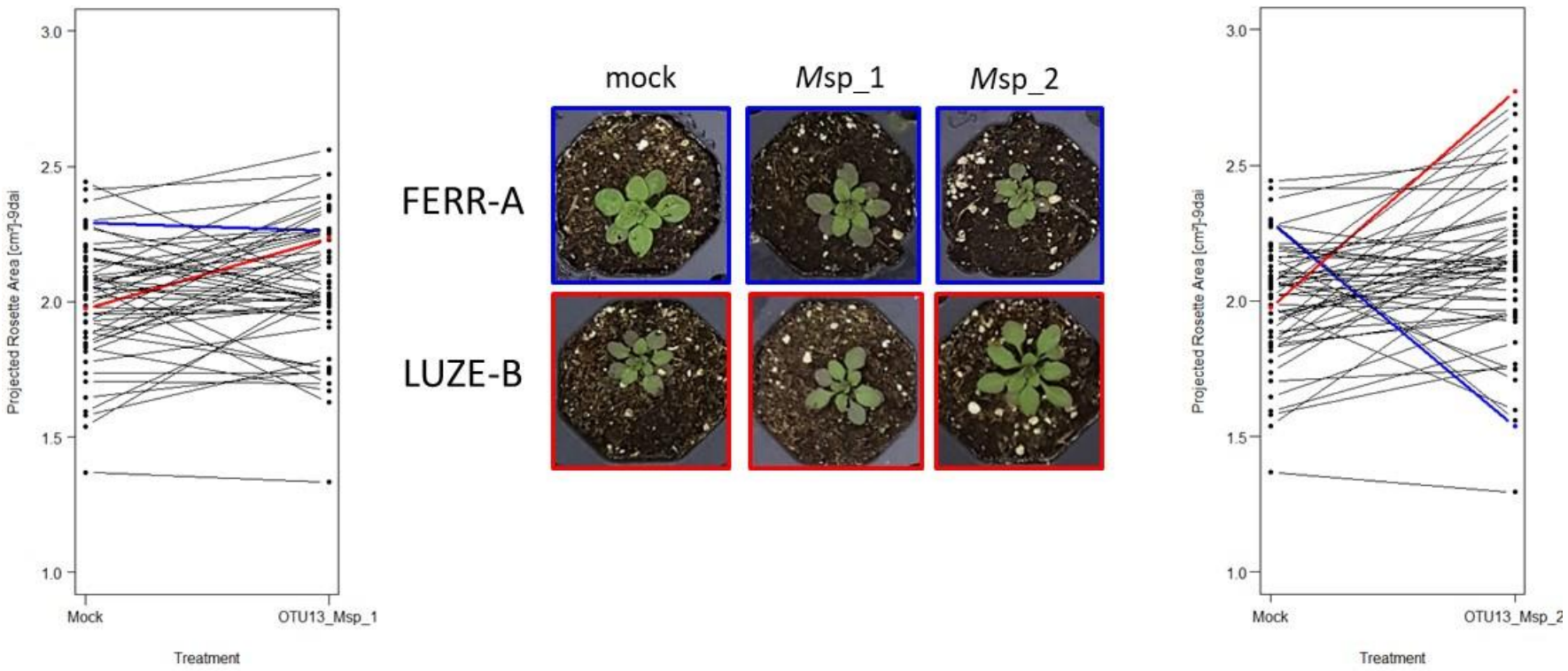
3.0
2.5
2.0
1.5
1.0
Projected Rosette Area [cm²]-9dai
Mock
OTU13_Msp_1
Treatment
mock
Msp_1
Msp_2
FERR-A
LUZE-B
3.0
2.5
2.0
1.5
1.0
Projected Rosette Area [cm²]-9dai
Mock
OTU13_Msp_2
Treatment

Figure 4

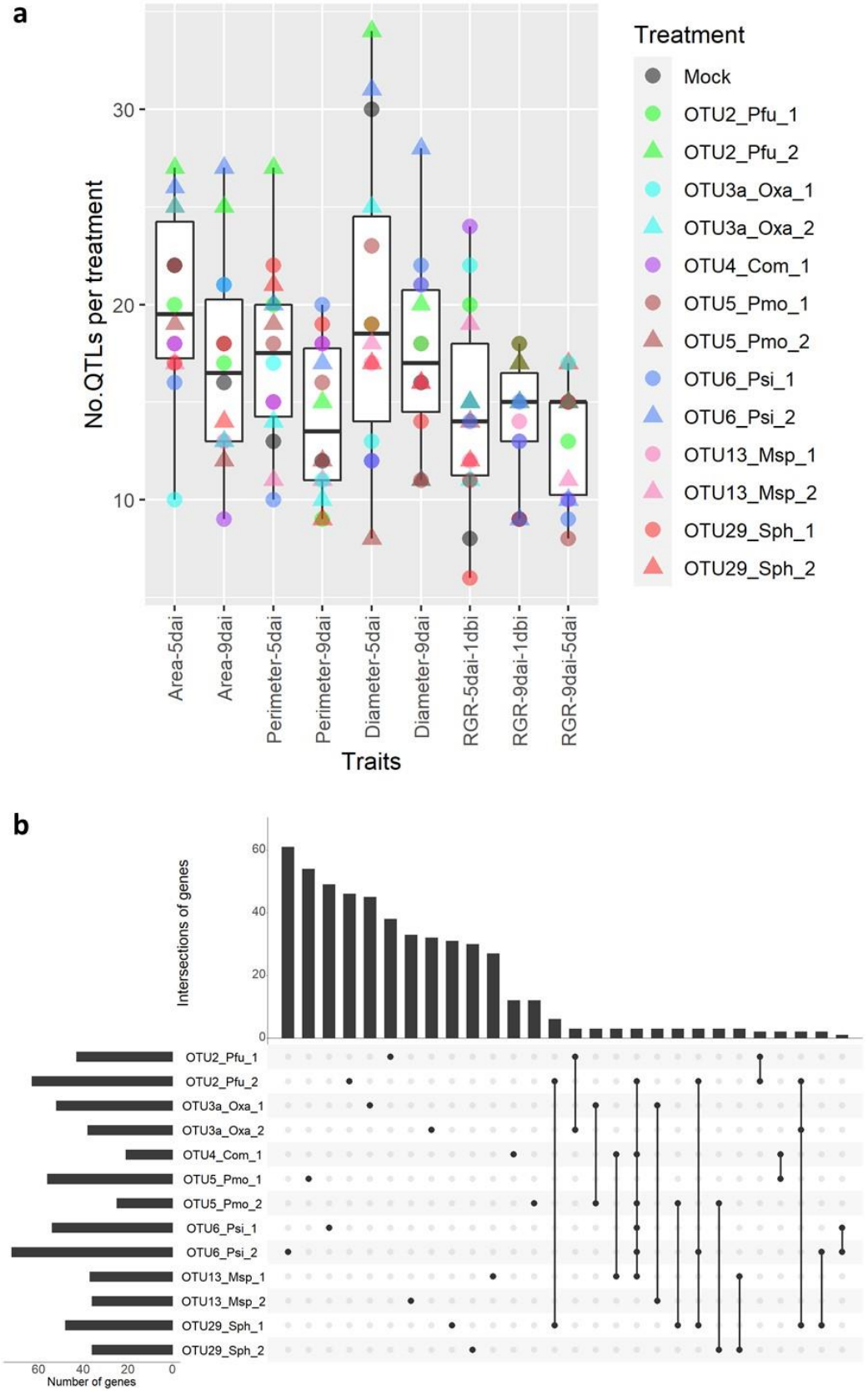

Figure 5

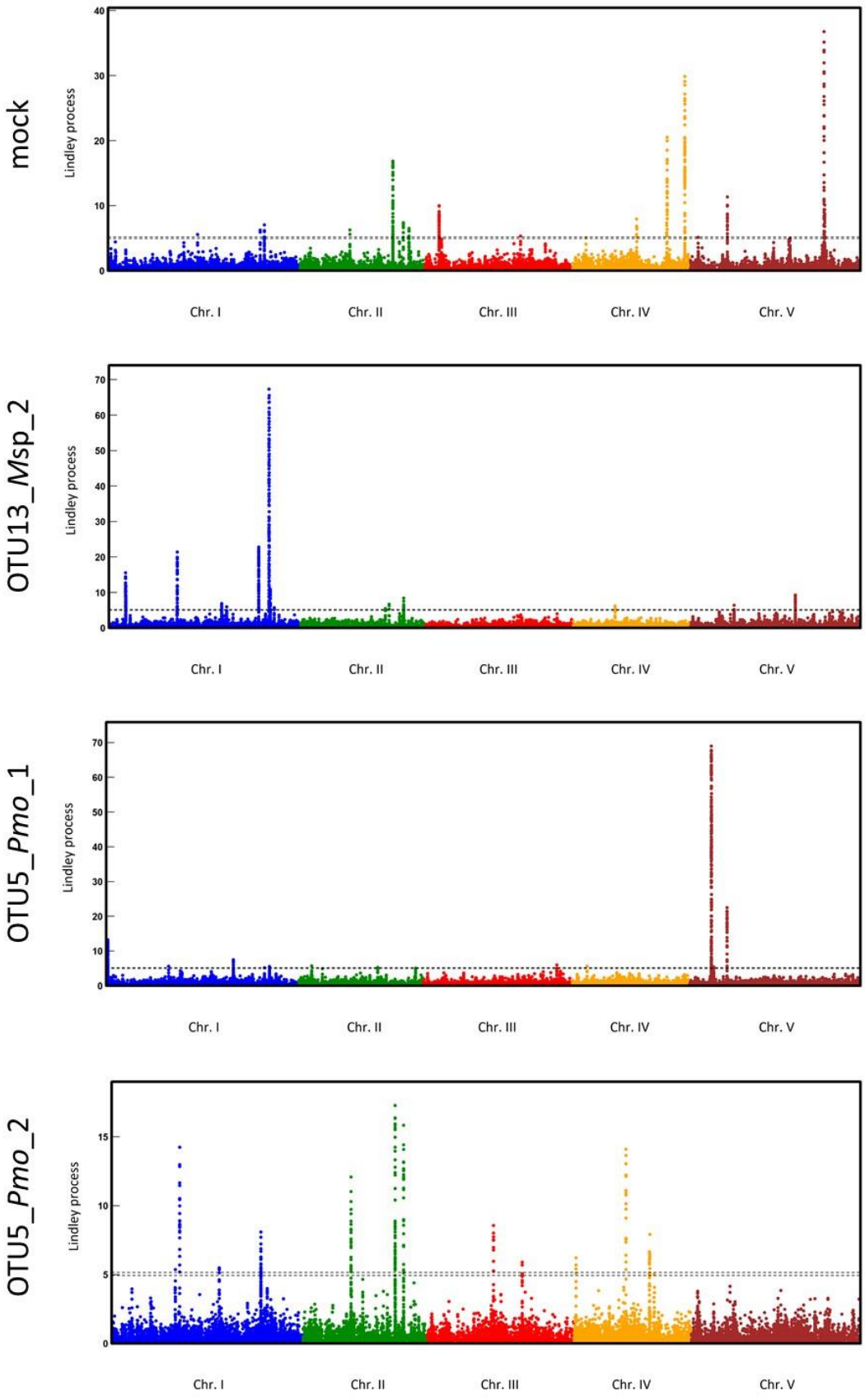

Figure 6