

ARTICLE OPEN



Commensal *Pseudomonas* protect *Arabidopsis thaliana* from a coexisting pathogen via multiple lineage-dependent mechanisms

Or Shalev¹, Haim Ashkenazy¹, Manuela Neumann¹ and Detlef Weigel¹✉

© The Author(s) 2021

Plants are protected from pathogens not only by their own immunity but often also by colonizing commensal microbes. In *Arabidopsis thaliana*, a group of cryptically pathogenic *Pseudomonas* strains often dominates local populations. This group coexists in nature with commensal *Pseudomonas* strains that can blunt the deleterious effects of the pathogens in the laboratory. We have investigated the interaction between one of the *Pseudomonas* pathogens and 99 naturally co-occurring commensals, finding plant protection to be common among non-pathogenic *Pseudomonas*. While protective ability is enriched in one specific lineage, there is also a substantial variation for this trait among isolates of this lineage. These functional differences do not align with core-genome phylogenies, suggesting repeated gene inactivation or loss as causal. Using genome-wide association, we discovered that different bacterial genes are linked to plant protection in each lineage. We validated a protective role of several lineage-specific genes by gene inactivation, highlighting iron acquisition and biofilm formation as prominent mechanisms of plant protection in this *Pseudomonas* lineage. Collectively, our work illustrates the importance of functional redundancy in plant protective traits across an important group of commensal bacteria.

The ISME Journal (2022) 16:1235–1244; <https://doi.org/10.1038/s41396-021-01168-6>

INTRODUCTION

The health of a plant depends to a large extent on its resident microbiota. The effect of individual microbes on plant health has been extensively investigated since the dawn of phytopathology [1], mainly focusing on pathogens, considering their profound effect on global agriculture and food supply [2]. The ability of phytopathogens to overpopulate plants, leading to disease, is reflected in the reciprocal ability of the plant to recognize and control these pathogens [3]. Nonetheless, in recent years there has been an increasing realization that the plant relies not only on its immune system but also on other resident microbes to suppress pathogens [4–7].

For example, in natural settings, the health of the ephemeral plant *Arabidopsis thaliana* is associated with many bacteria that reduce disease threats due to filamentous microbes [5]. Similar patterns have also been found in controlled settings, with suppression of bacterial pathogens by other bacteria [8, 9]. These protective agents can have several modes of action, including (i) activation of systemic defences that spread throughout the plant [10], (ii) outcompeting pathogens over nutrients [11, 12], and (iii) direct antibiosis [13, 14]. These mechanisms are non-exclusive and can operate simultaneously.

These studies have greatly advanced our understanding of protective interactions in plants, even if the investigated pathogens and protective microbes do not always coexist in the wild. A notable exception is a recent study in which individual members of native tomato microbiomes were tested for suppression of the

soil-borne pathogen *Ralstonia solanacearum* [15]. This report focused on siderophore production and competition for iron as a known mechanism for microbe-microbe competition, revealing a link to pathogen inhibition in the plant rhizosphere. The work also highlighted how insights can be gained from investigating pathogen protection in a phylogenetic framework [15].

Substantial diversity among *Pseudomonas*, a bacterial genus that includes pathogens, commensals and mutualists of plants (Haas and Défago 2005; Baltrus et al. 2017), was found in a survey in southwest Germany, based on genome sequencing of 1524 isolates isolated from wild *A. thaliana* plants [16]. One cryptically pathogenic lineage dominated this collection, although other, apparently commensal strains were found as well. By employing synthetic communities of commensal and pathogenic *Pseudomonas* from this collection, commensal strains were found to protect *A. thaliana* from pathogenic *Pseudomonas* strains that co-occur in the wild [17]. In a parallel study with a *Pseudomonas* collection from an *A. thaliana* relative, commensal strains were shown to often outcompete pathogenic *Pseudomonas syringae* isolates sampled from the same plant [18].

These studies point to the importance of interactions among wild *Pseudomonas* strains in maintaining plant health. Here, we leveraged a local *Pseudomonas* collection from *A. thaliana* [16] to examine the extent of protection conferred by commensal *Pseudomonas* against co-existing pathogenic *Pseudomonas*, and to discover some of the underlying mechanisms. Using a high throughput image-based assay, we monitored outcomes of

¹Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany. ✉email: weigelt@tue.mpg.de

Received: 10 August 2021 Revised: 19 November 2021 Accepted: 26 November 2021

Published online: 11 December 2021

systematic co-infections of a diverse set of commensal *Pseudomonas* with a focal *Pseudomonas* pathogen, covering the entire phylogeny of commensal *Pseudomonas* isolates in this collection. We found that protection by commensals was a common feature, although it was enriched in a specific lineage. We discovered bacterial genes for plant protection using genome-wide association and comparative genomics. Using knockout mutants, we experimentally validated the role of several candidate genes in plant protection, establishing a link to iron uptake and biofilm formation in the mitigation of phyllosphere pathogens.

RESULTS

Systemic co-infections of commensal *Pseudomonas* with an individual pathogen

To examine the ability of commensal *Pseudomonas* strains to protect host plants from members of a pathogenic *Pseudomonas* lineage, we made use of a local isolate collection [16]. We henceforth refer to an operational taxonomic unit (OTU) as reported in that study as “ATUE” (isolates from Around TUEbingen), and following previous findings [16, 17], we refer to the lineage ATUE5 as pathogenic, and to all non-ATUE5 lineages as commensals.

We grew plants on MS agar and monitored plant growth and health by extracting the number of green pixels from images over time (illustration in Fig. 1A). Green pixel count and rosette fresh weight were strongly correlated (Supplementary Fig. S1; $R^2 = 0.92$, p value $< 2.2e-16$), validating the use of green pixels as a proxy for the plant biomass.

We use a conservative definition of plant protection, with protective strains leading to near-normal plant growth (comparable with uninfected plants) in the presence of a pathogen. To estimate how common the ability of non-ATUE5 strains to protect against the impact of pathogenic *Pseudomonas* is, we infected plants with a panel of non-ATUE5 isolates in the presence of a common ATUE5 representative, strain p4.C9, which hereafter will be referred to as “the pathogen” or “the ATUE5 pathogen”. This isolate was chosen because it is dominant over other ATUE5 strains but at the same time highly susceptible to the presence of non-ATUE5 strains, correlating with plant protection [17]. We excluded highly similar isolates, with Jaccard distances ≥ 0.99 in gene content. From a total of 151 non-ATUE5 strains in our local collection [16], we initially chose a subset of 127 isolates, from which we were able to revive 99 (Fig. 1B; Supplementary Table S1). We included three isolates from the pathogenic ATUE5 clade as control (Supplementary Table S1), resulting in 102 strains that were tested in co-infections with the pathogen p4.C9.

One of the 102 strains, p5.F2, was known to suppress ATUE5 strains inside plants [17] and was used as a “protective control”. To confirm that protection was due to bacterial activity and not merely a host response to the inoculum, e.g., PAMP-triggered immunity, we carried out infections with a heat-killed preparation of our protective control. As expected, co-infection with the protective strain resulted in normal plant growth, while treatment with the pathogen by itself or co-inoculation with the heat-killed strain impaired growth (Fig. 1C; Supplementary Fig. S2; mean growth [$\Delta 7\text{dpi}-0\text{dpi}$ green pixels]: control 29,227 [15,721, 40,662], pathogen 3328 [−10,350, 14,759], pathogen + heat-killed protective −5222 [−16,410, 6211] and protective 29,585 [16,323, 40,812], at 95% confidence interval).

Plant protection is phylogenetically widespread, but also lineage-specific

Co-inoculation of non-ATUE5 strains with the ATUE5 pathogen led to a range of outcomes for plant health, from reduced to enhanced growth in comparison to uninfected plants, with protection being common (Fig. 2A).

Protection was unevenly distributed among the three highly sampled ATUE groups—ATUE2, ATUE3, and ATUE4—with ATUE2 being the most protective mean growth ($\Delta[7 \text{ dpi}]-[0 \text{ dpi}]$ green pixels, 95% CI): ATUE2 22,231 [19,522, 24,908], ATUE3 14,120 [10,780, 17,389], ATUE4 10,868 [6844, 14,937], and bacteria-free 27,069 [19,882, 34,264] (Fig. 2B).

Within each ATUE group, there was considerable variation in protective ability, even among strains with highly similar genomes (Fig. 2C). In some cases, closely related strains had contrasting activities, with one providing robust protection and the other having no effect, e.g., ATUE2 strains p11.F1 and p12.H7: mean growth after coinfection with p11.F1 being 25,554 pixels [8303, 42,567], but with p12.H7 being −28,914 pixels [−45,789, −11,837]).

Even within ATUE5, some isolates had the protective ability. We had chosen the three ATUE5 strains because in a previous set of experiments [17], they had appeared to be less competitive than our focal ATUE5 pathogen p4.C9, with lower abundance in the context of *A. thaliana* infections with synthetic communities of ATUE5 and non-ATUE5 strains. Surprisingly, two of them also mitigated the pathogen effect, causing normal plant growth. This further exemplifies the differential functions within ATUE groups (Fig. 2C).

ATUE2 protective genes are lineage-specific

That functional variation could not be explained by phylogeny, considering both topology or branch length, suggested that variation in gene content, possibly due to horizontal gene transfer, was causal for protection. We, therefore, were curious whether the presence/absence (P/A) of gene orthology groups could explain the propensity for plant protection.

There were 32,753 gene orthology groups, that were found in more than one strain, but not shared by all ATUE strains. We examined the correlation between plant growth and the presence of each of the individual orthology groups using treeWAS [19], a tool for genome-wide association studies (GWAS) in bacteria. Because treeWAS accounts for population structure, it will remove true positives that are highly correlated with population structure [20]. To address this issue, we used Spearman’s rank correlation coefficient (SRCC) to search for genes that were correlated with protection. SRCC should also be useful for differentiating a global from a lineage-specific signal. To this end, we calculated SRCCs separately among the highly sampled ATUE2, ATUE3, and ATUE4 groups, considering the difference in median green pixels between 0 and 7 dpi. For treeWAS, we used three additional plant growth metrics: the median change in green pixels between the last day of the experiment (7 days post-infection) to the day of infection (“median growth”), a Bayesian-derived approximation of the ‘median growth’ (“cdI50”, i.e., the posterior distribution median of the ‘median growth’), a binary categorization based on cdI50 and the area under the curve while accounting for all sampled time points (Supplementary Table S2; Methods). The extent of agreement between the four phenotypes was used as another indicator for the robustness of each association (Supplementary Table S2). We overlapped the results from all four analyses and removed genes with a negative SRCC (SRCC < 0 , which implies negative effects on plant health in addition to those of the pathogen rather than protection from the pathogen), leaving us with 14 strong candidates for plant protection (Supplementary Table S2). The nine genes with the highest SRCC values (Rho 0.37–0.46) were unique to ATUE2 (Fig. 3A).

The existence of a gene set unique to one lineage implies that plant protection by *Pseudomonas* is not driven by a universal mechanism in this genus, but rather by lineage-specific processes. To test this assumption, we ran the same GWAS analysis on each of the main ATUE groups separately, and independently analyzed the subsets of strains related to ATUE2, ATUE3, and ATUE4. In total,

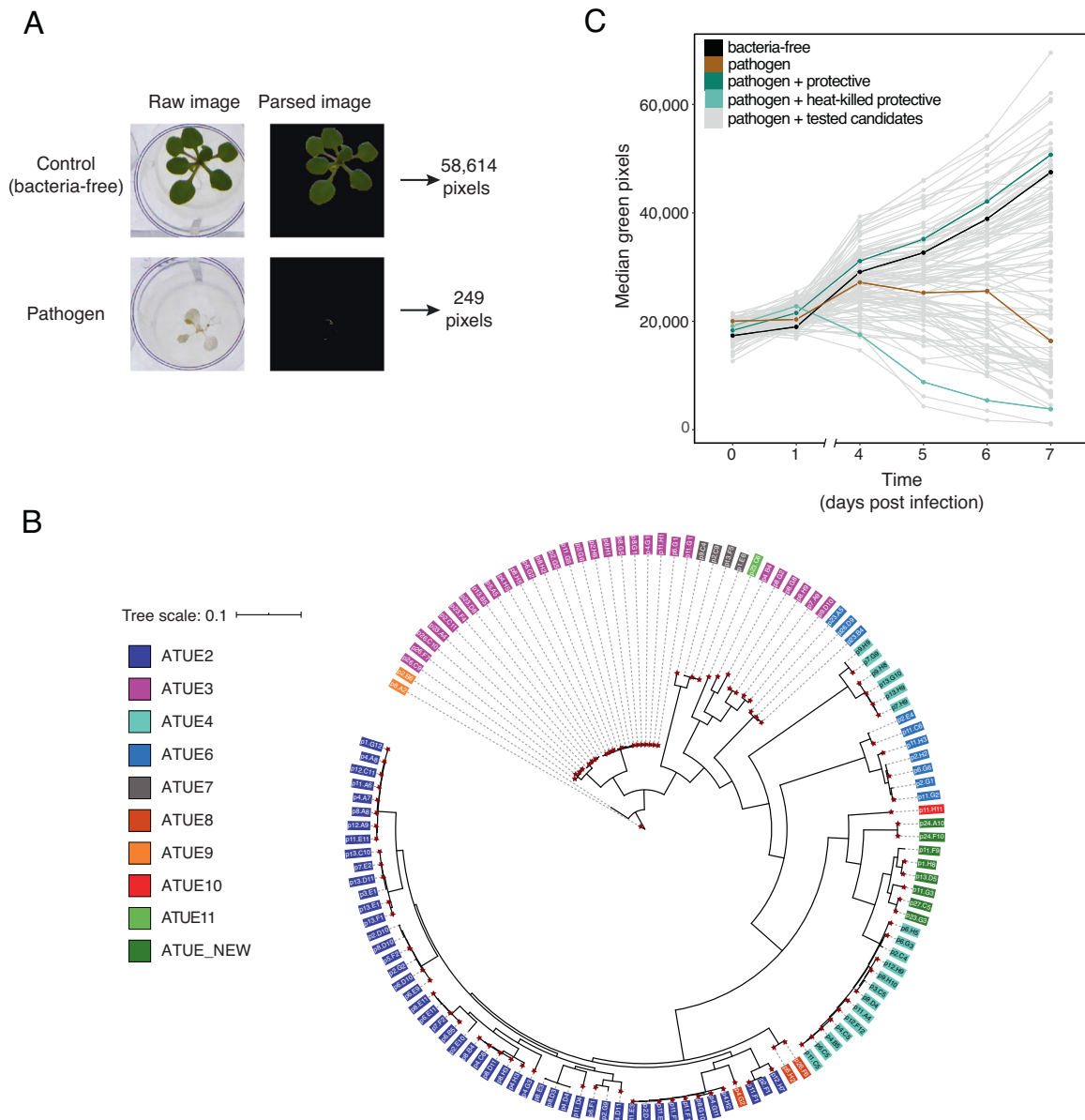


Fig. 1 Panel of potentially protective *Pseudomonas* isolates and experimental design. **A** Illustration of image processing to enumerate plant green pixels, approximating plant biomass. **B** Phylogenetic tree of 127 representatives, putatively commensal non-ATUE5 strains sampled from southwest Germany [16]. All other non-ATUE5 isolates in this collection are represented in this core collection by a strain with which they share $\geq 99\%$ of genes. Colors indicate the ATUE group [16], and asterisks the 99 strains used here. **C** Daily median of plant green pixels among the different treatments. Control treatments: Bacteria-free buffer, pathogen only, and co-infections of the pathogen with the protective strain and the heat-killed protective strain. Plants were imaged daily. $n = 8$ replicates per treatment.

95 genes with positive SRCCs were significantly associated with plant protection: 14 genes from the full set of strains, 46 from ATUE4, 35 from ATUE3, and none from ATUE2. Except for a single gene, we found no overlap among hits from the different ATUE groups, providing further evidence for lineage specificity (Supplementary Fig. S3).

Half of all protection-associated genes, 47 out of 95, were annotated as encoding ‘uncharacterized protein’ or had no-hit in TrEMBL and Swiss-Prot databases (Supplementary Table S2). Out of the well-annotated genes, a few seemed to be likely to be directly related to microbe-microbe interactions: three iron-uptake-related genes unique to ATUE2 (GC00000450_7: ‘TonB_C domain-containing protein’, GC00000032_87: ‘Putative iron(III) dicitrate sensor protein FecR’ and GC00000050_54: ‘Probable RNA polymerase sigma factor FecI’), a gene related to resistance to

antimicrobial peptides in ATUE3 (GC00000089_55: ‘UDP-4-amino-4-deoxy-L-arabinose-oxoglutarate aminotransferase’ [*arnB*] [21]), and an antitoxin gene in ATUE4 (GC00007392_r1_1: ‘Antitoxin FitA’) alongside mobility-related genes (GC00002204_5: ‘Twitching motility protein PilT’, GC0000715_r1_r1_2: ‘Pilus assembly protein PilW’). These putative functions suggest not only a diverse set of genes but also diverse protective mechanisms among commensal *Pseudomonas*.

Several of these genes had identical SRCC values in a specific ATUE group and turned out to be physically linked (Supplementary Table S2; details in Methods). There was also evidence for genetic elements associated with horizontal gene transfer, such as phage genes (clusters 3 and 4 in Supplementary Table S2), consistent with these clusters being genomic islands. Cumulatively, these results suggest a scenario in which plant protection

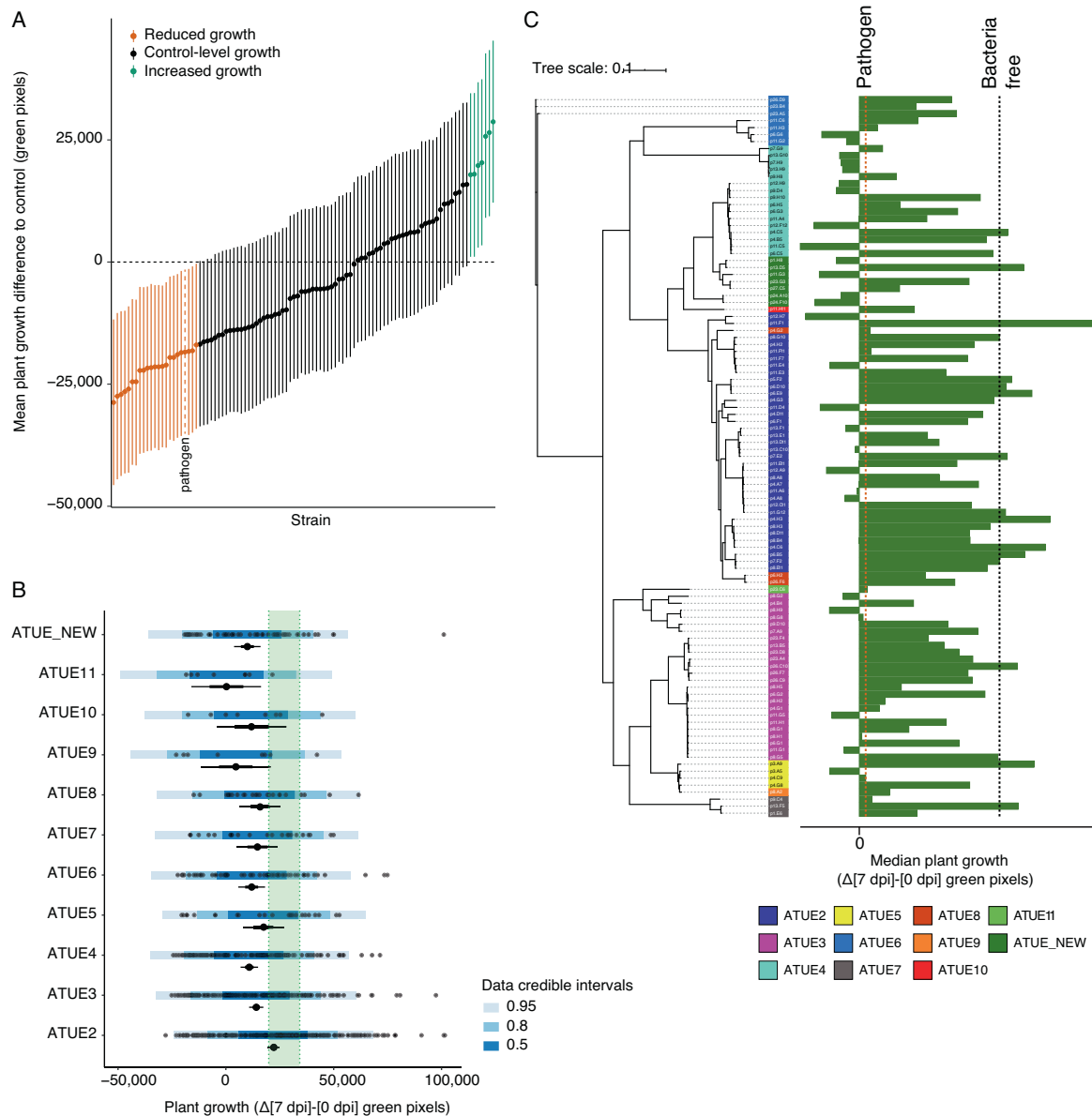


Fig. 2 Protection by different *Pseudomonas* strains is common and enriched in the ATUE2 lineage. **A** Mean difference in plant growth to control, after co-infection with different *Pseudomonas* strains and the focal p4.C9 pathogen. Growth was measured as the change in green pixels between days 0 and 7 after infection. Vertical lines indicate 95% Bayesian credible intervals and dots indicate the median estimate. The dashed vertical line signifies the positive control, that is, the pathogen-only treatment and the dashed horizontal line signifies the negative control, that is, the baseline average growth in the absence of bacteria. $n = 8$ replicates each. **B** Plant growth after co-infections, binned by ATUE group. In each ATUE group, raw data for individual replicates are shown with dots; the overlain shades of blue indicate the posterior predictive intervals, as presented on the bottom right. The mean growth for each ATUE group is also shown; the dot indicates the median, the thin horizontal line the 95% credible interval, and the thick line the 67% credible interval. Green shade indicates the 95% credible interval of the negative control, growth in the absence of bacteria. See Supplementary Table S1 for the number of strains in each ATUE group. **C** Median plant growth after co-infections, ordered by phylogeny. Colors indicate ATUE groups [16]. Medians of plant growth in mono-association with pathogen p4.C9 or without infection indicated by dashed vertical lines.

by commensal *Pseudomonas* is driven by multiple, clade-specific mechanisms that were horizontally transferred.

As described above, of the 14 protection-associated genes found among all strains, nine were unique to ATUE2 (Fig. 3A). Although there were no significant hits when we analyzed the ATUE2 subset on its own, strains carrying these nine genes were more protective not only when compared among all strains, but also within ATUE2 (Fig. 3B, C). For ATUE2 strains with at least one of these genes, plant growth was 29,299 pixels (22,567, 35,983, 95% confidence interval). In contrast, plant mean growth for

ATUE2 strains without any of these genes was only 12,780 pixels (6891, 18,821), and for non-ATUE2 strains without any of these genes, it was only 10,791 pixels (7541, 14,078). Moreover, we found that the effect of these nine genes was additive within ATUE2, and plant growth was correlated with the number of genes present in a given *Pseudomonas* strain (Supplementary Fig. S4; $R^2 = 0.55$, p value = 0.0004). These results highlight the importance of these nine genes in plant protection both among all *Pseudomonas* strains in our collection, and within ATUE2, despite the lack of statistical signal in the ATUE2-subset alone.

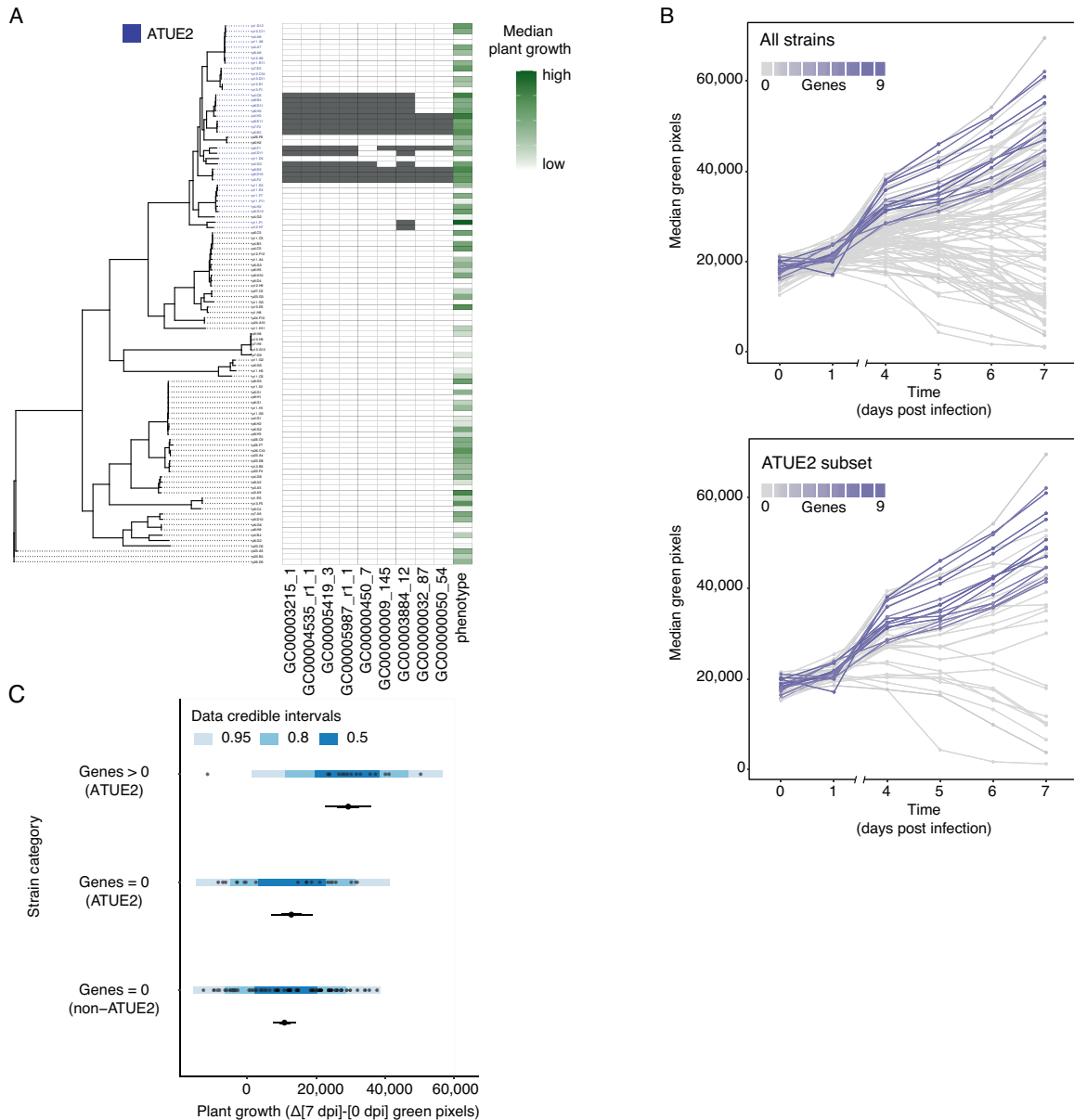


Fig. 3 Genes most strongly associated with protection are unique to ATUE2. **A** Presence/absence variation of nine genes that are most strongly associated with protection. Strains are ordered by their phylogeny. Strains belonging to the ATUE2 group are indicated in magenta. Gene presence is indicated in grey, absence in white. Growth was measured as the change in green pixels between days 0 and 7 after infection, with the scale indicated on the right. **B** Daily median of plant green pixels among all strains (top panel) or the ATUE2 subset (bottom panel). Shades of magenta indicate the number of protective genes out of 9 present in each strain, as indicated by the scale on the bottom. Eight replicates per treatment. **C** Plant growth after co-infections with the pathogen p4.C9, binned by (i) the presence of at least one gene from the set of the nine protective genes (presented in panel **A**) in a given commensal strain, and (ii) membership in the ATUE2 group. In each group, raw data for individual replicates are shown as dots; the overlain shades of blue indicate the posterior predictive intervals, as indicated on the bottom right. The mean growth for each group is also shown; the dot indicates the median, the thin horizontal line the 95% credible interval, and the thick line the 67% credible interval. $n = 8$ replicates per strain, with the number of strains in the different categories as follows: non-ATUE2 with 0/9 genes = 67; ATUE2 with 0/9 genes = 20, and ATUE2 with at least 1/9 genes = 16.

Plant protection by ATUE2 linked to iron acquisition and biofilm formation

We selected the nine ATUE2-lineage specific candidates for gene deletion in the protective strain p5.F2 as a representative for the ATUE2 clade, to validate the role of these genes in plant protection. Since two genes were found next to each other (GC00000032_87 and GC00000050_54), we included them in one deletion, resulting in eight mutations. Because our knockout mutants were marked with gentamicin resistance, we generated a wild type strain that also expresses gentamicin (as well as the lux

operon), ensuring a similar metabolic burden (hereafter “wild type” or “WT”; see Methods).

Out of the eight knockout mutants, three lost their ability to mitigate the effect of the pathogen p4.C9 following co-infection, and we call them henceforth “non-protective mutants” (Fig. 4A; Supplementary Fig. S5A; results of an additional experiment with the three non-protective mutants in Supplementary Fig. S6A, B). These three non-protective mutants did not affect plant weight when tested individually, similarly to the wild type (Supplementary Fig. S7A, B), implying that the weight reduction

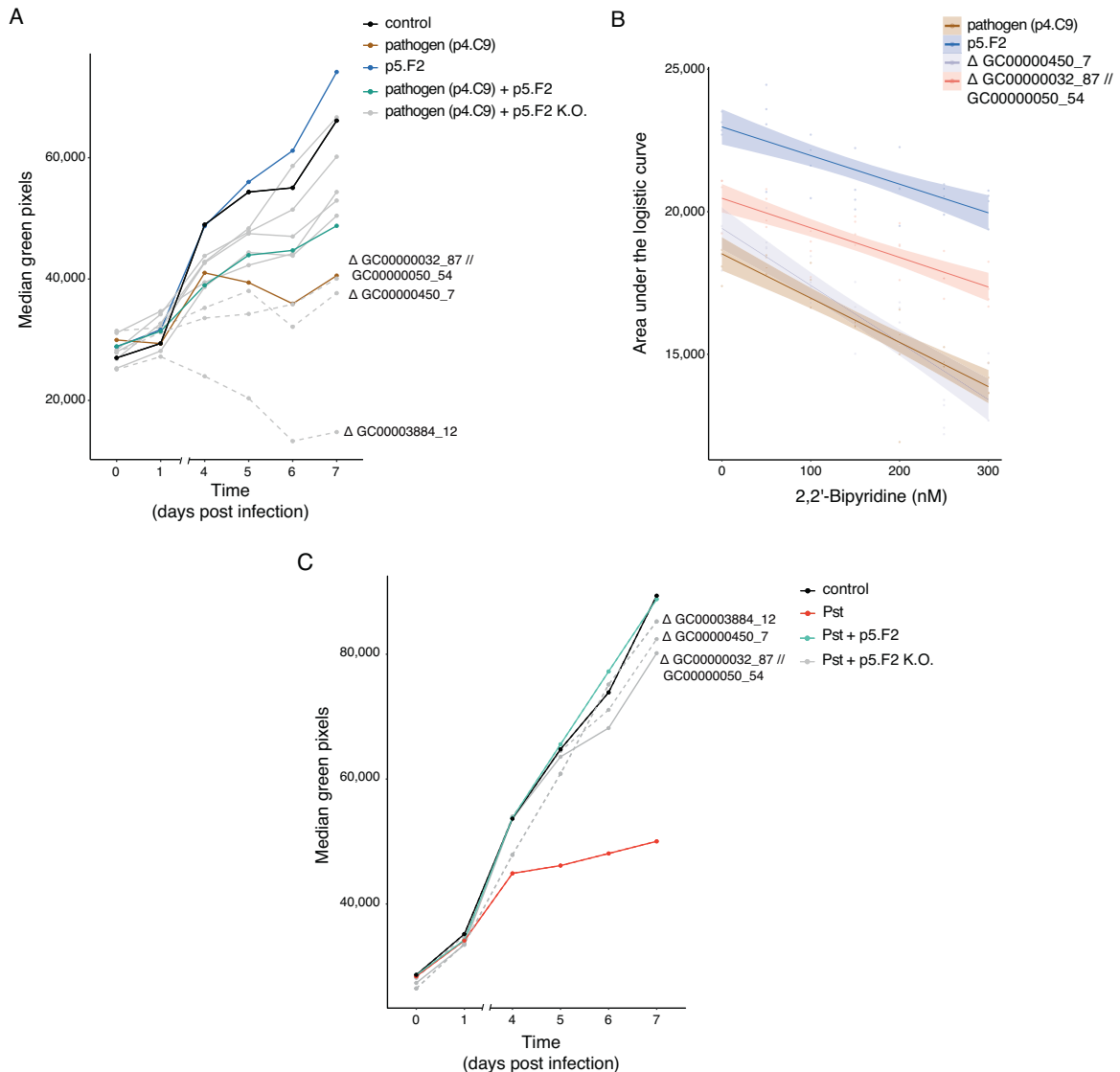


Fig. 4 Functional analysis of genes implicated in plant protection. **A** Daily median of plant green pixels after treatment with control, p4.C9 pathogen, the protective strain p5.F2, a mixture of the pathogen and p5.F2, and mixtures of the pathogen and each of eight p5.F2 knockout mutants, indicated as p5.F2 K.O. Dashed lines indicate three treatments with p5.F2 knockout mutants that lost their ability to protect the plant (Δ GC00000032_87 // GC00000050_54, Δ GC00000450_7, and Δ GC00003884_12), as analyzed in Supplementary Fig. S5A. $n = 20$ replicates. A second experiment gave similar results, as detailed in Supplementary Fig. S6A, B. **B** In vitro growth of the pathogen, p5.F2 and two knockout mutants (Δ GC00000032_87 // GC00000050_54 and Δ GC00000450_7) as a function of 2,2'-bipyridine concentration, analyzed in six 50 nM increments from 0 to 300 nM. OD_{600} was monitored for 10 h (see also Supplementary Fig. S5 and **Methods**). The logistic area under the growth curve was extracted as a proxy for bacterial growth. The shaded area indicates 95% confidence intervals of the regression curve. $n = 4$ for each bacterial strain grown in the corresponding 2,2'-bipyridine concentration. **C** Daily median of plant green pixels after treatment with control, *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*), mixture of *Pst* and p5.F2, and mixture of *Pst* and each of the three tested p5.F2 knockout mutants, indicated as p5.F2 K.O. Dashed lines indicate two treatments with p5.F2 knockout mutants that lost their ability to protect the plant (Δ GC00000032_87 // GC00000050_54 and Δ GC00000450_7), as analyzed in Supplementary Fig. S5B. $n = 20$ replicates. A second experiment gave similar results, as detailed in Supplementary Fig. S6C, D.

after co-infection was due to the pathogen or to the interaction with the pathogen, rather than the knockout mutants themselves.

Out of the four genes in the three mutants that had lost their ability to protect from pathogens, three adjacent genes are annotated as iron-related: one encodes a TonB_C domain-containing protein (GC00000450_7) [22, 23], one a putative iron (III) dicitrate sensor related to FecR (GC00000032_87) [24], and one a probable RNA polymerase sigma factor related to FecI (GC00000050_54) [25] (SupplementaryTable S2).

The fourth gene, GC00003884_12, is annotated as encoding an 'uncharacterized protein' (Supplementary Table S2). When Δ GC00003884_12 was grown in tubes while shaking, clumps

formed, suggesting the aberrant formation of cellular aggregates. To investigate this further, we grew the wild type and Δ GC00003884_12 overnight, and then let the cell suspension settle on the benchtop. After 1 h, cell sediment had formed in the Δ GC00003884_12 tube, with a clear liquid forming at the top and aggregates at the bottom (Supplementary Fig. S8), while the wild-type cells had stayed in suspension and the culture still had retained its homogenous opaque appearance. This phenotype implies that at least in the tested conditions, GC00003884_12 suppresses cell-cell aggregation and biofilm formation.

To validate the role of GC00000032_87 // GC00000050_54 and GC00000450_7 in iron uptake, we performed an in vitro

iron-deficiency growth assay. The wild-type strain, the $\Delta\text{GC00000032_87}$ // GC00000050_54 and $\Delta\text{GC00000450_7}$ mutants and the pathogen were grown in LB with increasing levels of the iron chelator 2,2'-dipyridyl. We also tested $\Delta\text{GC00003884_12}$ in the same assay, but noted an unusual growth curve shape for this strain, regardless of chelator levels (Supplementary Fig. S9), in agreement with the formation of cell aggregates we had observed (Supplementary Fig. S8). Consequently, we excluded the mutant $\Delta\text{GC00003884_12}$ from further analysis.

Both $\Delta\text{GC00000032_87}$ // GC00000050_54 and $\Delta\text{GC00000450_7}$, as well as the pathogen p4.C9, grew more slowly than the wild-type commensal strain in LB without chelator (Fig. 4B; Supplementary Figs. S9, S10A; mean difference to control: pathogen -4424 AUC [$-5299, -3551$, 95% confidence interval], $\Delta\text{GC00000450_7}$ -3526 AUC [$-4412, -2634$] and $\Delta\text{GC00000032_87}$ // GC00000050_54 $-2,469$ AUC [$-3337, -1587$]). This confirms that both loci have a role in growth that is independent of iron availability. Increasing chelator levels led to a reduction in the growth of the mutants, the pathogen and the wild-type commensals, with the pathogen and $\Delta\text{GC00000450_7}$ being more sensitive to the iron deficiency than the other three strains (Fig. 4B; Supplementary Figs. S9, S10B; mean slope difference to control: pathogen -5.6 AUC [$-10.5, -0.7$, 95% confidence interval], $\Delta\text{GC00000450_7}$ -10.1 AUC [$-15.1, -5.3$] and $\Delta\text{GC00000032_87}$ // GC00000050_54 -0.4 AUC [$-5.3, 4.3$]). These findings provide evidence for the involvement of iron acquisition in plant protection by ATUE2 strains and specifically imply that ATUE2 strains protect the plant by outcompeting the pathogen over iron.

Taken together, these results suggest that ATUE2 members can protect *A. thaliana* from coexisting pathogenic ATUE5 isolates via mechanisms related to iron uptake, and apparently also to biofilm formation. To test whether these mechanisms act specifically against the local ATUE5 strains or whether they are also effective against other *Pseudomonas* pathogens, we co-infected the protective wild-type commensal and the three non-protective mutants with the model pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst*) [26]. *Pst*-like strains were rare in the local *A. thaliana* populations from which ATUE2 was sampled [16], hence it is unknown whether ATUE2-like strains and *Pst*-like strains co-exist. The wild-type commensal protected against *Pst* infection, greatly reducing the effects of *Pst* on plant growth (Fig. 4C; Supplementary Fig. S5B). While this suggests that ATUE2 protective ability is not restricted to the pathogenic lineage ATUE5, the exact modes of protection against the two pathogens seem to differ, since our mutants that had lost protection against the ATUE5 pathogen were still able to protect against *Pst* infection (Fig. 4C; Supplementary Fig. S5B; replicated in an additional experiment, shown in Supplementary Fig. S6C, D).

DISCUSSION

Here, we systematically tested the ability of non-pathogenic *Pseudomonas* strains to protect their natural host from a pathogenic *Pseudomonas* lineage, using a set of commensal and pathogenic isolates that co-exist in nature. We found that plant protection is a common function among non-pathogenic *Pseudomonas* lineages, but with significant differences even between closely related strains. A genome-wide association study (GWAS) pointed to lineage-specific gene sets associated with plant protection. Focusing on a small set of candidate genes, we demonstrated that iron acquisition and biofilm formation appear to play an important role in helping protective *Pseudomonas* strains to mitigate the impact of pathogenic *Pseudomonas* in the phyllosphere.

While protective and non-protective strains are found in several ATUE clades, the fraction of either type varies between lineages. Similar patterns were described in a recent study focused on

bacterial commensals that protect from the effects of *Ralstonia solanacearum*, a tomato pathogen [15]. In addition, a few strains enhanced disease outcomes in co-infection trials. One plausible explanation is that some strains that are closely related to commensals facilitate the activity of the pathogen, as described for fungi that modify the response of *Populus trichocarpa* to the leaf rust pathogen *Melampsora* [27]. Alternatively, some ATUE strains might actually be pathogenic by themselves despite clustering phylogenetically with true commensals. This would not be without precedence, since a pathogenic lineage within a commensal clade has been described for plant-associated *Pseudomonas* [28]. The two scenarios—pathogenesis-facilitating commensals and pathogenic strains within a commensal clade—do not have to be mutually exclusive, and they clearly warrant further investigation to reveal the underlying genes, physiological mechanisms and evolutionary forces.

The involvement of iron uptake suggests competition between commensal and pathogenic *Pseudomonas* over this micronutrient as a potential mechanism for plant protection. This is in agreement with the importance of iron competition in microbial communities [15, 29–32] and the link between iron availability and plant pathogenesis [33, 34]. Iron, an essential micronutrient [35], has limited bioavailability [36], driving competition for iron and evolutionary diversification of iron uptake instruments [37]. Therefore, it is not unexpected that iron competition dictates microbial interactions within the plant, which in turn can affect host health. In particular, TonB receptors, which are involved in siderophore-mediated iron uptake [38], are important vehicles for nutrient uptake in phyllosphere bacteria [39]. This further supports our finding of an iron-uptake TonB receptor gene as responsible for plant protection.

A new link that we discovered is a connection between bacterial biofilm formation and plant protection. Bacterial aggregation affects many bacterial activities, including altered competition with other bacteria [40]. Therefore, it is perhaps not surprising that biofilm formation can influence the competition between commensal and pathogenic *Pseudomonas* in the plant. More insight into the role of biofilms in pathogen protection will come from microscopic imaging of bacterial growth and bacteria-bacteria interactions during infections.

We have used comparative genomics to pinpoint protective genes among various *Pseudomonas* commensal lineages. Our results imply that the underlying protective mechanisms are lineage-specific, based on the annotation of GWAS hits. An additional complexity is that different genetic and physiological mechanisms appear to be responsible for protection from different pathogenic *Pseudomonas* lineage. This observation in turn warrants further investigation of the exact processes by which the natural pathogen ATUE5 pathogen and the model pathogen *Pst* reduce growth and/or cause disease in *A. thaliana*. Our observations underscore the benefits of studying diverse natural pathogens, commensals and mutualists of wild plants. From an applied perspective, more knowledge of the complete set of tools by which plants are being attacked by bacteria and other microbes may also lead to new avenues for how to protect crop plants from their enemies.

With the candidate protective genes that we discovered from GWAS, one can now return to the natural *A. thaliana* populations from with the commensal and pathogenic *Pseudomonas* strains were sampled [16], and investigate the prevalence of these genes, their co-occurrence, and also their impact on levels of colonization by pathogenic *Pseudomonas* in the real world.

METHODS

Plant material

The plant genotype Ey15-2 was originally collected from Eyach (Germany), near the sites in which *Pseudomonas* was isolated [16]. It represents a genetic background common to this region [41]. Seeds were sterilized by

overnight incubation at -80°C , followed by shaking of seeds for 5–15 min in a solution containing 75% EtOH and 0.5% Triton-X-100, washing with 95% EtOH and drying in a laminar flow hood). Seeds were stratified in the dark at 4°C for seven days, on $\frac{1}{2}$ MS-agar media including vitamins and MES buffer (M0255; Duchefa, Germany). Plants were grown in 1.8 ml $\frac{1}{2}$ MS-agar in 24-well plates. All plants were grown on long days (16 h of light) at 23°C , in plant-growth chambers (CLF CU-36L5, CLF Plant Climatics, Germany).

Plant protection assay

Pseudomonas strains were from a local collection [16], except for *Pseudomonas syringae* pv. *tomato* DC3000, which was taken from lab stocks. All bacterial treatments were diluted to a concentration of $\text{OD}_{600} = 0.01$ per strain. Thus, in co-infections with two strains the total bacterial concentration was $\text{OD}_{600} = 0.02$, while in infections with individual strains the total concentration was $\text{OD}_{600} = 0.01$. Bacteria-free control was 10 mM MgSO_4 .

Bacterial inocula were prepared as follows: The relevant isolates were grown overnight in Lysogeny broth (LB) and 10 mg/ml nitrofurantoin, an antibiotic to which our *Pseudomonas* isolates are resistant. Cultures were diluted 1:10 in the following morning and grown for 3 additional hours until they entered the log phase. Subsequently, bacteria were pelleted at 3500 g, resuspended in 10 mM MgSO_4 , pelleted again at 3500 g to remove residual LB, and resuspended again in 10 mM MgSO_4 to a concentration of $\text{OD}_{600} = 0.02$, creating a stock solution for each isolate for subsequent mixtures with either 10 mM MgSO_4 or a solution containing another strain, in both cases resulting in $\text{OD}_{600} = 0.01$ per strain.

All plants were treated with the relevant inocula 10 days after stratification. Infections were done by drip-inoculation, pipetting 100 μl onto the whole rosette, and plates were subsequently sealed using Micropore tape (3 M, Germany). The plants were photographed on the day of infection (before the actual infection), one day post-infection (day 1), and consecutively from day 4 to day 7 post-infection, using a tripod-mounted Canon PowerShot G12 digital camera and built-in flash. Lids were kept closed to maintain sterility, while a backlight was used to avoid reflections due to the flashlight. Green pixel counts were approximated as described [42]. In brief, individual plants were segmented from the background using lab color space thresholds, followed by morphology-based noise removal. Lastly, GrabCut postprocessing was applied, resulting in a list of plant IDs and the corresponding green pixel count. The script was written in Python 3.6, bash using OpenVN 3.1.0 and scikit-image 0.13.0.

Gene knockout in *Pseudomonas*

Genes were deleted in strain p5.F2 [16]. First, 300 bp flanking the relevant genes were extracted from the published p5.F2 genome [16]. Restriction sites (XhoI/AvrII/PmeI/PmlI) were added *in silico* between the two fragments of 300 bp flanking sequences, and the resulting sequence was synthesized and cloned into the pDEST2T18ms vector (<https://www.addgene.org/72647/>) by Twist Bioscience (USA). The full sequence list can be found in Supplementary Table S3.

The Gm^{R} marker for gentamicin resistance was amplified from the template pUC18-mini-Tn7T-Gm-lux (<https://www.addgene.org/64963/>) with primers 5' -ccgagctcatgcatgacg-3' and 5' -ccggacgatcgaattgggg-3' in a 25 μl reaction containing 0.25 μl Q5 high-fidelity DNA polymerase (New England Biolabs, USA), 1x Q5 5x reaction buffer, 0.08 μM forward and 0.16 μM of reverse (tagging) primer and 200 μM dNTP. PCR was carried out for 98°C for 30 s, followed by 30 cycles of 98°C for 15 s, 58°C for 20 s, 72°C for 2 min, and a final extension step of 72°C for 2 min. The amplified fragment was gel-purified (GeneJET Gel Extraction Kit; Thermo Scientific, USA), and ligated into the PmeI site between the synthesized 300 bp flanking regions using T4 DNA-Ligase (Thermo Fisher Scientific, USA), following the standard restriction-ligation protocol as instructed by the manufacturer. The resulting fragment for each gene was 300 bp 5' sequence- Gm^{R} -300 bp 3' sequence, cloned into a pDEST2T18ms vector. DH5 α competent *E. coli* cells were transformed with the ligation product and subsequently plated on selective Lysogeny broth (LB) agar (1.75%) with gentamicin (10 $\mu\text{g}/\text{ml}$) and tetracycline (5 ng/ml). Bacterial colonies were picked and used as a template for colony PCR using primers matching the foreign DNA (primers in Supplementary Table S4) to identify successful transformants. PCR reactions of 25 μl contained 0.25 μl Q5 high-fidelity DNA polymerase (New England Biolabs, USA), 1x Q5 5x reaction buffer, 0.08 μM forward and 0.16 μM of reverse (tagging) primer and 200 μM dNTP. PCR was carried out for 98°C for 30 s, followed by 30 cycles of 98°C for 30 s, appropriate

annealing temperature for the primers used (Supplementary Table S4) for 30 s, 72°C for 2 min, and a final extension at 72°C for 2 min. The same conditions were used for all following colony PCRs.

Recipient strain p5.F2 was subjected to triparental mating with positive DH5 α donors, carrying the pDEST2T18ms with the relevant insert for each gene, and HB101 *E. coli*, carrying the plasmid pRK2013 that facilitates mobilization, as helper strain. Cultures of the donor, recipient and helper strains were grown overnight with shaking, at 37°C for donor and helper, and 28°C for the recipient, in LB with the relevant antibiotics (10 $\mu\text{g}/\text{ml}$ gentamicin and 5 ng/ml tetracycline for the donor, 100 $\mu\text{g}/\text{ml}$ nitrofurantoin for the recipient and 50 $\mu\text{g}/\text{ml}$ kanamycin for the helper). The following morning, cultures were diluted 1:10 and grown for 2–4 additional hours. The three strains were then mixed in a single tube, followed by immediate centrifugation (3000 g), removal of supernatant and resuspension in 1 ml LB. The washing procedure was repeated, and cells were resuspended in 100 μl LB. The cell mixture was inoculated onto LB plates to allow for mating and incubated for 72 h at 28°C . The resulting cell mass was scraped off into 500 μl 10 mM MgSO_4 and centrifuged for 2 min at 3500 g. The supernatant was removed, and the cells were resuspended in 1 ml of 10 mM MgSO_4 . A 100 μl aliquot of the resulting mixture was plated onto LB-agar selective media including 100 $\mu\text{g}/\text{ml}$ nitrofurantoin, gentamicin 10 $\mu\text{g}/\text{ml}$ and tetracycline 5 ng/ml. The plates were incubated at 28°C for 2–3 days until visible colonies appeared. Colony-PCR (primers in Supplementary Table S4) was used to identify successful p5.F2 transformants, and to validate that the native gene was replaced by the Gm^{R} cassette by homologous recombination. Successful knockout was deduced from the size of the PCR product (Supplementary Table S4), the amplified fragment was gel-purified (GeneJET Gel Extraction Kit; Thermo Scientific, USA) and Sanger-sequenced to confirm the presence of the expected Gm^{R} marker. Positive transformants were plated on media with 5% sucrose and 10 $\mu\text{g}/\text{ml}$ gentamicin, to eliminate the pDEST2T18ms plasmid, which carries the sucrose counterselection marker SacB. Loss of SacB was confirmed by colony-PCR (same reaction and conditions as for Gm^{R} PCR, primers in Supplementary Table S4) and gel electrophoresis, and re-examined for the absence of the native gene and for the presence of Gm^{R} .

As the knockout mutants are Gm^{R} , we also generated a Gm^{R} wild-type derivative, ensuring a similar metabolic burden as in the mutants. *Pseudomonas* strain p5.F2 was transformed with the plasmid pUC18T-mini-Tn7T-Gm-lux (<https://www.addgene.org/64963/>), using the mini-Tn7 transformation protocol [43]. In brief, bacteria went through several sucrose washes, and subsequently were transformed with the plasmid by electroporation, together with the helper plasmid pTNS2 (<https://www.addgene.org/64968/>). Insertion is a single copy into a locus considered as neutral [43].

Iron sensitivity assay

Growth sensitivity to iron levels was assayed as described [44]. Briefly, each strain was grown in 200 μl LB in the presence of increasing 2,2'-dipyridyl concentration, ranging from 0 to 300 nM. Bacteria were diluted to a starting concentration of $\text{OD}_{600} = 0.05$ in 96-well plates (Greiner Bio One, Austria). Plates were incubated in a plate reader (Robot Tecan Infinite M200; Tecan Life Sciences, Switzerland) at 28°C while shaking for 10 h, and OD_{600} was measured in 30 min intervals.

Phylogenetic analysis

All maximum likelihood phylogenies were based on the core genome (ortholog presence > 70%) of the relevant strains in the analysis, constructed with IQtree (v1.6.10; using the parameters: -mset LG -st AA -nt 16 -ntmax 16) [45]. Trees were visualized with iTOL [46], except for the combined tree-heatmap plots, in which the R function *ggtree* from the package *ggtree* was used [47].

Regression analysis

Posterior distributions of the relevant predictors were approximated using the function *stan_glm* in the R package *rstanarm* [48]. Default settings including prior distribution were used, unless more iterations were required (Rhat > 1.1). The number of iterations was increased until a sufficient number was reached (Rhat < 1.1). In all figures, the median and 95% credible intervals from 2.5% to 97.5% of the posterior distribution are presented. The response variable and predictors, as well as the reference for comparison, are explained in the figure legend for each analysis.

Bacterial genome-wide association with plant protection

Gene orthology clusters as well as the phyletic pattern of gene presence/absence (P/A) for each strain used in this study, were from [16]. Associations between gene orthologs and plant growth used treeWAS [19], a tool for bacterial GWAS. Phyletic pattern and phenotypic measurements for all strains can be found in Supplementary Data 1. The core-genome-based phylogeny used for treeWAS can be found in Supplementary Data 2.

Four different sets of strains were used, in four separate analyses: All strains; ATUE2; ATUE3; ATUE4. In each set, the resulting maximum likelihood phylogeny was used in treeWAS. To gain higher sensitivity, each set of strains was analyzed using four different metrics of plant growth. All significant hits per set of strains were combined, and the number of tests in which an ortholog resulted as significant was counted (Supplementary Table S2). The four plant growth metrics that we used were median change in green pixels between the end of the experiment 7 days post-infection and the day of infection (median growth); a Bayesian-derived approximation of median growth ($cdl50$, the posterior distribution median of median growth); a binary categorization based on $cdl50$ ($cdl < 0$, i.e., zero growth or decline, and $cdl > 0$, i.e., any amount of growth); and the area under the curve while accounting for all sampled time points (0–1, 4–7 days post-infection), as the value of auc_I after running the R function *growthcurver* [49]. The four metrics are detailed in the legend of Supplementary Table S2.

Heatmaps were produced and combined with phylogenetic trees using the R function *ghetmap* from the package *ggtree* [47].

To test for physical proximity between significant genes, that is, whether they potentially formed genomic islands, we examined the location of the corresponding genes in strains that had all of these genes. We accounted for gene ID provided by Prokka [50], which includes the order of genes within a contig. For example, *gene_1* and *gene_3* are separated on contig_X only by *gene_2*. Only genes that were found on the same contig and that were separated by no more than one gene were considered as consecutive. All genes behaved in a congruent manner across strains examined. We further validated the proximity of genes by counting the number of bases differing from each consecutive pair of genes, confirming that they were at most hundreds of bp apart.

Gene ortholog annotation

To assign the potential function to each of the proteins encoded by the studied genes we first used the multiple sequence alignment of the protein orthology group to build a representative HMM model using HMMER (version 3.2.1). Next, the resulting HMMs were used to search for homologous proteins using *hmmsearch* against the UniProt Knowledgebase (Swiss-Prot+TrEMBL); release 2020_05. The function of the best significant hit, which covered at least 80% of the query protein, was transferred to the query protein. Annotations of each orthology group can be found in Supplementary Data 3.

Analysis of growth in iron-sensitivity assay

The growth of all bacteria was analyzed using the R function *SummarizeGrowthByPlate* from the *Growthcurver* R package [49]. The logarithmic area under the curve, auc_I , was extracted and compared between the wild-type, mutant strains of interest and the focal pathogen.

Statistical analysis

Statistical analyses were performed using R (version 3.5.1), unless mentioned otherwise.

DATA AVAILABILITY

Knockout mutants are freely available from the authors.

CODE AVAILABILITY

The code for all analyses including the GWAS, is available at https://github.com/orshalevsk/GWAS_project.

REFERENCES

- Galloway BT. Plant pathology; a review of the development of the Science in the United States. *Agric Hist.* 1928;2:49–60.

- Strange RN, Scott PR. Plant disease: a threat to global food security. *Annu Rev Phytopathol.* 2005;43:83–116.
- Dangl JL, Horvath DM, Staskawicz BJ. Pivoting the plant immune system from dissection to deployment. *Science.* 2013;341:746–51.
- Hacquard S, Spaepen S, Garrido-Oter R, Schulze-Lefert P. Interplay between innate immunity and the plant microbiota. *Annu Rev Phytopathol.* 2017;55:565–89.
- Durán P, Thiergart T, Garrido-Oter R, Agler M, Kemen E, Schulze-Lefert P, et al. Microbial interkingdom interactions in roots promote *Arabidopsis* survival. *Cell.* 2018;175:973–83.e14.
- Berendsen RL, Pieterse CMJ, Bakker PAHM. The rhizosphere microbiome and plant health. *Trends Plant Sci.* 2012;17:478–86.
- Compant S, Samad A, Faist H, Sessitsch A. A review on the plant microbiome: ecology, functions, and emerging trends in microbial application. *J Advert Res.* 2019;19:29–37.
- Haas D, Défago G. Biological control of soil-borne pathogens by fluorescent *Pseudomonas*. *Nat Rev Microbiol.* 2005;3:307–19.
- Innerebner G, Knief C, Vorholt JA. Protection of *Arabidopsis thaliana* against leaf-pathogenic *Pseudomonas syringae* by *Sphingomonas* strains in a controlled model system. *Appl Environ Microbiol.* 2011;77:3202–10. https://link.springer.com/chapter/10.1007%2F978-3-030-41870-0_20.
- Kamle M, Borah R, Bora H, Jaiswal AK, Singh RK, Kumar P. Systemic acquired resistance (SAR) and induced systemic resistance (ISR): Role and mechanism of action against phytopathogens. In: Hesham AE-L, Upadhyay RS, Sharma GD, Manohararachary C, Gupta VK, editors. *Fungal Biotechnology and Bioengineering*. Cham: Springer International Publishing; 2020. pp. 457–70.
- Spadaro D, Drobny S. Development of biocontrol products for postharvest diseases of fruit: the importance of elucidating the mechanisms of action of yeast antagonists. *Trends Food Sci Technol.* 2016;47:39–49.
- Raaijmakers JM, Sluis L van der, Bakker PAHM, Schippers B, Koster M, Weisbeek PJ. Utilization of heterologous siderophores and rhizosphere competence of fluorescent *Pseudomonas* spp. *Can J Microbiol.* 1995;41:126–35.
- Raaijmakers JM, Mazzola M. Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annu Rev Phytopathol.* 2012;50:403–24.
- Bernal P, Allsopp LP, Filloux A, Llamas MA. The *Pseudomonas putida* T6SS is a plant warden against phytopathogens. *ISME J* 2017;11:972–87.
- Gu S, Wei Z, Shao Z, Friman V-P, Cao K, Yang T, et al. Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. *Nat Microbiol.* 2020;5:1002–10.
- Karasov TL, Almario J, Friedemann C, Ding W, Giolai M, Heavens D, et al. *Arabidopsis thaliana* and *Pseudomonas* pathogens exhibit stable associations over evolutionary timescales. *Cell Host Microbe.* 2018;24:168–79.e4.
- Shalev O, Karasov TL, Lundberg DS, Ashkenazy H, Weigel D. Protective host-dependent antagonism among *Pseudomonas* in the *Arabidopsis* phyllosphere. *bioRxiv.* 2021. p. 2021.04.08.438928. <https://doi.org/10.1101/2021.04.08.438928>
- Humphrey PT, Satterlee TT, Whiteman NK. Competitive hierarchies, antibiosis, and the distribution of bacterial life history traits in a microbiome. *bioRxiv.* 2020. p. 2020.06.15.152272. <https://doi.org/10.1101/2020.06.15.152272>
- Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. *PLoS Comput Biol.* 2018;14:e1005958.
- Didelot X, Lawson D, Darling A, Falush D. Inference of homologous recombination in bacteria using whole-genome sequences. *Genetics* 2010;186:1435–49.
- Segev-Zarko L-A, Kapach G, Josten M, Klug YA, Sahl H-G, Shai Y. Deficient lipid A remodeling by the *arnB* gene promotes biofilm formation in antimicrobial peptide susceptible *Pseudomonas aeruginosa*. *Biochemistry.* 2018;57:2024–34.
- Shultis DD, Purdy MD, Banchs CN, Wiener MC. Outer membrane active transport: structure of the BtuB:TonB complex. *Science.* 2006;312:1396–9.
- Braun V, Hantke K. Recent insights into iron import by bacteria. *Curr Opin Chem Biol.* 2011;15:328–34.
- Van Hove B, Staudenmaier H, Braun V. Novel two-component transmembrane transcription control: regulation of iron dicitrate transport in *Escherichia coli* K-12. *J Bacteriol.* 1990;172:6749–58.
- Mahren S, Braun V. The Fecl extracytoplasmic-function sigma factor of *Escherichia coli* interacts with the beta' subunit of RNA polymerase. *J Bacteriol.* 2003;185:1796–802.
- Xin X-F, Kvitko B, He SY. *Pseudomonas syringae*: what it takes to be a pathogen. *Nat Rev Microbiol.* 2018;16:316–28.
- Busby PE, Peay KG, Newcombe G. Common foliar fungi of *Populus trichocarpa* modify *Melampsora* rust disease severity. *N. Phytol.* 2016;209:1681–92.
- Melnyk RA, Hossain SS, Haney CH. Convergent gain and loss of genomic islands drive lifestyle changes in plant-associated *Pseudomonas*. *ISME J.* 2019;13:1575–88.

29. Cordero OX, Ventouras L-A, DeLong EF, Polz MF. Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. *Proc Natl Acad Sci USA*. 2012;109:20059–64.
30. Andersen SB, Marvig RL, Molin S, Krogh Johansen H, Griffin AS. Long-term social dynamics drive loss of function in pathogenic bacteria. *Proc Natl Acad Sci USA*. 2015;112:10756–61.
31. Barber MF, Elde NC. Buried treasure: evolutionary perspectives on microbial iron piracy. *Trends Genet*. 2015;31:627–36.
32. Kramer J, Özkaya Ö, Kümmerli R. Bacterial siderophores in community and host interactions. *Nat Rev Microbiol*. 2020;18:152–63.
33. Expert D, Franza T, Dellagi A. Iron in plant–pathogen interactions. In: Expert D, O'Brian MR, editors. *Molecular aspects of iron metabolism in pathogenic and symbiotic plant-microbe associations*. Dordrecht: Springer Netherlands; 2012. pp. 7–39.
34. Aznar A, Chen NWG, Thomine S, Dellagi A. Immunity to plant pathogens and iron homeostasis. *Plant Sci*. 2015;240:90–97.
35. Andrews SC, Robinson AK, Rodríguez-Quinones F. Bacterial iron homeostasis. *FEMS Microbiol Rev*. 2003;27:215–37.
36. Emerson D, Roden E, Twining BS. The microbial ferrous wheel: iron cycling in terrestrial, freshwater, and marine environments. *Front Microbiol*. 2012;3:383.
37. Niehus R, Picot A, Oliveira NM, Mitri S, Foster KR. The evolution of siderophore production as a competitive trait. *Evolution*. 2017;71:1443–55.
38. Moeck GS, Coulton JW. TonB-dependent iron acquisition: mechanisms of siderophore-mediated active transport. *Mol Microbiol*. 1998;28:675–81.
39. Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R, et al. Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc Natl Acad Sci*. 2009;106:16428–33.
40. Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. Biofilms: an emergent form of bacterial life. *Nat Rev Microbiol*. 2016;14:563–75.
41. Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, et al. Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. *PLoS Genet*. 2010;6:e1000890.
42. Karasov TL, Neumann M, Duque-Jaramillo A, Kersten S, Bezrukov I, Schröppel B, et al. The relationship between microbial biomass and disease in the *Arabidopsis thaliana* phyllosphere. *Cold Spring Harbor Laboratory*. 2019. p. 828814. <https://doi.org/10.1101/828814>
43. Choi K-H, Schweizer HP. mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. *Nat Protoc*. 2006;1:153–61.
44. Watson RJ, Millichap P, Joyce SA, Reynolds S, Clarke DJ. The role of iron uptake in pathogenicity and symbiosis in *Photobacterium luminescens* TT01. *BMC Microbiol*. 2010;10:177.
45. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32:268–74.
46. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res*. 2016;44:W242–5.
47. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *McInerney G, editor. Methods Ecol Evol*. 2017;8: 28–36.
48. Goodrich B, Gabry J, Ali I, Brilleman S. rstanarm: Bayesian applied regression modeling via Stan. 2020. <https://cran.r-project.org/web/packages/rstanarm/citation.html>.
49. Sprouffske K, Wagner A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinforma*. 2016;17:172.
50. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9.

AUTHOR CONTRIBUTIONS

OS conceived the research, performed the experiments, and analyzed the results; MN generated the knockout mutants; DW supervised the research; OS, HA, and DW discussed and interpreted the results of this study; OS wrote the first draft of the manuscript. The final draft was written by OS and DW with input from all authors.

FUNDING INFORMATION

This work was supported by fellowships from DAAD (OS) and Alexander von Humboldt Foundation (HA), and by DFG SPP2125 DECrypT and the Max Planck Society. Open Access funding enabled and organized by Projekt DEAL.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41396-021-01168-6>.

Correspondence and requests for materials should be addressed to Detlef Weigel.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021