Experimental-Evolution-Driven Identification of *Arabidopsis* Rhizosphere Competence Genes in *Pseudomonas protegens*

Erqin Li,[a]* Hao Zhang,[a] Henan Jiang,[a] Corné M. J. Pieterse,[a] Alexandre Jousset,[b] Peter A. H. M. Bakker,[a] Ronnie de Jonge[a]

[a]Plant-Microbe Interactions, Department of Biology, Science4Life, Utrecht University, Utrecht, The Netherlands
[b]Ecology and Biodiversity, Department of Biology, Science4Life, Utrecht University, Utrecht, The Netherlands

Erqin Li and Hao Zhang contributed equally to this article. The author order was determined by contributions to writing the manuscript.

ABSTRACT Beneficial plant root-associated microorganisms carry out a range of functions that are essential for plant performance. Establishment of a bacterium on plant roots, however, requires overcoming several challenges, including competition with neighboring microorganisms and host immunity. Forward and reverse genetics have led to the identification of mechanisms that are used by beneficial microorganisms to overcome these challenges, such as the production of iron-chelating compounds, the formation of strong biofilms, or the concealment of characteristic microbial molecular patterns that trigger the host immune system. However, how such mechanisms arose from an evolutionary perspective is much less understood. To study bacterial adaptation in the rhizosphere, we employed experimental evolution to track the physiological and genetic dynamics of root-dwelling *Pseudomonas protegens* in the *Arabidopsis thaliana* rhizosphere under axenic conditions. This simplified binary one plant/one bacterium system allows for the amplification of key adaptive mechanisms for bacterial rhizosphere colonization. We identified 35 mutations, including single-nucleotide polymorphisms, insertions, and deletions, distributed over 28 genes. We found that mutations in genes encoding global regulators and in genes for siderophore production, cell surface decoration, attachment, and motility accumulated in parallel, underlining the finding that bacterial adaptation to the rhizosphere follows multiple strategies. Notably, we observed that motility increased in parallel across multiple independent evolutionary lines. All together, these results underscore the strength of experimental evolution in identifying key genes, pathways, and processes for bacterial rhizosphere colonization and a methodology for the development of elite beneficial microorganisms with enhanced root-colonizing capacities that can support sustainable agriculture in the future.

IMPORTANCE Beneficial root-associated microorganisms carry out many functions that are essential for plant performance. Establishment of a bacterium on plant roots, however, requires overcoming many challenges. Previously, diverse mechanisms that are used by beneficial microorganisms to overcome these challenges were identified. However, how such mechanisms have developed from an evolutionary perspective is much less understood. Here, we employed experimental evolution to track the evolutionary dynamics of a root-dwelling pseudomonad on the root of *Arabidopsis*. We found that mutations in global regulators, as well as in genes for siderophore production, cell surface decoration, attachment, and motility, accumulate in parallel, emphasizing these strategies for bacterial adaptation to the rhizosphere. We identified 35 mutations distributed over 28 genes. All together, our results demonstrate the power of experimental evolution in identifying key pathways for rhizosphere colonization and a methodology for the development of elite beneficial microorganisms that can support sustainable agriculture.

Address correspondence to Erqin Li, erqinli22@gmail.com, or Ronnie de Jonge, r.dejonge@uu.nl.

* Present address: Erqin Li, Freie Universität Berlin, Institut für Biologie, Dahlem Center of Plant Sciences, Plant Ecology, Berlin, Germany.

Plants are associated with complex microbial communities assembled into a functional microbiome that safeguards optimal plant performance under harsh environmental conditions (1). The rhizosphere is a particularly interesting hot spot of plant-microbe interactions. Plants deposit up to 44% of their photosynthetically fixed carbon into the rhizosphere, fueling a specific microbial community (2). The microbial species pool in the bulk soil is the source from which the root microbiome is recruited, and plant genotype, immune responses, and environmental factors are postulated to affect this process (3–6). The establishment of beneficial microbial associations requires a high degree of coordination of both plant and microbial responses by means of a continuous molecular dialogue (7, 8). Plant-associated microorganisms can improve plant yield by protecting the plant from abiotic stresses (9), improving plant nutrition and growth (10–12), antagonizing soilborne pathogens (13), or stimulating plant immunity (14). To exert their beneficial effects on plant performance, bacteria must colonize the roots efficiently and establish significant populations. For example, bacterial population densities above 10^5 cells per gram of root are required for efficient suppression of soilborne plant pathogens by *Pseudomonas* spp. (15, 16). Therefore, bacterial adaptation to the plant root environment may be essential for the successful implementation of microbiome services in agriculture in order to support plant health.

Among all root-dwelling organisms, fluorescent *Pseudomonas* spp. are well characterized in terms of the traits required for their growth in the rhizosphere and for the establishment of beneficial plant-microbe interactions (17, 18). To study bacterial traits involved in efficient root colonization, mutants defective in specific traits suspected of being involved in colonization are compared to the parental strains for their ability to colonize plant roots. Such studies have highlighted a range of bacterial traits involved in efficient root colonization, including flagella (19), surface lipopolysaccharides (LPS) (20), and amino acid synthesis (21). Using random mutagenesis and by determining the fitness of each mutant in competition with its parental strain in the rhizosphere, many other important traits for rhizosphere competence in *Pseudomonas* were discovered (17). Recently, random mutagenesis in *Pseudomonas capeferrum* WCS358 led to the identification of two genes that are important for gluconic acid (GA) biosynthesis. GA, in turn, is essential for the suppression of local, flagellin-induced root immune responses (22). Such suppression was shown to be important for rhizosphere competence, as GA-deficient mutants maintain reduced populations in the rhizosphere (22). In another recent study, genome-wide saturation mutagenesis of *Pseudomonas simiae* WCS417r revealed that 2% of the protein-coding genes are important for successful root colonization (23). Mutations that negatively affect rhizosphere competence and mutations that confer a competitive advantage were identified in this study. The identification of mutations that can lead to increased root colonization (23) suggests that there is room for improvement of bacterial fitness in the rhizosphere.

In the present study, we used an experimental-evolution approach (24) to study how *Pseudomonas protegens* CHA0 evolves during repeated colonization of the rhizosphere of the model plant *Arabidopsis thaliana*. The model biological control agent CHA0 displays broad-spectrum antagonistic activity against several plant-pathogenic fungi and bacteria (11), and its complete genome sequence is available (25). We performed highly controlled experimental evolution in a gnotobiotic and carbon-free sand system in which bacteria depend solely on the plant for their supply of carbon. Following inoculation and establishment on the roots, bacterial populations were transferred to new plants, and this cycle was repeated eight times. We hypothesized that the repeated colonization and establishment of the bacterial population on the plant root would create an environment in which selection pressure drives the accumulation of better colonizers. *In vitro* characterization of individual bacterial colonies from these populations combined with sequencing analysis led to the identification of

several evolutionary trajectories involving 35 distinct mutations that impact social traits representing interpopulation communication and cooperation, carbon source utilization, motility, or biocontrol activity. By combining experimental evolution with whole-genome resequencing, we created a powerful screen for the identification of adaptive mutations with positive effects on rhizosphere colonization.

RESULTS

Mutational events in independent evolutionary lines. We previously studied five experimental-evolutionary populations, referred to as lines, of CHA0 evolving in the rhizosphere of *Arabidopsis* in a gnotobiotic system. Independent populations were introduced on the roots, and after 4 weeks of plant growth, the populations were transferred to new plants. This cycle of transferring was repeated eight times, and we performed extensive characterizations up until cycle 6 to account for feasibility (26). In short, for each line, after every cycle, we plated a fraction of the population on culture media and randomly picked 16 colonies for extensive phenotypic assessment of bacterial life history traits (26). In order to study adaptation at the genetic level, we selected six colonies from each line at cycles 2, 4, and 6, such that they represented most of the observed phenotypic diversity among the 16 colonies that were initially picked and characterized. These colonies, as well as six colonies from the ancestral population that was initially introduced on the roots, were resequenced to an average depth of 25-fold coverage (minimum, 10; maximum, 70) and used for the identification of single nucleotide polymorphisms (SNPs), as well as small and large insertions or deletions (indels). In total, we thus set out to acquire genetic data for 96 bacterial colonies (5 lines × 3 cycles × 6 colonies + 6 ancestral colonies). Unfortunately, we were unable to acquire sufficient sequencing data for two colonies from line 4 at cycle 4, yielding a final set of 94 (88 evolved, 6 ancestral) resequenced colonies. The six ancestral colonies were all identical, indicating that there was no genetic variation in the starting population and that all observed mutations were *de novo* mutations. In total, one or more mutations were detected in 64 evolved colonies, which collectively represent 73% of the 88 characterized bacterial colonies (see Table S1 in the supplemental material). We identified 5 synonymous substitutions, 20 nonsynonymous substitutions, and 4 deletions ranging in length from 1 bp to about 400 bp, distributed over 22 genes, and 6 additional mutational events located in intergenic regions (Table 1). Mutations located in intergenic regions possibly affect the transcription of nearby genes via affecting their regulatory protein binding sites and making subsequent changes to their promoter activity (27, 28). Several mutations were found to be clustered in select genes and/or regions in the CHA0 genome, e.g., those in the response regulator *gacA* gene (PFLCHA0_RS17965; GenBank accession no. NC _021237, 4,039,113–4,039,754) or in the *OBC3* gene cluster involved in LPS biosynthesis (accession no. NC_021237, 2,173,707–2,196,028) (29), but the majority of the mutations were spread across the 6.1-Mbp CHA0 genome (Fig. 1). Functional characterization of the mutated genes by analyzing their cluster of orthologous groups (COG) annotation revealed that the majority of these genes are involved in transcription (COG term K), signal transduction mechanisms (COG term T), amino acid transport and metabolism (COG term E), and cell wall/membrane/envelope biogenesis (COG term M) (Fig. 1; Table 1).
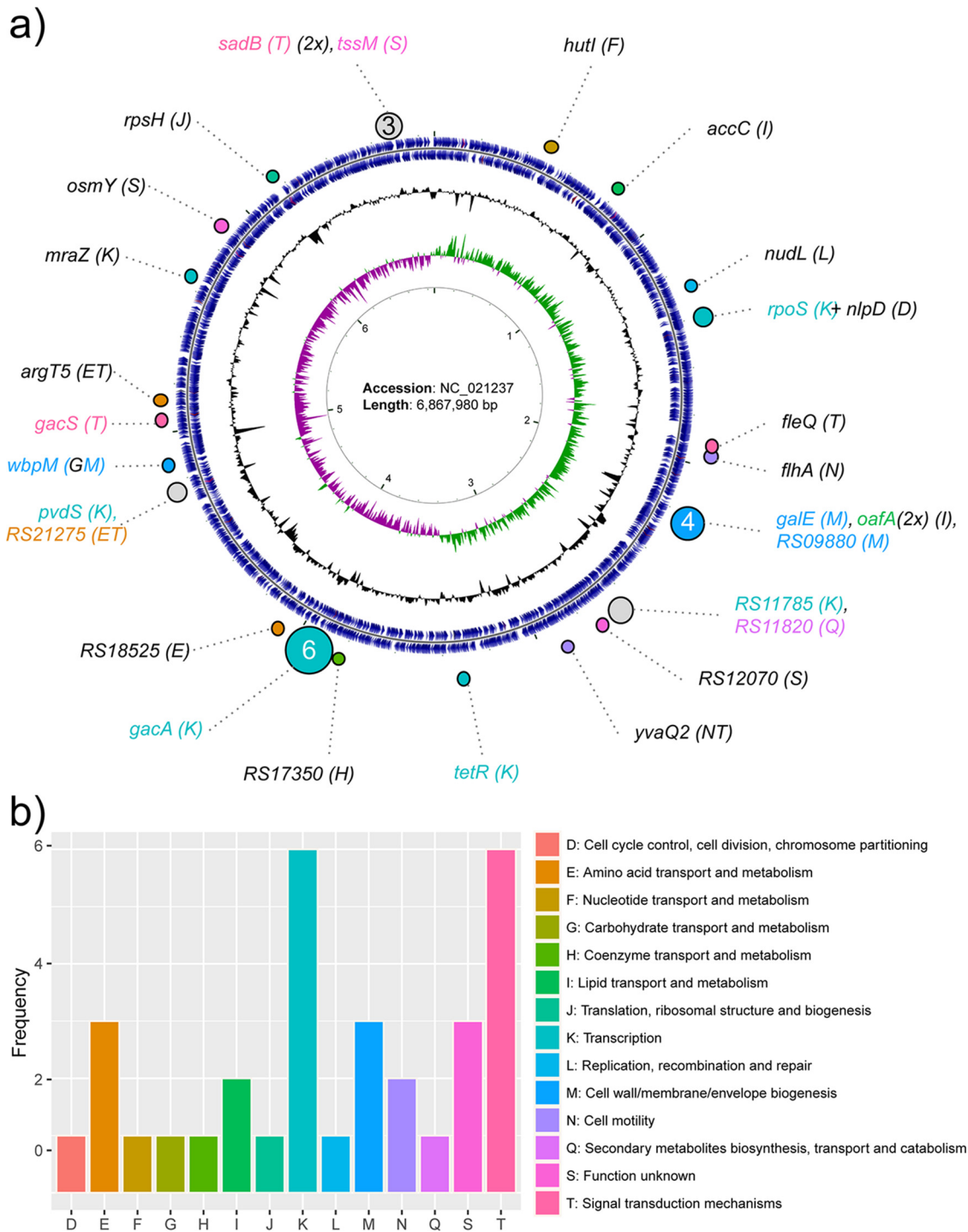
Identification of potential root colonization genes. Bacterial genes that are involved in the colonization of plant roots can be revealed by identifying beneficial mutations that evolve during adaption of bacteria to the rhizosphere environment and have positive effects on root colonization. Over time, such mutants will outcompete the ancestral strain and become dominant in the bacterial population. The observation that only a limited number of mutations accumulated relative to the total number of genes within the genome across the entire experiment makes it highly unlikely that the same gene would acquire several changes by chance in independent evolutionary lines. Nevertheless, we observed recurrent mutations in several of the same genes and/or pathways (Table 1), which is a strong indication for adaptive evolution. Genes or pathways that acquired mutations in multiple independent CHA0 populations

mBio®

TABLE 1 Mutations that occurred in 28 genes during *Arabidopsis* rhizosphere adaptation over five replicate CHA0 populations (lines)

Locus tag[a]	Gene name	Description	COG[b]	Line(s)	Total no. of alleles	Nonsynonymous	Synonymous	Deletions	No. of intergenic mutations
						No. of mutations or deletions in protein-coding genes			
PFLCHA0_RS02080	hutI	Imidazolonepropionase	F	4	1	0	1	0	0
PFLCHA0_RS03400	accC	Biotin carboxylase, acetyl-CoA carboxylase subunit	I	2	1	1	0	0	0
PFLCHA0_RS05510	nudL	Coenzyme A pyrophosphatase/nudix hydrolase NudL	L	4	1	0	1	0	0
PFLCHA0_RS06125	rpoS	RNA polymerase sigma factor RpoS	K	5	1	1	0	0	0
PFLCHA0_RS08340	fleQ	Sigma 54-dependent Fis family transcriptional regulator	T	5	1	1	0	0	0
PFLCHA0_RS08490	flhA	Flagellar biosynthesis protein FlhA	N	4	1	0	0	1	0
PFLCHA0_RS09880	RS09880	Glycosyltransferase (GT)	M	5	1	1	0	0	0
PFLCHA0_RS09890	oafA	O-antigen acetylase	I	1, 3	2	1	0	1	0
PFLCHA0_RS09920	galE	UDP-glucose 4-epimerase	M	2	1	1	0	0	0
PFLCHA0_RS11785	RS11785	LysR family transcriptional regulator	K	4	1	1	0	0	0
PFLCHA0_RS11820	RS11820	PaaI family thioesterase	Q	4	1	0	1	0	0
PFLCHA0_RS12070	RS12070	2OG-Fe(II) oxygenase superfamily	S	5	1	0	1	0	0
PFLCHA0_RS13000	yvaQ2	Methyl-accepting chemotaxis protein	NT	2	1	0	0	0	1
PFLCHA0_RS14960	tetR	TetR/AcrR family transcriptional regulator	K	5	1	1	0	0	0
PFLCHA0_RS17350	RS17350	Methyltransferase domain-containing protein	H	1	1	0	0	1	0
PFLCHA0_RS17965	gacA	UvrY/SirA/GacA family response regulator transcription factor	K	1, 2, 4	6	5	0	0	1
PFLCHA0_RS18525	RS18525	ABC transporter substrate-binding protein	E	5	1	1	0	0	0
PFLCHA0_RS21265	pvdS	RNA polymerase factor sigma 70	K	1	1	0	0	0	1
PFLCHA0_RS21275	RS21275	Transporter substrate-binding domain-containing protein	ET	1	1	1	0	0	0
PFLCHA0_RS21855	wbpM	Polysaccharide biosynthesis protein/NDP-sugar epimerase	GM	1	1	1	0	0	0
PFLCHA0_RS22600	gacS	Hybrid sensor histidine kinase/response regulator	T	3	1	1	0	0	0
PFLCHA0_RS22950	argT5	ABC transporter substrate-binding protein/lysine-arginine-ornithine-binding periplasmic protein ArgT	ET	4	1	0	0	1	0
PFLCHA0_RS25175	mraZ	Division/cell wall cluster transcriptional repressor MraZ	K	2	1	0	0	0	1
PFLCHA0_RS26215	osmY	Osmotically inducible protein OsmY/BON domain-containing protein	S	2	1	0	0	0	1
PFLCHA0_RS27515	rpsH	30S ribosomal protein S8	J	4	1	0	0	0	1
PFLCHA0_RS30075	sadB	Surface attachment defective (SadB) ortholog/HDOD domain-containing protein	T	1	2	2	0	0	0
PFLCHA0_RS30120	tssM	Type VI secretion system membrane subunit TssM	S	2	1	0	1	0	0
PFLCHA0_RS31060	nlpD	Lipoprotein NlpD/LppB/LysM domain-containing protein	D	3	1	1	0	0	0
Total	28				35	20	5	4	6

FIG 1 Genomic distribution of *Pseudomonas protegens* CHA0 evolutionary adaptations. (a) Rings from inside to outside: 1, nucleotide position indicator; 2, green/purple GC skew (−/+); 3, %GC; 4, protein-coding genes; 5, distribution of identified mutations arisen during the evolutionary experiment. Functional annotations of mutated genes are indicated by colors representing clusters of orthologous groups (COGs), the key for which can be seen in panel b. (b) Frequency of mutations per COG class, highlighting enrichment in the classes T (signal transduction), K (transcription), M (cell wall), and E (amino acid transport and metabolism).

included *gacA* and *gacS*, the *OBC3* gene cluster (*oafA*; *galE*; *PFLCHA0_RS09880* [29]), and a putative pyoverdine siderophore biosynthesis cluster (*pvdS*; *PFLCHA0_RS21275*). Other genes, like *sadB*, were targeted more than once in the same population. Because these genes were repeatedly identified in the CHA0 evolution experiment, they can be
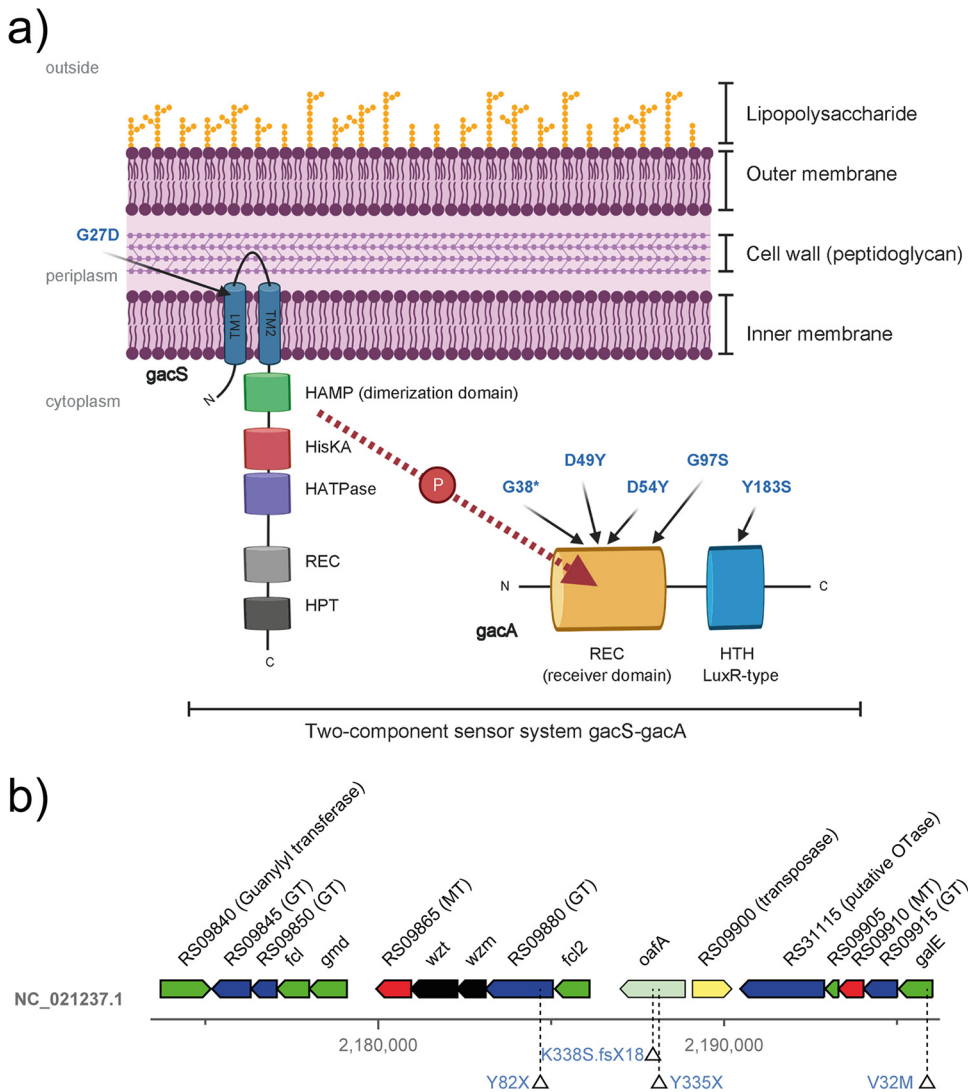
FIG 2 Localization of *gac* and *OBC3* mutations identified in the evolutionary experiment. (a) Schematic view of the cell wall of the Gram-negative bacterium *Pseudomonas protegens* highlighting the presence of the lipopolysaccharide (LPS) on the outside and the presence of the two-component regulator system GacS/GacA on the inner membrane and in the cytoplasm. The arrows indicate the locations of the amino acid substitutions that were identified in this study in GacS (G27D) and GacA (G38X, D49Y, D54Y, G97S, and Y183S). Note the travel of a phosphate group (P) from GacS to GacA upon signal perception, which is accepted by either the Asp49 or Asp54 residue in GacA, both of which are found mutated in two different mutants. (b) Genomic region harboring the *OBC3* gene cluster responsible for the synthesis of long O-polysaccharide (O-PS) on the LPS and indicated mutations that were identified in a glycosyl transferase (RS09880; Y82X), the O-antigen acetyltransferase *oafA* (K338S.fsX18 and Y335X), and the UDP-glucose 4-epimerase (*galE*; V32M).

assumed to contribute significantly to bacterial fitness in the rhizosphere (Table 1; Fig. 2). Moreover, our findings suggest that the independent evolutionary lines converge on similar evolutionary trajectories involving overlapping biological processes and molecular mechanisms.

To illustrate the diverse evolutionary trajectories, we constructed phylogenetic trees of each replicate population (line) using all mutations that were detected (Fig. 3). In these lines, both the numbers and the depths of branches were very different (Fig. 3; Table 2). In lines 1 and 3, the populations were swept early by *oafA* mutants and later by specific *gac* mutants, while in lines 2 and 4, the coexistence of multiple genotypes characterized by *gac* mutations was observed, indicative of clonal interference between lineages (Fig. 3) (30). Notably, the relative abundance of each genotype in the population
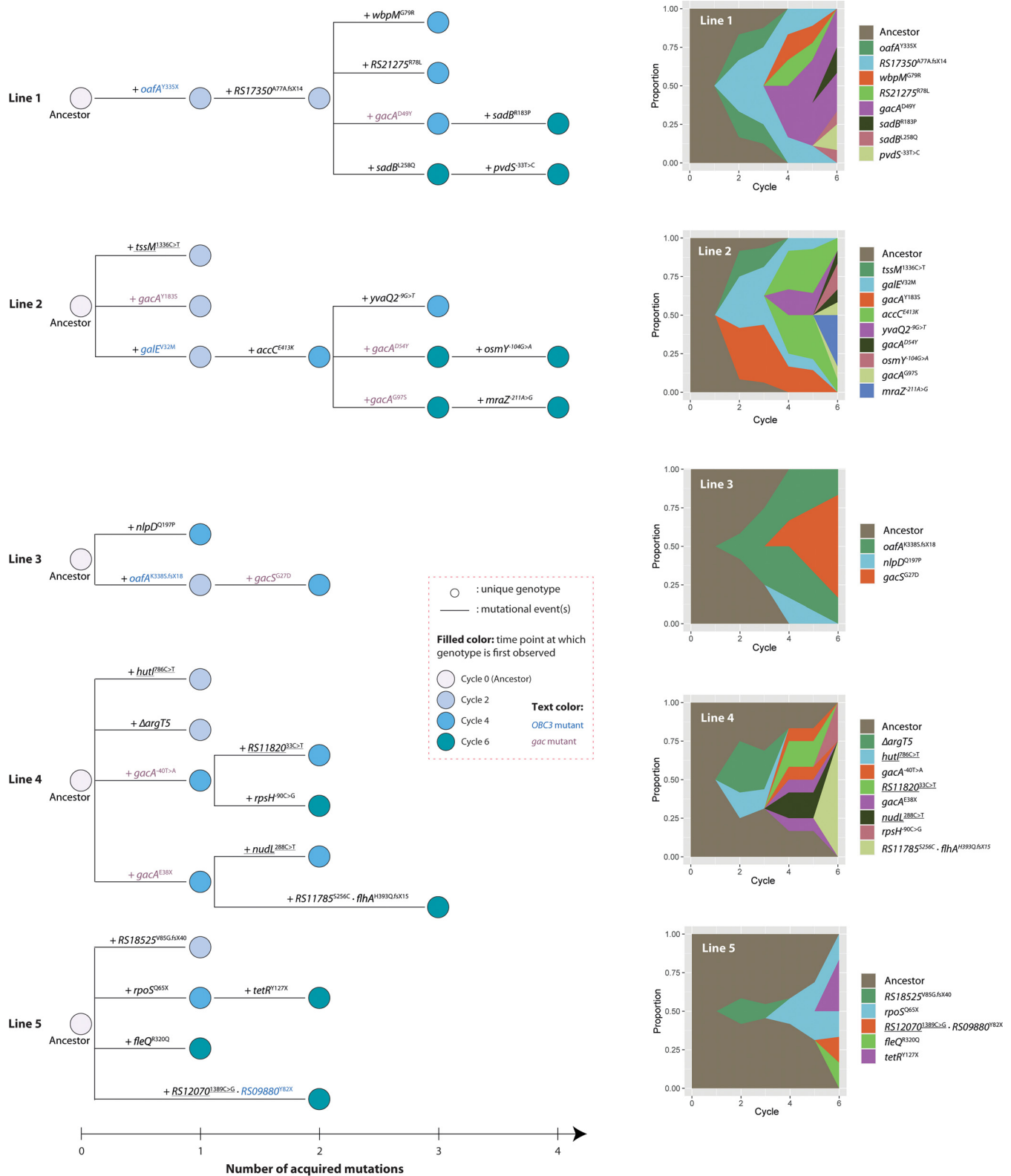
FIG 3 Phylogenetic trees and Muller plots of five independently evolving *P. protegens* CHA0 populations. (Left) Phylogenies for 18 genomes from each population (16 for the population in experimental line 4), based on the sequential appearance of mutations, are shown. Synonymous mutations are underlined. Circles represent unique bacterial genotypes. The color of the circle fill represents the time point at which a genotype was first detected. *OBC3* and *gac* mutants are highlighted in blue and purple, respectively. (Right) Muller plots depicting the dynamics of mutant alleles during the evolutionary experiment. The Muller plots show the estimated frequencies, by their heights, of 33 mutations in the respective population over 6 experimental cycles. Descendant genotypes are shown emerging from inside of their respective ancestors. The frequency of each mutation in the respective population can also be found in Table 2. Muller plots are prepared using *ggmuller* in R (102).

TABLE 2 Nomenclature for CHA0 mutations and variants and their frequency/cycle in each respective population

Line	Locus tag	Gene name	DNA sequence change[a]	Amino acid change (three-letter code)[b]	Standard nomenclature (used in text)	Frequency in the population (x/6)[c]			
						Cycle 2	Cycle 4	Cycle 6	HRM
1	PFLCHA0_RS09890	oafA	c.1005C>A	p.Tyr335X	Y335X	4/6	6/6	6/6	
1	PFLCHA0_RS17350	RS17350	c.230delC	p.Ala77Ala.fsX14	A77A.fsX14	2/6	6/6	6/6	
1	PFLCHA0_RS17965	gacA	c.145G>T	p.Asp49Tyr	D49Y		2/6	4/6	
						9.2%	19.1%	46.7%	x
1	PFLCHA0_RS21855	wbpM	c.235G>C	p.Gly79Arg	G79R		1/6		
1	PFLCHA0_RS21275	RS21275	c.233G>T	p.Arg78Leu	R78L		1/6		
1	PFLCHA0_RS21265	pvdS	c.-33T>C	NA	-33T>C			1/6	
1	PFLCHA0_RS30075	sadB	c.548G>C	p.Arg183Pro	R183P			1/6	
1			c.773T>A	p.Leu258Gln	L258Q			2/6	
2	PFLCHA0_RS17965	gacA	c.548A>C	p.Tyr183Ser	Y183S	2/6	1/6		
2	PFLCHA0_RS09920	galE	c.94G>A	p.Val32Met	V32M	2/6	5/6	6/6	
2	PFLCHA0_RS30120	tssM	c.1336C>T	p.Leu446Leu	1336C>T	1/6			
2	PFLCHA0_RS03400	accC	c.1237G>A	p.Glu413Lys	E413K		4/6	6/6	
2	PFLCHA0_RS13000	yvaQ2	c.-9G>T	NA	-9G>T		1/6		
2	PFLCHA0_RS17965	gacA	c.160G>T	p.As54Tyr	D54Y			2/6	
						5.5%	4.5%	15.5%	x
2	PFLCHA0_RS17965	gacA	c.289G>A	p.Gly97Ser	G97S			3/6	
2	PFLCHA0_RS25175	mraZ	c.-211A>G	NA	-211A>G			2/6	
2	PFLCHA0_RS26215	osmY	c.-104G>A	NA	-104G>A			1/6	
3	PFLCHA0_RS09890	oafA	c.1013delA	p.Lys338Ser.fsX18	K338S.fsX18	1/6	5/6	6/6	
3	PFLCHA0_RS22600	gacS	c.80G>A	p.Gly27Asp	G27D		1/6	4/6	
						3.8%	19.9%	75.6%	x
3	PFLCHA0_RS31060	nlpD	c.590A>C	p.Gln197Pro	Q197P		1/6		
4	PFLCHA0_RS22950	argT5	c.120-533del	p.Lys40-Leu185del.fsX6	ΔargT5	2/6			
4	PFLCHA0_RS02080	hutI	c.786C>T	p.Phe262Phe	786C>T	1/6			
4	PFLCHA0_RS17965	gacA	c.-40T>A	NA	-40T>A		2/6	1/4	
4	PFLCHA0_RS17965	gacA	c.112G>T	p.E38X	E38X		2/6	3/4	
4	PFLCHA0_RS05510	nudL	c.288C>T	p.Ser96Ser	288C>T		1/6		
4	PFLCHA0_RS11820	RS11820	c.33C>T	p.Ala11Ala	33C>T		1/6		
4	PFLCHA0_RS08490	flhA	c.1176delT	p.His393Gln.fsX15	H393Q.fsX15			3/4	
4	PFLCHA0_RS11785	RS11785	c.766A>T	p.Ser256Cys	S256C			3/4	
4	PFLCHA0_RS27515	rpsH	c.-90C>G	NA	-90C>G			1/4	
5	PFLCHA0_RS18525	RS18525	c.254delT	p.Val85Gly.fsX40	V85G.fsX40	1/6			
5	PFLCHA0_RS06125	rpoS	c.193C>T	p.Gln65X	Q65X		1/6	4/6	
5	PFLCHA0_RS08340	fleQ	c.959G>A	p.Arg320Gln	R320Q			1/6	
5	PFLCHA0_RS14960	tetR	c.381C>G	p.Tyr127X	Y127X			2/6	
5	PFLCHA0_RS09880	RS09880	c.246C>A	p.Tyr82X	Y82X			1/6	
5	PFLCHA0_RS12070	RS12070	c.1389C>G	p.Pro463Pro	1389C>G			1/6	

[a]DNA sequence change positions are relative to the cDNA (c). del, deletion. Mutations upstream of the transcription start site (TSS) are preceded by a dash to indicate the position relative to the TSS.

[b]X, stop codon (at its relative position in case of a shifted frame); fs, frameshift; p., amino acid substitution relative to the protein; NA, not applicable.

[c]Percentages represent the frequency of the selected (indicated by x) mutation in the population as determined by PCR-based high-resolution melt (HMR) analysis.

is estimated from a small set of sequenced individuals and therefore is expected to lack precision. The evolved isolates contain one to four mutations in general, and in each line, 3 (line 3) to 9 (lines 2 and 4) mutations were identified in total. All mutations were unique for each line, as was expected for independently evolving populations.

As the frequency of each mutation (Fig. 3; Table 2) was determined from only six bacterial colonies that were isolated and sequenced from each evolutionary line at cycles 2, 4, and 6, we further investigated and accurately measured the population-level frequency at the end of each experimental cycle. We determined the frequency of three gac mutations, i.e., gacA^{D49Y}, gacA^{D54Y}, and gacS^{G27D}, with increased accuracy, i.e., by PCR-based high-resolution melting (HRM) analysis incorporating mutation-specific HRM probes (Table S2), and increased sampling depth, i.e., at the end of experimental cycles 1 to 8. The HRM methodology allows for the accurate quantification of mutant frequencies across a wide range for all three mutations ($P < 0.001$) (Fig. S1). Using this method, we

found that the population-level frequency of these three mutations in the respective evolutionary lines corresponded remarkably well with the frequency previously obtained from the cultured and sequenced isolates (Table 2; Fig. S2). These findings corroborate the culture-based quantification of mutant frequency, suggesting that they provide a reasonable measure for population-level frequency. Intriguingly, it can be seen that the *gacS*^{G27D} mutation reaches fixation after eight experimental cycles, while the *gacA*^{D49Y} and *gacA*^{D54Y} mutants stabilize at around 50% and 25%, respectively (Fig. S2). Stabilization of mutations is indicative of frequency-dependent (FD) selection putatively reinforced by clonal interference with coexisting lineages carrying beneficial mutations in *sadB* (L258Q) and *gacA* (G97S) (Fig. 3). FD selection describes the phenomenon that the fitness of a particular genotype or phenotype is dependent on its frequency. Such context dependency has been linked to cheating behavior in which microbial cells that lack or have limited production of certain costly compounds benefit from other cells that do produce these compounds. When a minimal amount of such a compound increases genotype or phenotype fitness, FD selection can occur, resulting in stabilization of the mutation frequency.

Early adaptations are driven by cell surface-related genes. In general, mutations that are fixed early on in the rhizosphere adaptation process tend to have a high selective advantage (31). Disruptive mutations in *oafA*, resulting in premature stops halfway in the coding region, were detected as the first acquired mutations in two independent evolutionary lines (lines 1 and 3) and appear to have swept the population in the following generations (Table 2). *oafA* is part of the O-polysaccharide (O-PS; O-antigen) biosynthesis cluster 3 (*OBC3*) (29) and encodes an O-acetyltransferase, which is postulated to acetylate the O-antigen component of the outer membrane LPS (32). Another *OBC3* mutation that accumulated early on in the rhizosphere adaptation process, *galE*$^{c.94G>A}$, leads to an amino acid substitution (V32M) in *galE*'s product, and this mutation swept through the population in evolutionary line 2, reaching fixation in cycle 6 (Table 2; Fig. 3). *galE* encodes an UDP-glucose 4-epimerase which is involved in O-antigen and LPS core biosynthesis (33–35). One colony with a mutation in a third *OBC3* cluster gene, *RS09880*, encoding a putative glycosyl transferase (GT), was found in cycle 6 of evolutionary line 5. Thus, in four out of the five evolutionary lines, mutations that likely affect bacterial LPS structure appeared during rhizosphere adaptation, and these mutations became dominant in three out of four evolutionary lines. These results strongly suggest that modifying bacterial cell surface structure is an important bacterial strategy in early adaptation to the rhizosphere.

Adaption driven by global regulators. In the present study, six mutations were detected in *gacA* in three out of five evolutionary lines, representing approximately 20% of all missense mutations. Notably, in evolutionary lines 2 and 4, multiple *gacA* alleles accumulated, some of which were detected in early experimental cycles (Table 2). Additionally, a *gacS* mutation accumulated in line 3. *gacA* and *gacS* encode the main constituents of the conserved GacA/GacS two-component regulator system, i.e., the hybrid sensor histidine kinase GacS and the cognate response regulator GacA (Fig. 2). In Gram-negative bacteria, activation of GacS results in cross-phosphorylation of GacA via phosphotransfer, which in turn leads to activation of the expression of the small RNA genes *rsmY* and *rsmZ* via its helix-turn-helix (HTH) domain-binding domain (36). In CHA0, this regulatory pathway is known to control quorum sensing as well as secondary metabolism and stress resistance (37–40). In the closely related strain *P. protegens* Pf-5, this pathway was shown to have a big impact on bacterial genome-wide gene expression, affecting the expression of more than 10% of the annotated genes (41). Similarly, *gacA* mutants that arose on the edge of swarming Pf-5 colonies showed dramatically altered genome-wide gene expression patterns (42).

Including *gacS* and *gacA*, about half of the mutated genes in this study are global regulators or sigma factors (Table 1; Fig. 1). This high frequency suggests that global regulator-controlled networks are evolvable and play a major role in rapid bacterial adaptation. Pleiotropic adaptive mutations in global regulator genes have been shown to be important for bacterial adaption both in the laboratory (43, 44) and in natural

settings (45, 46). Remodeling and continuous optimization of existing regulatory networks by single mutations are an important strategy for bacterial adaptation to the host (47).

Bacterial motility. Bacterial motility is an important trait for rhizosphere competence, mediating colonization of distal parts of the root system (48), and both LPS and the GacS/GacA two-component regulator system are known to affect this trait. The O-antigen side chain of the LPS was reported to contribute to swimming and swarming motility in the plant-pathogenic bacterium *Erwinia amylovora* (49, 50). The GacS/GacA two-component regulator system controls bacterial motility, for example, by affecting the transcription of genes related to flagellum and biosurfactant biosynthesis (42, 48, 51). We also identified several other mutations across the various evolutionary lines that can be linked to bacterial motility. For instance, a disruptive mutation in the flagellar biosynthesis protein-coding gene *flhA* (H393Q.fsX15) that is involved in the biogenesis of the bacterial flagellum (Table 1) appeared in evolutionary line 4, reaching up to a predicted frequency of 75% (Table 2; Fig. 3). In *P. fluorescens* Pf-5 and in *Pseudomonas aeruginosa* PAO1, FlhA is reported to be essential for swimming motility (41).

Furthermore, we identified one amino acid substitution, R320Q, in FleQ and two in SadB, R183P and L258Q, that based on sequence similarity to well-studied, homologous proteins in other bacteria can be linked to motility in addition to several other bacterial traits. FleQ is a σ^{54}-dependent Fis family transcriptional regulator which regulates flagellar motility and biofilm formation as well as Pel exopolysaccharide (EPS) production in response to cellular c-di-GMP levels in *P. aeruginosa* (52). *P. protegens* FleQ shares 84% sequence identity and 98% sequence coverage with *P. aeruginosa* FleQ, and like *P. aeruginosa* FleQ, it is comprised of the N-terminal flagellar regulatory FleQ domain (PF06490), a central AAA+/ATPase σ^{54} interaction domain (PF00158), and a C-terminal Fis-type HTH DNA-binding domain (PF02954). The R320Q substitution that we identified here is found in the AAA+/ATPase σ^{54}-interacting domain in between the arginine (Arg) finger (amino acids 300 to 303) and a c-di-GMP-binding motif, ExxxR (amino acids 330 to 334) (52). The conserved arginine residues in FleQ, including the here-mutated Arg^{320}, are thought to be important for protein oligomerization, and substitution of any of these residues abolishes ATPase activity in *Vibrio cholerae* EpsE completely (53). Finally, *sadB* encodes an HD-related (named after the conserved histidine [H] aspartic acid [D] doublet of predicted catalytic residues [104]) output domain (HDOD)-containing protein (Table 1) that shares 77% sequence identity and 99% sequence coverage with *P. aeruginosa* SadB. In *P. aeruginosa*, SadB stimulates Pel EPS production and the chemotaxis-like cluster CheIV, which in turn affects flagellar motility as well as biofilm formation (54). In *Pseudomonas fluorescens* F113, SadB and FleQ control flagellar motility, both dependent upon and independent of the GacS/GacA two-component regulator system (55, 56).

Parallelism of targeted mutations on this functional-motility pathway impelled us to assess bacterial motility and track its dynamics across all evolutionary lines. We selected all *OBC3* and *gac* mutants as well as *sadB*, *fleQ*, and *flhA* mutants. Additionally, we included two *gacA* mutant progenitors with mutations in *accC* and *RS17350*, encoding a biotin carboxylase and a methyltransferase, respectively, plus two *gacA* mutant descendants with mutations in *osmY* and *mraZ* that encode an osmotically inducible protein and a transcriptional repressor, respectively (Table 1). Altogether, we assessed the swimming and swarming motility of strains with 17 distinct genotypes (Fig. 4; Fig. S3). We found that *OBC3* mutants, the *gacA* progenitors with mutations in *accC* and *RS17350*, and the *gacA* descendants with mutations in *osmY* and *mraZ* were unaltered compared to their respective ancestors when both swimming and swarming motility were considered, with the exception of a small yet significant increase in the swarming motility of the *oafA*[Y335X] mutant. However, *gac* mutants themselves were significantly affected, both in swimming, which is generally enhanced, and in swarming, which is repeatedly decreased (Fig. 4). *sadB* mutants, like *gac* mutants, display enhanced swimming and worsened swarming compared to those of their respective progenitors. Oddly, both
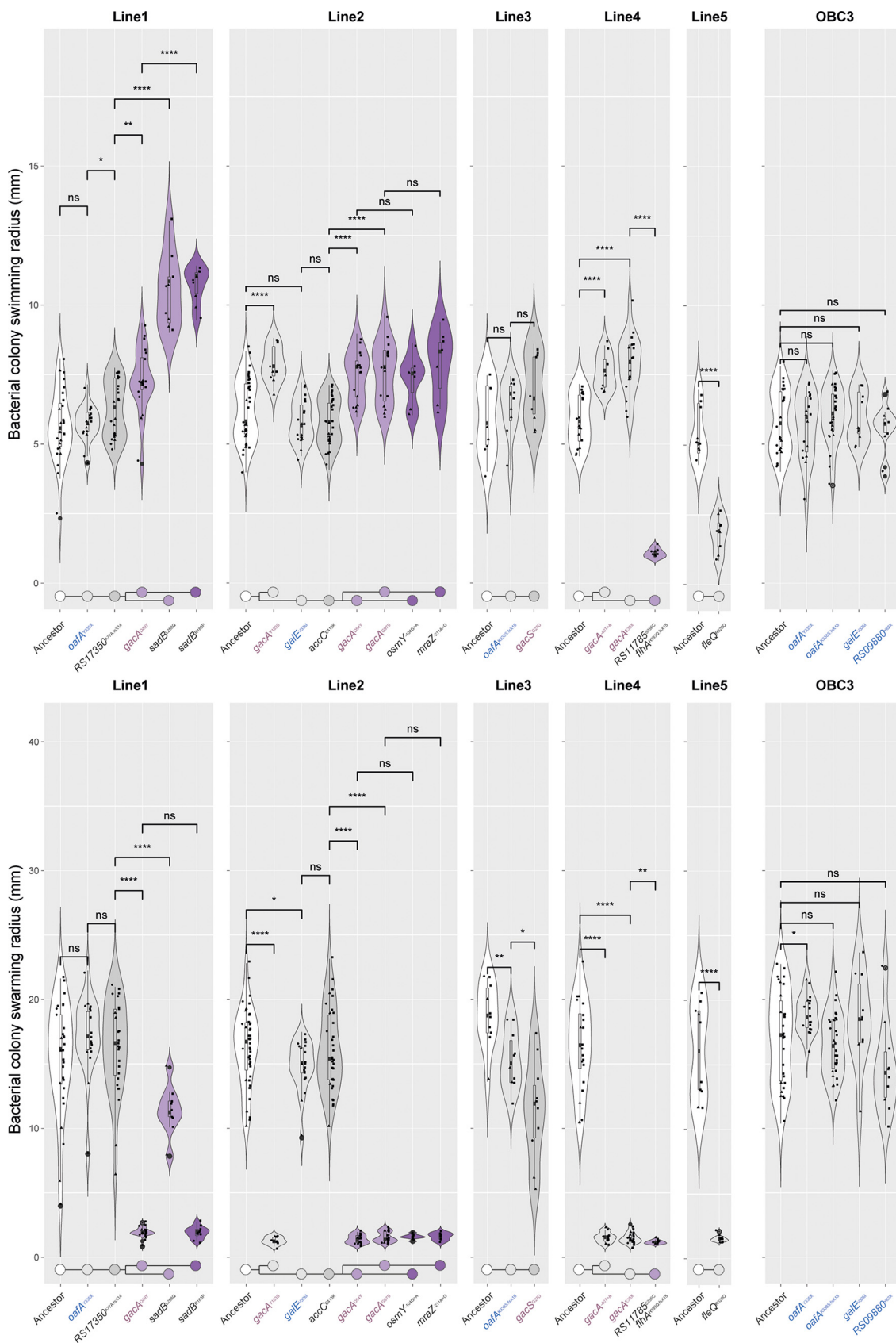
FIG 4 Evolved, root-competent mutant *P. protegens* CHA0 strains are characterized by enhanced swimming and impaired swarming. The accumulative effects of parallel and sequential mutations on swimming motility (top panel) and swarming motility

the *fleQ* and the *flhA* mutant displayed severely reduced swimming and swarming motility, thus representing two examples of an alternative evolutionary route toward adaptation in the rhizosphere in which motility is reduced. Loss of motility might coincide with another trait in these cases, such as EPS production and/or biofilm formation, regulated via shared yet oftentimes opposing mechanisms. Also, the frequency of the *fleQ* mutation was low; i. e., only one out of six isolates from cycle 6 of experimental line 5 carried this mutation, and therefore, whether this mutant is truly beneficial is unclear. On the other hand, the *flhA* mutation was found in three out of the four cultured and sequenced isolates from cycle 6 of line 4 and therefore may be present in a significant proportion of the total population.

Dynamics of global phenotypic change. Since natural selection eventually operates at the phenotypic level, revealing bacterial global phenotypic evolutionary dynamics can help us to identify traits that are under selection. Moreover, beneficial genetic mutations can be predicted if they are linked to well-known root colonization traits. A broad range of 30 bacterial traits, including different aspects of bacterial life/history traits, were examined for the sequenced isolates, which allows for genotype-phenotype association analysis.

As shown in Fig. 5a, the 30 bacterial traits separated into four clusters that share similar patterns across the different mutant genotypes, and this clustering is supported by model-based clustering analysis (Fig. S4). The growth of the bacteria in 1/3-strength King's B (KB) medium was positively correlated with siderophore and exoprotease production, tryptophan side oxidase activity, and growth inhibition of the bacterial plant pathogen *Ralstonia solanacearum*. Thus, this cluster, designated cluster 1, contains traits associated with bacterial social behavior related to microbe-microbe communication and cooperation, such as the production of public goods. In a principal-component analysis (PCA) (Fig. S5a), the first principal component (PC1) is strongly correlated with all five traits, with a total explanation of 64.5% for all variables. Cluster 2 contains traits linked to carbon source utilization. For this cluster, PC1 is strongly correlated with all carbon source usage-related traits, with a total explanation of 83.9% for all variables (Fig. S5b). A third cluster was observed for the bacterial ability to form a biofilm, to produce indole-3-acetic acid (IAA), and to inhibit the growth of two fungal plant pathogens, with a total explanation of 82.4% by PC1 (Fig. S5c). Finally, cluster 4 contains all seven traits that are related to bacterial resistance to biotic and abiotic stresses. The PC1 is strongly correlated with all seven traits, with a total explanation of 53.9% for all variables (Fig. S5d). For these four clusters, the PC1 (or −PC1) value is used as a proxy to present the general performance of all the traits that clustered together.

As is shown in Fig. 5b, all five evolutionary lines showed parallel trends of accumulative declines in social traits. Mutations in *gacA*, *gacS*, and *rpoS* resulted in significantly decreased bacterial social traits, and this was also observed for the $RS11785^{S256C}$ $flhA^{H393Q.fsX15}$ double mutant relative to the social traits of a strain with the background $gacA^{E38X}$ mutation. *RS11785*, like *gacA*, encodes a LysR-type transcriptional regulator. In addition to the mutations in these global regulators, earlier mutations in *oafA*, *galE*, and $RS17350^{A77A.fsX14}$, encoding a putative methyltransferase, also resulted in a significant but relatively small decrease in social traits. This parallel decline of the bacterial

FIG 4 Legend (Continued)

(bottom panel) are shown. For each experimental line, strains carrying mutations that can be connected to bacterial motility or that are the ancestor or descendants of such mutant strains were studied using typical swimming and swarming assays on Cook's Cytophaga (CC) medium (96). The colony area as a measure for bacterial motility was determined using ImageJ. The combined data from three independent experiments are shown, and data points from each experiment can be discerned by their respective shape. The sample size, *n*, varies per genotype, as related strains are combined on each plate, and we report the minimum and maximum *n* per replicate experiment. For swimming assays, circles represent replicate 1 ($2 \leq n \leq 6$) (excludes $osmY^{-104G>A}$ and $mraZ^{-211A>G}$), triangles replicate 2 ($3 \leq n \leq 15$), and squares replicate 3 ($4 \leq n \leq 20$). For swarming assays, circles represent replicate 1 ($1 \leq n \leq 3$) (excludes $osmY^{-104G>A}$ and $mraZ^{-211A>G}$), triangles replicate 2 ($2 \leq n \leq 6$), and squares replicate 3 ($7 \leq n \leq 35$). Obvious outliers were removed after pooling. Significant differences in motility between a genotype and its respective progenitor were determined by unpaired *t* test analysis (swimming, $7 \leq n \leq 41$; swarming, $7 \leq n \leq 44$; *, $\alpha = 0.05$; **, $\alpha = 0.01$; ***, $\alpha = 0.005$; ****, $\alpha = 0.001$; ns, nonsignificant), and the result is shown above each comparison. The genealogy of the mutations is shown below each experimental line, highlighting both parallel (branching) and sequential mutations. Colors depict numbers of acquired mutations relative to the ancestor (white), and these are light gray (1 mutation), gray (2 mutations), light purple (3 mutations), and purple (4 mutations).

a)
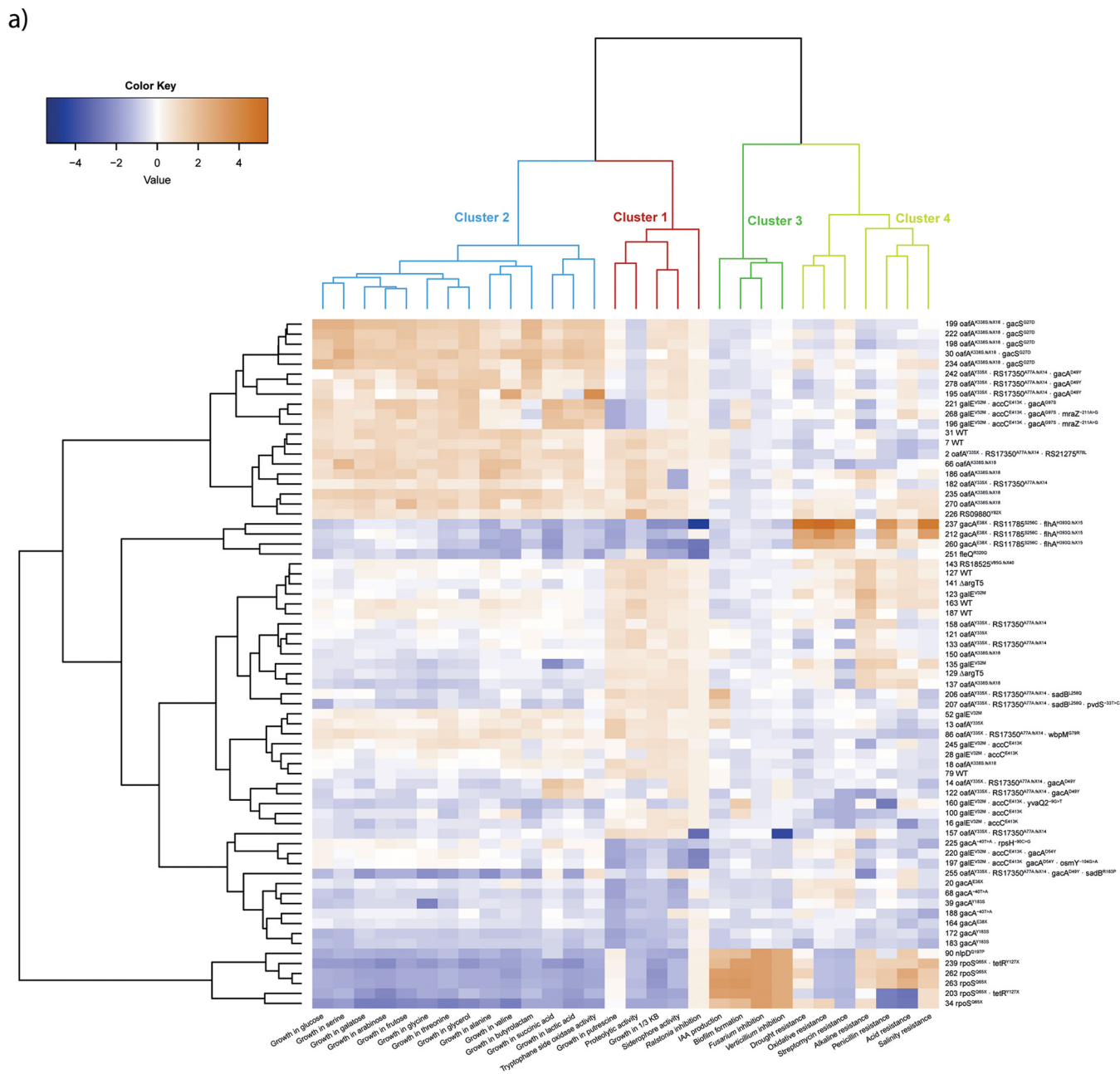
FIG 5 Genotype-phenotype association analysis of evolved bacterial populations. Isolates without mutations or with only synonymous mutations were excluded for this association analysis. (a) Heatmap representing the performance of 30 life history traits of six ancestral and 63 evolved isolates. The ward. D2 method was used for hierarchical clustering, based on the dissimilarities of the set of objects being clustered. The four clusters depicted on top are model predicted. Each column illustrates one bacterial life history trait, indicated at the bottom of the figure. Each row represents an isolate, indicated by its sample ID number and mutational events. (b to e) Principal-component analysis was applied to generate a general symbolic index for each model-predicted cluster. The proxy values of each cluster were normalized; higher values represent better performance. To illustrate better the accumulative effects of each mutational event, ANOVA was applied to reveal the phenotypic change between each genotype and its identified nearest ancestor in each line (this was applied only for genotypes that were detected more than once [$n \geq 2$]). Asterisks alongside the mutational events indicate significant differences (*, $\alpha = 0.05$; **, $\alpha = 0.01$; ***, $\alpha = 0.001$). Colored circles represent bacterial genotypes, with different colors representing independent evolving populations, and the size denotes the number of replicates (n). The dashed line represents the average value of the ancestor.

social trait index suggests a negative selection of bacterial social behavior, especially the production of costly public goods, such as siderophores and exoproteases.

All traits related to the utilization of 14 different carbon sources, which were selected based on their reported presence in root exudates of *Arabidopsis* (57), grouped in one cluster (cluster 2) (Fig. 5a). This suggests that carbon source utilization is coregulated. Different mutations in *gacS* and *gacA* resulted in contrasting bacterial carbon source
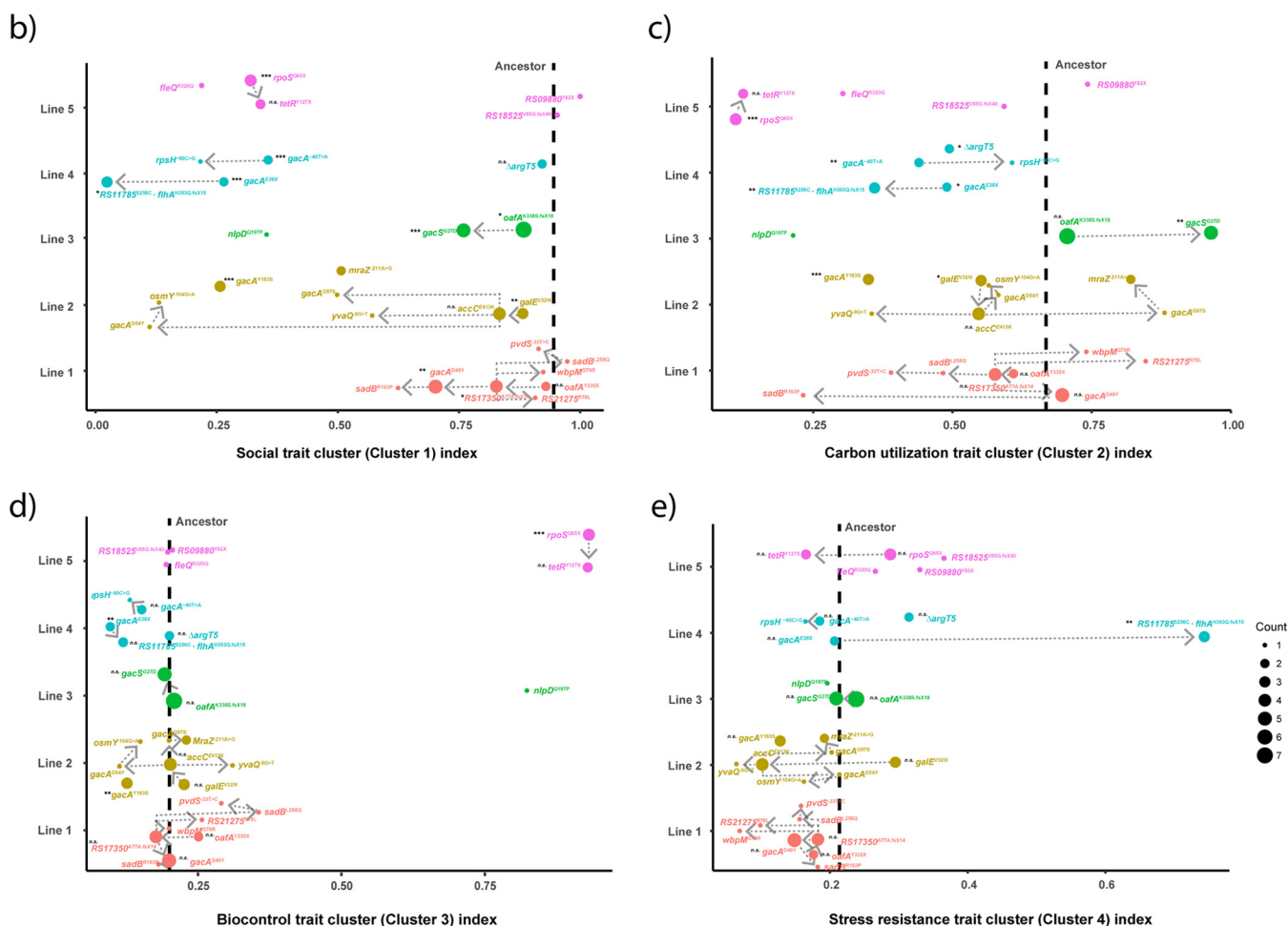
FIG 5 (Continued)

utilizations. A mutation in the first N-terminal transmembrane domain of GacS (G27D) (Fig. 2) resulted in significant enhancement of carbon source utilization (Fig. 5a and c). A similar trend was observed for the *gacA*^{G97S} mutant. In contrast, the majority of evolved genotypes, including most mutations in *gacA* and *rpoS*, showed a reduced ability to utilize carbon sources (Fig. 5a and c).

Traits in clusters 3 (Fig. 5d) and 4 (Fig. 5e), unlike in the previous two clusters, were more stable, as most genotypes behave like the ancestor. Only two *gacA* mutants, the *gacA*^{E38X} and *gacA*^{Y183S} mutants, were associated with a significant decline of bacterial traits associated with biocontrol, i.e., antifungal activity and biofilm formation, while the *rpoS*^{Q65X} mutation resulted in a significant increase in these traits. The double mutation *RS11785*^{S256C} *flhA*$^{H393Q.fsX15}$, in the background of the disruptive *gacA*^{E38X} mutation (Fig. 5e), was the only mutational event that led to a significant increase of general resistance to various environmental stresses. *flhA* encodes the flagellar biosynthesis protein FlhA and is linked to bacterial motility. The LysR-type transcriptional regulator RS11785 comprises an N-terminal HTH DNA-binding domain (PF00126) and a C-terminal substrate-binding domain (PF03466). This substrate-binding domain is linked to aromatic compound degradation and resembles that of the type 2 periplasmic binding proteins, or PBP2s. PBP2s are responsible for the binding and uptake of several substrates, including polysaccharides and several amino acids. We used HHpred (58) for protein remote homology detection and three-dimensional structure analysis to assess the possible consequence of the S256C amino acid substitution. This analysis identified OxyR from *Escherichia coli* as the best template for modeling (E value, 2.9e−30; score,

166.45). OxyR represents the master peroxide sensor in Gram-negative bacteria and operates via intramolecular disulfide bond formation in the regulatory PBP2-like domain. RS11785 is unlikely to represent the master peroxide sensor in *P. protegens* CHA0 because another CHA0 protein encoded by *PFLCHA0_RS030065* is much more similar to OxyR (89% pairwise sequence identity versus 29% for RS11785). Nevertheless, it is tempting to speculate that the serine (S, Ser)-to-cysteine (C, Cys) amino acid substitution (S256C) that we observed here might influence regulatory activity by altering intramolecular disulfide bond formation, as such bonds are governed by pairs of cysteine residues. Lastly, RS11785 bears resemblance (37% sequence identity, 93% sequence coverage) to the LysR-type PBP2 domain-containing regulator *alsR*, which regulates the activity of the acetoin operon (*alsSD*) in response to several signals, such as glucose and acetate, in *Bacillus amyloliquefaciens*. Acetoin (3-hydroxy-2-butanone, a precursor of 2,3-butanediol) can elicit induced systemic resistance (59) and is linked to general plant growth promotion in bacilli. In the absence of singular mutations in *flhA* and *RS11785*, the molecular mechanism underlying the observed enhanced environmental stress resistance in this double mutant remains to be clarified, but we hypothesize that altered regulatory activity by RS11785^{256C} is the most likely cause.

DISCUSSION

The rhizosphere is a nutrient-rich environment for root-associated bacteria. However, to access the available nutrients, bacteria must overcome various challenges, including the plant immune system, the presence of competing and/or predatory microorganisms, and abiotic stresses. In the gnotobiotic binary system used in this study, we tracked changes in *P. protegens* CHA0 in the rhizosphere of *Arabidopsis* under reproducible and controlled conditions without interference of complex interactions with other microbes. Mutations affecting global regulators, bacterial cell surface structure, and motility accumulated in parallel in our evolutionary experiment, revealing at least three important strategies of bacterial adaptation to the rhizosphere.

Global regulators and rhizosphere adaptation. The GacS/GacA two-component regulator system controls the production of antimicrobial secondary metabolites, exoenzymes, and siderophores but also biofilm formation, stress responses, motility, and quorum sensing (13, 38, 41, 60, 61). In the present study, mutations in the GacS/GacA two-component regulator system caused dramatic changes in several bacterial phenotypic traits, including motility, carbon source utilization, and social traits (Fig. 4 and 5). The overall gain in swimming motility is not unexpected, as swimming motility, driven by the flagellum apparatus, has been repeatedly reported to be an important root colonization trait in root-associated bacteria, including several *Pseudomonas* spp. (17, 62–64). In line with this, genome-wide transposon disruption mutant analysis by Cole and coworkers showed that the majority of motility-related genes in *Pseudomonas simiae* WCS417 are positively associated with *Arabidopsis* root colonization (23).

In contrast with the observed enhanced swimming motility across the evolutionary selection lines, swarming motility was found to be severely hampered throughout, and this appears to be the case especially for *gac* mutants. Swarming, like swimming, is driven by the flagellum but in addition depends on the production of several compounds, including quorum-sensing molecules and biosurfactants. *P. protegens* CHA0 is a known producer of the biosurfactant orfamide A, and the GacS/GacA two-component system is known to be an important regulator for its biosynthesis (65). These results fit observations by Song and coworkers (42) in the closely related *P. protegens* strain Pf-5. During swarming motility, *gac* mutants that lack production of the surfactant orfamide emerged. These mutants cannot swarm, but they coswarm with orfamide-producing cells (42).

Remarkably, we also identified two completely nonmotile mutants in our experiments with disruptions in the *fleQ* and *flhA* genes. Possibly, these mutants adhere much better to the root surface or might be able to form a biofilm more rapidly or strongly, as these traits

are often inversely correlated with bacterial motility (66–68). Future experiments should reveal whether such a tradeoff between bacterial motility and root adherence indeed underlies our observations in this evolutionary experiment.

The decreased production of public goods, such as siderophores and exoproteases, which were observed in all independent evolutionary lines (Fig. 5), may be beneficial for bacterial fitness by saving energy and primary metabolites. For example, adaptation of *Pseudomonas aeruginosa* to the human host through mutations in regulators was accompanied by losses of siderophore production, secreted proteases, and biofilm formation (69, 70).

Cell surface structures as bacterial adaptation targets. Bacterial cell surface components are the first line of defense against environmental stress and interplay with hosts (71–73). LPS is a central outer membrane component for Gram-negative bacteria and exhibits structural adaptability that is contributed especially by its O-antigen part (71, 73). Bacterial O-antigen structure modification plays an important role in the evasion of host immunity (74) and has the potential to change host-bacterium interactions (71, 73, 75). In plant-pathogenic bacteria, LPS components are important virulence determinants (76, 77) that can activate a variety of defense-related responses (76, 78–80). In *P. fluorescens*, the O-antigen component has been implicated to induce systemic resistance in radish (81) and *Arabidopsis* (82).

We observed parallel mutations in genes that are involved in LPS biosynthesis and structure modification. In three out of the five evolutionary lines, the first fixed mutations were identified in *oafA* and *galE*, which are annotated as related to O-antigen biosynthesis and structure modification (Table 1). *oafA* encodes an O-acetyltransferase, which is postulated to modify the O antigen by acetylation (32). The enzyme GalE, UDP-galactose 4-epimerase, is involved in the interconversion of UDP-glucose to UDP-galactose, an essential intermediate for LPS core and O-antigen structures (33–35). Inactivation of *galE* in *Porphyromonas gingivalis* resulted in shortening of the O antigen (83), and in *Bradyrhizobium japonicum*, disruption of *galE* resulted in the complete absence of O antigen (33). Thus, it is tempting to speculate that evasion of the plant's immune response plays a role in the adaptation of CHA0 to the rhizosphere of *Arabidopsis*.

In conclusion, the observed bacterial genetic and phenotypic adaption dynamics emphasize important roles for global regulators, motility, and cell surface structure in bacterial adaptation to its host. The parallel emergence of mutations in similar genes resulted in specific fitness advantages for mutants in the rhizosphere, suggesting that this evolutionary process is driven by the rhizosphere environment.

MATERIALS AND METHODS

Experimental setup. We set up an experimental evolution experiment with *Arabidopsis thaliana* ecotype Col-0 as the host plant and *Pseudomonas protegens* CHA0 as the evolving bacterial strain. CHA0 (84) is a model strain originally isolated from roots of tobacco plants grown in soil naturally suppressive to black root rot (85). CHA0 was chromosomally tagged with green fluorescent protein (GFP) and a kanamycin resistance cassette (86) to enable consistent tracking of the strain and identification of contaminations. We previously described the setup of the evolutionary experiment in great detail (26). In brief, the ancestral bacterial population (10^6 cells) was inoculated on *Arabidopsis* roots grown under gnotobiotic conditions inside ECO2 boxes in carbon-free silver sand. For each cycle, *Arabidopsis* seeds were surface sterilized using chlorine gas, germinated on modified Hoagland's agar medium, and grown for 2 weeks until transplantation into the ECO2 boxes, each containing two plants (26). Inoculated 2-week-old seedlings were then grown for an additional 4 weeks, after which the root-associated bacteria were collected in 10 mM MgSO$_4$ and subjected to fluorescence-based cell counting by flow cytometry, yielding on average 10^7 cells/root. Cells of the evolved bacterial populations (10^6) were then transferred to new plants, and this cycle was repeated eight times. After each cycle, a small fraction of each population was plated on general-purpose, nonselective medium, 3 g/liter tryptic soy agar (TSA), to assess for contaminations and to verify that all colonies carried the *GFP* marker gene, as observed under UV light, on the one hand, and to select individual isolates for phenotypic characterization on the other hand.

We previously picked 16 random isolates from each of five experimental lines at cycles 2, 4, and 6 plus 16 isolates from the ancestor population, yielding a total set of 256 isolates (26). These 256 isolates were subsequently grouped into five distinct phenotypes based on their performance on a variety of bacterial life history traits (26). In the current study, we selected 6 of these 16 isolates per line, and per cycle, for genome sequence analysis. Isolates were selected to represent the breadth of phenotypic diversity observed previously (26). We set out to obtain genome sequences of 96 isolates in total, together with 6 isolates from the ancestor population, using the NextSeq-500 Illumina platform (2×75-bp paired-end reads). Sequencing of two isolates from line 4, cycle 4, however, failed, and thus a final set of 94

genomes were retrieved. We then used the snippy pipeline (https://github.com/tseemann/snippy), integrating reference genome-based mapping of the Illumina reads by BWA-MEM, variant calling by SAMtools and FreeBayes, and variant impact analysis using SnpEff (87), to identify single nucleotide polymorphisms (SNPs) and small insertions and deletions (indels). Larger indels were identified by calculating the breadths of coverage of the mapped Illumina reads on the reference genome in a sliding window using bedtools (88). Regions with reduced coverage (<99%) were manually inspected in the Integrative Genome Viewer (IGV). Phylogenetic trees for each line were constructed manually with Illustrator based on all detected mutations, with the lengths of the branches representing the numbers of mutations. The genealogy and frequency of each lineage are shown in the Muller plots that are prepared with the R package ggmuller.

Bacterial life history traits. For the 94 sequenced isolates, a variety of bacterial life history traits reflecting various aspects of bacterial physiological processes were measured previously as part of all 256 isolates initially collected (26). Briefly, we monitored optical density (OD) at a wavelength of 600 nm to estimate the bacterial yield after 72 h of growth under different growth conditions in 96-well microplates. We measured bacterial growth yield, resistance to various stresses, including acidic (pH 5) and alkaline (pH 9) conditions, oxidative stress in 0.0025% H_2O_2, water potential stress (15% polyethylene glycol 6000 [PEG-6000]), and salt stress (2% NaCl), and resistance to the antibiotics streptomycin (1 μg · ml^{-1}), tetracycline (1 μg · ml^{-1}), and penicillin (5 μg · ml^{-1}). In this study, stress resistance was defined as the ratio of bacterial growth in the stressed relative to that in the nonstressed control treatment.

Bacterial carbon source utilization was quantified as growth yield in modified Ornston and Stanier (OS) minimal medium (89) supplemented with single-carbon sources that have been reported to be abundant in *Arabidopsis* root exudates (57). These included the following carbon sources: alanine, arabinose, butyrolactam, fructose, galactose, glucose, glycerol, glycine, lactic acid, putrescine, serine, succinic acid, threonine, and valine, which were added to a final concentration of 0.5 g · liter^{-1}. In addition, we measured bacterial auxin (indole-3-acetic acid, or IAA) production with a colorimetric test (90), iron-chelating ability using a modified chrome azurol S (CAS) assay (91), proteolytic activity by the adapted assay from Smeltzer et al. (92), tryptophan side chain oxidase activity using a colorimetric assay (93), and biofilm formation using a modified crystal violet staining assay (94). We measured the OD values reflecting the color intensities at specific wavelengths to quantify these traits. We further assessed bacterial antimicrobial activity by quantifying their effect on the growth of the fungi *Verticillium dahliae* and *Fusarium oxysporum* and on the bacterium *Ralstonia solanacearum*.

Motility assays. Motility assays were undertaken in round petri dish plates containing Cook's Cytophaga medium (CC medium) (0.3% agar for swimming, 0.5% agar for swarming) (95), using typical swimming and swarming assays as described by Déziel et al. (96). All tested strains were grown on King's B medium agar plates for 24 h before inoculation. Swim and swarm plates were inoculated with the tested strains with a sterile toothpick. For swimming plates, the inoculum was introduced by gently piercing the agar such that the motility within the semisolid agar could be evaluated. For swarming plates, the inoculum was introduced on the agar surface, enabling visualization of motility across the agar surface. Both swimming and swarming plates were imaged after 18 h of incubation at 21°C with the right side up. The radii of swimming and swarming motility were determined from the photographs by ImageJ by examining the inner circular turbid zone inside the 0.3% agar for swimming and the outer circular zone on top of the 0.5% agar surface for swarming.

Hierarchical and model-based clustering of bacterial traits. (i) Hierarchical clustering. A heatmap to illustrate the association patterns of bacterial genotypes and their measured traits was constructed in R using the ggplot2 package. Isolates without mutations or with only synonymous mutations were excluded from the association analysis. Hierarchical clustering was performed using the Ward.D2 method, which is based on the squared dissimilarities of the set of objects being clustered (97).

(ii) Model-based clustering. We applied a model-based clustering method to reveal the best-fitting structures of trait covariance patterns. For example, some traits might be either directly or indirectly coregulated by the same gene, which is expected for global regulators, particularly those that can coregulate thousands of genes. We used the mclust package in R to run the model simulation (98). This method assumes that the input data are distributed as a mixture of two or more clusters. The advantage of the model-based clustering method is that it avoids heuristic assumptions and uses a soft assignment that every data point has a possibility of falling to each cluster, which facilitates the best clustering solution. The so-called Bayesian information criterion (BIC) was used to select the best model. A large BIC score indicates a better fit of the model.

This result is in line with the outcome of hierarchical clustering with adjusted Rand index (ARI) set as 1 and *k* set as 4 in ward.D2, as indicated in Fig. 5a. ARI is usually used to evaluate the match degree of a given clustering solution compared to the model-based clustering result, with 0 reflecting a random partition and 1 the boundary of accuracy of a certain clustering solution (98).

Genotype-phenotype association analysis. Bacterial traits within each model-predicted cluster have similar data distribution patterns and vary with the definition of the clustering method. Thus, we applied a linear-regression-based method, i.e., principal-component analysis (PCA), to reduce the dimensionality of data and generate a proxy for each model-predicted cluster. These proxies were later used as the *x* axis values in Fig. 5b to e. We applied the package ggbiplot in R to generate the PCA plots and PC1 index from the normalized data sets. The proxies were normalized for further analysis.

To examine the accumulative effects of each mutation on bacterial phenotype, analysis of variance (ANOVA) was used to compare cluster proxies of evolved genotypes with their direct ancestors. Only genotypes identified more than once ($n \geq 2$) were included in this analysis.

Relative quantification of mutant frequency using HRM profile analysis. We used PCR-based high-resolution melting (HRM) profile analysis with integrated LunaProbes to quantify the ratio of mutant to wild type genotypes (99–101). The probes and primers used in this study are listed in Table S2. Primers were designed using Primer3. Probes were designed with the SNP located in the middle of the sequence, and the 3′ end was blocked by carbon spacer C-3. The primer asymmetry was set to 2:1 (excess primer, limiting primer) in all cases. Pre-PCR was performed in a 10-μl reaction system, with 0.25 μM excess primer, 0.125 μM limiting primer, 0.25 μM probe, 0.5 μl bacterial sample culture (with a 100-fold-diluted saved sample, the OD at 600 nm [OD_{600}] is about 0.01), and 1× LightScanner master mix (BioFire Defense). Dimethyl sulfoxide (DMSO) at a final concentration of 5% was added to all reaction mixtures to ensure that the targeted melting domains were within the detection limit of the LightScanner (Idaho Technology Inc.). Finally, Milli-Q (MQ) water was used to supplement up to 10 μl. A 96-well black microtiter plate with white wells was used to minimize background fluorescence. Before amplification, 25 μl mineral oil was loaded in each well to prevent evaporation, and the plate was covered with a foil seal to prevent the degradation of fluorescent molecules. Amplification was initiated by a holding the plate at 95°C for 3 min, followed by 55 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 30 s; the plate was then kept at 72°C for 10 min. After amplification, samples were heated in a thermal cycler (Bio-Rad) shortly to 95°C for 30 s to denature all double-stranded structures, followed by a rapid cooling to 25°C for 30 s to facilitate successful hybridization between probes and the target strands. The plate was then transferred to a LightScanner (Idaho Technology Inc.). Melting profiles of each well were collected by monitoring the continuous loss of fluorescence with a steady increase of the temperature from 35°C to 97°C, with a ramp rate of 0.1°C/s. The relative quantification was based on the negative first derivative plots using MATLAB software. The areas of probe-target duplex melting peaks were auto-calculated with the AutoFit Peaks I Residuals function in PeakFit software (SeaSolve Software Inc.). The mutant frequency, X, was calculated using the equation $X = \text{area}_{mutant}/(\text{area}_{mutant} + \text{area}_{WT})$, where WT is the wild type. To validate the HRM method, standard curves were generated by measuring mixed samples with the following known proportions of mutant templates: 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100%. Measurements for each sample were done in triplicate. The linear-regression formula of each mutant between actual frequencies and measured frequencies is shown in Fig. S1. The high R^2 values and slope values of these equations that were nearly equal to 1 confirmed that the HRM method can accurately detect a mutant's frequency in a mixed population.

Data availability. The *P. protegens* CHA0-GFP reference genome is deposited on GenBank under accession no. RCSR00000000.1. Raw sequencing data used in this study are deposited in the NCBI database under BioProject no. PRJNA473919. A conversion table for the CHA0-GFP–to–CHA0 gene annotations, including recent NCBI accession codes, is available at the following URL: https://doi.org/10.6084/m9.figshare.13295828.v2.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.
FIG S1, JPG file, 0.1 MB.
FIG S2, JPG file, 0.1 MB.
FIG S3, JPG file, 0.4 MB.
FIG S4, JPG file, 0.1 MB.
FIG S5, JPG file, 0.1 MB.
TABLE S1, DOCX file, 0.02 MB.
TABLE S2, DOCX file, 0.01 MB.

REFERENCES

1. Bakker PAHM, Berendsen RL, Van Pelt JA, Vismans G, Yu K, Li E, Van Bentum S, Poppeliers SWM, Sanchez Gil JJ, Zhang H, Goossens P, Stringlis IA, Song Y, de Jonge R, Pieterse CMJ. 2020. The soil-borne identity and microbiome-assisted agriculture: looking back to the future. Mol Plant 13:1394–1401. https://doi.org/10.1016/j.molp.2020.09.017.
2. Bais HP, Weir TL, Perry LG, Gilroy S, Vivanco JM. 2006. The role of root exudates in rhizosphere interactions with plants and other organisms. Annu Rev Plant Biol 57:233–266. https://doi.org/10.1146/annurev.arplant.57.032905.105159.
3. Stringlis IA, Yu K, Feussner K, de Jonge R, Van Bentum S, Van Verk MC, Berendsen RL, Bakker PAHM, Feussner I, Pieterse CMJ. 2018. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. Proc Natl Acad Sci U S A 115:E5213–E5222. https://doi.org/10.1073/pnas.1722335115.

4. Bakker PAHM, Pieterse CMJ, de Jonge R, Berendsen RL. 2018. The soil-borne legacy. Cell 172:1178–1180. https://doi.org/10.1016/j.cell.2018.02.024.

5. Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH. 2013. Going back to the roots: the microbial ecology of the rhizosphere. Nat Rev Microbiol 11:789–799. https://doi.org/10.1038/nrmicro3109.

6. Sasse J, Martinoia E, Northen T. 2018. Feed your friends: do plant exudates shape the root microbiome? Trends Plant Sci 23:25–41. https://doi.org/10.1016/j.tplants.2017.09.003.

7. Bais HP, Park SW, Weir TL, Callaway RM, Vivanco JM. 2004. How plants communicate using the underground information superhighway. Trends Plant Sci 9:26–32. https://doi.org/10.1016/j.tplants.2003.11.008.

8. Zamioudis C, Pieterse CMJ. 2012. Modulation of host immunity by beneficial microbes. Mol Plant Microbe Interact 25:139–150. https://doi.org/10.1094/MPMI-06-11-0179.

9. Yang J, Kloepper JW, Ryu C-M. 2009. Rhizosphere bacteria help plants tolerate abiotic stress. Trends Plant Sci 14:1–4. https://doi.org/10.1016/j.tplants.2008.10.004.

10. Browne P, Rice O, Miller SH, Burke J, Dowling DN, Morrissey JP, O'Gara F. 2009. Superior inorganic phosphate solubilization is linked to phylogeny within the *Pseudomonas fluorescens* complex. Appl Soil Ecol 43:131–138. https://doi.org/10.1016/j.apsoil.2009.06.010.

11. Lugtenberg B, Kamilova F. 2009. Plant growth promoting rhizobacteria. Annu Rev Microbiol 63:541–556. https://doi.org/10.1146/annurev.micro.62.081307.162918.

12. Ryu C-M, Farag MA, Hu C-H, Reddy MS, Wei H-X, Paré PW, Kloepper JW. 2003. Bacterial volatiles promote growth in *Arabidopsis*. Proc Natl Acad Sci U S A 100:4927–4932. https://doi.org/10.1073/pnas.0730845100.

13. Haas D, Défago G. 2005. Biological control of soil-borne pathogens by fluorescent pseudomonads. Nat Rev Microbiol 3:307–319. https://doi.org/10.1038/nrmicro1129.

14. Pieterse CMJ, Zamioudis C, Berendsen RL, Weller DM, Van Wees SCM, Bakker PAHM. 2014. Induced systemic resistance by beneficial microbes. Annu Rev Phytopathol 52:347–375. https://doi.org/10.1146/annurev-phyto-082712-102340.

15. Raaijmakers JM, Bonsall RF, Weller DM. 1999. Effect of population density of *Pseudomonas fluorescens* on production of 2,4-diacetylphloroglucinol in the rhizosphere of wheat. Phytopathology 89:470–475. https://doi.org/10.1094/PHYTO.1999.89.6.470.

16. Raaijmakers JM, Leeman M, van Oorschot MMP, van der Sluis I, Schippers B, Bakker PAHM. 1995. Dose-response relationships in biological control of Fusarium wilt of radish by *Pseudomonas* spp. Phytopathology 85:1075–1081. https://doi.org/10.1094/Phyto-85-1075.

17. Lugtenberg BJJ, Dekkers L, Bloemberg GV. 2001. Molecular determinants of rhizosphere colonization *Pseudomonas*. Annu Rev Phytopathol 39:461–490. https://doi.org/10.1146/annurev.phyto.39.1.461.

18. Pieterse CMJ, Berendsen RL, de Jonge R, Stringlis IA, Van Dijken AJH, Van Pelt JA, Van Wees SCM, Yu K, Zamioudis C, Bakker PAHM. 2021. *Pseudomonas simiae WCS417*: star track of a model beneficial rhizobacterium. Plant Soil. Plant Soil 461:245–263. https://doi.org/10.1007/s11104-020-04786-9.

19. de Weger LA, van der Vlugt CI, Wijffjes AH, Bakker PA, Schippers B, Lugtenberg B. 1987. Flagella of a plant-growth-stimulating *Pseudomonas fluorescens* strain are required for colonization of potato roots. J Bacteriol 169:2769–2773. https://doi.org/10.1128/jb.169.6.2769-2773.1987.

20. de Weger LA, van Loosdrecht MC, Klaassen HE, Lugtenberg B. 1989. Mutational changes in physiochemical cell surface properties of plant-growth-stimulating *Pseudomonas* spp. do not influence the attachment properties of the cells. J Bacteriol 171:2756–2761. https://doi.org/10.1128/jb.171.5.2756-2761.1989.

21. Simons M, Permentier HP, de Weger LA, Wijffelman CA, Lugtenberg BJJ. 1997. Amino acid synthesis is necessary for tomato root colonization by *Pseudomonas fluorescens* strain WCS365. Mol Plant Microbe Interact 10:102–106. https://doi.org/10.1094/MPMI.1997.10.1.102.

22. Yu K, Liu Y, Tichelaar R, Savant N, Lagendijk E, van Kuijk SJL, Stringlis IA, van Dijken AJH, Pieterse CMJ, Bakker PAHM, Haney CH, Berendsen RL. 2019. Rhizosphere-associated *Pseudomonas* suppress local root immune responses by gluconic acid-mediated lowering of environmental pH. Curr Biol 29:3913–3920.e4. https://doi.org/10.1016/j.cub.2019.09.015.

23. Cole BJ, Feltcher ME, Waters RJ, Wetmore KM, Mucyn TS, Ryan EM, Wang G, Ul-Hasan S, McDonald M, Yoshikuni Y, Malmstrom RR, Deutschbauer AM, Dangl JL, Visel A. 2017. Genome-wide identification of bacterial plant colonization genes. PLoS Biol 15:e2002860. https://doi.org/10.1371/journal.pbio.2002860.

24. Kawecki TJ, Lenski RE, Ebert D, Hollis B, Olivieri I, Whitlock MC. 2012. Experimental evolution. Trends Ecol Evol 27:547–560. https://doi.org/10.1016/j.tree.2012.06.001.

25. Jousset A, Schuldes J, Keel C, Maurhofer M, Daniel R, Scheu S, Thuermer A. 2014. Full-genome sequence of the plant growth-promoting bacterium *Pseudomonas protegens* CHA0. Genome Announc 2:e00322-14. https://doi.org/10.1128/genomeA.00322-14.

26. Li E, de Jonge R, Liu C, Jiang H, Friman V-P, Ahm Bakker P, Jousset A. 2020. Rapid evolution of bacterial mutualism in the plant rhizosphere. bioRxiv https://doi.org/10.1101/2020.12.07.414607.

27. Oren Y, Smith MB, Johns NI, Kaplan Zeevi M, Biran D, Ron EZ, Corander J, Wang HH, Alm EJ, Pupko T. 2014. Transfer of noncoding DNA drives regulatory rewiring in bacteria. Proc Natl Acad Sci U S A 111:16112–16117. https://doi.org/10.1073/pnas.1413272111.

28. Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell 134:25–36. https://doi.org/10.1016/j.cell.2008.06.030.

29. Kupferschmied P, Chai T, Flury P, Blom J, Smits THM, Maurhofer M, Keel C. 2016. Specific surface glycan decorations enable antimicrobial peptide resistance in plant-beneficial *pseudomonads* with insect-pathogenic properties. Environ Microbiol 18:4265–4281. https://doi.org/10.1111/1462-2920.13571.

30. Maddamsetti R, Lenski RE, Barrick JE. 2015. Adaptation, clonal interference, and frequency-dependent interactions in a long-term evolution experiment with Escherichia coli. Genetics 200:619–631. https://doi.org/10.1534/genetics.115.176677.

31. Barrett RDH, Maclean RC, Bell G. 2006. Mutations of intermediate effect are responsible for adaptation in evolving *Pseudomonas fluorescens* populations. Biol Lett 2:236–238. https://doi.org/10.1098/rsbl.2006.0439.

32. Allison GE, Verma NK. 2000. Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. Trends Microbiol 8:17–23. https://doi.org/10.1016/s0966-842x(99)01646-7.

33. Chang W-S, Park K-M, Koh S-C, So J-S. 2008. Characterization of the *Bradyrhizobium japonicum galE* gene: its impact on lipopolysaccharide profile and nodulation of soybean. FEMS Microbiol Lett 280:242–249. https://doi.org/10.1111/j.1574-6968.2008.01066.x.

34. Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SN, Wang Q, Reeves PR, Wang L. 2008. Structure and genetics of *Shigella* O antigens. FEMS Microbiol Rev 32:627–653. https://doi.org/10.1111/j.1574-6976.2008.00114.x.

35. Samuel G, Reeves P. 2003. Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. Carbohydr Res 338:2503–2519. https://doi.org/10.1016/j.carres.2003.07.009.

36. Brencic A, McFarland KA, McManus HR, Castang S, Mogno I, Dove SL, Lory S. 2009. The GacS/GacA signal transduction system of *Pseudomonas aeruginosa* acts exclusively through its control over the transcription of the RsmY and RsmZ regulatory small RNAs. Mol Microbiol 73:434–445. https://doi.org/10.1111/j.1365-2958.2009.06782.x.

37. Brown D. 2010. A mathematical model of the Gac/Rsm quorum sensing network in *Pseudomonas fluorescens*. Biosystems 101:200–212. https://doi.org/10.1016/j.biosystems.2010.07.004.

38. Heeb S, Blumer C, Haas D. 2002. Regulatory RNA as mediator in GacA/RsmA-dependent global control of exoproduct formation in *Pseudomonas fluorescens* CHA0. J Bacteriol 184:1046–1056. https://doi.org/10.1128/jb.184.4.1046-1056.2002.

39. Valverde C, Heeb S, Keel C, Haas D. 2003. RsmY, a small regulatory RNA, is required in concert with RsmZ for GacA-dependent expression of biocontrol traits in *Pseudomonas fluorescens* CHA0. Mol Microbiol 50:1361–1379. https://doi.org/10.1046/j.1365-2958.2003.03774.x.

40. Heeb S, Valverde C, Gigot-Bonnefoy C, Haas D. 2005. Role of the stress sigma factor RpoS in GacA/RsmA-controlled secondary metabolism and resistance to oxidative stress in *Pseudomonas fluorescens* CHA0. FEMS Microbiol Lett 243:251–258. https://doi.org/10.1016/j.femsle.2004.12.008.

41. Hassan KA, Johnson A, Shaffer BT, Ren Q, Kidarsa TA, Elbourne LDH, Hartney S, Duboy R, Goebel NC, Zabriskie TM, Paulsen IT, Loper JE. 2010. Inactivation of the GacA response regulator in *Pseudomonas fluorescens* Pf-5 has far-reaching transcriptomic consequences. Environ Microbiol 12:899–915. https://doi.org/10.1111/j.1462-2920.2009.02134.x.

42. Song C, Kidarsa TA, Van De Mortel JE, Loper JE, Jos M. 2016. Living on the edge: emergence of spontaneous gac mutations in *Pseudomonas protegens* during swarming motility. Environ Microbiol 18:3453–3465. https://doi.org/10.1111/1462-2920.13288.

43. Hunter P. 2008. The great leap forward. EMBO Rep 9:608–611. https://doi.org/10.1038/embor.2008.115.

44. Zambrano MM, Siegele DA, Almirón M, Tormo A, Kolter R. 1993. Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. Science 259:1757–1760. https://doi.org/10.1126/science.7681219.

45. Sumby P, Whitney AR, Graviss EA, DeLeo FR, Musser JM. 2006. Genome-wide analysis of group A streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathog 2:e5. https://doi.org/10.1371/journal.ppat.0020005.

46. Giraud A, Arous S, De Paepe M, Gaboriau-Routhiau V, Bambou JC, Rakotobe S, Lindner AB, Taddei F, Cerf-Bensussan N. 2008. Dissecting the genetic components of adaptation of *Escherichia coli* to the mouse gut. PLoS Genet 4:e2. https://doi.org/10.1371/journal.pgen.0040002.

47. Damkiær S, Yang L, Molin S, Jelsbak L. 2013. Evolutionary remodeling of global regulatory networks during long-term bacterial adaptation to human hosts. Proc Natl Acad Sci U S A 110:7766–7771. https://doi.org/10.1073/pnas.1221466110.

48. Sánchez-Contreras M, Martín M, Villacieros M, O'Gara F, Bonilla I, Rivilla R. 2002. Phenotypic selection and phase variation occur during alfalfa root colonization by *Pseudomonas fluorescens* F113. J Bacteriol 184:1587–1596. https://doi.org/10.1128/JB.184.6.1587-1596.2002.

49. Bowden SD, Hale N, Chung JCS, Hodgkinson JT, Spring DR, Welch M. 2013. Surface swarming motility by *Pectobacterium atrosepticum* is a latent phenotype that requires O antigen and is regulated by quorum sensing. Microbiology 159:2375–2385. https://doi.org/10.1099/mic.0.070748-0.

50. Berry MC, McGhee GC, Zhao Y, Sundin GW. 2009. Effect of a *waaL* mutation on lipopolysaccharide composition, oxidative stress survival, and virulence in *Erwinia amylovora*. FEMS Microbiol Lett 291:80–87. https://doi.org/10.1111/j.1574-6968.2008.01438.x.

51. Lapouge K, Schubert M, Allain FH-T, Haas D. 2008. Gac/Rsm signal transduction pathway of γ-proteobacteria: from RNA recognition to regulation of social behaviour. Mol Microbiol 67:241–253. https://doi.org/10.1111/j.1365-2958.2007.06042.x.

52. Matsuyama BY, Krasteva PV, Baraquet C, Harwood CS, Sondermann H, Navarro MVAS. 2016. Mechanistic insights into c-di-GMP-dependent control of the biofilm regulator FleQ from *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A 113:E209–E218. https://doi.org/10.1073/pnas.1523148113.

53. Patrick M, Korotkov KV, Hol WGJ, Sandkvist M. 2011. Oligomerization of EpsE coordinates residues from multiple subunits to facilitate ATPase activity. J Biol Chem 286:10378–10386. https://doi.org/10.1074/jbc.M110.167031.

54. Caiazza NC, O'Toole GA. 2004. SadB is required for the transition from reversible to irreversible attachment during biofilm formation by *Pseudomonas aeruginosa* PA14. J Bacteriol 186:4476–4485. https://doi.org/10.1128/JB.186.14.4476-4485.2004.

55. Muriel C, Blanco-Romero E, Trampari E, Arrebola E, Durán D, Redondo-Nieto M, Malone JG, Martín M, Rivilla R. 2019. The diguanylate cyclase AdrA regulates flagellar biosynthesis in *Pseudomonas fluorescens* F113 through SadB. Sci Rep 9:8096. https://doi.org/10.1038/s41598-019-44554-z.

56. Martínez-Granero F, Navazo A, Barahona E, Redondo-Nieto M, Rivilla R, Martín M. 2012. The Gac-Rsm and SadB signal transduction pathways converge on AlgU to downregulate motility in *Pseudomonas fluorescens*. PLoS One 7:e31765. https://doi.org/10.1371/journal.pone.0031765.

57. Chaparro JM, Badri DV, Bakker MG, Sugiyama A, Manter DK, Vivanco JM. 2013. Root exudation of phytochemicals in *Arabidopsis* follows specific patterns that are developmentally programmed and correlate with soil microbial functions. PLoS One 8:e55731. https://doi.org/10.1371/journal.pone.0055731.

58. Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J Mol Biol 430:2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007.

59. Rudrappa T, Biedrzycki ML, Kunjeti SG, Donofrio NM, Czymmek KJ, Paré PW, Bais HP. 2010. The rhizobacterial elicitor acetoin induces systemic resistance in *Arabidopsis thaliana*. Commun Integr Biol 3:130–138. https://doi.org/10.4161/cib.3.2.10584.

60. Haas D, Keel C. 2003. Regulation of antibiotic production in root-colonizing *Pseudomonas* spp. and relevance for biological control of plant disease. Annu Rev Phytopathol 41:117–153. https://doi.org/10.1146/annurev.phyto.41.052002.095656.

61. Raaijmakers JM, De Bruijn I, Nybroe O, Ongena M. 2010. Natural functions of lipopeptides from *Bacillus* and *Pseudomonas*: more than surfactants and antibiotics. FEMS Microbiol Rev 34:1037–1062. https://doi.org/10.1111/j.1574-6976.2010.00221.x.

62. Knights HE, Jorrin B, Haskett TL, Poole PS. 4 February 2021. Deciphering bacterial mechanisms of root colonization. Environ Microbiol Rep https://doi.org/10.1111/1758-2229.12934.

63. Colaianni NR, Parys K, Lee H-S, Conway JM, Kim NH, Edelbacher N, Mucyn TS, Madalinski M, Law TF, Jones CD, Belkhadir Y, Dangl JL. 2021. A complex immune response to flagellin epitope variation in commensal communities. Cell Host Microbe 29:635–649.e9. https://doi.org/10.1016/j.chom.2021.02.006.

64. Parys K, Colaianni NR, Lee H-S, Hohmann U, Edelbacher N, Trgovcevic A, Blahovska Z, Lee D, Mechtler A, Muhari-Portik Z, Madalinski M, Schandry N, Rodríguez-Arévalo I, Becker C, Sonnleitner E, Korte A, Bläsi U, Geldner N, Hothorn M, Jones CD, Dangl JL, Belkhadir Y. 2021. Signatures of antagonistic pleiotropy in a bacterial flagellin epitope. Cell Host Microbe 29:620–634.e9. https://doi.org/10.1016/j.chom.2021.02.008.

65. Sobrero PM, Muzlera A, Frescura J, Jofré E, Valverde C. 2017. A matter of hierarchy: activation of orfamide production by the post-transcriptional Gac-Rsm cascade of Pseudomonas protegens CHA0 through expression upregulation of the two dedicated transcriptional regulators. Environ Microbiol Rep 9:599–611. https://doi.org/10.1111/1758-2229.12566.

66. Romling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev 77:1–52. https://doi.org/10.1128/MMBR.00043-12.

67. Caiazza NC, Merritt JH, Brothers KM, O'Toole GA. 2007. Inverse regulation of biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. J Bacteriol 189:3603–3612. https://doi.org/10.1128/JB.01685-06.

68. Simm R, Morr M, Kader A, Nimtz M, Römling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Mol Microbiol 53:1123–1134. https://doi.org/10.1111/j.1365-2958.2004.04206.x.

69. Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, D'Argenio DA, Miller SI, Ramsey BW, Speert DP, Moskowitz SM, Burns JL, Kaul R, Olson MV. 2006. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. Proc Natl Acad Sci U S A 103:8487–8492. https://doi.org/10.1073/pnas.0602138103.

70. Winstanley C, O'Brien S, Brockhurst MA. 2016. *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. Trends Microbiol 24:327–337. https://doi.org/10.1016/j.tim.2016.01.008.

71. Lerouge I, Vanderleyden J. 2002. O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions. FEMS Microbiol Rev 26:17–47. https://doi.org/10.1111/j.1574-6976.2002.tb00597.x.

72. Whitfield C, Trent MS. 2014. Biosynthesis and export of bacterial lipopolysaccharides. Annu Rev Biochem 83:99–128. https://doi.org/10.1146/annurev-biochem-060713-035600.

73. Maldonado RF, Sá-Correia I, Valvano MA. 2016. Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. FEMS Microbiol Rev 40:480–493. https://doi.org/10.1093/femsre/fuw007.

74. Murray GL, Attridge SR, Morona R. 2006. Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of *Salmonella enterica* serovar Typhimurium with macrophages and complement. J Bacteriol 188:2735–2739. https://doi.org/10.1128/JB.188.7.2735-2739.2006.

75. Ranf S. 2016. Immune sensing of lipopolysaccharide in plants and animals: same but different. PLoS Pathog 12:e1005596. https://doi.org/10.1371/journal.ppat.1005596.

76. Newman M-A, von Roepenack E, Daniels M, Dow M. 2000. Lipopolysaccharides and plant responses to phytopathogenic bacteria. Mol Plant Pathol 1:25–31. https://doi.org/10.1046/j.1364-3703.2000.00004.x.

77. Di Lorenzo F, Silipo A, Andersen Gersby LB, Palmigiano A, Lanzetta R, Garozzo D, Boyer C, Pruvost O, Newman MA, Molinaro A. 2017. *Xanthomonas citri* pv. *citri* pathotypes: LPS structure and function as microbe-associated molecular patterns. Chembiochem 18:772–781. https://doi.org/10.1002/cbic.201600671.

78. Newman MA, Daniels MJ, Dow JM. 1995. Lipopolysaccharide from *Xanthomonas campestris* induces defense-related gene expression in *Brassica campestris*. Mol Plant Microbe Interact 8:778–780. https://doi.org/10.1094/mpmi-8-0778.

79. Barton-Willis PA, Wang MC, Holliday MJ, Long MR, Keen NT. 1984. Purification and composition of lipopolysaccharides from *Pseudomonas syringae*. pv. *glycinea*. Physiol Plant Pathol 25:387–398. https://doi.org/10.1016/0048-4059(84)90045-6.

80. Erbs G, Newman M-A. 2003. The role of lipopolysaccharides in induction of plant defence responses. Mol Plant Pathol 4:421–425. https://doi.org/10.1046/j.1364-3703.2003.00179.x.

81. Leeman M, van Pelt JA, den Ouden FM, Heinsbroek M, Bakker PAHM, Schippers B. 1995. Induction of systemic resistance against *Fusarium wilt* of radish by lipopolysaccharides of *Pseudomonas fluorescens*. Phytopathology 85:1021–1027. https://doi.org/10.1094/Phyto-85-1021.

82. Van Wees SC, Pieterse CM, Trijssenaar A, Van 't Westende YA, Hartog F, Van Loon LC. 1997. Differential induction of systemic resistance in *Arabidopsis* by biocontrol bacteria. Mol Plant Microbe Interact 10:716–724. https://doi.org/10.1094/MPMI.1997.10.6.716.

83. Nakao R, Senpuku H, Watanabe H. 2006. *Porphyromonas gingivalis galE* is involved in lipopolysaccharide O-antigen synthesis and biofilm formation. Infect Immun 74:6145–6153. https://doi.org/10.1128/IAI.00261-06.

84. Ramette A, Frapolli M, Saux MF-L, Gruffaz C, Meyer J-M, Défago G, Sutra L, Moënne-Loccoz Y. 2011. *Pseudomonas protegens* sp. nov., widespread plant-protecting bacteria producing the biocontrol compounds 2,4-diacetylphloroglucinol and pyoluteorin. Syst Appl Microbiol 34:180–188. https://doi.org/10.1016/j.syapm.2010.10.005.

85. Stutz EW, Défago G, Kern H. 1986. Naturally occurring fluorescent pseudomonads involved in suppression of black root rot of tobacco. Phytopathology 76:181–185. https://doi.org/10.1094/Phyto-76-181.

86. Jousset A, Lara E, Wall LG, Valverde C. 2006. Secondary metabolites help biocontrol strain *Pseudomonas fluorescens* CHA0 to escape protozoan grazing. Appl Environ Microbiol 72:7083–7090. https://doi.org/10.1128/AEM.00557-06.

87. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain *w1118; iso-2; iso-3*. Fly (Austin) 6:80–92. https://doi.org/10.4161/fly.19695.

88. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033.

89. Højberg O, Schnider U, Winteler HV, Sørensen J, Haas D. 1999. Oxygen-sensing reporter strain of *Pseudomonas fluorescens* for monitoring the distribution of low-oxygen habitats in soil. Appl Environ Microbiol 65:4085–4093. https://doi.org/10.1128/AEM.65.9.4085-4093.1999.

90. Glickmann E, Dessaux Y. 1995. A critical examination of the specificity of the Salkowski reagent for indolic compounds produced by phytopathogenic bacteria. Appl Environ Microbiol 61:793–796. https://doi.org/10.1128/AEM.61.2.793-796.1995.

91. Alexander DB, Zuberer DA. 1991. Use of chrome azurol S reagents to evaluate siderophore production by rhizosphere bacteria. Biol Fertil Soils 12:39–45. https://doi.org/10.1007/BF00369386.

92. Smetzler MS, Hart ME, Iandolo JJ. 1993. Phenotypic characterization of *xpr*, a global regulator of extracellular virulence factors in *Staphylococcus aureus*. Infect Immun 61:919–925. https://doi.org/10.1128/IAI.61.3.919-925.1993.

93. Oberhänsli T, Dfago G, Haas D. 1991. Indole-3-acetic acid (IAA) synthesis in the biocontrol strain CHA0 of *Pseudomonas fluorescens*: role of tryptophan side chain oxidase. J Gen Microbiol 137:2273–2279. https://doi.org/10.1099/00221287-137-10-2273.

94. Moskowitz SM, Foster JM, Emerson J, Burns JL. 2004. Clinically feasible biofilm susceptibility assay for isolates of *Pseudomonas aeruginosa* from patients with cystic fibrosis. J Clin Microbiol 42:1915–1922. https://doi.org/10.1128/JCM.42.5.1915-1922.2004.

95. Christensen PJ, Cook FD. 1972. The isolation and enumeration of cytophagas. Can J Microbiol 18:1933–1940. https://doi.org/10.1139/m72-299.

96. Déziel E, Comeau Y, Villemur R. 2001. Initiation of biofilm formation by *Pseudomonas aeruginosa* 57RP correlates with emergence of hyperpiliated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. J Bacteriol 183:1195–1204. https://doi.org/10.1128/JB.183.4.1195-1204.2001.

97. Murtagh F, Legendre P. 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J Classif 31:274–295. https://doi.org/10.1007/s00357-014-9161-z.

98. Scrucca L, Fop M, Murphy TB, Raftery AE. 2016. Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. R J 8:289–317. https://doi.org/10.32614/RJ-2016-021.

99. Lee HK, Lee CK, Loh TP, Tang JW-T, Tambyah PA, Koay ES-C. 2011. High-resolution melting approach to efficient identification and quantification of H275Y mutant influenza H1N1/2009 virus in mixed-virus-population samples. J Clin Microbiol 49:3555–3559. https://doi.org/10.1128/JCM.01087-11.

100. Zhong W-L, Wang L, Wu X, Zhang J, Chen X-F, Zhang W, Dou X, Yu B. 2016. Development of unlabeled probe based high-resolution melting analysis for detection of filaggrin gene mutation c.3321delA. J Clin Lab Anal 30:892–896. https://doi.org/10.1002/jcla.21953.

101. Capper RL, Jin YK, Lundgren PB, Peplow LM, Matz MV, van Oppen MJH. 2015. Quantitative high resolution melting: two methods to determine SNP allele frequencies from pooled samples. BMC Genet 16:62. https://doi.org/10.1186/s12863-015-0222-z.

102. Noble R. 2019. ggmuller. Create Muller plots of evolutionary dynamics [R package version 0.5.4]. https://rdrr.io/cran/ggmuller/.

103. Galperin MY, Makarova KS, Wolf YI, Koonin EV. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res 43:D261–D269. https://doi.org/10.1093/nar/gku1223.

104. Aravind L, Koonin EV. 1998. The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem Sci 23:469–472. https://doi.org/10.1016/S0968-0004(98)01293-6.