





# Horizontally Acquired Quorum-Sensing Regulators Recruited by the PhoP Regulatory Network Expand the Host Adaptation Repertoire in the Phytopathogen *Pectobacterium brasiliense*

 Daniel Bellieny-Rabelo,<sup>a,b</sup> Ntombikayise Precious Nkomo,<sup>a,b</sup> Divine Yufetar Shyntum,<sup>a,b</sup>  Lucy Novungayo Moleleki<sup>a,b</sup>

<sup>a</sup>Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, Gauteng, South Africa

<sup>b</sup>Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, Gauteng, South Africa

Daniel Bellieny-Rabelo and Ntombikayise Precious Nkomo contributed equally to this work. Author order was determined in order of decreasing seniority.

**ABSTRACT** In this study, we examine the impact of transcriptional network rearrangements driven by horizontal gene acquisition in PhoP and SlyA regulons using as a case study a phytopathosystem comprised of potato tubers and the soft-rot pathogen *Pectobacterium brasiliense* 1692 (Pb1692). Genome simulations and statistical analyses uncovered the tendency of PhoP and SlyA networks to mobilize lineage-specific traits predicted as horizontal gene transfer at late infection, highlighting the prominence of regulatory network rearrangements in this stage of infection. The evidence further supports the circumscription of two horizontally acquired quorum-sensing regulators (*carR* and *expR1*) by the PhoP network. By recruiting *carR* and *expR1*, the PhoP network also impacts certain host adaptation- and bacterial competition-related systems, seemingly in a quorum sensing-dependent manner, such as the type VI secretion system, carbapenem biosynthesis, and plant cell wall-degrading enzymes (PCWDE) like cellulases and pectate lyases. Conversely, polygalacturonases and the type III secretion system (T3SS) exhibit a transcriptional pattern that suggests quorum-sensing-independent regulation by the PhoP network. This includes an uncharacterized novel phage-related gene family within the T3SS gene cluster that has been recently acquired by two *Pectobacterium* species. The evidence further suggests a PhoP-dependent regulation of carbapenem- and PCWDE-encoding genes based on the synthesized products' optimum pH. The PhoP network also controls *slyA* expression *in planta*, which seems to impact carbohydrate metabolism regulation, especially at early infection, when 76.2% of the SlyA-regulated genes from that category also require PhoP to achieve normal expression levels.

**IMPORTANCE** Exchanging genetic material through horizontal transfer is a critical mechanism that drives bacteria to efficiently adapt to host defenses. In this report, we demonstrate that a specific plant-pathogenic species (from the *Pectobacterium* genus) successfully integrated a population density-based behavior system (quorum sensing) acquired through horizontal transfer into a resident stress-response gene regulatory network controlled by the PhoP protein. Evidence found here underscores that subsets of bacterial weaponry critical for colonization, typically known to respond to quorum sensing, are also controlled by PhoP. Some of these traits include different types of enzymes that can efficiently break down plant cell walls depending on the environmental acidity level. Thus, we hypothesize that PhoP's ability to elicit regulatory responses based on acidity and nutrient availability fluctuations has strongly impacted the fixation of its regulatory connection with quorum sensing. In addition, another global gene regulator, known as SlyA, was found under the PhoP regulatory network. The SlyA regulator controls a series of carbohydrate metabolism-related traits, which also seem to be regulated by PhoP. By centralizing quorum sensing and *slyA* under PhoP scrutiny, *Pectobacterium*

**Citation** Bellieny-Rabelo D, Nkomo NP, Shyntum DY, Moleleki LN. 2020. Horizontally acquired quorum-sensing regulators recruited by the PhoP regulatory network expand the host adaptation repertoire in the phytopathogen *Pectobacterium brasiliense*. mSystems 5:e00650-19. <https://doi.org/10.1128/mSystems.00650-19>.

**Editor** Matthew F. Traxler, University of California, Berkeley

**Copyright** © 2020 Bellieny-Rabelo et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Lucy Novungayo Moleleki, [lucy.moleleki@up.ac.za](mailto:lucy.moleleki@up.ac.za).

**Received** 8 October 2019

**Accepted** 25 December 2019

**Published** 28 January 2020

cells added an advantageous layer of control over those two networks that potentially enhances colonization efficiency.

**KEYWORDS** *Pectobacterium*, carbapenems, gene regulatory networks, horizontal gene transfer, quorum sensing, type III secretion systems, type VI secretion systems

Highly specialized colonization traits are constantly selected in phytopathogenic Gram-negative bacteria to overcome severe obstacles imposed by plant apoplasts. The apoplastic space comprises a nutrient-poor milieu that harbors an extensive inventory of toxic and defense-related molecules, such as plant antimicrobial peptides, reactive oxygen species, and plant organic compounds (1, 2). In addition, the apoplastic space is generally acidic (ranging between pH 4.5 and 6.5) due to the presence of organic acids such as malic and citric acids, which also enforces the anti-growth strategy against pathogenic invasions (3). In this context, a crucial regulatory system in bacteria, which is frequently associated with response to acidic stress, is the PhoQ/PhoP two-component system (4). This response-based system is highly conserved and widely studied across bacterial lineages because of its prominent role as a global transcriptional regulator.

The simultaneous flow of transcription/translation in bacterial cells inhabiting such competitive environments requires highly optimized regulatory mechanisms to ensure accurate responses during colonization. Thus, to cope with frequent gene gains and losses that result mainly from horizontal transfer events, bacterial regulatory circuits must efficiently accommodate newly acquired genes (5). Horizontal gene transfer is perceived as a critical driving force in the evolution of bacterial genomes from which a number of pathogenicity themes emerge, including some secretion systems and prophages (6, 7). Hence, it is not surprising that the influence of important virulence regulators such as SlyA (8) or PecS (9) over regions incorporated through horizontal gene transfer (HGT) have been previously reported. Both SlyA and PecS belong to the MarR family of transcriptional regulators, which encompasses a range of genes involved in virulence and antibiotic resistance control (10–12).

Among Gram-negative bacteria, one specific group, commonly referred to as soft-rot *Pectobacteriaceae* (SRP) (13, 14) (formerly known as soft-rot *Enterobacteriaceae*), has increasingly gained attention over the last few decades as causative agents of wilt/blackleg diseases in economically important crops worldwide (15, 16). This group is most prominently represented by *Pectobacterium* and *Dickeya* genera. The SRPs are opportunistic Gram-negative pathogens capable of producing distinctly larger amounts of pectinolytic enzymes than other pectolytic bacteria (e.g., *Clostridium* spp., *Bacillus* spp., and *Pseudomonas* spp.) (17). These plant cell wall-degrading enzymes (PCWDEs) encompass a variety of families that concertedly promote disease through tissue maceration (18). While some PCWDE classes exhibit close to neutral or high optimum pH, such as cellulases and pectate and pectin lyases (Cel, Pel, and Pnl), others function at low optimum pH, namely, polygalacturonases (Peh) (19, 20). In this sense, the expression of different groups of PCWDEs tends to be regulated according to the pH within the plant tissue, which is acidic in the apoplast at first and then becomes basic as the disease progresses (17). In this context, one of the best-characterized mechanisms of PCWDE regulation in SRP pathogens is quorum sensing (QS) (21). This is a crucial regulatory circuit controlling population density-based behavior in SRP as well as in a large spectrum of bacteria (22, 23). Quorum-sensing networks in Gram-negative bacteria rely on recognition of autoinducer molecules, such as acyl-homoserine lactones (AHL) and other products synthesized using *S*-adenosylmethionine, by specialized receptors such as ExpR/LuxR homologs (24, 25). The complexed form, ExpR/LuxR-autoinducer, then can bind specific DNA regions to participate in transcriptional regulation (26). One of the regions regulated by the ExpR-autoinducer complex is the *rsmA* promoter region. Thereafter, the transcriptional repression of *rsmA* prevents RsmA-induced inactivation of PCWDEs (27). In addition to PCWDEs, QS has been shown to regulate other traits involved in virulence and interbacterial competition, such as the

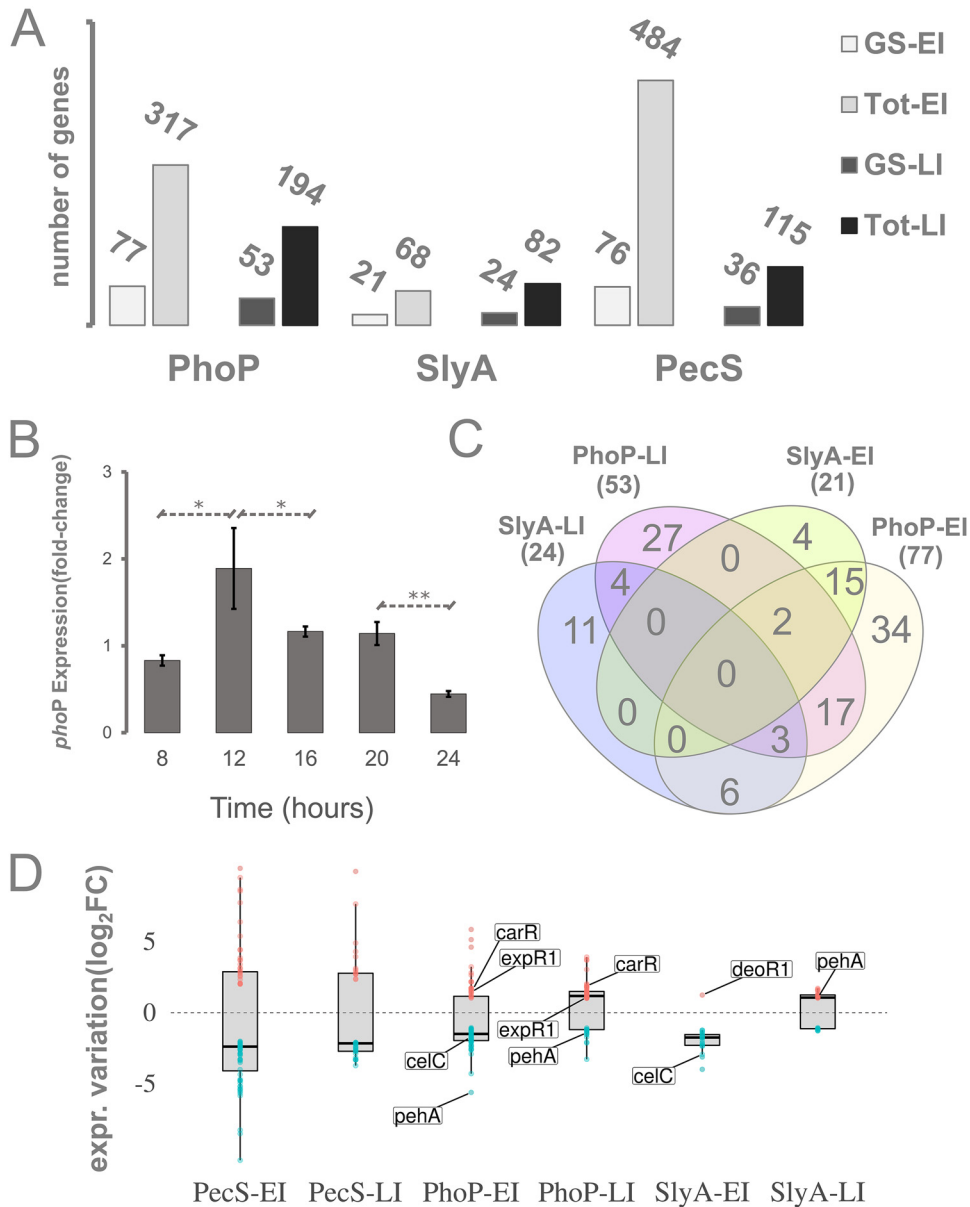
carbapenem antibiotic (*car* genes) and the type III and VI secretion systems (T3SS and T6SS) in *Pectobacterium* spp. (21, 28).

In the present report, we investigate the impact of transcriptional network rearrangements in two global transcriptional regulators (PhoP and SlyA) on the *in planta* regulation of crucial host adaptation- and fitness-oriented systems. Since regulatory network rearrangement events have the potential to give rise to phenotypic differences in pathogenesis development among these organisms, understanding these processes is crucial to unveiling how these networks shape pathogenicity in different lineages.

## RESULTS

**RNA-Seq mapping and the reach of PhoP and SlyA regulons *in planta*.** To examine both PhoP and SlyA regulatory networks *in planta*, we engineered Pb1692 $\Delta$ *slyA* and Pb1692 $\Delta$ *phoP* mutants using the lambda recombination technique (see Materials and Methods for details). The integrity of the mutants and complement strains were confirmed by PCR analyses (see Fig. S1A and B in the supplemental material) as well as by DNA sequencing. Additionally, an *in vitro* growth assay demonstrated that the deletion of neither the *phoP* nor *slyA* gene impaired the growth of the mutant strains (Fig. S1C and D). The global impact caused by *phoP* or *slyA* deletion in the Pb1692 genome toward *in planta* transcriptional profiles next was analyzed in original whole-transcriptome data sets comprising ~124 million transcriptome sequencing (RNA-Seq) reads (Table S1). By comparing significant gene expression changes between the wild type and either the Pb1692 $\Delta$ *slyA* or Pb1692 $\Delta$ *phoP* mutant, we infer transcriptional regulation affected by these regulatory networks (see Materials and Methods for details). Additional validation by quantitative reverse transcription-PCR (qRT-PCR) of the differentially expressed genes of particular interest for this study was also conducted (Fig. S2). The samples selected for this study depicted two stages of plant infection, 12 and 24 h postinfection (hpi), here denoted early (EI) and late (LI) infection, respectively. The data set exhibited good quality, as 89.5% (~111 million) of the reads were mapped on the Pb1692 reference genome (Table S1). Of these, 89% (~99 million reads) were uniquely mapped on the reference genome. Transcriptional variations between wild-type Pb1692 and mutant strains during infection enabled collective identification of 480 genes affected (up-/downregulation; see Materials and Methods for details) either by PhoP or SlyA regulatory networks at the two time points (Table S2). In this context, the prominent transcriptional reach of PhoP compared to that of SlyA during plant infection was observed, as the former directly or indirectly controls 410 genes but the latter controls 145.

From the PhoP regulon, we detected a 36.1% reduction in the total network size between early and late infection (Fig. 1A). This phenomenon could result from a variety of non-mutually exclusive events, among which the most predictable would be (i) increased number of active transcriptional regulators at 24 hpi competing with PhoP for promoter binding regions, (ii) PhoP dephosphorylation by PhoQ following the absence of triggering conditions, or (iii) decreased transcription of the *phoP* gene. Thus, to further inspect this shift in PhoP network sizes between EI and LI, we set out to evaluate the transcriptional variation of the *phoP* gene *in planta* by the Pb1692 wild type. First, Pb1692 was inoculated into potato tubers and samples were harvested at 8, 12, 16, 20, and 24 hpi. Next, qRT-PCR was used to analyze the expression of the *phoP* gene in those samples. The results showed that the relative expression of the *phoP* gene nearly doubled between 8 and 12 hpi (Fig. 1B). This increase was followed by a significant decrease ( $P = 0.042$ ) between 12 and 16 hpi. After keeping a constant expression level between 16 and 20 hpi, Pb1692 cells drastically decrease *phoP* gene expression at 24 hpi ( $P = 0.004$ ). This converges with the shrinkage observed in the network size. Hence, since PhoP is frequently associated with transcriptional responses to acidic environments and low  $Mg^{2+}$  concentration, the elevated *phoP* transcription in the early stages of infection is arguably a direct response to the nutrient-poor environment found in the apoplast. Conversely, after a certain time point between 20 and 24 hpi, *phoP* transcription decreases, which is consistent with the reduced cellular



**FIG 1** Genome-wide regulatory impact of three global regulators in Pb1692 and Dd3937 and the transcriptional levels of *phoP* regulator during plant infection. (A) Genus-specific (GS) versus total gene (Tot) amounts in PhoP, SlyA, and PecS regulons *in planta* both at early and late infection (EI and LI) are represented in the bar plot. PhoP and SlyA regulons were extracted from Pb1692 infection samples. PecS regulons were extracted from Dd3937 samples. (B) Transcriptional levels of the *phoP* gene measured by qRT-PCR in 4-h intervals in potato tubers infected by Pb1692 across the first 24 hpi. Consecutive time points were tested for statistically significant (*t* test) differences in their expression levels. Significant results are marked with asterisks in the graph (\*,  $P < 0.05$ ; \*\*,  $P < 0.005$ ). (C) Venn diagrams depicting genus-specific genes found in PhoP and SlyA *in planta* regulons at EI and LI assessed in Pb1692. The total number of genus-specific genes found within each regulon is highlighted under each regulon label. (D) Gene transcriptional variations ( $\log_2$  fold change) are depicted specifically for genus-specific genes. Transcriptional variation data were extracted from PhoP, SlyA, and PecS regulons either at early or late infection (EI and LI). The box plots are overlaid with dot plots, in which each dot represents a single gene expression value. In PhoP and SlyA regulons, specific genes are labeled for polygalacturonase A (*pehA*, AED-0002061), quorum-sensing regulators (*carR*, AED-0003542; *expR1*, AED-0000069), a transcriptional regulator (*deoR1*, AED-0001096), and a PTS permease gene (*celC*, AED-0001290).

demand for this regulator following environmental alkalization due to progressive host cell lysis.

**Determining the genus-specific content in PhoP, SlyA, and PecS *in planta* regulons.** The prevalence of distinctive traits (genus specific) next was inspected in the full range of transcriptionally regulated genes by three global regulators frequently associated with virulence and/or stress responses. This was achieved by adding the PecS regulon data set to this analysis, which was the only publicly available *in planta* whole-transcriptome data set obtained from an SRP mutant strain lacking a transcriptional regulator (9). The PecS regulon was obtained from *Dickeya dadantii* strain 3937 (Dd3937). We then analyzed 6 distinct regulons obtained from those 3 global regulators in two different stages (i.e., early and late) of plant infection. Hence, this analysis focuses on three early infection regulons, i.e., PhoP-, SlyA-, and PecS-EI, along with three late infection regulons, i.e., PhoP-, SlyA-, and PecS-LI. The identification of distinctive systems between the two genera was achieved by predicting genus-specific genes in *Pectobacterium* and *Dickeya* genera based on protein sequence orthology. This analysis was underpinned by an extensive correlational database previously generated that included 39 and 61 strains from *Dickeya* and *Pectobacterium* genera, respectively (6). By using this database, we inquired, for each gene in a given strain, how many orthologs could be detected in other strains within the genus as well as in the opposing genus. Those gene products for which no orthologous counterparts exist in the opposing genus were assigned, given the scope of this study, as GS in the SRP context, i.e., *Pectobacterium* exclusive or *Dickeya* exclusive.

The presence of GS protein-coding genes next was surveyed within each of the *in planta* regulons. The preliminary aim was to measure the relative presence of GS genes in these regulons in specific infection stages. Our screening revealed that across the three early infection regulons (from PhoP, SlyA, and PecS), 15 to 31% of the regulated genes are GS. Interestingly, at late infection, two out of three regulons (except for SlyA) exhibited increased relative GS content, ranging between 27 and 31% (Fig. 1A). Notably, the PecS regulons exhibited the most drastic shift in the GS content proportion between EI (15%) and LI (31%). Furthermore, by comparing the GS gene sets of SlyA and PhoP regulons from Pb1692, the presence of GS genes that are exclusive to individual regulons (regulon exclusive) was more conspicuous in those of late infection. Thus, in PhoP and SlyA data sets, 46 to 51% (27/53 and 11/24) of the GS genes regulated at LI are regulon exclusive (Fig. 1C). Conversely, 19 to 44% of the GS (4/21 and 34/77) were regulon exclusive among EI samples. This analysis revealed that at late infection, PhoP, SlyA (in Pb1692), and PecS (in Dd3937) tend to mobilize transcription of an equal or greater proportion of GS genes than during early infection. Additionally, it also showed that PhoP and SlyA exhibit larger relative amounts of GS regulon-exclusive genes at late infection than during early infection.

In the SlyA-EI data set, the only GS gene displaying increased expression was a MarR transcriptional regulator homolog to *deoR1* (AED-0001096), which is frequently associated with repression of carbohydrate metabolism (Fig. 1D). Interestingly, the *deoR1* homolog in Pb1692 is regulon exclusive from SlyA-EI. Furthermore, besides promoting activation of *deoR1*, the absence of SlyA also causes significant repression over several sugar permeases, including the GS gene *celC* (AED-0001290). All of these permeases also seem to be affected by *phoP* deletion in early stages of infection. This indicates a particular demand for the transcriptional regulation of certain carbohydrate metabolism-related systems controlled by *slyA* and *deoR1* regulators in specific stages of infection.

Moreover, a *Pectobacterium*-exclusive gene encoding a polygalacturonase (Peh) was found within PhoP *in planta* regulons (Fig. 1D). Converging with the evidence for the transcriptional regulation of this PCWDE-encoding gene, especially by PhoP, we also found two QS regulators under PhoP control, namely, *expR1* and *carR* (Fig. 1D). These QS regulators are GS genes in *Pectobacterium* genomes. While ExpR1 is often, but not exclusively, associated with transcriptional regulation of pectinolytic enzymes (21, 29), CarR is associated with regulation of carbapenem antibiotic biosynthesis (28). The



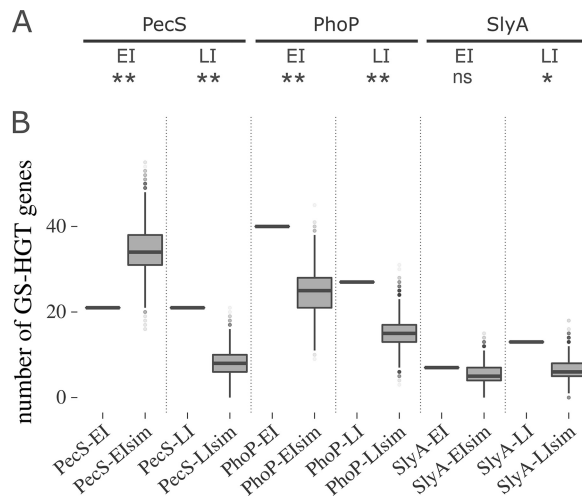
presence of *luxR-expR* homologs within the PhoP network might be directly linked to the presence of two bacterial secretion systems. This type of association may directly impact fitness, which will be inspected in a subsequent section.

**Horizontal acquisitions in genus-specific contents and their regulation by PhoP, SlyA, and PecS.** While the presence of GS genes across the six regulons was observed, their origin was still unclear. The GS genes observed here may have emerged through different processes, such as (i) gene loss in the neighboring lineage, (ii) intralinear functional divergence following gene duplication, or (iii) gene acquisition via HGT. Hence, aiming to evaluate the prevalence of transcriptional network rearrangement processes, especially in Pb1692 networks, we set out to predict HGT candidates using parametric methods. Per previous observations (30), we chose a gene-based method (GC content at third codon position [GC3]) and a window-based method (dinucleotide frequencies [DINT\_KL]) to evaluate sequence composition bias in Pb1692 and Dd3937 genes (see Materials and Methods for details). These predictions will be further compared to the orthology-based prediction of genus-specific genes as an additional layer of evidence.

By examining the entire set of GS genes in Pb1692 (787 genes), we found that 36.2 and 19.3%, respectively, were predicted as HGT candidates by GC3 or DINT\_KL methods. Importantly, 53% of the high-confidence HGT candidates, meaning those supported by both GC3 and DINT\_KL, belong to the GS gene set. On the other hand, in Dd3937 37.8% of the high-confidence HGT candidates are GS genes. This preliminary inquiry revealed that the GS content in Pb1692 tends to include a larger portion of horizontally acquired genes than Dd3937. Furthermore, the inspection of PecS *in planta* regulons at early and late infection in Dd3937 revealed that 6.6% and 44.4%, respectively, of the GS genes are high-confidence HGT candidates. On the other hand, by examining PhoP and SlyA *in planta* regulons in Pb1692, the high-confidence candidates within the GS genes accounted for 24.7% and 14.3%, respectively, at early infection and 30.2% and 37.5%, respectively, at late infection. In this context, the genus-specific HGT candidate (GS-HGT) genes most likely arise from recent gene acquisitions in a given genus. Hence, this preliminary analysis supports that the regulation of some distinctive traits observed among *Pectobacterium* and *Dickeya* genera involves recent rearrangements in PhoP, SlyA, and PecS regulatory networks driven by HGT.

To inspect whether these regulatory networks are biased toward or against controlling GS-HGT candidate genes, we conducted a statistical enrichment analysis. The results from Fisher's exact tests showed that, for the three LI regulons analyzed (from PhoP, SlyA, and PecS), there is an overrepresentation of GS-HGT genes (Fig. 2A). Among them, PhoP is the only regulator exhibiting overrepresentation of GS-HGT genes both on EI and LI regulatory networks. Curiously, the PecS-EI regulon exhibits a negative correlation with GS-HGT genes. We next aimed to further confirm these results using a different approach. Computational simulations were performed to shuffle the genomes from each organism (i.e., Pb1692 and Dd3937), generating pseudogenomes. This strategy aims to empirically evaluate the frequency at which GS-HGT genes should be expected by chance in these regulons and compare this with the real data. For each of the six regulons analyzed, this experiment generated 10,000 shuffled pseudogenomes (see Materials and Methods for details). The results corroborated, by empirical means, what statistical analysis had previously established (Fig. 2B). Overall, these analyses showed that transient shifts, in this case due to temporal variation, in regulon profiles can alter their bias toward the presence of recently acquired genes. Although at late infection all regulons exhibited a consistent bias toward the regulation of GS-HGT genes, only one of them (PhoP) displayed similar bias at early infection. On the contrary, the PecS network seemed to avoid transcriptional interactions with GS-HGT genes at early infection. Overall, this evidence suggests a particular requirement for expression of distinctive traits seemingly acquired through HGT in these organisms at late stages of infection.

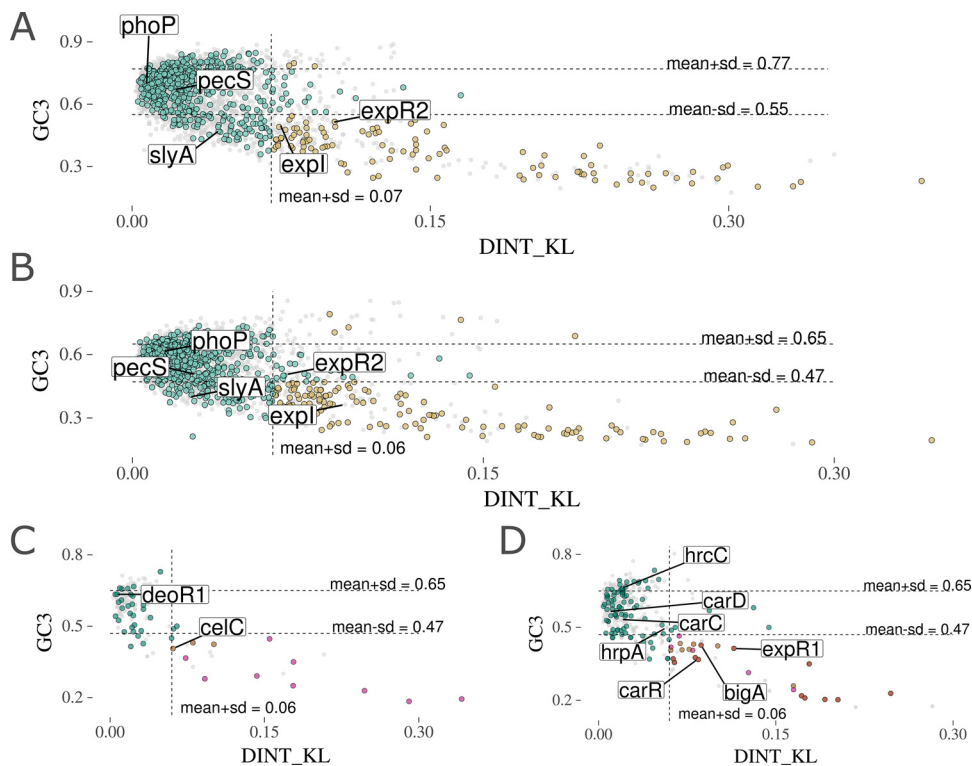
**Host adaptation and regulatory systems acquired by PhoP and SlyA networks.** Highly conserved genes among the SRPs, such as *phoP* (AED-0004376 in Pb1692,



**FIG 2** Overrepresentation of genus-specific HGT candidate genes in PhoP, SlyA, and PecS regulons in Pb1692 and Dd3937. (A) The significance of genus-specific HGT candidate (GS-HGT) gene occurrence in each regulon was analyzed through two-tailed Fisher exact tests using the respective species genomes as the background set (\*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; nonsignificant [ns],  $P > 0.05$ ). (B) Each section of the plot (separated by dotted lines) represents the comparison between the actual number of GS-HGT genes found in each regulon as a single horizontal line and the distribution of GS-HGT gene amounts found in each of the 10,000 computationally shuffled versions of the respective regulons (see Materials and Methods for details). The PecS data sets were obtained from Dd3937, whereas PhoP and SlyA data sets were obtained from Pb1692. The simulated data sets are labeled with the Sim postfix in the graph.

*DDA3937\_RS11500* in Dd3937) or *pecS* (*AED-0003935* and *DDA3937\_RS20885*), exhibited overall sequence composition similar to that of their resident genomes (Dd3937 or Pb1692) (Fig. 3A and B and Table S3). Unlike these two regulators, *slyA* (*AED-0001000* and *DDA3937\_RS12595*) homologs consistently exhibited deviant GC3 indexes compared to those of the two resident genomes analyzed, although in terms of dinucleotide frequencies a significant distance was not observed. Thus, considering the conservation of *slyA* across all SRP genomes, it is likely that the acquisition of this gene occurred at least in the last common ancestor between *Pectobacterium* and *Dickeya* genera. We next observed that the SlyA-regulated GS transcriptional regulator *deoR1* (*AED-0001096*) had no support from HGT prediction (Fig. 3C). This indicates that this gene was lost in the *Dickeya* lineage, whereas most of the *Pectobacterium* retained it. In contrast, the SlyA-regulated gene involved in carbohydrate uptake, *celC* (*AED-0001290*), was predicted as an HGT candidate with high confidence. This indicates that the *Pectobacterium* genus benefits from increasing the regulatory complexity of carbohydrate metabolism either through gene gain or by simply keeping ancestral genes in the lineage. This notion is further corroborated by the enrichment of genes associated with carbohydrate metabolism in the SlyA-EI regulon (Fig. S3).

As for quorum-sensing regulators, the *expR2-virR* homologs showed stronger support for horizontal acquisition in Dd3937 (*DDA3937\_RS20730*), supported by both GC3 and DINT\_KL, than in Pb1692 (*AED-0002741*), which indicates that *Pectobacterium* and *Dickeya* lineages acquired *expR2* independently (Fig. 3A and B). This hypothesis will be further discussed in a subsequent section. Moreover, the *Pectobacterium*-exclusive PhoP-regulated transcriptional regulator *expR1* (*AED-0000069*) was strongly supported as a horizontal acquisition by both parametric methods in Pb1692 (Fig. 3D). This result is further corroborated by a similar profile observed in the Pb1692 *expl* gene (*AED-0000062*) (Fig. 3A and B). Furthermore, the carbapenem biosynthesis transcriptional regulator *carR* (*AED-0003542*) is another GS gene in Pb1692 that has been successfully supported as an HGT candidate with high confidence (Fig. 3D). In addition, *carR* not only apparently has been horizontally transferred into the Pb1692 genome but also is the only gene in the carbapenem biosynthesis gene cluster supported by both parametric methods for HGT prediction (Table S3). *carR* is also the only GS gene in the *car*



**FIG 3** Horizontal gene transfer predictions for genus-specific genes in Dd3937 and Pb1692. (A) Two sequence composition measurements methods are represented for each gene in the Dd3937 genome as GC content at the third codon position index (GC3) in the y axis and the Kullback-Leibler distance of dinucleotide frequency distributions (DINT\_KL) relative to the Dd3937 genome in the x axis. The coding sequences from Dd3937 comprise the background set, colored gray. From this set, specific genes are colored according to both the orthologous conservation in SRP genomes and sequence composition distance to the Dd3937 genome in both GC3 and DINT\_KL metrics. Those genes exhibiting above-threshold (means  $\pm$  SD) values are represented in light brown if they are genus specific (orthologs found in *Dickeya* but not in *Pectobacterium* genomes). Those genes exhibiting below-threshold values in either GC3 or DINT\_KL are represented in green. Five genes of interest in this study are the following: *pecS* (DDA3937\_RS20885), *phoP* (DDA3837\_RS11500), *slyA* (DDA3937\_RS12595), *expR2* (DDA3937\_RS20730), and *expl* (DDA3937\_RS20725). (B) The graph follows the same representation as that described for panel A. All analyses presented above for Dd3937 and the *Dickeya* genus are mirrored for Pb1692 and the *Pectobacterium* genus in panel B. Different functional classes are labeled as transcriptional regulators *pecS* (AED-0003935), *phoP* (AED-0004376), *slyA* (AED-0001000), *expR2* (AED-0002741), and *expl* (AED-0000062). (C) The entire set of genes regulated *in planta* (at early or late infection) by SlyA in Pb1692 comprises the background set, colored gray. Genus-specific genes are highlighted in green if below threshold, light brown if above threshold and regulated at early infection, pink if above threshold and regulated at late infection, or dark brown if above threshold and regulated at both early and late infection. Genes associated with carbohydrate metabolism regulation are *celC* (AED-0001290) and *deoR1* (AED-0001096). (D) The graph follows the representation described for panel C but depicting PhoP regulons (early and late infection) in Pb1692. Genes from different Pb1692 operons are the following: type III secretion operon, *hrpA* (AED-0001214), *hrcC* (AED-0003292), and *bigA* (AED-0003294); carbapenem biosynthesis associated, *carR* (AED-0003542), *carD* (AED-0003538), and *carC* (AED-0003539); and the transcriptional regulator *expR1* (AED-0000069).

gene cluster (Table S2). These results strongly suggest that *carR* was acquired by Pb1692 independently (probably at a later stage) from *carABCDEF* genes. The presence of two QS regulators (*expR1* and *carR*), along with two conspicuous genomic regions encoding bacterial secretion systems (i.e., T3- and T6SS) under PhoP regulation in Pb1692, also may be correlated. Hence, we conducted an in-depth investigation of the PhoP-dependent regulation of these secretion systems and the phenotypic outcomes that this regulatory association may give rise to.

**PhoP-dependent regulation of secretion systems and their effects on virulence and fitness.** The PhoP-dependent regulation of *Pectobacterium*-exclusive QS regulators, along with type III and VI secretion systems, prompted us to measure the ability of Pb1692 $\Delta$ *phoP* to (i) outcompete other SRP *in planta* and (ii) cause tissue maceration in potato tubers. The Pb1692 $\Delta$ *slyA* strain was also included for comparative means.



Therefore, we first performed *in planta* bacterial competition assays to determine the relative contribution of PhoP to bacterial competition. For this experiment, Dd3937 was selected because it is a well-studied niche competitor of Pb1692 in potato tubers. The growth of Dd3937 when coinoculated *in planta* with the Pb1692 $\Delta$ *phoP* mutant displayed a 5-fold reduction (in CFU per milliliter), as opposed to a 3-fold reduction when coinoculated with wild-type Pb1692 (Fig. 4A). Conversely, the Pb1692 $\Delta$ *slyA* strain inhibition against Dd3937 exhibited no significant difference from that of wild-type Pb1692. The results imply a direct outcome of the overexpression of competition-related mechanisms following *phoP* deletion. Arguably, one of the major players in this increased *in planta* competition aggressiveness is the extensive array of 23 T6SS-related genes overexpressed in the absence of PhoP.

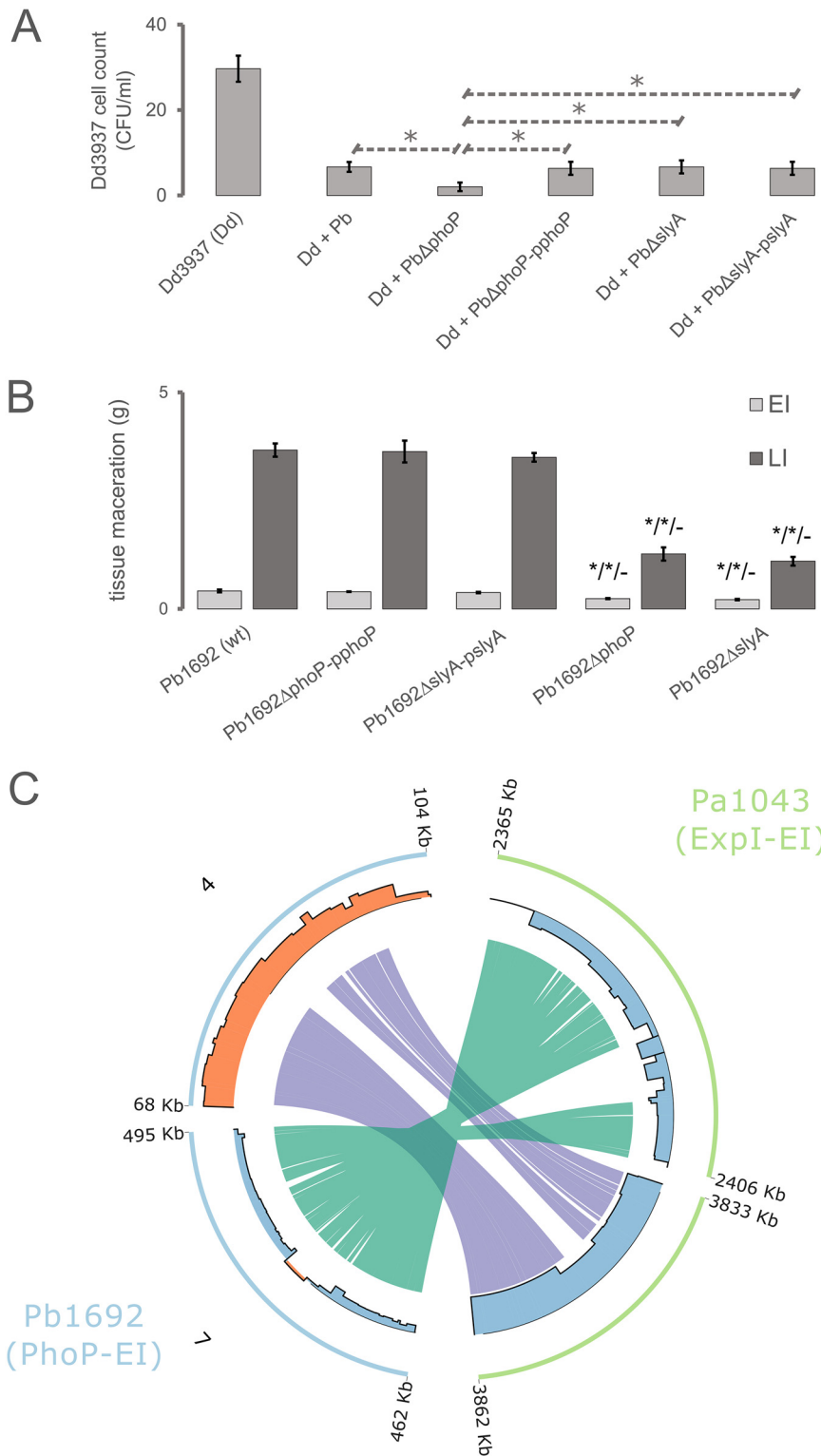
Both Pb1692 $\Delta$ *slyA* and Pb1692 $\Delta$ *phoP* strains next had their virulence against potato tubers examined by measuring the extent of maceration in early and late infection. The results showed that despite the positive regulation of T6SS- and PCWDE-encoding genes in the Pb1692 $\Delta$ *phoP* mutant, this strain was attenuated in virulence, similar to the Pb1692 $\Delta$ *slyA* strain, with no statistically significant difference between them (Fig. 4B). Compared to the bacterial competition assay, tissue maceration comprises a complex phenotype that encompasses a wide range of necessary factors. Thus, the transcriptional disruption of a large array of genes in the four regulons (PhoP-EI, SlyA-EI, PhoP-LI, and SlyA-LI), ranging between 68 and 317 protein-coding genes, may be responsible for the attenuated tissue maceration in these mutants (Table S2).

Aiming to inspect the possible PhoP-regulated mechanisms involved in this increased *in planta* growth inhibition by Pb1692 over that of Dd3937, we will next emphasize secretion system regulation. The T6SS is a protein complex assembled through both bacterial membranes, featuring a contractile sheath and an injectisome-like structure, composed of hemolysin-coregulated proteins (Hcp), that delivers specialized effectors into target cells (31). Both T6 structure- and effector-encoding genes were found under PhoP regulation, totaling 23 T6SS-related genes consistently overexpressed upon *phoP* deletion across the early and late infection (Table S2). By comparing evidence from the PhoP network in Pb1692 with that from previously reported analyses on a QS mutant strain of *P. atrosepticum* (Pa1043) during infection on potato tubers (21), some important observations were made. For this analysis, the fundamental difference between the two mutant strains, i.e., Pb1692 $\Delta$ *phoP* and Pa1043 $\Delta$ *expl* strains, is that the *expR1* regulator is overexpressed in the first and repressed in the second. In the Pa1043 $\Delta$ *expl* strain, 23 T6SS genes exhibited a strong transcriptional decrease in Pa1043 upon the disruption of *expl-expR1* (Fig. 4C). On the contrary, Pb1692 overexpresses *expR1* along with 23 T6SS genes in the absence of *phoP*. This pattern shows the direct interference of QS disruption over the transcriptional modulation of T6SS genes in Pa1043. Additionally, it strongly suggests that the PhoP-dependent regulation of the T6SS gene cluster in Pb1692 primarily depends on the *expR1* regulator.

Moreover, we analyzed T3SS regulation in the two mutants (Fig. 4C). Curiously, both *phoP* and *expl* mutants exhibited a similar pattern of extensive downregulation of the T3SS genes. In Pb1692, the PhoP transcriptional impact encompasses 20 genes. Of these, 11 have detectable orthologs in Pa1043 under the transcriptional influence of QS. In this context, since *expR1* homologs are regulated in opposite directions in each mutant, the PhoP network in Pb1692 seemingly regulates several elements from the T3SS in a QS-independent manner.

#### **Characterizing new gene families within the T3SS operon in *Pectobacterium*.**

During the secretion system regulation analyses, we found two GS genes unannotated by BLASTP-based methods within a T3SS-encoding cluster in Pb1692 (AED-0001197 and AED-0003294). While AED-0003294 was successfully detected under PhoP regulation (Table S2), AED-0001197 exhibited slightly above-threshold false discovery rate (FDR) values and was not listed among the differentially expressed genes. Nonetheless, the *in planta* differential expression of AED-0001197 in the *phoP* mutant strain could be confirmed through qRT-PCR (Fig. S2). By examining the sequences from their products,



**FIG 4** *In planta* bacterial competition and virulence of Pb1692 strains and the transcriptional variation of type III and VI secretion systems in *phoP* and *expl* mutants in Pb1692 and Pa1043, respectively. (A) The graph depicts the survival of *Dickeya dadantii* strain 3937 (Dd3937 or Dd) in potato tubers when coinoculated with different strains of *Pectobacterium brasiliense* strain 1692 (*Pb*), *Pb $\Delta$ phoP*, *Pb $\Delta$ slyA* (mutants), *P. brasiliense* (wild type), *Pb $\Delta$ phoP-pphoP*, and *Pb $\Delta$ slyA-pslyA* (complemented mutant strains), measured in CFU per milliliter. Significant differences between different samples were analyzed using Student's *t* test (\*, *P* < 0.05). (B) The weighed macerated tissue from potato tubers infected with Pb1692 strains was measured and the mass is expressed in grams. The strains (labels) analyzed are Pb1692 wild type (wt), Pb1692 *phoP* mutant (Pb1692 $\Delta$ phoP), Pb1692 *slyA* mutant (Pb1692 $\Delta$ slyA), Pb1692 *phoP* mutant (Continued on next page)

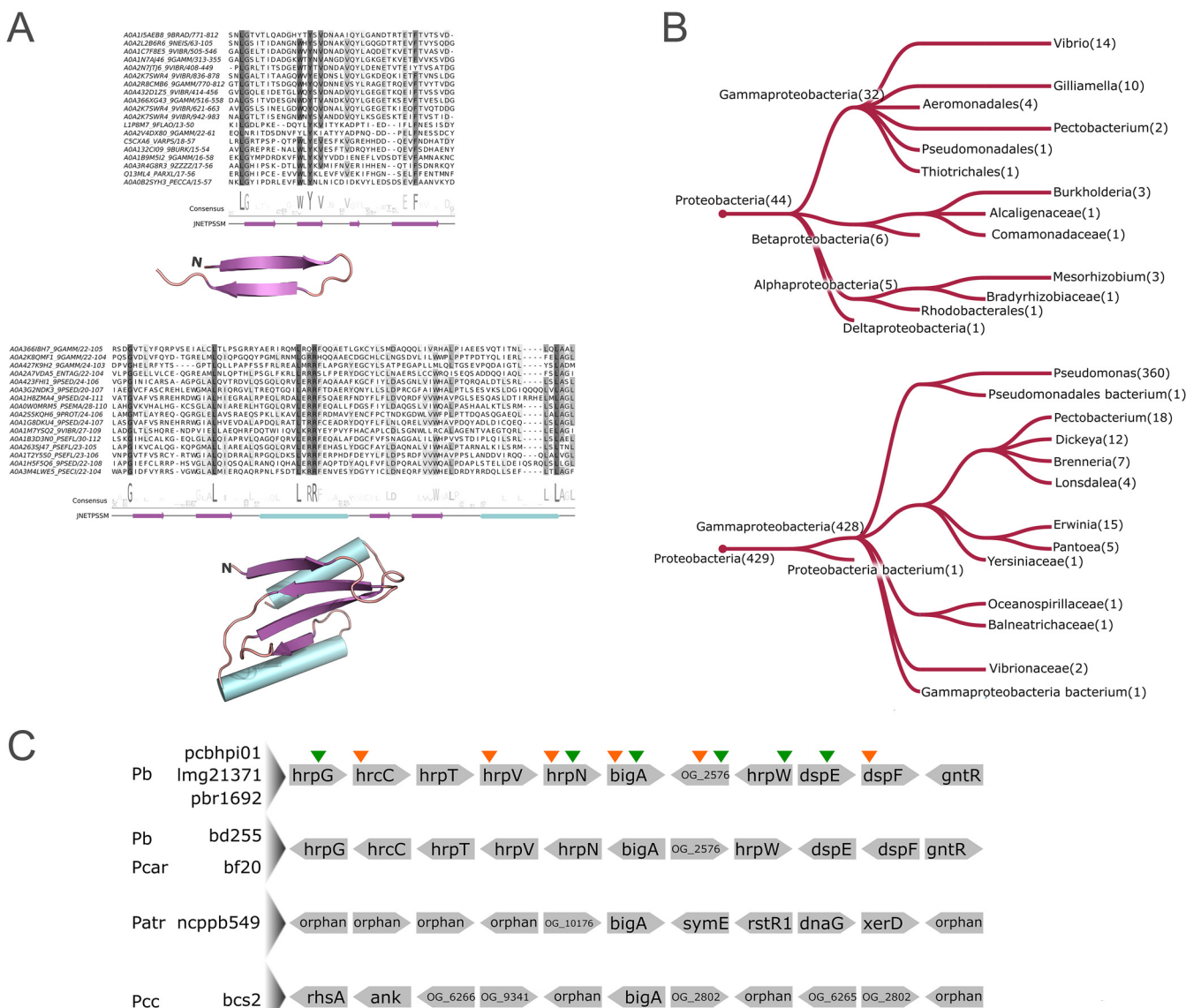
we found no detectable functional domains that indicated their biological role. Hence, aiming to classify these unannotated genes, we implemented sensitive sequence comparison techniques (see Materials and Methods for details) to examine their primary structure and search for conserved motifs. This analysis led to the identification of one novel motif in each protein, which are associated with (i) bacterial Ig-like (Big) and cadherin domains and (ii) *hrp*-related chaperones.

After extensive iterative searches against the UniprotKB online database via the Hmmer search tool (32) and sequence alignment inspections, we gathered 538 and 429 similar sequence sets for the Big-associated motif and the *hrp*-related chaperones, respectively. The Big-associated motif, found in AED-0003294, exhibits highly conserved Tyr and Leu residues in the second  $\beta$ -sheet and prior to the first  $\beta$ -sheet, which may represent core residues in their structure (Fig. 5A). This conserved motif was found primarily in sequences exhibiting no companion domains, similar to those from *Pectobacterium* genomes. Representative genes encoding this exact architecture are most common in *Gammaproteobacteria*, encompassing diverse lifestyles such as animal and plant pathogens (e.g., *Vibrio* and *Pectobacterium* genera) or insect gut symbionts (e.g., *Gilliamella* genus). Additionally, *Alpha*- and *Betaproteobacteria* also carry genes encoding the same architecture, including organisms from *Mesorhizobium* and *Burkholderia* genera (Fig. 5B). Aside from the instances in which these Big-associated motifs are found solo in a given sequence, they are even more frequently found in long (~1,000 amino acids) proteins constituted by numerous domain repeats. These repeats often include combinations of Cadherin and/or Big domains (Pfam entries PF00028 and PF02369), along with the novel Big-associated motif. These families are known for their typical association with phage-like structures and have been linked to carbohydrate recognition on the cell surface, enabling phage adsorption (33), which will be further discussed below.

In the second unannotated gene product (AED-0001197), we found a conserved  $\alpha_2\beta_4$  structure, which was underpinned by 429 similar sequences gathered during the iterative searches described above (Fig. 5B). Approximately 45% of the significant hits found were annotated as HrpV, which form chaperone heterodimers with HrpG to regulate T3SS expression (34). Importantly, besides this newly found  $\alpha_2\beta_4$  motif, no other annotated domains could be detected in any of those 429 HrpV-like sequences. We next investigated the structural conservation of those sequences using HHpred (35), which detected similarities with type III chaperones, such as those from SchA or CesT (PDB entries 4G6T and 5Z38). Thus, although the primary structure does not match any currently annotated domain, the  $\alpha_2\beta_4$  fold found in these sequences was recognized as being similar to type III chaperone-related by HHpred. The impossibility of annotating the HrpV sequence via BLASTP-based methods may be a consequence of extensive sequence variations occurring in different organisms. This notion is supported by our orthology analysis, which has split the HrpV sequences from *Pectobacterium* (represented in 50/61 genomes) and *Dickeya* (represented in 39/39 genomes) into two separate groups (Table S4). Nonetheless, the hidden Markov model (hmm) profile consolidated in this analysis (see Materials and Methods for details) was successfully detected in both groups.

#### FIG 4 Legend (Continued)

complemented with *phoP* (Pb1692 $\Delta$ *pphoP*), and Pb1692 *slyA* mutant complemented with *slyA* (Pb1692 $\Delta$ *pslyA*). Above the mutant's measurement bars, the sequence of three Student's *t* test results is depicted in which the mutants' macerated masses are compared with that of the wild type, respective complement, and other mutant (the Pb1692 $\Delta$ *phoP* mutant is compared to the Pb1692 $\Delta$ *slyA* mutant and vice versa). (C) Genomic regions from Pb1692 (light blue) and *P. atrosepticum* strain scri1043 (*Pa*1043) (light-green) are represented in circular ideograms. Coordinates of 5' and 3' ends of each genomic region are displayed in kilobases. In the Pb1692 ideograms, respective contig numbers are labeled outside the ideogram. In the inner radius, the bar plot indicates, for each region, the transcriptional variation level ( $\log_2$  fold change) found in each transcriptome study (either Liu et al. [21] for *Pa*1043 or this study for Pb1692). The bars are colored to highlight upregulation (orange) or downregulation (blue). The inner links binding genomic regions represent the homologous range of type VI (purple) and III (green) secretion systems conserved between Pb1692 and *Pa*1043 genomes.



**FIG 5** Multiple-sequence alignments, structural scaffolds, and phylogenetic distributions of two conserved motifs in type III secretion system-related putative effectors found in Pb1692. (A) The degree of conservation of each amino acid column is highlighted in increasingly darker shades of gray. Below each alignment, the amino acid logo depicts the amino acid consensus among the representative sequences. Secondary structure predictions next are displayed in both linear and folded (obtained through sequence modeling) configurations. Purple arrows represent  $\beta$ -sheets, and cyan cylinders represent  $\alpha$ -helices. (B) Taxonomic groups in which the *hrpV*-borne (top) and *hrpN*-borne (bottom) motifs are found in sequences exhibiting the same domain architecture as that in Pb1692 are depicted in the phylogenetic trees. Taxonomic distributions were assessed through the Hmsearch online tool (32). The branches are labeled with the corresponding number of genes found in the group. (C) Genomic architectures of seven *Pectobacterium* strains are represented, depicting five genes up- and downstream in *bigA* gene neighborhoods. Gene names are annotated according to sequence searches on the eggNOG database, where applicable. Sequences unannotated by eggNOG are either represented by their respective orthologous group labels (see Materials and Methods for details) or regarded as orphans in case no orthologous sequences were identified across the SRP genomes. Different species are abbreviated on the left: Pb (*P. brasiliense*), Pcar (*P. carotovorum*), Patr (*P. atrosepticum*), and Pcc (*P. carotovorum* subsp. *carotovorum*). Species names are adjacent to strain names. Arrows above the top architecture indicate up-/downregulation (up or down arrow, respectively) of the respective genes by PhoP in Pb1692 during either early infection (orange) or late infection (green) on potato tubers.

We next inspected gene neighborhoods of *hrpV* and the Big-associated (*bigA*) genes across SRP genomes. As expected, all the 89 *hrpV* genes from both ortholog groups, i.e., OG\_3591 (*Pectobacterium*) and OG\_3957 (*Dickeya*), are consistently surrounded by T3SS-related genes (Table S5). However, this pattern is not observed in *bigA* orthologs. Gene neighborhood screening of seven out of ten genomes carrying *bigA* orthologs (three genomes were not suitable for the analysis due to incompleteness of genome assembly) revealed three different types of neighborhoods. Four *Pectobacterium brasil-*

*iense* strains and one strain of *P. carotovorum* exhibit *bigA* homologs within T3SS gene clusters (Fig. 5C). In one strain of *P. atrosepticum*, the *bigA* homolog is adjacent to phage elements, including homologs of the transcriptional repressor *rstR1*, the DNA primase *dnaG*, and the site-specific recombinase *xerD*. The strains of *P. carotovorum* subsp. *carotovorum* have their *bigA* homolog surrounded by unannotated genes and orphan genes that could not be clustered with any other sequences from SRP during the orthology analysis. Furthermore, the newly identified Big-associated gene *bigA*, described above, was predicted as a horizontally transferred gene (Fig. 3D). By inspecting the surrounding genes within the Pb1692 T3SS cluster, the *bigA* gene is the only one that exhibits full support from the HGT prediction (Fig. 3D and Table S3). This result suggests that *bigA* acquisition is posterior to the consolidation of the T3SS in the Pb1692 genome, which is consistent with the above-described genus-specific prediction for *bigA*. Together, these results elucidated the presence of HrpV and the HGT candidate *bigA* under PhoP regulation. Moreover, the *in planta* transcriptional regulation of a known T3SS regulatory element (i.e., *hrpV*) by PhoP points to an additional layer of control over T3SS, which will be addressed in the Discussion.

## DISCUSSION

**The influence of lineage-specific rearrangements in the PhoP and SlyA regulatory networks on host adaptation.** It has been reiterated that the birth of novel functions in prokaryotes is mainly driven by HGT instead of by duplication and divergence (36, 37). For instance, it has been estimated that in representatives of both *Escherichia* and *Salmonella* genera, more than 98% of protein family expansions arise from HGT (37). Moreover, even gene copy increments in bacterial genomes can often be a consequence of HGT rather than duplication events (7, 38). As a consequence, the fixation of newly acquired genes requires that existing regulatory circuits reshape accordingly, allowing those new elements to fit properly into specific transcriptional programs. This often results in the recruitment of new genes into existing transcriptional networks (5). Our results showed PhoP, SlyA, and PecS exhibited a consistent bias toward the regulation of GS-HGT candidate genes at late plant infection. This indicates that transcriptional mobilization of recently acquired genes by these regulators plays a particularly important role in the late stages of infection in these organisms. These findings may even point to a broader trend in which these distinctive traits are preferentially recruited in late stages of the soft-rot disease. At this point in time, with the increasing availability of nutrients in consequence of plant cell lysis (14), the demand for mobilization of distinctive traits in closely related lineages could be specifically focused on interspecies competition.

The participation of PhoP in the regulation of GS genes was also reported in the *Salmonella* genus, as ~50% of the genes found under PhoP regulation have no orthologous counterparts outside this group (39). This evidence is in accordance with the overrepresentation of GS genes found under PhoP control in Pb1692. It also reinforces the PhoP tendency toward regulating GS loci, despite the contrasting lifestyles between most *Salmonella* and *Pectobacterium* representatives. A similar pattern has been reported in *Yersinia pestis* and *Y. enterocolitica* in the RovA regulatory network (40). That report uncovered horizontally acquired genes within the RovA regulon, including genes associated with disease development (40). The *rovA* gene belongs to the MarR family of regulators, from which the best-characterized members, i.e., RovA and SlyA, have been linked to direct binding site competition with the histone-like nucleoid structuring regulator (H-NS) (41, 42). The H-NS role in the repression of horizontally acquired regions is widely characterized in Gram-negative bacteria, as it typically works through the recognition of AT-rich regions (43, 44). These observations converge with the results presented here, as 26.5 and 41.5% of all genes regulated by SlyA at early or late infection, respectively, were successfully predicted, by at least one parametric method, as HGT candidates.

Of particular importance, our results shed light on the relevance of transcriptional network rearrangement processes for PhoP network expansion, which includes the

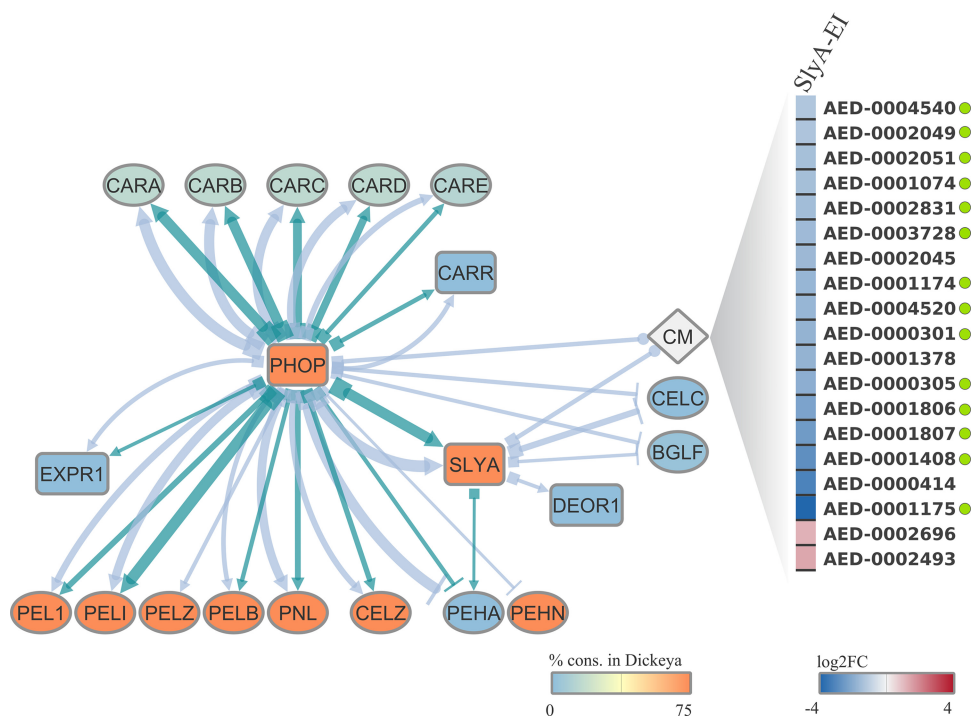


acquisition of other transcriptional regulators, such as *carR* and *expR1*. The evidence found here for the horizontal acquisition of *carR* and *expR1* is in accordance with previous phylogeny-based predictions of *luxI-luxR* horizontal transfer within the *Proteobacteria* phylum (45), as well as between Gram-negative and -positive bacteria (46). These two regulatory connections, i.e., *phoP-carR* and *phoP-expR*, represent an important innovation in the SRP group that impacts the regulation of several host adaptation-related systems in Pb1692 and possibly other species in the *Pectobacterium* genus.

**The *deoR1* transcription factor and the prominent role of SlyA in carbohydrate metabolism regulation.** As previously highlighted, the transcriptional regulator *deoR1* homologs spread across 53 out of 61 *Pectobacterium* genomes, whereas the *Dickeya* lineage apparently lost this gene. DeoR-type regulators have been recognized for a long time as transcriptional repressors of carbohydrate metabolism-related genes (47, 48). In the Gram-positive soilborne bacterium *Corynebacterium glutamicum*, the DeoR-type regulator SugR was shown to repress the transcription of phosphoenolpyruvate-dependent phosphotransferase system (PTS) genes, such as *ptsG*, *ptsS*, and *ptsF* (49). In this context, our data set unveiled an expression pattern following *slyA* deletion that corroborates this notion. While *deoR1* displays increased expression, four PTS permease-encoding genes (*AED-0001174*, *AED-0001290*, *AED-0002702*, and *AED-0003729*) consistently show decreased transcription in Pb1692 (see Table S2 and Fig S2 in the supplemental material). This could account for a DeoR1-dependent repression of those permease-encoding genes. Another DeoR family member, termed UlaR, also was identified as a repressor of the L-ascorbate gene cluster (*ula*) in *Escherichia coli*; this repressor was shown to release the *ula* promoter region upon binding to the L-ascorbate-6-phosphate molecule (50). These reports are strikingly supported by our data set, in which, upon the overexpression of *deoR1* in the *slyA* mutant, 19 out of 21 genes annotated with the carbohydrate metabolism KEGG term (09101) within the SlyA regulatory network were consistently downregulated at early infection (Fig. 6). These observations suggest that SlyA directly or indirectly represses the *deoR1* homolog in wild-type Pb1692, which in turn may release DeoR1 from repressing an array of carbohydrate metabolism-associated genes. Importantly, 76.2% (16/21) of the carbohydrate metabolism-related genes repressed in SlyA-EI were also repressed in the PhoP-EI regulon (Fig. 6 and Table S2). Thus, since Pb1692 $\Delta$ *phoP* and Pb1692 $\Delta$ *slyA* mutant strains exhibit opposite patterns of *slyA* regulation, the evidence indicates that normal expression levels of both *phoP* and *slyA* are required for the expression of these carbohydrate metabolism-associated genes *in planta*.

**PhoP/QS interplay *in planta* in Pb1692: transcriptional regulation of carbapenem biosynthesis- and PCWDE-encoding genes.** Two out of three quorum-sensing regulator homologs from the Pb1692 genome were found to be under PhoP transcriptional regulation, *expR1* and *carR*. The absent *expR* homolog in the PhoP regulons is the *expR2-virR* gene, which has been shown to regulate several virulence themes, such as iron uptake, motility, and expression of PCWDEs in *P. atrosepticum* (51, 52). The mechanistic synergy between ExpR1 and ExpR2 has been previously described by Sjoblom et al. (29), in which both regulators were implicated in the control of virulence-related genes in *Pectobacterium*. In that study, they also observed the ability of ExpR2 to sense a broader range of autoinducer (AI) molecules than ExpR1 (29). Notably, QS has been previously reported to be controlled by other global regulators. For example, both in *Pseudomonas syringae* (53) and in Pb1692 (54), the relationship between the ferric uptake regulator (Fur) and the QS system was determined through contrasting concentrations of N-acylhomoserine-lactone produced by the *fur* mutant and wild-type strains. The regulation of QS by PhoQ/PhoP observed in *P. fluorescens* also seems to be affected by the Mg<sup>2+</sup> concentration level in the medium (55).

The biosynthesis of the carbapenem antibiotic is an important QS-subordinate system. CarR-dependent QS regulation in *Pectobacterium carotovorum* subsp. *carotovorum* relies on the presence of N-(3-oxohexanoyl)-L-homoserine lactone (OHHL) ligand (28). Interestingly, the stability of both carbapenem and the OHHL molecules are affected by pH variations (56, 57). In this context, the PhoPQ two-component system



**FIG 6** Regulatory interplay between PhoP, SlyA, and QS systems in Pb1692. Three subsets of host adaptation and/or virulence themes were extracted from both PhoP and SlyA *in planta* regulons. Transcriptional regulator genes in the network are highlighted in square shapes. The diamond shape represents an array of genes annotated by the KEGG term 09101, carbohydrate metabolism (CM). Each node is colored according to relative presence in the *Dickeya* genus according to the color scale at the bottom of the network. The target link shape depicts the link relationship between source and target nodes: upregulation (arrow), downregulation (perpendicular line), or mixed (circular shape). The source link shape is represented as a square to indicate “in the absence of,” which is the direct inference based on the RNA-Seq experiment performed using the two Pb1692 $\Delta$ phoP and Pb1692 $\Delta$ slyA mutants. Increasingly thicker links in the network represent higher  $\log_2$  fold change ( $\log_2$ FC) values. Light-blue and dark-green links indicate regulation at early and late infection, respectively, by one of the two analyzed regulators (PhoP or SlyA). On the right side of the network, a one-column heatmap depicts the differential expression ( $\log_2$  fold change) in genes associated with carbohydrate metabolism regulated by SlyA at early infection. Genes are represented by the respective locus tags according to the NCBI database. Green circles adjacent to individual genes highlight those for which PhoP-dependent regulation at early infection was also found.

ability to respond to pH variations (58, 59) may (at least partially) explain the success of recruiting *carR* and other *car* genes into the PhoP regulatory network. Thus, in the absence of *phoP*, *carR* and five other *car* genes are overexpressed *in planta*, which, by inference, means that the PhoP network suppresses their transcription *in planta* in the wild type. Hence, PhoP activation, arguably as a consequence of milieu acidification, is able to prevent carbapenem biosynthesis under pH conditions that are unfavorable for the antibiotic stability (Fig. 6).

A similar logic might also apply to the PhoP regulation of *expR1* and several PCWDEs. Indeed, we observed a marked contrast in PhoP regulation patterns over distinct PCWDE classes exhibiting (i) neutral/alkaline optimum pH (~6.8 to 8), such as pectate/pectin lyases, and a cellulase (19, 20), and (ii) acidic optimum pH (~6), such as polygalacturonases (Fig. 6) (60). This regulatory pattern does not correspond to the unidirectional regulation of PCWDE-encoding genes through the QS-RsmA system, which has been repeatedly reported in the *Pectobacterium* genus (16, 27). Thus, the fact that the PhoP network coerces the expression of different sets of PCWDE in opposite directions at the same stage implies that the QS-RsmA cascade is not the only regulatory mechanism controlling those genes in Pb1692. Therefore, although at first glance our results suggest a direct PhoP-ExpR1-RsmA-PCWDE regulatory hierarchy, an additional fine-tuning step enforced either directly or indirectly by PhoP seems to provide an alternative control system over PCWDE transcription. Moreover, this alter-

native pathway seems to be RsmA independent, which follows a previously proposed model of QS regulation (29), since neither *rsmB* nor *rsmA* transcription are influenced by PhoP absence in Pb1692. In this context, the regulation of PCWDE by PhoP in *Dickeya dadantii* has been previously reported. However, PhoP seemed to cause a unidirectional impact on pectate lyases and polygalacturonases (4), in contrast to what was observed in Pb1692. Such contrast should not be surprising, since the *Dickeya* spp. lack the *expR1* QS element, which may be an important part of the PhoP-dependent regulation of PCWDE in Pb1692. Thus, the different patterns found in these two genera may occur as a result of the PhoP-ExpR1 interplay in the regulation of PCWDE that seems to take place in *Pectobacterium* and not in *Dickeya* spp.

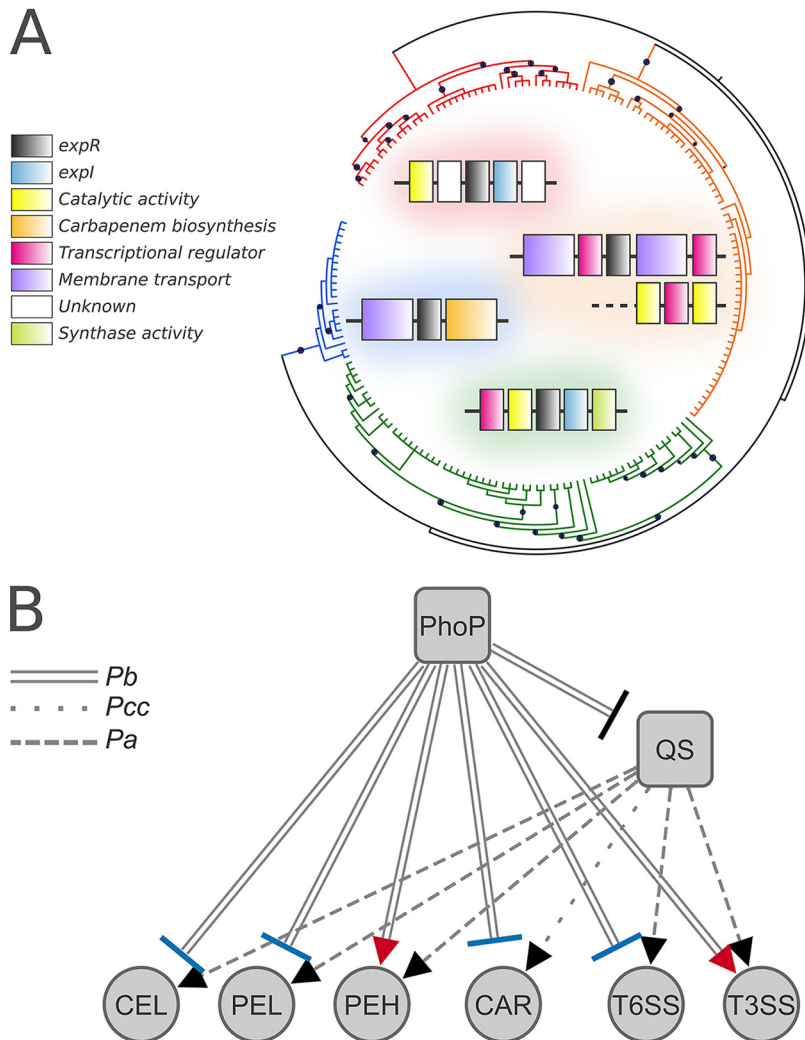
Intriguingly, gene neighborhood analyses revealed that the *expR1* position in *Pectobacterium* genomes is locked upstream of *expl*, which is the exact same pattern as that observed in the *expR2-virR* homologs across the *Dickeya* genus (Fig. 7A). In terms of conservation, *expR2-virR* can be found in 100% of the SRP strains analyzed, and in *Pectobacterium* genomes, it is mostly surrounded by membrane transport systems and other transcriptional regulators, whereas the third and evolutionary more distant homolog, *carR*, exists in only 21 out of 61 *Pectobacterium* strains analyzed and is consistently adjacent to the carbapenem biosynthetic cluster in the genus, as previously reported by Shyntum et al. (61) (Fig. 7A). This agrees with a recent estimate that ~76% of ExpR/LuxR genes exist in genomes without an adjacent Expl/LuxI homolog (25). Also, according to the HGT predictions, the acquisition of *expR1* and *carR* by Pb1692 appears to have occurred through horizontal transfer in the *Pectobacterium* lineage. Although some species, such as *P. parmentieri* and *P. wasabiae*, lost the entire carbapenem biosynthesis cluster, as previously communicated by Shyntum et al. (61), *expR1* remains prevalent in the *Pectobacterium* genus. Thus, the evidence suggests that each of the three *expR-luxR* homologs found in *Dickeya* and *Pectobacterium* genera have been acquired horizontally in an independent fashion. Furthermore, horizontal acquisition of *expR-luxR* homologs has been predicted to occur either as a regulatory cassette that includes *expl-luxI* or individually (45, 62). This notion converges with the results observed in (i) *expR2-virR* or *carR* found in *Pectobacterium* genomes, in which individual acquisition is the most parsimonious assumption, and (ii) *expR2-virR* in *Dickeya* or *expR1* in *Pectobacterium* genomes, which most likely were acquired as a regulatory cassette (Fig. 7A).

Based on previous reports on how QS regulators modulate either PCWDE (21, 63) or carbapenem biosynthesis (28, 64) in SRPs, the possible PhoP-QS interplay can be inferred from our results. Hence, for pectate lyases (Pel), cellulase (Cel), and the *car* genes, the observed regulation of QS by PhoP agrees with those from previous reports, which reinforces the idea of QS regulatory mediation for those genes (Fig. 7B). Conversely, since two polygalacturonase-encoding genes exhibit the opposite transcriptional behavior, they may be controlled by a QS-independent mechanism under the PhoP network. This raises the possibility of a cascade within the PhoP network, providing additional pH-dependent control over the transcription of some PCWDE, including *Pectobacterium*-exclusive PCWDE, which seems to override QS regulation.

#### **The PhoP interplay with QS in the transcriptional regulation of T3SS and T6SS.**

The role of T3SS in pathogenicity has been widely explored in a range of taxa, including some from the SRP group. In *Pectobacterium atrosepticum*, a significant reduction in virulence toward potatoes was observed in mutant strains lacking both T3SS structural genes *hrcC* and *hrcV* (65). In that same study, Holeva et al. (65) also observed that upon the deletion of either the T3SS-related effector *dspE* or the helper gene *hrpN*, virulence was also reduced. Among the 12 T3SS-related genes found under PhoP network control in Pb1692 at early infection, *hrcC*, *hrcV*, and *hrpN* are present (Table S2). This correspondence implies that the results observed in our virulence assays, in which the Pb1692 $\Delta$ *phoP* mutant exhibited attenuated virulence compared to that of the wild type, were (at least partially) impacted by the repression of those T3SS-related genes.

The transcriptional regulation of T3SS involves complex coordination by several regulatory systems, and some have been reported in different bacterial lineages. This



**FIG 7** Evolution of ExpR subfamilies in distinct genomic contexts in SRP and the summary of PhoP-QS interplay in Pb1692. (A) The phylogenetic reconstruction of the autoinducer recognition module found in the ExpR/LuxR homologs from *Pectobacterium* and *Dickeya* was inferred by approximately maximum likelihood. The tree is represented in unscaled branches, with bootstrap values displayed as black circles. The different ExpR/LuxR orthologous groups previously predicted through OrthoMCL, all of which were clustered exclusively with sequences from the same genus, are highlighted in colors: ExpR2/VirR from *Dickeya* (red) and *Pectobacterium* (orange) genera, and ExpR1 (green) and CarR (blue) from *Pectobacterium*. For each group, the dominant gene neighborhoods found across the genus are depicted with the corresponding clade color in the background. (B) The summary network is displayed based on indirect inference from RNA-Seq: if a given gene/system is repressed in the absence of PhoP, then this gene/system is activated in the presence of PhoP. Blue and red target shapes in the network links represent correspondence and conflict between the observed regulation by PhoP in Pb1692 and the quorum-sensing regulation *in planta* reported in other *Pectobacterium* species, respectively. Regulated pathogenicity themes (abbreviations) are cellulases (CEL), pectate/pectin lyases (PEL), polygalacturonases (PEH), carbapenem biosynthesis (CAR), type VI and III secretion systems (T6SS and T3SS, respectively), and quorum-sensing regulators (QS). Link patterns represent the species on which the inference is based, *Pectobacterium carotovorum* subsp. *carotovorum* (Pcc), *Pectobacterium atrosepticum* (Pa), and Pb1692 (Pb).

includes the QS system (21) as well as other regulators, such as SlyA (66), PhoP (67), PecS (9), and Fur (68). Specifically in *Salmonella enterica*, the T3SS termed Spi/Ssa was observably under indirect PhoP control through the SsrB/SpiR system (67, 69). Otherwise, within the SRP group, previous efforts have concluded that PhoP deletion in Dd3937 has no impact on the regulation of T3SS under *in vitro* conditions (70). In contrast with this observation, our results show a wide impact exerted *in planta* by PhoP on T3SS transcription. This contrast indicates that PhoP regulation over T3SS

transcription depends on specific cues that cannot be met *in vitro*. However, since a PhoP-T3SS regulatory link has not been assessed *in planta* in Dd3937 so far, which would allow a direct comparison with the results observed in Pb1692, there is not enough data currently available to predict if (i) the PhoP-dependent regulation of T3SS *in planta* is orthologous among Pb1692 and Dd3937 or (ii) this mechanism evolved at some point within the *Pectobacterium* lineage. Nonetheless, PhoP control over T3SS transcription *in planta* observed in Pb1692 had not been reported in phytopathogens so far. Moreover, our SlyA regulon analysis also shows that contrary to what was reported in Dd3937 by Zou et al. (66), Pb1692 does not rely on SlyA to control T3SS transcription.

Unlike the T3SS, the role of the T6SS in plant colonization processes is not deeply understood in phytopathogens and specifically within the SRP group. However, it has been reported that the T6SS, similar to the T3SS, is under the regulatory control of the QS, as observed in *Pseudomonas aeruginosa* (71) and in the *P. atrosepticum* mutant strain lacking the *expl* gene (21). The T6SS has been functionally implicated in both host defense manipulation and interbacterial competition. However, these two functionalities are not characteristic of any of the five T6SS phylogenetic clades recently reported and, thus, cannot be used to separate T6SS phylogenetic groups (72). As an example, the ability of a T6SS-related Hcp protein to facilitate tumorigenesis in potatoes has been reported in the phytopathogenic *Agrobacterium tumefaciens* DC58 (73). In the same species, the homologous T6SS was later implicated in intra- and interspecies competition *in planta* (74). Furthermore, T6SS has been specifically reported as a decisive asset in interspecies competition *in planta* for Pb1692 (61). In agreement with previous reports, our results from *in planta* bacterial competition indicate that the increased transcription of T6SS as a consequence of *phoP* deletion in Pb1692 is an important factor to boost competitive advantage against closely related competitors (Fig. 4A and C).

The observed influence of QS over T6SS transcription in *P. atrosepticum* spans 23 genes, which exhibited decreased transcription intensity upon QS disruption (21). Similarly, T3SS transcription was also impacted by the disruption of QS, as 21 genes were consistently repressed (21). We verified that in the absence of PhoP, the QS receptor *expR1* is overexpressed, and so are several T6SS-related genes, which strongly suggests that QS mediates PhoP regulation over T6SS during infection. Indeed, this hypothesis is supported by results from Liu et al. (21) in *P. atrosepticum*, in which T6SS transcription decreases as the QS system is disrupted (Fig. 7B). On the other hand, since T3SS exhibits the opposite regulatory pattern, similar to the one observed in *P. atrosepticum* following QS disruption, it is likely that PhoP does not depend on QS to regulate the T3SS. Hence, PhoP must be able to override QS regulation of T3SS-encoding genes and control their expression. This could result either from PhoP directly binding to their promoter regions or through an alternative QS-independent regulatory cascade.

**New T3SS-related families and the transcriptional regulation of the *hrp* gene cluster by PhoP.** The bacterial Ig-like domains are highly promiscuous entities that have been described in structures of (i) adhesins, such as invasins and intimins (75, 76), (ii) phage-tail proteins (77), and (iii) bacterial surface glycohydrolases (78). The presence of Big-encoding genes in bacteria has been associated with horizontal transfers as a result of its frequent presence in phage genomes (78), which was strongly corroborated by our HGT prediction. Although the precise function of Big domains remains elusive, several lines of evidence point to a role in surface carbohydrates recognition (77, 78). Indeed, from the guilt by association standpoint, this could be the case for these newly found BigA proteins, since they are located in a region where T3SS-related products are typically encoded. Also, the analyses conducted in this study support the posterior horizontal acquisition of the *bigA* gene in Pb1692 independently from the rest of the T3SS gene cluster.

The other newly found motif in T3SS-related proteins from Pb1692 is comprised of an  $\alpha_2\beta_4$  structure, which was found to be similar to that of bacterial HrpV. Curiously,



this motif has been identified before in *Erwinia amylovora*, although it has not been deposited in any public domain databases (34). HrpV is known as a negative regulator of the *hrp-hrc* gene cluster in *P. syringae* (79). One of the important protein interactions involving HrpV is the ability to bind HrpS. This interaction forms the HrpV-HrpS complex, which blocks the formation of HrpR-HrpS heterohexamers, subsequently hindering HrpS's ability to bind and activate the *hrp-hrc* promoter (80, 81). This negative regulation imposed by HrpV was shown to be attenuated by its interaction with HrpG, which then generates a double-negative regulatory circuit controlling the expression of T3SS elements (82). We also found that, apart from *hrpV*, none of the other transcriptional regulators of the T3SS region (i.e., *hrpG*, *hrpS*, *hrpR*, and *hrpL*) are affected by *phoP* deletion, yet 20 *hrp-hrc* genes exhibit differential expression in the PhoP mutant strain. Thus, the expected result of having the main repressor of *hrp-hrc* (i.e., *hrpV*) transcription repressed should be the increased expression of several genes in this region. However, the opposite was observed, as 20 T3SS genes are consistently repressed across early and late infection stages. These observations imply that an alternative regulatory mechanism employed by the PhoP network coordinates the expression of T3SS elements, one that is independent of HrpG/S/R/L. As previously discussed, QS does not seem to be the mechanism responsible for this regulation, based on the results from PhoP regulon analyses combined with past observations (21). This means that PhoP is either directly or indirectly involved in the regulation of T3SS, apparently independently from Hrp-borne regulators or the QS system. Indeed, this hypothesis adds complexity to the topic if taken together with the conclusions reported by Bijlsma and Groisman (67) in *Salmonella enterica*, in which PhoP is responsible for the post-transcriptional regulation of a T3SS gene cluster.

By elucidating the mechanisms of network expansion through transcriptional rearrangement in an important phytopathosystem, we found a wide range of host adaptation- and environmental fitness-related traits being coopted by SlyA and especially by PhoP. Here, we uncover the success of the centralizing strategy that recruited carbohydrate metabolism regulation along with virulence-related systems controlled by QS in Pb1692 into the stress response regulator PhoP. This seems to provide optimized control over specific systems that may be sensitive to certain environmental conditions that can be recognized by the PhoQ/PhoP two-component system.

## MATERIALS AND METHODS

**Growth conditions and construction of *Pectobacterium brasiliense* 1692 mutant strains.** The strains and plasmids used are listed in Table S6 in the supplemental material. Luria-Bertani (LB) broth and agar plates were used for growing all bacterial strains at 37°C. Different antibiotics were used to supplement the media, i.e., kanamycin (50 µg/ml) and ampicillin (100 µg/ml) (Sigma-Aldrich), when needed. All reagents were used according to the manufacturer's instructions. To generate the mutant strains, the ASAP database BLASTN tool was used to identify the Pb1692 *phoP* and *slyA* genes (AED-0004376 and AED-0001000). Both Pb1692Δ*phoP* and Pb1692Δ*slyA* mutants were generated using a strategy developed by Datsenko and Wanner (83). Briefly, the up- and downstream regions flanking the target genes (approximately 1,000 bp) were amplified using specific primers (Table S7) using PCR. The kanamycin resistance gene cassette was amplified from the pKD4 plasmid. The resulting upstream fragment, kanamycin cassette, and downstream fragments were fused using primers denoted as previously described by Shyntum et al. (61). The fused PCR product was purified and electroporated into electrocompetent Pb1692 cells with pKD20. The resulting transformants were selected on nutrient agar supplemented with 50 µg/ml kanamycin. The HiFi HotStart PCR kit (KAPA system) was used in all PCRs. The PCR conditions were set as initial denaturation at 95°C for 3 min, followed by 25 cycles of denaturing at 98°C for 30 s, annealing at 60 to 65°C (depending on the primer set), extension at 72°C for 2 min, and a final extension at 72°C for 2 min. The Pb1692Δ*phoP* and Pb1692Δ*slyA* mutant strains were verified by PCR analyses and nucleotide sequencing.

**Complementation of *phoP* and *slyA* mutants.** The p-JET1.2/blunt cloning vector was used for the complementation of mutant strains. The *phoP* and *slyA* genes from Pb1692, including ~500 nucleotides of downstream sequence containing the putative promoter sequence, were amplified by PCR using complementation primers listed in Table S7. The corresponding fragments were excised from an agarose gel and purified using the Thermo Scientific gel extraction kit according to the manufacturer's instructions. The fragments were cloned into p-JET-T to generate pJET-*phoP* and pJET-*slyA* (Table S6). These plasmids were electroporated into electrocompetent Pb1692Δ*phoP* and Pb1692Δ*slyA* mutant strains, transformants (Pb1692Δ*phoP*-*pphoP* and Pb1692Δ*slyA*-*pslyA*) were selected on agar plates supplemented with 100 µg/ml ampicillin, and the cloned *phoP* and *slyA* were confirmed using PCR and sequencing.

**Tissue maceration assay and total RNA extraction from potato tubers.** *Solanum tuberosum* (cv. Mondial, a susceptible cultivar) potato tubers were sterilized with 10% sodium hypochlorite, rinsed twice with double-distilled water, air dried, and then stabbed with a sterile pipette to a 1-cm depth. A 10- $\mu$ l aliquot of the bacterial cells at an optical density at 600 nm ( $OD_{600}$ ) equivalent to 1 were pipetted into generated holes. As a control, 10 mM  $MgSO_4$  was inoculated into potato tubers. Holes were sealed with petroleum jelly. Potato tubers were then placed in moist plastic containers and incubated for 12 and 24 h at 25°C. The macerated tissue was scooped and weighed 12 and 24 h postinoculation to quantify the extent of tuber maceration caused by Pb1692 wild-type and Pb1692 $\Delta$ *phoP* and Pb1692 $\Delta$ *slyA* mutant strains, as well as Pb1692 $\Delta$ *phoP*-*pphoP* and Pb1692 $\Delta$ *slyA*-*pslyA* complemented mutant strains. This experiment was performed in triplicates three independent times.

For RNA extraction, potato tubers inoculated with Pb1692 wild-type, Pb1692 $\Delta$ *phoP*, and Pb1692 $\Delta$ *slyA* strains were incubated at 25°C with high humidity in plastic containers for 12 and 24 h maximum. Sampling was done by scooping out macerated tissue. Macerated potato tissue from each inoculated site was scooped out and homogenized in double-distilled water. Bacterial cells were recovered by grinding the scooped macerated potato tissues in 20 ml of double-distilled water using an autoclaved pestle and mortar. Starch material was removed by centrifuging the ground tissues at 10,000 rpm for 1 min. The supernatant was removed and transferred into new sterilized 50-ml Falcon tubes containing RNA stabilization buffer (Qiagen, Hilden, Germany). The experiments were performed using three biological replicates, with three tubers per replicate.

**Determination of total RNA quality.** The concentration and purity of each extracted total RNA sample were evaluated using spectrophotometric analysis (NanoDrop ND-1000; NanoDrop Technologies, Wilmington, DE) at a ratio of 230 to 260 nm. The integrity of the RNA was further confirmed using 1% (wt/vol) agarose gel electrophoresis using 1% Tris-acetate-EDTA buffer at 100 V for 30 min and visualized, and an image was developed using a Gel Doc EZ system (Bio-Rad Laboratories, Berkeley, CA, USA). Using an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.), total RNA sample concentration, RNA integrity number (RIN), and 28S/18S ratio were determined.

**RNA sequencing, read mapping, differential expression analysis, and genome-wide functional annotation.** RNA samples were sequenced in the Novogene facility (Chula Vista, CA) using an Illumina NovaSeq 6000 machine. Reads sequencing quality analysis was carried out by utilizing FastQC software (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Low-quality segments were then trimmed by Trimmomatic v 0.36 (84). Trimmed reads were next supplied to *hisat2* v 2.1.0 (85), which performed read alignment to the reference genome of Pb1692 (WPP1692 version 3), obtained from the ASAP database (<https://asap.genetics.wisc.edu/>). The number of aligned reads was subsequently computed by the *featureCounts* package (86), and differential expression was analyzed by utilizing the EdgeR package (87), both within the R environment (<https://www.r-project.org/>). Genes exhibiting transcriptional variation of  $\log_2$  fold change of  $>1$  (upregulation) or  $<-1$  (downregulation), with FDR of  $<0.05$ , were designated differentially expressed. Comparisons of mutant samples to wild-type ones were done at 12 and 24 hpi. Pb1692 sequences were functionally annotated by using the eggNOG-mapper tool (88). The annotation provided by eggNOG was next used to retrieve higher annotation levels in the KEGG library hierarchy (KEGG B and A) (89) through custom scripts written in the Perl language (<https://www.perl.org/>). Enrichment of KEGG terms was determined by one-tailed Fisher exact test, and subsequent *P* value adjustment by FDR (*q* value of  $<0.05$ ) was performed by custom R scripts using the Pb1692 genome as the background data set. An additional level of annotation was provided by conserved domain inspection on the protein sequences carried out using HMMER3 (90) and the Pfam database (91). Venn diagrams were created using the InteractiVenn online tool (92).

**Genus-specific contents, enrichment in specific regulons, and genome simulations.** The framework used to identify orthologs among strains from *Pectobacterium* and *Dickeya* genomes was previously described (6). Briefly, genomic and proteomic information was acquired from the NCBI online database for 61 and 39 *Pectobacterium* and *Dickeya* strains, respectively. All protein sequences were clustered by implementing the OrthoMCL pipeline (93). The presence of representatives from each genus was analyzed throughout the 10,635 orthologous groups using custom Perl scripts. Those gene products that (i) belong to clusters populated exclusively with sequences from one of the genera or (ii) are orphans (not clustered) then are considered genus specific. To evaluate the possibility of overrepresentation of GS genes within individual regulons, we used both statistical analysis and computational simulations. Statistical verification performed in the R environment (<https://www.r-project.org/>) was based on a two-tailed Fisher exact test to assess the correlation between GS occurrence in a given regulon, using the respective strain genome as the background. Further, the computational simulations were performed to generate 10,000 shuffled copies of the respective genomes (Pb1692 or Dd3937 pseudogenomes) for each regulon comparison. In these comparisons for a single regulon, gene positions relative to each regulated gene were retrieved. These exact positions next were checked in each of the 10,000 pseudogenomes for the occurrence of GS genes. The total number of GS genes found in this assessment for each pseudogenome is computed and compared to the real data.

**Gene expression analysis through qRT-PCR.** To validate the differential expression analysis of genes from the RNA-Seq data, a qRT-PCR analysis was performed using 12 selected genes. RNA samples used in the qRT-PCR analysis were the same as those used in RNA sequencing and were prepared using an RNeasy minikit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The first-strand cDNA was reverse transcribed from 5  $\mu$ g of total RNA with the SuperScript IV first-strand synthesis system (Invitrogen). qRT-PCR was performed using a QuantStudio 12K Flex real-time PCR system (Applied Biosystems). Primers used in this study are listed in Table S7. The following cycling parameters were used: 95°C for 3 min, followed by 40 cycles of 95°C for 60 s, 55°C for 45 s, and 72°C for 60 s, and then 72°C for

7 min. *ffh* and *gyrA* (housekeeping gene control) were used to normalize gene expression. The gene expression levels of the following genes were analyzed: *pehA* (AED-0002061), *pehN* (AED-0000675), *expR1* (AED-0000069), *carR* (AED-0003542), *deoR1* (AED-0001096), *tssE* (AED-0001996), *tssC* (AED-0001994), *bigA* (AED-0003294), *hrpV* (AED-0001197), *celC* (AED-0001290), *bglF* (AED-0003729), and AED-0001174. The same method was applied to measure the expression levels of the *phoP* gene in the Pb1692 wild type. Overnight cultures of the wild type were inoculated into sterile potato tubers and extracted every 4 h for a period of 24 h (see "Tissue maceration assay and total RNA extraction from potato tubers," above). For statistical analysis of relative gene expression, the threshold cycle ( $C_T$ ) method was used (94). A *t* test was used to determine statistical significance ( $P < 0.05$ ).

**In planta competition assays in potato tubers.** Competition assays were performed in potato tubers as described by Axelrood et al. (95). In summary, *Solanum tuberosum* (cv. Mondial, a susceptible cultivar) tubers were sterilized with 10% sodium hypochlorite, rinsed twice with double-distilled water, air dried, and then stabbed with a sterile pipette to 1-cm depth. Overnight bacterial cultures of Pb1692 wild-type, Pb1692 $\Delta$ *phoP*, Pb1692 $\Delta$ *phoP*-*pphoP*, Pb1692 $\Delta$ *slyA*, and Pb1692 $\Delta$ *slyA*-*pslyA* strains and *Dickeya dadantii* at an OD<sub>300</sub> equivalent to 0.3 were mixed in a 1:1 ratio and inoculated into surface-sterilized potato tubers. Holes were sealed with petroleum jelly. Potato tubers were then placed in moist plastic containers and incubated for 24 h at 25°C. Macerated tuber tissue was scooped out, and the CFU per milliliter of surviving targeted bacteria was determined by serial dilutions on LB supplemented with gentamicin (15  $\mu$ g/ml). The number of colonies observed was then converted to CFU per milliliter. This experiment was performed in triplicates three independent times.

**Sequence analysis and protein domain characterization.** Orthologous sequences of unannotated proteins (Big-associated and HrpV) from SRP strains were aligned using Clustal Omega (96). The resulting alignments then were used in iterative searches against the UniprotKB database ([www.uniprot.org](http://www.uniprot.org)) through Hmmersearch (32). The aligned results from the iterative searches were manually curated through Jalview alignment viewer (97), and the conserved areas in the alignments were selected. Those selected conserved blocks next were analyzed for secondary structure prediction and programmatic gap removals by Jpred (98). The resulting concise alignments were submitted to HHpred (35) to identify appropriate templates for structure modeling. This strategy managed to retrieve several PDB structures to be used as templates. These templates were selected for subsequent modeling according to the predicted probability (>40%). The following PDB entries were selected: for the Big-associated sequences, 5K8G, 4N58, 4G75, 4FZL, 5K8G, 4FZM, and 2XMX; for HrpV, 4G6T, 3KXY, 5Z38, 3EPU, 4GF3, 1JYA, 1S28, 5WEZ, 4JMF, and 1XKP. The models were predicted by Modeller (99), and visualization of these models was generated through PyMOL (100).

**Horizontal gene transfer prediction and regulatory network analyses.** Parametric methods are aimed at evaluating sequence composition bias in genes or genomic regions and measure their distance to the overall trend observed in the respective genome. These methods are especially powerful when applied in recent HGT candidate predictions, contrary to phylogenetic methods (101). Predictions were made with the support of two different methods, namely, GC content at the third codon position (GC3) and dinucleotide frequencies (DINT). The method choice was adapted based on the conclusions drawn in the benchmark conducted by Becq et al. (30). Genome-wide prediction of GC3 indexes of Pb1692 and Dd3937 coding sequences was made through the codonW tool (<http://codonw.sourceforge.net/>). Dinucleotide frequency analyses were performed using the *fasta2kmercontent* script from the CGAT package (102). The Kullback-Leibler distance (KLD) (103) was used to measure the difference between dinucleotide frequencies of individual coding genes and the respective genomes (averaged value for all coding sequences). KLD is calculated according to the following formula:

$$KLD[p(g) \parallel p(G)] = \sum_i^N p(g_i) \log \frac{p(g_i)}{p(G_i)}$$

$p(g)$  is the probability distribution of dinucleotide frequencies in an individual coding gene, and  $p(G)$  is the probability distribution of averaged dinucleotide frequencies from all coding genes in that same genome. KLD values were computed in the R statistical environment. These two metrics (GC3 and DINT\_KL) are then combined with the information on the genus specificity of genes previously collected from orthologous clustering to predict HGT candidates. The above-threshold genes that were also genus specific according to the previous analysis are predicted to be HGT candidates. Gene networks were analyzed using Cytoscape (104) and the heatmap generated with Gitoools (105).

**Data availability.** The data sets generated in this study are available at the Sequence Read Archive (SRA; <https://www.ncbi.nlm.nih.gov/sra>) under the accession number PRJNA565562.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, TIF file, 1.2 MB.

**FIG S2**, TIF file, 0.4 MB.

**FIG S3**, TIF file, 0.4 MB.

**TABLE S1**, XLSX file, 0.01 MB.

**TABLE S2**, XLSX file, 0.1 MB.

**TABLE S3**, XLSX file, 0.01 MB.

**TABLE S4**, XLSX file, 0.02 MB.

**TABLE S5**, XLSX file, 0.02 MB.

**TABLE S6**, XLSX file, 0.01 MB.

**TABLE S7**, XLSX file, 0.01 MB.

## ACKNOWLEDGMENTS

We thank Collins K. Tanui (former Ph.D. candidate at the Department of Biochemistry Genetics and Microbiology) for assistance in RNA extractions.

This research study was funded by the National Research Foundation (NRF), South Africa, through Competitive Funding for Rated Researchers (CFRR 98993) and NRF Bioinformatics and Functional Genomics (BFG 93685). D.Y.S. received an NRF BFG postdoctoral fellowship. D.B.-R. received a University of Pretoria postdoctoral fellowship. N.P.N. was funded by NRF BFG PhD Bursary and Potato South Africa (PSA) Transformation Fund. Any opinion, findings, conclusions, or recommendations expressed in this material are those of the authors, and the NRF does not accept any liability in this regard.

## REFERENCES

- García-Olmedo F, Molina A, Alamillo JM, Rodríguez-Palenzuela P. 1998. Plant defense peptides. *Peptide Sci* 47:479–491. [https://doi.org/10.1002/\(SICI\)1097-0282\(1998\)47:6<479::AID-BIP6>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1097-0282(1998)47:6<479::AID-BIP6>3.0.CO;2-K).
- Alamillo JM, García-Olmedo F. 2001. Effects of urate, a natural inhibitor of peroxynitrite-mediated toxicity, in the response of *Arabidopsis thaliana* to the bacterial pathogen *Pseudomonas syringae*. *Plant J* 25: 529–540. <https://doi.org/10.1046/j.1365-3113x.2001.00984.x>.
- Grignon C, Sentenac A. 1991. pH and ionic conditions in the apoplast. *Annu Rev Plant Physiol Plant Mol Biol* 42:103–128. <https://doi.org/10.1146/annurev.pp.42.060191.000535>.
- Llama-Palacios A, López-Solanilla E, Poza-Carrion C, García-Olmedo F, Rodríguez-Palenzuela P. 2003. The *Erwinia chrysanthemi* *phoP*-*phoQ* operon plays an important role in growth at low pH, virulence and bacterial survival in plant tissue. *Mol Microbiol* 49:347–357. <https://doi.org/10.1046/j.1365-2958.2003.03583.x>.
- Perez JC, Groisman EA. 2009. Evolution of transcriptional regulatory circuits in bacteria. *Cell* 138:233–244. <https://doi.org/10.1016/j.cell.2009.07.002>.
- Bellieny-Rabelo D, Tanui CK, Miguel N, Kwenda S, Shyntum DY, Moleleki LN, Bellieny-Rabelo D, Tanui CK, Miguel N, Kwenda S, Shyntum DY, Moleleki LN. 2019. Transcriptome and comparative genomics analyses reveal new functional insights on key determinants of pathogenesis and interbacterial competition in *Pectobacterium* and *Dickeya* spp. *Appl Environ Microbiol* 85:e02050-18. <https://doi.org/10.1128/AEM.02050-18>.
- Eme L, Doolittle WF. 2016. Microbial evolution: xenology (apparently) trumps paralogy. *Curr Biol* 26:R1181–R1183. <https://doi.org/10.1016/j.cub.2016.09.049>.
- Haque MM, Kabir MS, Aini LQ, Hirata H, Tsuyumu S. 2009. *SlyA*, a *MarR* family transcriptional regulator, is essential for virulence in *Dickeya dadantii* 3937. *J Bacteriol* 191:5409–5418. <https://doi.org/10.1128/JB.00240-09>.
- Pedron J, Chapelle E, Alunni B, Van Gijsegem F. 2017. Transcriptome analysis of the *Dickeya dadantii* *PecS* regulon during the early stages of interaction with *Arabidopsis thaliana*. *Mol Plant Pathol* 19:647–663. <https://doi.org/10.1111/mpp.12549>.
- Alekshun MN, Kim YS, Levy SB. 2000. Mutational analysis of *MarR*, the negative regulator of *marRAB* expression in *Escherichia coli*, suggests the presence of two regions required for DNA binding. *Mol Microbiol* 35:1394–1404. <https://doi.org/10.1046/j.1365-2958.2000.01802.x>.
- Hommais F, Oger-Desfeux C, Van Gijsegem F, Castang S, Ligori S, Expert D, Nasser W, Reverchon S. 2008. *PecS* is a global regulator of the symptomatic phase in the phytopathogenic bacterium *Erwinia chrysanthemi* 3937. *J Bacteriol* 190:7508–7522. <https://doi.org/10.1128/JB.00553-08>.
- Zhou JN, Zhang HB, Lv MF, Chen YF, Liao LS, Cheng YY, Liu SY, Chen SH, He F, Cui ZN, Jiang ZD, Chang CQ, Zhang LH. 2016. *SlyA* regulates phytotoxin production and virulence in *Dickeya zeae* EC1. *Mol Plant Pathol* 17:1398–1408. <https://doi.org/10.1111/mpp.12376>.
- Adeolu M, Alnajjar S, Naushad S, Gupta RS. 2016. Genome-based phylogeny and taxonomy of the “Enterobacterales”: proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov. *Int J Syst Evol Microbiol* 66:5575–5599. <https://doi.org/10.1099/ijsem.0.001485>.
- Charkowski AO. 2018. The changing face of bacterial soft-rot diseases. *Annu Rev Phytopathol* 56:269–288. <https://doi.org/10.1146/annurev-phyto-080417-045906>.
- Moleleki LN, Pretorius RG, Tanui CK, Mosina G, Theron J. 2017. A quorum sensing-defective mutant of *Pectobacterium carotovorum* ssp. *brasiliense* 1692 is attenuated in virulence and unable to occlude xylem tissue of susceptible potato plant stems. *Mol Plant Pathol* 18:32–44. <https://doi.org/10.1111/mpp.12372>.
- Charkowski A, Blanco C, Condemine G, Expert D, Franza T, Hayes C, Hugouvieux-Cotte-Pattat N, Lopez Solanilla E, Low D, Moleleki L, Pirhonen M, Pitman A, Perna N, Reverchon S, Rodriguez Palenzuela P, San Francisco M, Toth I, Tsuyumu S, van der Waals J, van der Wolf J, Van Gijsegem F, Yang CH, Yedidia I. 2012. The role of secretion systems and small molecules in soft-rot Enterobacteriaceae pathogenicity. *Annu Rev Phytopathol* 50:425–449. <https://doi.org/10.1146/annurev-phyto-081211-173013>.
- Pérombelon M. 2002. Potato diseases caused by soft rot erwinias: an overview of pathogenesis. *Plant Pathol* 51:1–12.
- Toth IK, Bell KS, Holeva MC, Birch PR. 2003. Soft rot erwiniae: from genes to genomes. *Mol Plant Pathol* 4:17–30. <https://doi.org/10.1046/j.1364-3703.2003.00149.x>.
- Collmer A, Keen NT. 1986. The role of pectic enzymes in plant pathogenesis. *Annu Rev Phytopathol* 24:383–409. <https://doi.org/10.1146/annurev.py.24.090186.002123>.
- Saarilahti HT, Henriks B, Palva ET. 1990. *CelS*: a novel endoglucanase identified from *Erwinia carotovora* subsp. *carotovora*. *Gene* 90:9–14. [https://doi.org/10.1016/0378-1119\(90\)90433-r](https://doi.org/10.1016/0378-1119(90)90433-r).
- Liu H, Coulthurst SJ, Pritchard L, Hedley PE, Ravensdale M, Humphris S, Burr T, Takle G, Brurberg MB, Birch PR, Salmond GP, Toth IK. 2008. Quorum sensing coordinates brute force and stealth modes of infection in the plant pathogen *Pectobacterium atrosepticum*. *PLoS Pathog* 4:e1000093. <https://doi.org/10.1371/journal.ppat.1000093>.
- Valente RS, Nadal-Jimenez P, Carvalho AFP, Vieira FJD, Xavier KB, Valente RS, Nadal-Jimenez P, Carvalho AFP, Vieira FJD, Xavier KB. 2017. Signal integration in quorum sensing enables cross-species induction of virulence in *Pectobacterium wasabiae*. *mBio* 8:e00398-17. <https://doi.org/10.1128/mBio.00398-17>.
- Waters CM, Bassler BL. 2005. Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* 21:319–346. <https://doi.org/10.1146/annurev.cellbio.21.012704.131001>.
- Barnard AM, Salmond GP. 2007. Quorum sensing in *Erwinia* species. *Anal Bioanal Chem* 387:415–423. <https://doi.org/10.1007/s00216-006-0701-1>.
- Papenfort K, Bassler BL. 2016. Quorum sensing signal-response systems in Gram-negative bacteria. *Nat Rev Microbiol* 14:576–588. <https://doi.org/10.1038/nrmicro.2016.89>.
- Andersson RA, Eriksson AR, Heikinheimo R, Mae A, Pirhonen M, Koiv V, Hyytiäinen H, Tuikkala A, Palva ET. 2000. Quorum sensing in the plant pathogen *Erwinia carotovora* subsp. *carotovora*: the role of *expr(Ecc)*. *Mol Plant Microbe Interact* 13:384–393. <https://doi.org/10.1094/MPMI.2000.13.4.384>.
- Chatterjee A, Cui Y, Liu Y, Dumenyo CK, Chatterjee AK. 1995. Inactivation of *rsmA* leads to overproduction of extracellular pectinases,



- cellulases, and proteases in *Erwinia carotovora* subsp. *carotovora* in the absence of the starvation/cell density-sensing signal, N-(3-oxohexanoyl)-L-homoserine lactone. *Appl Environ Microbiol* 61: 1959–1967. <https://doi.org/10.1128/AEM.61.5.1959-1967.1995>.
28. Welch M, Todd DE, Whitehead NA, McGowan SJ, Bycroft BW, Salmond GP. 2000. N-acyl homoserine lactone binding to the CarR receptor determines quorum-sensing specificity in *Erwinia*. *EMBO J* 19:631–641. <https://doi.org/10.1093/emboj/19.4.631>.
  29. Sjoblom S, Brader G, Koch G, Palva ET. 2006. Cooperation of two distinct ExpR regulators controls quorum sensing specificity and virulence in the plant pathogen *Erwinia carotovora*. *Mol Microbiol* 60: 1474–1489. <https://doi.org/10.1111/j.1365-2958.2006.05210.x>.
  30. Becq J, Churlaud C, Deschavanne P. 2010. A benchmark of parametric methods for horizontal transfers detection. *PLoS One* 5:e9989. <https://doi.org/10.1371/journal.pone.0009989>.
  31. Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ. 2006. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the Dictyostelium host model system. *Proc Natl Acad Sci U S A* 103:1528–1533. <https://doi.org/10.1073/pnas.0510322103>.
  32. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update. *Nucleic Acids Res* 46:W200–W204. <https://doi.org/10.1093/nar/gky448>.
  33. Fraser JS, Maxwell KL, Davidson AR. 2007. Immunoglobulin-like domains on bacteriophage: weapons of modest damage? *Curr Opin Microbiol* 10:382–387. <https://doi.org/10.1016/j.mib.2007.05.018>.
  34. Gazi AD, Charova S, Aivaliotis M, Panopoulos NJ, Kokkinidis M. 2015. HrpG and HrpV proteins from the type III secretion system of *Erwinia amylovora* form a stable heterodimer. *FEMS Microbiol Lett* 362:1–8. <https://doi.org/10.1093/femsle/fnu011>.
  35. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, Gabler F, Soding 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>.
  36. Koonin EV. 2015. The turbulent network dynamics of microbial evolution and the statistical tree of life. *J Mol Evol* 80:244–250. <https://doi.org/10.1007/s00239-015-9679-7>.
  37. Treangen TJ, Rocha EP. 2011. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet* 7:e1001284. <https://doi.org/10.1371/journal.pgen.1001284>.
  38. Hehemann JH, Arevalo P, Datta MS, Yu X, Corzett CH, Henschel A, Preheim SP, Timberlake S, Alm EJ, Polz MF. 2016. Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes. *Nat Commun* 7:12860. <https://doi.org/10.1038/ncomms12860>.
  39. Perez JC, Shin D, Zwir I, Latifi T, Hadley TJ, Groisman EA. 2009. Evolution of a bacterial regulon controlling virulence and Mg(2+) homeostasis. *PLoS Genet* 5:e1000428. <https://doi.org/10.1371/journal.pgen.1000428>.
  40. Cathelyn JS, Ellison DW, Hinchliffe SJ, Wren BW, Miller VL. 2007. The RovA regulons of *Yersinia enterocolitica* and *Yersinia pestis* are distinct: evidence that many RovA-regulated genes were acquired more recently than the core genome. *Mol Microbiol* 66:189–205. <https://doi.org/10.1111/j.1365-2958.2007.05907.x>.
  41. Heroven AK, Nagel G, Tran HJ, Parr S, Dersch P. 2004. RovA is auto-regulated and antagonizes H-NS-mediated silencing of *invasin* and *rovA* expression in *Yersinia pseudotuberculosis*. *Mol Microbiol* 53: 871–888. <https://doi.org/10.1111/j.1365-2958.2004.04162.x>.
  42. Wyborn NR, Stapleton MR, Norte VA, Roberts RE, Grafton J, Green J. 2004. Regulation of *Escherichia coli* hemolysin E expression by H-NS and *Salmonella* SlyA. *J Bacteriol* 186:1620–1628. <https://doi.org/10.1128/jb.186.6.1620-1628.2004>.
  43. Gordon BR, Li Y, Cote A, Weirauch MT, Ding P, Hughes TR, Navarre WW, Xia B, Liu J. 2011. Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc Natl Acad Sci U S A* 108:10690–10695. <https://doi.org/10.1073/pnas.1102544108>.
  44. Oshima T, Ishikawa S, Kurokawa K, Aiba H, Ogasawara N. 2006. *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res* 13: 141–153. <https://doi.org/10.1093/dnares/dsl009>.
  45. Gray KM, Garey JR. 2001. The evolution of bacterial LuxI and LuxR quorum sensing regulators. *Microbiology* 147:2379–2387. <https://doi.org/10.1099/00221287-147-8-2379>.
  46. Rajput A, Kumar M. 2017. In silico analyses of conservational, functional and phylogenetic distribution of the LuxI and LuxR homologs in Gram-positive bacteria. *Sci Rep* 7:6969. <https://doi.org/10.1038/s41598-017-07241-5>.
  47. Mortensen L, Dandanell G, Hammer K. 1989. Purification and characterization of the deoR repressor of *Escherichia coli*. *EMBO J* 8:325–331. <https://doi.org/10.1002/j.1460-2075.1989.tb03380.x>.
  48. Zeng X, Saxild HH, Switzer RL. 2000. Purification and characterization of the DeoR repressor of *Bacillus subtilis*. *J Bacteriol* 182:1916–1922. <https://doi.org/10.1128/jb.182.7.1916-1922.2000>.
  49. Engels V, Wendisch VF. 2007. The DeoR-type regulator SugR represses expression of *ptsG* in *Corynebacterium glutamicum*. *J Bacteriol* 189: 2955–2966. <https://doi.org/10.1128/JB.01596-06>.
  50. Garces F, Fernández FJ, Gómez AM, Pérez-Luque R, Campos E, Prohens R, Aguilar J, Baldomà L, Coll M, Badía J, Vega MC. 2008. Quaternary structural transitions in the DeoR-type repressor UlaR control transcriptional readout from the L-ascorbate utilization regulon in *Escherichia coli*. *Biochemistry* 47:11424–11433. <https://doi.org/10.1021/bi800748x>.
  51. Burr T, Barnard AM, Corbett MJ, Pemberton CL, Simpson NJ, Salmond GP. 2006. Identification of the central quorum sensing regulator of virulence in the enteric phytopathogen, *Erwinia carotovora*: the VirR repressor. *Mol Microbiol* 59:113–125. <https://doi.org/10.1111/j.1365-2958.2005.04939.x>.
  52. Monson R, Burr T, Carlton T, Liu H, Hedley P, Toth I, Salmond GP. 2013. Identification of genes in the VirR regulon of *Pectobacterium atrosepticum* and characterization of their roles in quorum sensing-dependent virulence. *Environ Microbiol* 15:687–701. <https://doi.org/10.1111/j.1462-2920.2012.02822.x>.
  53. Cha JY, Lee JS, Oh JI, Choi JW, Baik HS. 2008. Functional analysis of the role of Fur in the virulence of *Pseudomonas syringae* pv. *tabaci* 11528: Fur controls expression of genes involved in quorum-sensing. *Biochem Biophys Res Commun* 366:281–287. <https://doi.org/10.1016/j.bbrc.2007.11.021>.
  54. Tanui CK, Shyntum DY, Priem SL, Theron J, Moleleki LN. 2017. Influence of the ferric uptake regulator (Fur) protein on pathogenicity in *Pectobacterium carotovorum* subsp. *brasiliense*. *PLoS One* 12:e0177647. <https://doi.org/10.1371/journal.pone.0177647>.
  55. Yan Q, Gao W, Wu XG, Zhang LQ. 2009. Regulation of the PcoI/PcoR quorum-sensing system in *Pseudomonas fluorescens* 2P24 by the PhoP/PhoQ two-component system. *Microbiology* 155:124–133. <https://doi.org/10.1099/mic.0.020750-0>.
  56. Byers JT, Lucas C, Salmond GP, Welch M. 2002. Nonenzymatic turnover of an *Erwinia carotovora* quorum-sensing signaling molecule. *J Bacteriol* 184:1163–1171. <https://doi.org/10.1128/jb.184.4.1163-1171.2002>.
  57. Parker WL, Rathnum ML, Wells JS, Jr, Trejo WH, Principe PA, Sykes RB. 1982. SQ 27,860, a simple carbapenem produced by species of *Serratia* and *Erwinia*. *J Antibiot* 35:653–660. <https://doi.org/10.7164/antibiotics.35.653>.
  58. Choi J, Groisman EA. 2016. Acidic pH sensing in the bacterial cytoplasm is required for *Salmonella* virulence. *Mol Microbiol* 101:1024–1038. <https://doi.org/10.1111/mmi.13439>.
  59. Gunn JS, Richards SM. 2007. Recognition and integration of multiple environmental signals by the bacterial sensor kinase PhoQ. *Cell Host Microbe* 1:163–165. <https://doi.org/10.1016/j.chom.2007.05.001>.
  60. Cooper VJ, Salmond GP. 1993. Molecular analysis of the major cellulase (CelV) of *Erwinia carotovora*: evidence for an evolutionary “mix-and-match” of enzyme domains. *Mol Gen Genet* 241:341–350. <https://doi.org/10.1007/bf00284687>.
  61. Shyntum DY, Nkomo NP, Shingange NL, Gricia AR, Bellieny-Rabelo D, Moleleki LN. 2019. The impact of type VI secretion system, bacteriocins and antibiotics on bacterial competition of *Pectobacterium carotovorum* subsp. *brasiliense* and the regulation of carbapenem biosynthesis by iron and the ferric-uptake regulator. *Front Microbiol* 10:2379. <https://doi.org/10.3389/fmicb.2019.02379>.
  62. Lerat E, Moran NA. 2004. The evolutionary history of quorum-sensing systems in bacteria. *Mol Biol Evol* 21:903–913. <https://doi.org/10.1093/molbev/msh097>.
  63. Pemberton CL, Whitehead NA, Sebahia M, Bell KS, Hyman LJ, Harris SJ, Matlin AJ, Robson ND, Birch PR, Carr JP, Toth IK, Salmond GP. 2005. Novel quorum-sensing-controlled genes in *Erwinia carotovora* subsp. *carotovora*: identification of a fungal elicitor homologue in a soft-rotting bacterium. *Mol Plant Microbe Interact* 18:343–353. <https://doi.org/10.1094/MPMI-18-0343>.
  64. McGowan S, Sebahia M, Jones S, Yu B, Bainton N, Chan PF, Bycroft B, Stewart GS, Williams P, Salmond GP. 1995. Carbapenem antibiotic production in *Erwinia carotovora* is regulated by CarR, a homologue of



- the LuxR transcriptional activator. *Microbiology* 141:541–550. <https://doi.org/10.1099/13500872-141-3-541>.
65. Holeva MC, Bell KS, Hyman LJ, Avrova AO, Whisson SC, Birch PR, Toth IK. 2004. Use of a pooled transposon mutation grid to demonstrate roles in disease development for *Erwinia carotovora* subsp. *atroseptica* putative type III secreted effector (DspE/A) and helper (HrpN) proteins. *Mol Plant Microbe Interact* 17:943–950. <https://doi.org/10.1094/MPMI.2004.17.9.943>.
  66. Zou L, Zeng Q, Lin H, Gyaneshwar P, Chen G, Yang CH. 2012. SlyA regulates type III secretion system (T3SS) genes in parallel with the T3SS master regulator HrpL in *Dickeya dadantii* 3937. *Appl Environ Microbiol* 78:2888–2895. <https://doi.org/10.1128/AEM.07021-11>.
  67. Bijlsma JJ, Groisman EA. 2005. The PhoP/PhoQ system controls the intramacrophage type three secretion system of *Salmonella enterica*. *Mol Microbiol* 57:85–96. <https://doi.org/10.1111/j.1365-2958.2005.04668.x>.
  68. Ellermeier JR, Schlauch JM. 2008. Fur regulates expression of the *Salmonella* pathogenicity island 1 type III secretion system through HiiD. *J Bacteriol* 190:476–486. <https://doi.org/10.1128/JB.00926-07>.
  69. Cordero-Alba M, Bernal-Bayard J, Ramos-Morales F. 2012. SrfJ, a *Salmonella* type III secretion system effector regulated by PhoP, RcsB, and IolR. *J Bacteriol* 194:4226–4236. <https://doi.org/10.1128/JB.00173-12>.
  70. Haque MM, Hirata H, Tsuyumu S. 2012. Role of PhoP–PhoQ two-component system in pellicle formation, virulence and survival in harsh environments of *Dickeya dadantii* 3937. *J Gen Plant Pathol* 78:176–189. <https://doi.org/10.1007/s10327-012-0372-z>.
  71. Sana TG, Hachani A, Bucior I, Soscia C, Garvis S, Termine E, Engel J, Filloux A, Blevess S. 2012. The second type VI secretion system of *Pseudomonas aeruginosa* strain PAO1 is regulated by quorum sensing and Fur and modulates internalization in epithelial cells. *J Biol Chem* 287:27095–27105. <https://doi.org/10.1074/jbc.M112.376368>.
  72. Bernal P, Llamas MA, Filloux A. 2018. Type VI secretion systems in plant-associated bacteria. *Environ Microbiol* 20:1–15. <https://doi.org/10.1111/1462-2920.13956>.
  73. Wu HY, Chung PC, Shih HW, Wen SR, Lai EM. 2008. Secretome analysis uncovers an Hcp-family protein secreted via a type VI secretion system in *Agrobacterium tumefaciens*. *J Bacteriol* 190:2841–2850. <https://doi.org/10.1128/JB.01775-07>.
  74. Ma L-S, Hachani A, Lin J-S, Filloux A, Lai E-M. 2014. *Agrobacterium tumefaciens* deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. *Cell Host Microbe* 16:94–104. <https://doi.org/10.1016/j.chom.2014.06.002>.
  75. Kelly G, Prasannan S, Daniell S, Fleming K, Frankel G, Dougan G, Connerton I, Matthews S. 1999. Structure of the cell-adhesion fragment of intimin from enteropathogenic *Escherichia coli*. *Nat Struct Biol* 6:313–318. <https://doi.org/10.1038/7545>.
  76. Luo Y, Frey EA, Pfuetzner RA, Creagh AL, Knoechel DG, Haynes CA, Finlay BB, Strynadka NC. 2000. Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. *Nature* 405:1073–1077. <https://doi.org/10.1038/35016618>.
  77. Pell LG, Gasmi-Seabrook GM, Morais M, Neudecker P, Kanelis V, Bona D, Donaldson LW, Edwards AM, Howell PL, Davidson AR, Maxwell KL. 2010. The solution structure of the C-terminal Ig-like domain of the bacteriophage lambda tail tube protein. *J Mol Biol* 403:468–479. <https://doi.org/10.1016/j.jmb.2010.08.044>.
  78. Fraser JS, Yu Z, Maxwell KL, Davidson AR. 2006. Ig-like domains on bacteriophages: a tale of promiscuity and deceit. *J Mol Biol* 359:496–507. <https://doi.org/10.1016/j.jmb.2006.03.043>.
  79. Preston G, Deng WL, Huang HC, Collmer A. 1998. Negative regulation of hrp genes in *Pseudomonas syringae* by HrpV. *J Bacteriol* 180:4532–4537. <https://doi.org/10.1128/JB.180.17.4532-4537.1998>.
  80. Jovanovic M, James EH, Burrows PC, Rego FG, Buck M, Schumacher J. 2011. Regulation of the co-evolved HrpR and HrpS AAA+ proteins required for *Pseudomonas syringae* pathogenicity. *Nat Commun* 2:177. <https://doi.org/10.1038/ncomms1177>.
  81. Jovanovic M, Waite C, James E, Synn N, Simpson T, Kotta-Loizou I, Buck M. 2017. Functional characterization of key residues in regulatory proteins HrpG and HrpV of *Pseudomonas syringae* pv. *tomato* DC3000. *Mol Plant Microbe Interact* 30:656–665. <https://doi.org/10.1094/MPMI-03-17-0073-R>.
  82. Wei CF, Deng WL, Huang HC. 2005. A chaperone-like HrpG protein acts as a suppressor of HrpV in regulation of the *Pseudomonas syringae* pv. *syringae* type III secretion system. *Mol Microbiol* 57:520–536. <https://doi.org/10.1111/j.1365-2958.2005.04704.x>.
  83. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97:6640–6645. <https://doi.org/10.1073/pnas.120163297>.
  84. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
  85. Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12:357–360. <https://doi.org/10.1038/nmeth.3317>.
  86. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923–930. <https://doi.org/10.1093/bioinformatics/btt656>.
  87. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. <https://doi.org/10.1093/bioinformatics/btp616>.
  88. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol Biol Evol* 34:2115–2122. <https://doi.org/10.1093/molbev/msx148>.
  89. Kanehisa M, Goto S. 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30. <https://doi.org/10.1093/nar/28.1.27>.
  90. Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F, Bateman A, Eddy SR. 2015. HMMER web server: 2015 update. *Nucleic Acids Res* 43:W30–W38. <https://doi.org/10.1093/nar/gkv397>.
  91. Finn RD, Mistry J, Tate J, Coghill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A. 2010. The Pfam protein families database. *Nucleic Acids Res* 38:D211–D222. <https://doi.org/10.1093/nar/gkp985>.
  92. Heberle H, Meirelles GV, da Silva FR, Telles GP, Minghim R. 2015. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 16:169. <https://doi.org/10.1186/s12859-015-0611-3>.
  93. Li L, Stoeckert CJ, Jr, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13:2178–2189. <https://doi.org/10.1101/gr.1224503>.
  94. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup> $\Delta\Delta CT$  method. *Methods* 25:402–408. <https://doi.org/10.1006/meth.2001.1262>.
  95. Axelrood PE, Rella M, Schroth MN. 1988. Role of antibiosis in competition of *Erwinia* strains in potato infection courts. *Appl Environ Microbiol* 54:1222–1229. <https://doi.org/10.1128/AEM.54.5.1222-1229.1988>.
  96. Sievers F, Higgins DG. 2018. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* 27:135–145. <https://doi.org/10.1002/pro.3290>.
  97. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>.
  98. Drozdetskiy A, Cole C, Procter J, Barton GJ. 2015. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* 43:W389–W394. <https://doi.org/10.1093/nar/gkv332>.
  99. Eswar N, Eramian D, Webb B, Shen MY, Sali A. 2008. Protein structure modeling with MODELLER. *Methods Mol Biol* 426:145–159. [https://doi.org/10.1007/978-1-60327-058-8\\_8](https://doi.org/10.1007/978-1-60327-058-8_8).
  100. Schrödinger L. 2015. The AxPyMOL molecular graphics plugin for Microsoft PowerPoint, version 1.8. Schrödinger, LLC, New York, NY.
  101. Ravenhall M, Škunca N, Lassalle F, Dessimoz C. 2015. Inferring horizontal gene transfer. *PLoS Comput Biol* 11:e1004095. <https://doi.org/10.1371/journal.pcbi.1004095>.
  102. Sims D, Ponting CP, Heger A. 2016. CGAT: a model for immersive personalized training in computational genomics. *Brief Funct Genomics* 15:32–37. <https://doi.org/10.1093/bfpg/elv021>.
  103. Kullback S, Leibler RA. 1951. On information and sufficiency. *Ann Math Statist* 22:79–86. <https://doi.org/10.1214/aoms/1177729694>.
  104. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>.
  105. Perez-Llamas C, Lopez-Bigas N. 2011. Gitoools: analysis and visualisation of genomic data using interactive heat-maps. *PLoS One* 6:e19541. <https://doi.org/10.1371/journal.pone.0019541>.