Comparative transcriptomics reveal a highly polymorphic *Xanthomonas* HrpG virulence regulon

Authors

Thomas Quiroz Monnens [a], Brice Roux [a], Sébastien Cunnac [d], Erika Charbit[b], Sébastien Carrère [a], Emmanuelle Lauber [a], Marie-Françoise Jardinaud [a], Armelle Darrasse [b], Matthieu Arlat [a], Boris Szurek [d], Olivier Pruvost [c], Marie-Agnès Jacques [b], Lionel Gagnevin [c,d,e], Ralf Koebnik [d], Laurent D. Noël [a] and Alice Boulanger [a]

[a] LIPME, Université de Toulouse, INRAE/CNRS/Université Paul Sabatier Toulouse 3, UMR 0441/2598, Castanet-Tolosan, France

[b] Univ. Angers, Institut Agro, INRAE, IRHS, SFR QUASAV, F-49000 Angers, France

[c] CIRAD, UMR PVBMT, F-97410 Saint-Pierre, La Réunion, France.

[d] PHIM, Université de Montpellier, IRD, CIRAD, INRAE, Institut Agro, Montpellier, France

[e] CIRAD, UMR PHIM, F-34398 Montpellier, France

Authors for correspondence

Alice Boulanger, alice.boulanger@inrae.fr

Laurent D. Noël, laurent.noel@inrae.fr

Manuscript word count: 5186

Abstract word count: 242

Importance word count: 145

ABSTRACT

Bacteria of the genus *Xanthomonas* cause economically significant diseases in various crops. Their virulence is dependent on the translocation of type III effectors (T3Es) into plant cells by the type III secretion system (T3SS), a process regulated by the master response regulator HrpG. Although HrpG has been studied for over two decades, its regulon across diverse *Xanthomonas* species, particularly beyond type III secretion, remains understudied. In this study, we conducted transcriptome sequencing to explore the HrpG regulons of 17 *Xanthomonas* strains, encompassing six species and nine pathovars, each exhibiting distinct host and tissue specificities. We employed constitutive expression of plasmid-borne *hrpG**, which encodes a constitutively active form of HrpG, to induce the regulon. Our findings reveal substantial inter- and intra-specific diversity in the HrpG* regulons across the strains. Besides 21 genes directly involved in the biosynthesis of the T3SS, the core HrpG* regulon is limited to only five additional genes encoding the transcriptional activator HrpX, the two T3E proteins XopR and XopL, a major facility superfamily (MFS) transporter, and the phosphatase PhoC. Interestingly, genes involved in chemotaxis and genes encoding enzymes with carbohydrate-active and proteolytic activities are variably regulated by HrpG*. The diversity in the HrpG* regulon suggests that HrpG-dependent virulence in *Xanthomonas* might be achieved through several distinct strain-specific strategies, potentially reflecting adaptation to diverse ecological niches. These findings enhance our understanding of the complex role of HrpG in regulating various virulence and adaptive pathways, extending beyond T3Es and the T3SS.

IMPORTANCE

In the decades since its discovery, HrpG and its role in the regulation of the type III secretion system (T3SS) and its associated type III effectors (T3Es) in *Xanthomonas* has been the

subject of extensive research. Despite notable progress in understanding its molecular regulatory mechanisms, the full spectrum of processes under control of HrpG, particularly beyond the T3SS and T3Es, and the degree of regulatory conservation across plant-pathogenic *Xanthomonas* species, remained unclear. To address this knowledge gap, we systematically compared the transcriptomes of 17 *Xanthomonas* strains, expressing a constitutively active form of HrpG, called HrpG*. We showed that HrpG* regulates different physiological processes other than the T3SS and T3Es and that this regulation shows substantial variation across the different strains. Taken together, our results provide new insights into *Xanthomonas*-plant interactions through the regulation of different metabolic and virulence pathways by the master response regulator HrpG.

Keywords: *Xanthomonas*, HrpG, type III secretion system, regulon diversity, transcriptome sequencing

INTRODUCTION

In phytopathogenic *Xanthomonas* bacteria, HrpG is a conserved response regulator belonging to the OmpR family of two-component regulatory systems (1). Two-component regulatory systems are one of the major mechanisms used by bacteria to perceive and adapt their physiology to changing environmental conditions (2). These systems typically rely on a phosphorelay between a sensor kinase and a response regulator, which becomes active upon phosphorylation. In *Xanthomonas*, HrpG regulates the expression of *hrpX*, which encodes a transcriptional regulator. HrpX binds to the plant-inducible promoter (PIP) box (3), a DNA sequence present in the cis-elements of various pathogenicity genes, including type III effector proteins (T3Es) and the *hrp* (hypersensitive response and pathogenicity) genes, which encode the type III secretion system (T3SS) (4, 5). This system is one of the main virulence factors of *Xanthomonas* and it enables the direct delivery of bacterial T3Es into the plant cellular environment, where they interfere with plant immunity and alter plant physiology to facilitate disease (6). Disruption of any key structural component of the T3SS results in a complete loss of pathogenicity for *Xanthomonas* species (4).

The *Xanthomonas* genus comprises 33 species of Gram-negative □-proteobacteria (7) causing diseases on more than 400 different monocot and dicot plants including economically-important crops (8). *Xanthomonas* species are further divided into pathovars that can exhibit distinct host specificities and tissue tropisms, causing a variety of symptoms such as wilting, necrosis or blight (8). Phylogenetic group-I *Xanthomonas* include monocot pathogens such as *X. translucens* pv. *translucens* (*Xtt*) (9), causal agent of bacterial leaf streak of barley. On the other hand, phylogenetic group-II *Xanthomonas* can infect both monocot and dicot plants. For example, *X. oryzae* pv. *oryzae* (*Xoo*) causes bacterial blight on rice, a monocot, while *X. phaseoli* pv. *phaseoli* (*Xpp*) and *X. citri* pv. *fuscans* (*Xcf*) belonging to phylogenetically distinct species are both able to cause common blight of common bean, a dicot. Other economically important *Xanthomonas* pathovars include *X. citri* pv. *mangiferaeindicae* (*Xcm*), causal agent of mango bacterial black spot, *X. citri* pv. *citri* (*Xci*), responsible for Asiatic citrus canker and *X. euvesicatoria* pv. *alfalfae* (*Xea*), which causes disease on legumes. The genus also comprises *X. campestris* pv. *campestris* (*Xcc*) and *X. campestris* pv. *raphani* (*Xcr*), the causal agents of black rot disease and bacterial spot disease of Brassicaceae respectively. *Xcc*, *Xoo* and *Xtt* are known to infect the plant vasculature through wounds or hydathodes while *Xci*, *Xcm*, *Xea* and *Xcr* colonize the mesophyll through stomata. *Xcf* and *Xpp* colonize their host plant by both ways (10–13).

Since its discovery, research has shown that HrpG regulates a broad array of genes beyond T3Es and the T3SS. This broader regulatory role was initially identified by comparing the transcriptomes of wild-type versus *hrpG* mutant strains in *hrp*-inducing media (14) or wild-type strains versus strains expressing an auto-active gain-of-function form of HrpG, such as HrpG^{E44K}, commonly named HrpG* (1, 15, 16). Additional studies have supported these claims, showing HrpG could regulate the expression of genes other than those associated with type III secretion, including genes associated with chemotaxis and motility, and genes encoding extracellular proteases and cell wall degrading enzymes associated with the Xps type II secretion system (T2SS) (17–20). However, this part of the HrpG regulon remains understudied and appears to be poorly conserved between *Xanthomonas* species.

To investigate the conservation of the HrpG regulon across *Xanthomonas* species, we performed a comprehensive transcriptomic and genomic analysis of 17 *Xanthomonas* strains transiently expressing the constitutively active *hrpG** variant (21). The HrpG* regulons identified in the different strains varied considerably in their size, ranging from 137 to 2,355 genes. Interestingly, the core HrpG* regulon across these 17 strains comprises only 26 genes, mainly involved in the biogenesis of the T3SS itself. Moreover, it was found that the full extent of genes and processes regulated by HrpG* is remarkably diverse across the 17 strains. These findings suggest that HrpG-dependent pathogenicity in *Xanthomonas* species can be achieved through diverse strategies.

RESULTS

Transcriptome-based structural annotation of 17 *Xanthomonas* genomes

17 *Xanthomonas* strains belonging to nine pathovars were selected to study the diversity and conservation of the HrpG regulon (Fig. 1A). These strains represent pathogens associated

with diverse host plants, were sampled over nearly a hundred years in ten countries, and have various lifestyles (Table S1A; Fig. 1B). Genome sequences of strains *Xtt*$_{\text{CFBP2054}}$,*Xcc*$_{\text{CN05}}$, *Xcm*$_{\text{LG56-10}}$ and *Xcm*$_{\text{LG81-27}}$ were determined using either short- (Illumina) or long-read (PacBio) sequencing (Table S1B; S1C). High-quality genome sequences were readily available for the 13 other strains.

Experimentally-based and homogeneous annotation of genomes is vital for cross-species comparative transcriptomic analyses. Such annotations were previously made for *Xcc*$_{8004}$ and *Xcr*$_{\text{CFBP5828R}}$, which were based on various transcriptomic datasets and the EuGene-PP pipeline (16, 22). The same strategy was used to annotate the remaining genomes in combination with different RNA-Seq libraries (Table S1E), including those from conjugates carrying an empty vector or a vector harboring *hrpG**. For *Xci* and *Xcm* strains, empty vector conjugates were not generated, and RNA-Seq libraries from wild-type strains were used for annotation instead. Small RNA sequencing was performed for several strains in order to identify small RNAs and determine transcriptional start sites. The newly annotated genomes consist of 4066 (*Xtt*$_{\text{CFBP2054}}$) to 5421 (*Xcf*$_{\text{CFBP7767}}$) protein-coding genes (Table S1C; Data S1 https://doi.org/10.57745/9OTYNJ). The 17 homogeneously annotated genomes were used for all downstream analyses.

The *Xanthomonas* core genome is composed of 2,483 orthogroups

To enable cross species comparative transcriptomics, an orthogroup (OG) database was built with the newly annotated genomes using Orthofinder. Collectively, the 17 genomes contained 88,480 genes, of which 81,490 were predicted to be protein-coding. Out of these, 78,225 were assigned to a total of 6,899 orthogroups (Fig. 1C). 2,483 orthogroups were represented by at least one ortholog in each strain, representing the *Xanthomonas* core genome in this study (Fig. 1D; Data S2 https://doi.org/10.57745/9OTYNJ). A phylogenetic tree including all strains was built using the STAG algorithm and rooted using STRIDE (Fig. 1A). As expected, group-I *Xanthomonas* strain *Xtt*$_{\text{CFBP2054}}$ was located at the root of the tree. To assess the overall similarity in orthogroup content among the strains, a heatmap depicting the number of shared orthogroups and a principal component analysis (PCA) based on a binary presence/absence matrix of all orthogroups in the dataset were used (Fig. 1EF). The high number of shared orthogroups between related strains in the heatmap, and the clustering of related strains in the PCA is consistent with the genetic relationships of the *Xanthomonas* strains.

The accessory HrpG* regulons are highly diverse

The HrpG* regulons were identified by sequencing cDNAs of either wild-type strains with or without an empty plasmid or a plasmid carrying *hrpG**. The biological reproducibility of each triplicate per genotype was evaluated through MultiQCs on mapping statistics, and through PCA, MA (M = log fold change, A = mean of normalized counts) and volcano plots of the DESeq2 output (Data S3, https://doi.org/10.57745/9OTYNJ). The reproducibility across all samples was consistent except for one *Xtt* replicate, which was excluded from subsequent differential expression analysis. As expected for a direct HrpG target, expression of the *hrpX* gene was significantly upregulated in all *hrpG** samples (Fig. 2AB). This demonstrates the presence of biologically active HrpG in all *hrpG** samples, thereby validating its use for the analysis of HrpG regulons.

The number of differentially-expressed genes (DEGs, AdjPval < 0.05) varied across species from 137 to 2,355 genes (Fig. 2C; Table S2A). Considering a threshold of $|\text{Log}_2\text{FC}| > 2$, the number of DEGs ranged from 59 to 429 (Fig. 2D). Notably, in all strains but *Xcf*$_{\text{CFBP7767}}$ and *Xcf*$_{\text{CFBP6996R}}$, the proportion of HrpG* regulated protein-coding genes in the core genome was significantly lower than in the accessory genome (Chi-square, BH FDR, AdjPval < 0.05, Fig.

2EF). Furthermore, the regulation of orthogroups by HrpG* within the core genome was highly variable, as highlighted by a PCA (Fig. 2G). In addition, the strains showed substantial differences in enriched Gene Ontology (GO) terms amongst DEGs considering a threshold of $|Log_2FC| > 2$ (Fig. 3). Notably, only GO terms associated with protein secretion and protein secretion by the type III secretion system were enriched in the regulons of all strains, illustrating that the functions of the accessory HrpG regulon are highly diverse (Fig. 3).

The HrpG* core regulon is limited to only 26 orthogroups

To determine the HrpG* core regulon, orthogroups of the core genome of which at least one ortholog was differentially regulated in *hrpG** samples across all strains were identified (Table 1). The identified HrpG* core regulon comprises one orthogroup coding for HrpX, 21 orthogroups encoding structural components of the T3SS and two orthogroups coding for the T3Es XopR and XopL. Additionally, the core regulon also includes two orthogroups coding for a putatively Sec/SPI-secreted acid phosphatase of the PAP2 superfamily (annotated as PhoC) and a major facilitator superfamily (MFS) transporter. Here, *hrpG* upregulation was ignored as sequence reads predominantly originated from plasmid-borne *hrpG** transcripts (Data S4 https://doi.org/10.57745/9OTYNJ). Thus, the identified HrpG* core regulon is almost exclusively associated with type III secretion.

There is compelling evidence that group-I *Xanthomonas* acquired the type III secretion system cluster independently of the group-II *Xanthomonas* species, potentially resulting in the existence of different HrpG regulons (23). As the majority of strains used in this study belong to group-II *Xanthomonas* species, their core regulon was also identified. The identified group-II core regulon includes an additional 12 orthogroups compared to the *Xanthomonas* core regulon (Table 2). These include orthogroups encoding several known virulence determinants, such as a chorismate mutase and LipA whose function is independent of the T3SS and represent conserved and ancestral virulence mechanisms (24).

Expression of most but not all T3Es encoding genes is controlled by HrpG*

Previous studies have shown that T3Es, both with and without a PIP-box motif in their promoters (TTCGB-N_{15}-TTCGB), can be under positive regulation by HrpG or HrpX (3, 14–16, 20). To investigate the regulation of T3Es by HrpG* in the 17 strains analysed, we used the Effectidor software (25) to identify orthogroups encoding T3Es. A total of 72 orthogroups were predicted to contain orthologs coding for T3Es (Table S3A). From these, 12 orthogroups corresponding to *xopA*, *hrpW*, *hpaA* and other T3SS associated genes, which are generally not considered T3Es, were excluded (Table S3C). Additionally, Effectidor identified 15 singletons that are predicted to encode T3Es. Consistent with previous findings, our observations show that the majority of T3E genes are under the control of HrpG* (Table S3B; Fig. 4). However, the expression of few predicted effectors was not HrpG*-dependent, as observed for *xopAW*. As for the core T3E gene *xopM*, its expression was HrpG*-dependent except for both *Xcm* strains. Interestingly, most orthologs encoding transcription activator-like effectors (TALEs, OG000009) were positively regulated by HrpG*, although *Xci* TALEs were previously reported as expressed in a HrpG/HrpX-independent manner (14, 20). We thus conclude that the expression of most, but not all of the predicted T3E genes is under the positive control of HrpG*.

Chemotaxis and motility are differentially regulated by HrpG* at both inter- and intra-specific levels

The PCA, based on the average HrpG*-dependent Log_2FC of expression of orthogroups within the core genome, indicated that the regulatory network of HrpG* within the core genome differed considerably between strains, even for those of the same pathovar (Fig. 2G). Orthogroups that correlated strongly with the first three principal components were enriched for GO terms associated with the biological process of chemotaxis (Table S4). We therefore investigated the regulation of orthogroups comprising genes encoding methyl-accepting chemotaxis proteins (MCPs), Che signalling genes and structural components of the flagellum and the type-IV pilus. Interestingly, the regulation of these orthogroups was HrpG*-dependent in half the strains and both HrpG*-dependent upregulation and downregulation of these processes could be observed, depending on the strain (Fig. 5). For example, *Xcm*$_{\text{LG56-10}}$ showed strong upregulation of genes in motility- and chemotaxis-related orthogroups while *Xcm*$_{\text{LG81-27}}$ did not. These findings highlight diverse scenarios of HrpG*-dependent regulation of chemotaxis and motility in the *Xanthomonas* genus.

Strain-dependent variability in HrpG*-mediated regulation of T2SS substrates

The HrpG* regulon of different strains exhibited significant enrichment in genes linked to GO terms associated with carbohydrate-active enzymatic activity and proteolytic activity (Fig. 3). Such genes play a crucial role in *Xanthomonas* virulence due to their involvement in the degradation of plant cell wall components, which facilitates nutrient acquisition and the efficient translocation of T3Es (19, 26). Typically, proteins encoded by these genes are characterized by the presence of a Sec/SPI signal peptide, indicative of secretion by the T2SS. Indeed, such signal peptides were predicted in 546 of the 3351 protein-coding genes of the pan HrpG* regulon, representing a significant enrichment for the presence of a signal peptide amongst regulated genes ($|\text{Log}_2\text{FC}| > 2$, Chi-square, $p < 0.0001$, Table S2C). We therefore hypothesized that HrpG* might also regulate the T2SS, as previously shown for *Xanthomonas euvesicatoria* pv. *euvesicatoria* (*Xev*) (19). However, genes encoding the T2SS were not under the control of HrpG* ($|\text{Log}_2\text{FC}| > 2$), except for *Xcm*$_{\text{LG56-10}}$ (Fig. S1).

Though HrpG is known to regulate the expression of genes encoding proteins with carbohydrate active enzymatic and/or proteolytic activity (15, 17–19, 27, 28), such regulation can either be positive or negative without any obvious pattern (29). We thus examined the regulation of those genes in the transcriptomes of the 17 *Xanthomonas* strains (Fig. 6 and Fig. 7). Once more, various HrpG*-dependent regulatory patterns were observed. For example, orthologs in OG0003431, encoding putative S53 family serine proteases, were generally upregulated except in *Xoo*$_{\mathrm{BAI3}}$. Conversely, expression of orthologs in OG0000136, encoding putative S1 family serine proteases, were mainly downregulated, while *Xoo*$_{\mathrm{BAI3}}$ strongly upregulated one specific ortholog. Genes encoding proteins with putative pectin-lyase activity were also variably regulated by HrpG*. For example, orthologs in OG0000700, corresponding to genes encoding putative secreted GH28 family pectin lyases, were strongly upregulated in all strains except *Xoo*$_{\mathrm{BAI3}}$ and *Xtt*$_{\mathrm{CFBP2054}}$. In contrast, orthologs in OG0000100 and OG0000105, corresponding to genes encoding putative secreted PL1 family pectin lyases, showed considerable downregulation, especially in *Xanthomonas campestris* pathovars, but not in others such as *Xpp*$_{\mathrm{CFBP6545R}}$. These results are in line with previous reports, highlighting the differential regulation of genes encoding proteins with specific carbohydrate active enzymatic and/or proteolytic activity by HrpG* across different *Xanthomonas* strains. Nonetheless, most orthogroups showed consistent HrpG*-dependent regulation across the majority of strains, evidencing the existence of conserved HrpG*-mediated regulatory patterns for genes encoding specific families of carbohydrate active enzymes and/or proteases in the *Xanthomonas* genus.

The HrpG* regulon members prepare *Xanthomonas* cells for the degradation of plant-derived phenolic compounds

The degradation of plant cell walls mediated by T2SS substrates in *Xanthomonas* species leads to the release of diverse phenolic compounds, including hydroxycinnamic acids, vanillic acid and 4-hydroxybenzoic acid (4-HBA). Interestingly, the HrpG* core regulon comprises an MFS transporter (OG0001168) involved in 4-HBA uptake in *Xcc*$_{8004}$ (Table 1) (30, 31). In addition, the group-II *Xanthomonas* regulon comprises orthogroups coding different proteins involved in the uptake (VanK/PcaK, OG0000558) and putative degradation (two subunits of a PCA dioxygenase, OG0001327, OG0001328) of phenolic compounds (30, 31) as well as a PcaQ-like transcriptional regulator (OG0001326), which regulates phenolic compound metabolism in other plant-associated bacteria (Table 2) (32, 33). Therefore, a broader survey of orthogroups relevant for phenolic compound metabolism was conducted (Fig. 8). We observed that expression of numerous genes involved in the import and degradation of such compounds was upregulated, while some genes putatively involved in their efflux were downregulated. Notably, many orthogroups involved in these processes were absent from the *Xtt*$_{\mathrm{CFBP2054}}$ genome. Collectively, these results suggest that HrpG* prepares the group-II *Xanthomonas* metabolism for the import and degradation of plant-derived phenolic compounds.

HrpG* induces the expression of genes involved in cytochrome C maturation in most *Xanthomonas* strains.

GO terms associated with “iron binding”, “heme transport” and “siderophore uptake” were significantly enriched in the HrpG* regulons of different strains (Fig. 3) and are commonly associated with iron homeostasis. Orthologs of TonB-dependent receptors involved in iron uptake in *Xcc* (34) were differentially regulated in some strains expressing HrpG* (Fig. S2) but did not explain the observed enrichment of those three GO terms in all strains. Upon further investigation we found that the enrichment in these GO terms originated mainly from orthogroups comprising genes putatively involved in cytochrome C maturation (*Ccm* genes). Cytochrome C maturation complexes can have multiple roles in bacteria, including respiration (35), resistance to antimicrobial phenazines(36), or virulence as shown for *Xcc* (37). Interestingly, the identified regulated *Ccm* genes were located in an evolutionary-conserved gene cluster composed of a hypothetical protein, the sigma-factor RpoE4, another hypothetical protein with a zinc-finger domain and a S8A protease with a predicted Sec/SPI secretion signal (Fig. 9). Most genes in this cluster were regulated positively by HrpG* except in *Xpp*$_{\mathrm{CFBP6546R}}$ and *Xtt*$_{\mathrm{CFBP2054}}$. These results indicate that HrpG* regulates cytochrome C maturation in most *Xanthomonas* strains.

DISCUSSION

This comparative transcriptomics study has provided a genus-wide overview of the evolutionarily-conserved processes regulated by HrpG*, as well as a glimpse into the intra- and inter-specific diversity of the regulon (Fig. 10). The regulons of the strains investigated here encompass hundreds, if not thousands of genes, including a variety of known virulence and adaptive pathways with notable enrichment for putative T2SS substrates. However, these regulons differ considerably, both within and between species. This variation is illustrated by a small HrpG* core regulon (26 orthogroups for the *Xanthomonas* genus and 38 for group-II *Xanthomonas*) and by the fact that only two gene ontology terms (GO0009306: protein secretion, and GO0030254: protein secretion by type III secretion system), were significantly enriched among the regulated genes across all strains. Additionally, our analysis revealed that the proportion of regulated protein-coding genes in the accessory genomes of most strains was significantly higher than that in the core genome. Collectively, these findings illustrate that the HrpG regulon from one *Xanthomonas* strain cannot be inferred from another related *Xanthomonas* strain. The diversity of HrpG regulon could be the results of long-term evolution process. Indeed, HrpG and HrpX were acquired prior to the split of *Xanthomonas* in group I and II and the acquisition of the T3SS (38) suggestive of an ancestral role in the regulation of other biological functions. How the T3SS became part of the core regulon of HrpG remains to be elucidated.

The acid phosphatase PhoC belongs to the *Xanthomonas* HrpG* core regulon

The acid phosphatase PhoC (OG0000410) from the PAP2 superfamily belongs to the HrpG* core regulon and exhibits high conservation across *Xanthomonas* species. Of the 657 publicly available *Xanthomonas* genomes that contain a *hrp* T3SS gene cluster (Data S6, https://doi.org/10.57745/9OTYNJ), 650 harbour a *phoC* ortholog. Among these, 606 orthologs have a predicted Sec/SPI secretion signal whilst all others have a Sec/SPII secretion signal, indicating a strong conservation of T2SS-mediated secretion (Table S5A). The ortholog of *phoC* in *Xev* has been shown to be upregulated during the interaction with tomato although knockout mutants were not affected in virulence (39). Additionally, the *Xcc phoC* ortholog was not important for fitness inside cauliflower hydathodes nor for *Xcc* virulence on cabbage (40, 41). Despite the strong conservation of this orthogroup, its molecular or enzymatic functions remain elusive. In contrast, the functions of the T3Es XopL and XopR, both member of the HrpG* core regulon, have extensively been studied (42). The *xopL* orthologs encode atypical E3 ubiquitin ligases which can contribute to *Xanthomonas* virulence in several distinct mechanistical ways, depending on the species (43–47). This, together with significant interspecific variation in amino acid sequence, *in planta* subcellular localizations and host specific cell death-inducing capability indicate that this ancestral effector has undergone significant diversification (47). In *Xoo,* XopR is an effector which localizes to the plasma membrane, where it associates with various receptor-like cytoplasmic kinases (48–50). In both *Xoo* and *Xanthomonas axonopodis* pv. *manihotis* XopR is thought to contribute to virulence by interfering with PTI (48, 51).

The group-II *Xanthomonas* HrpG* core regulon mediates ancestral T3SS-independent virulence mechanisms

The regulation of 12 orthogroups was identified to be HrpG*-dependent specifically in the group-II *Xanthomonas* species tested in this study. Among them, orthogroups encoding VirK, a chorismate mutase, and the LipA/LesK lipase have been previously identified by comparative genomic analyses to be potential conserved virulence determinants in *Xanthomonas* species (24). Orthologs of *virK* are conserved across various lineages of plant-associated bacteria and are present in the genomes of all 657 *Xanthomonas* strains with a T3SS (Data S6, https://doi.org/10.57745/9OTYNJ). While the function of VirK has not yet been experimentally addressed, *in silico* analyses suggest that it interacts with structural components of both the T3SS and the Xps T2SS, and it is hypothesized to be secreted by the T2SS (24). Orthologs of the chorismate mutase are also found in all genomes of *Xanthomonas* species with a T3SS, showing high levels of amino acid sequence similarity, including a Sec/SPI/SPII signal secretion signal (Table S5B; Data S6, https://doi.org/10.57745/9OTYNJ). Although its molecular functions remain undefined, it contributes to *Xoo* virulence, possibly by interfering with salicylate-dependent immunity (24, 52) (Degrassi *et al.*, 2010; Assis *et al.*, 2017). LipA/LesK orthologs are conserved with a Sec/SPI secretion signal in nearly all *Xanthomonas* species with a T3SS but absent from non-pathogenic *Xanthomonas* (Table S5C; Data S6, https://doi.org/10.57745/9OTYNJ) (24). Interestingly, while an ortholog is present in *Xtt*$_{\text{CFBP2054}}$ genome, it lacks a PIP-box promoter motif, likely explaining the HrpG*-independent expression of *LipA/LesK* in this Group-I *Xanthomonas*. *Xev*, *Xci* and *Xoo* LipA/LesK contribute to virulence on tomato (53), citrus (24) and rice (54), respectively.

Importantly, the ortholog of LipA/LesK is also a virulence mechanism for *Xylella fastidiosa* on grapevine (55) suggestive of an ancestral virulence function in the genera *Xanthomonas*, *Xylella* and *Burkholderia* (24) before its co-optation in the HrpG regulon of Group-II *Xanthomonas*.

The group-II *Xanthomonas* HrpG* regulon relies on a transcriptional regulatory cascade

In addition to HrpX, the group-II core regulon comprises three additional transcriptional regulators, namely a MarR transcription factor, HpaR and a PcaQ-like transcriptional activator. The *Xci* MarR ortholog is essential for pathogenicity on Rangpur lime (56) although its gene targets are unknown. The transcriptional regulator HpaR is essential for *Xcc* virulence and thought to regulate the expression of various virulence mechanisms including extracellular proteases secreted by the Xps T2SS (17). As for the PcaQ-like transcriptional activator, it is hypothesized to be the main regulator of genes involved in protocatechuate degradation pathway, which is known to be important for *Xcc* virulence (30, 31). Interestingly, in *Xtt*$_{\mathrm{CFBP2054}}$, the PcaQ-like transcriptional activator is absent and the expression of the MarR transcription factor and *hpaR* is HrpG*-independent. This could explain the great divergence observed between the regulons in *Xtt*$_{\mathrm{CFBP2054}}$ and those in group-II *Xanthomonas*. These results thus suggest that a complex cascade defines the full extent of HrpG* regulons in group-II *Xanthomonas*, which could be refined by determining the transcriptomes of the corresponding single and multiple mutants in a *hrpG** mutant background.

HrpG* regulates motility and chemotaxis in unpredictable ways.

Motility and chemotaxis are known virulence determinants in *Xanthomonas*, primarily associated with the initial stages of infection (57–60). As HrpG expression gradually increases during infection (22, 61), it seems likely that HrpG could mediate the suppression of motility and chemotaxis, thereby potentially limiting PTI induced by flagellar components (29). So far, reports in *Xci* have indeed shown that HrpG negatively regulates motility and chemotaxis, likely in a HrpX-independent manner (14). However, we report the variable regulation of these processes. For instance, HrpG*-dependent downregulation was observed in *Xoo*$_{\mathrm{BAI3}}$, while no regulation was observed in *Xcm*$_{\mathrm{LG81\text{-}27}}$ and, remarkably, upregulation was measured in *Xcm*$_{\mathrm{LG56\text{-}10}}$. These findings are intriguing, especially considering the crucial role of motility and chemotaxis in the virulence of *Xanthomonas* species, and suggest the presence of complex inter- and intra-specific virulence mechanisms that require further investigation.

In conclusion, we demonstrate that *Xanthomonas* species not only differ by their gene contents (Fig. 1), but also by their gene expression profiles, which are diverse even at the intra-pathovar scale. These observations are consistent with a general evolutionary context where virulence factors important in some hosts are often directly or indirectly recognised as immune elicitors in other hosts and therefore undergo diversifying selection. The observed differential expression within the HrpG regulon may limit recognition of some of these genes while diversifying the strains encountered by plants. Such adaptative transcriptional regulation would have the advantage of being transient, plastic and environment-dependent thus facilitating the emergence of novel virulence properties.

MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed (Table S1). *Xanthomonas* strains were grown at 28 °C in MOKA medium (34). *Escherichia coli* cells were grown on LB medium at 37 °C. For solid media, agar was added at a final concentration of 1.5 % (w/v). Antibiotics were used when appropriate at the following concentrations: 50 µg/mL kanamycin, 50 µg/mL rifampicin, 40 µg/mL spectinomycin.

Genome sequencing and assembly

Genomic DNA was extracted from bacterial cells grown overnight in MOKA-rich medium using the Wizard genomic DNA purification kit (Promega) or as described (62) for short- or long-read sequencing, respectively. Shotgun sequencing of genomic DNA was performed either on HiSeq2000 Illumina platform (63) or PacBio (64). Genome assembly was performed as described for short read assemblies (63). FLYE Assembler (version 2.9) was used for long-read assemblies (65). The GenBank accession numbers for the genomes generated in this study are given in the Table S1.

Cloning of pBBR-hrpG* plasmid

The *hrpG** (E44K) coding sequences were amplified from *Xcc* strain 8004 *hrpG** (66) and *Xanthomonas vesicatoria* pv. *euvesicatoria* strain 85-10 *hrpG* (1) and cloned into pBBR1-MCS-2 (67) as described (16) (Table S1D). Plasmids were introduced into *E. coli* by electroporation. The plasmids were introduced into the different *Xanthomonas* strains by triparental mating using pRK2073 as helper plasmid (68, 69) (Table S1D).

RNA extraction, rRNA and tRNA depletion and cDNA pyro-sequencing

RNAs were extracted from *Xanthomonas* strains grown at exponential phase ($OD_{600\ nm}$ between 0.5 and 0.7) in MOKA rich medium. For each strain, RNA was extracted from at least three independent replicates of both the wild-type or empty vector conjugates and the HrpG* conjugates as described (22). Specific *Xanthomonas* probes were employed to deplete rRNA and tRNA molecules (16) after which RNAs were fractionated into short (<200 nt) and long RNAs using Zymo Research RNA Clean & Concentrator TM-5 columns (Proteigene). Strand-specific RNA sequencing was performed on an Ilumina HiSeq2000 platform, as described (22). A detailed overview of all generated and used RNA-Seq libraries and their NCBI Sequence Read Archive (SRA) accessions is provided in Table S1E.

Structural annotation of *Xanthomonas* genomes

For each strain, the different RNA-Seq libraries available were merged in order to experimentally support the structural annotation of the genomes using Eugene-PP (70). The used RNA-Seq libraries are detailed in Table S1. The annotations used and generated in this study are provided in Data S1 (https://doi.org/10.57745/9OTYNJ).

Differential expression analyses of the RNA-Seq results

RNA-Seq reads were pseudo mapped onto the reference genomes using Salmon (version 1.4.0) with standard parameters (71). An overview of which RNA-Seq datasets were used is given in Table S1E. Overall mapping quality was assessed using multiQC (72). DEGs were identified using DESeq2 with Salmon's count tables using standard parameters (73). To assess sample reproducibility, PCA and MA plots were generated using DESeq2's plotPCA and plotMA functions. Volcano plots for each strain were made using EnhancedVolcano (74). The PCA on the expression of all orthogroups in the core genome across all strains was based on a matrix with a single average Log_2FC value for all genes within an orthogroup. Genes with non-significant Log_2FC were assigned a Log_2FC of 0. Session information, appropriate raw data and relevant R scripts to reproduce the results are available in a code repository (Data S5, https://doi.org/10.57745/9OTYNJ).

Identification of the core HrpG* regulon

An orthogroup database constructed using Orthofinder (version 2.4.0) with standard parameters (75) was used to compare the regulation of genes by HrpG* across different strains. The species trees in Figure 1 were made using Orthofinder's STAG algorithm (76) and rooted using STRIDE (77). Orthogroups, for which across all strains, at least one ortholog was differentially regulated in the wild-type or empty vector strains compared to the *hrpG** strains (AdjPval <0.05) were considered orthogroups part of the core regulon. To investigate the regulation of the endogenous *hrpG* by HrpG*, we investigated whether reads mapping to *hrpG* originated from the endogenous *hrpG* gene or the exogenous *hrpG** gene by mapping RNA-Seq reads to both endogenous and HrpG* sequences. Reads that did not map to these two sequences or which mapped multiple times were then removed. The coverage at the polymorphic site of *hrpG/hrpG** for the remaining reads was then assessed. A Chi-square test was used to test for equal proportions of HrpG*-dependent regulation of predicted protein-coding genes within the core and accessory genomes. Multiple testing correction was performed using the Benjamini–Hochberg procedure (78).

Conservation of the core HrpG regulon across all publically available genomes

To investigate the conservation of orthogroups identified in this study across plant pathogenic *Xanthomonas* strains with a T3SS, an additional orthogroup database was built using publically available genomes of *Xanthomonas* species with a T3SS (Data S6, https://doi.org/10.57745/9OTYNJ). This database was built on available genomes having less than 206 contigs and a Busco score of over 94.9% (657 genomes in total representing 25 *Xanthomonas* species and at least 67 different pathovars, Data S6). The conservation of relevant orthogroups identified in this study was studied for each corresponding orthogroup in the orthology analysis constructed with the 657 genomes.

Gene ontology and signal peptide prediction

Enrichment of specific GO terms in various gene lists was investigated using the TOPGO R package (79). Signal peptides were predicted by signalP6.0 (80). A Chi-square test was used to test for an enrichment of genes with Sec/SPI amongst regulated genes. When appropriate, promoter sequences were scanned for PIP-box motifs using the TTCGB-N_{15}-TTCGB consensus (3) allowing for one mismatch.

Regulation of type three effector gene expression

T3Es were predicted using Effectidor (25) using the advanced mode, which makes use of the genomes GFF3 files to more accurately predict T3Es. Each orthogroup of which at least one ortholog was annotated as a true T3E by Effectidor was considered a true T3E orthogroup. True T3E orthogroups were further manually annotated and curated by integrating BLASTP, paperblast (81) and the EUROXANTH database (42). Because genes encoding TALEs are often poorly assembled due to their repetitive sequences (82), only TALE orthologs for which the complete N-terminal domain upstream of the repeat region was present were considered.

Regulation of genes relevant for motility, carbohydrate active enzymes and proteases

Orthogroups involved in motility were identified using the Eugene-PP genome annotations and manually curated. Orthogroups encoding proteins with carbohydrate active enzyme activity were predicted using dbCAN3 (83) and further curated and annotated using the Carbohydrate-Active enZYmes (CAZY) Database (84). Genes encoding enzymes with protease activity were identified using the Eugene-PP genome annotations and manually curated and annotated using the MEROPS database (85).

Visualization of regulation against phylogeny

Several figures were generated in order to explore a potential relationship between strain phylogeny and the regulation of specific orthogroups under HrpG*. The HrpG*-dependent Log_2FC of expression all genes within an orthogroup were plotted against the phylogenetic tree as made by Orthofinder. The Log_2FC of genes for which either Log_2FC or AdjPval were assigned NA values by DESeq were set to 0. Figures were built using GGtree with equal branch length for the species tree (86).

Availability of data and materials

All genomes sequenced and assembled in this study have been made publicly accessible. For a comprehensive list of their DOIs, please see Table S1. The transcriptome sequencing reads are available in the NCBI Sequence Read Archive (SRA), with the specific SRA accession numbers provided in Table S1. Supplementary data files can be found on recherche.data.gouv.fr under the DOI: https://doi.org/10.57745/9OTYNJ. This includes generated genome annotation for the 17 strains as Data S1, and comparative statistics and orthogroup analyses for these genomes, performed using OrthoFinder, as Data S2. R scripts utilized for analyzing the data, are stored in a code repository in Data S5. Comparative statistics and orthogroup information for 675 publicly available *Xanthomonas* genomes which include a type T3SS are detailed in Data S6, along with their respective genome accession numbers.

ACKNOWLEDGEMENTS

We wish to thank Claudine Zischek for the construction of *Xcc* strains, Aude Cerutti for helpful advice on ribodepletion, Caroline Bellenot for sharing curated *Xanthomonas* genomics resources and Naama Waagner for running Effectidor analyses.

DATA AVAILABILITY

All genomes sequenced and assembled in this study have been made publicly accessible at bbric.toulouse.inra.fr. For a comprehensive list of their DOIs, please see Additional File 1. The transcriptome sequencing reads are available in the NCBI Sequence Read Archive (SRA), with the specific SRA accession numbers provided in Additional File 1. Supplementary data files can be found on data.gouv.fr under the DOI: https://doi.org/10.57745/9OTYNJ. This includes generated genome annotation for the 17 strains as Data S1, and comparative statistics and orthogroup analyses for these genomes, performed using OrthoFinder, as Data S2. R scripts utilized for analyzing the data, are stored in a code repository in Data S5. Comparative statistics and orthogroup information for 675 publicly available *Xanthomonas* genomes which include a type T3SS are detailed in Data S6, along with their respective genome accession numbers.

FUNDINGS

This work was supported by grants from the “Agence Nationale de la Recherche” projects XANTHOMIX (ANR-2010-GENM-013-02 to BR, SC, EC, SC, EL, AD, MA, BS, OP, MAJ, LG, RK and LDN) and XBOX (ANR-19-CE20-JCJC-0014-01 to TQM, LDN and AB). All co-authors are members of the French Network on Xanthomonads (FNX) supported by the INRAE Plant Health division. This study is set within the framework of the “Laboratoires d’Excellences” (LABEX) TULIP (ANR-10-LABX-41) and of the “Ecole Universitaire de Recherche” (EUR) TULIP-GS (ANR-18-EURE-0019).

AUTHORS CONTRIBUTION

The initial XANTHOMIX proposal was written under coordination of MA, OP, MAJ, LDN and RK. BR, SCU, BS, EC, AD, LDN, EL and LG were involved in the sequencing of four *Xanthomonas* genomes, the construction of the different *Xanthomonas* conjugates and the preparation and sequencing of cDNAs. SCA handled genome annotations, transcriptomic pseudomapping, and Orthofinder analyses. MFJ performed the initial transcriptomic analyses. TQM performed quality controls on the raw data, conducted the in-depth comparative transcriptomic analyses, prepared all the figures and drafted the manuscript. AB and LDN supervised the analyses, interpreted of the data and wrote the manuscript. All authors read, commented and approved the final manuscript.

COMPETING INTEREST

The authors declare that they have no competing interests.

REFERENCES

1. Wengelnik K. 1996. HrpG, a Key *hrp* Regulatory Protein of *Xanthomonas campestris* pv. *vesicatoria* Is Homologous to Two-Component Response Regulators. MPMI 9:704.

2. Stock AM, Robinson VL, Goudreau PN. 2000. Two-Component Signal Transduction. Annu Rev Biochem 69:183–215.

3. Koebnik R, Krüger A, Thieme F, Urban A, Bonas U. 2006. Specific binding of the *Xanthomonas campestris* pv. *vesicatoria* AraC-type transcriptional activator HrpX to plant-inducible promoter boxes. Journal of bacteriology 188:7652–7660.

4. Büttner D, Bonas U. 2010. Regulation and secretion of *Xanthomonas* virulence factors. FEMS microbiology reviews 34:107–133.

5. Tampakaki AP, Skandalis N, Gazi AD, Bastaki MN, Sarris PF, Charova SN, Kokkinidis M, Panopoulos NJ. 2010. Playing the “harp”: Evolution of our understanding of *hrp*/*hrc* Genes. Annual Review of Phytopathology 48:347-70.

6. White FF, Potnis N, Jones JB, Koebnik R. 2009. The type III effectors of *Xanthomonas*. Molecular Plant Pathology 10:749–766.

7. Parte AC, Carbasse JS, Meier-Kolthoff JP, Reimer LC, Göker M. 2020. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. Int J Syst Evol Microbiol 70:5607–5612.

8. Ryan RP, Vorhölter F-J, Potnis N, Jones JB, Van Sluys M-A, Bogdanove AJ, Dow JM. 2011. Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. Nature Reviews Microbiology 9:344–355.

9. Young JM, Park D-C, Shearman HM, Fargier E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. Systematic and Applied Microbiology 31:366–377.

10. Zaumeyer WJ, Thomas HR, Zaumeyer WJ, Thomas HR. 1957. A Monographic Study of Bean Diseases and Methods for Their Control. https://doi.org/10.22004/AG.ECON.169625.

11. Jacques M-A, Josi K, Darrasse A, Samson R. 2005. *Xanthomonas axonopodis* pv. *phaseoli* var. fuscans Is Aggregated in Stable Biofilm Population Sizes in the Phyllosphere of Field-Grown Beans. Appl Environ Microbiol 71:2008–2015.

12. De O. Carvalho A, Da Cunha M, Rodrigues R, Sudré CP, Santos IS, Fernandes KVS, Rabelo GR, Gomes VM. 2011. Ultrastructural changes during early infection of Vigna unguiculata and Phaseolus vulgaris leaves by *Xanthomonas axonopodis* pv. *phaseoli* and

an unexpected association between chloroplast and mitochondrion. Acta Physiol Plant 33:2025–2033.

13. Almeida IMGD, Rodrigues LMR, Beriam LOS. 2015. *Xanthomonas fuscans subsp. fuscans* causing wilt symptoms in bean plants (Phaseolus vulgaris) in Brazil. Arq Inst Biol 82.

14. Guo Y, Figueiredo F, Jones J, Wang N. 2011. HrpG and HrpX play global roles in coordinating different virulence traits of *Xanthomonas axonopodis* pv. *citri*. Molecular Plant-Microbe Interactions 24(6):649-61.

15. Noel L, Thieme F, Nennstiel D, Bonas U. 2001. cDNA-AFLP analysis unravels a genome-wide hrpG-regulon in the plant pathogen *Xanthomonas campestris* pv. *vesicatoria*. Mol Microbiol 41:1271–1281.

16. Roux B, Bolot S, Guy E, Denancé N, Lautier M, Jardinaud M-F, Fischer-Le Saux M, Portier P, Jacques M-A, Gagnevin L, Pruvost O, Lauber E, Arlat M, Carrère S, Koebnik R, Noël LD. 2015. Genomics and transcriptomics of *Xanthomonas campestris* species challenge the concept of core type III effectome. BMC Genomics 16, 975.

17. Wei K, Tang D-J, He Y-Q, Feng J-X, Jiang B-L, Lu G-T, Chen B, Tang J-L. 2007. *hpaR* , a Putative *marR* Family Transcriptional Regulator, Is Positively Controlled by HrpG and HrpX and Involved in the Pathogenesis, Hypersensitive Response, and Extracellular Protease Production of *Xanthomonas campestris* Pathovar campestris. J Bacteriol 189:2055–2062.

18. Wang L, Rong W, He C. 2008. Two *Xanthomonas* Extracellular Polygalacturonases, PghAxc and PghBxc, Are Regulated by Type III Secretion Regulators HrpX and HrpG and Are Required for Virulence. MPMI 21:555–563.

19. Szczesny R, Jordan M, Schramm C, Schulz S, Cogez V, Bonas U, Büttner D. 2010. Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv *vesicatoria*. New Phytologist 187:983–1002.

20. Kogenaru S, Yan Q, Guo Y, Wang N. 2012. RNA-seq and microarray complement each other in transcriptome profiling. BMC Genomics 13:629.

21. Wengelnik K, Bonas U. 1996. HrpXv, an AraC-type regulator, activates expression of five of the six loci in the *hrp* cluster of *Xanthomonas campestris* pv. *vesicatoria*. Journal of Bacteriology 178(12):3462-9.

22. Luneau JS, Cerutti A, Roux B, Carrère S, Jardinaud MF, Gaillac A, Gris C, Lauber E, Berthomé R, Arlat M, Boulanger A, Noël LD. 2022. *Xanthomonas* transcriptome inside cauliflower hydathodes reveals bacterial virulence strategies and physiological adaptations at early infection stages. Molecular Plant Pathology 23(2):159-174.

23. Merda D, Briand M, Bosis E, Rousseau C, Portier P, Barret M, Jacques MA, Fischer-Le Saux M. 2017. Ancestral acquisitions, gene flow and multiple evolutionary trajectories of the type three secretion system and effectors in *Xanthomonas* plant pathogens. Molecular Ecology 26:5939–5952.

24. Assis RDAB, Polloni LC, Patané JSL, Thakur S, Felestrino ÉB, Diaz-Caballero J, Digiampietri LA, Goulart LR, Almeida NF, Nascimento R, Dandekar AM, Zaini PA, Setubal JC, Guttman DS, Moreira LM. 2017. Identification and analysis of seven effector protein families with different adaptive and evolutionary histories in plant-associated members of the *Xanthomonadaceae*. Sci Rep 7:16133.

25. Wagner N, Avram O, Gold-Binshtok D, Zerah B, Teper D, Pupko T. 2022. Effectidor: an automated machine-learning-based web server for the prediction of type-III secretion system effectors. Bioinformatics 38:2341–2343.

26. Alvarez-Martinez CE, Sgro GG, Araujo GG, Paiva MRN, Matsuyama BY, Guzzo CR, Andrade MO, Farah CS. 2021. Secrete or perish: The role of secretion systems in *Xanthomonas* biology. Computational and Structural Biotechnology Journal 19:279–302.

27. Yamazaki A, Hirata H, Tsuyumu S. 2008. HrpG regulates type II secretory proteins in *Xanthomonas axonopodis* pv. *citri*. J Gen Plant Pathol 74:138–150.

28. Furutani A, Takaoka M, Sanada H, Noguchi Y, Oku T, Tsuno K, Ochiai H, Tsuge S. 2009. Identification of novel type III secretion effectors in *Xanthomonas oryzae* pv. *oryzae*. Molecular Plant-Microbe Interactions 22(1):96-106.

29. Teper D, Pandey SS, Wang N. 2021. The HrpG/HrpX Regulon of *Xanthomonads*—An Insight to the Complexity of Regulation of Virulence Traits in Phytopathogenic Bacteria. Microorganisms 9:187.

30. Wang J-Y, Zhou L, Chen B, Sun S, Zhang W, Li M, Tang H, Jiang B-L, Tang J-L, He Y-W. 2015. A functional 4-hydroxybenzoate degradation pathway in the phytopathogen *Xanthomonas campestris* is required for full pathogenicity. Sci Rep 5:18456.

31. Chen B, Li R, Zhou L, Qiu J, Song K, Tang J, He Y. 2020. The phytopathogen *Xanthomonas campestris* utilizes the divergently transcribed *pobA* / *pobR* locus for 4-hydroxybenzoic acid recognition and degradation to promote virulence. Molecular Microbiology 114:870–886.

32. Parke D. 1996. Characterization of PcaQ, a LysR-type transcriptional activator required for catabolism of phenolic compounds, from *Agrobacterium tumefaciens*. J Bacteriol 178:266–272.

33. MacLean AM, Haerty W, Golding GB, Finan TM. 2011. The LysR-type PcaQ protein regulates expression of a protocatechuate-inducible ABC-type transport system in *Sinorhizobium meliloti*. Microbiology 157:2522–2533.

34. Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, Vasse J, Lauber E, Arlat M. 2007. Plant Carbohydrate Scavenging through TonB-Dependent Receptors: A Feature Shared by Phytopathogenic and Aquatic Bacteria. PLoS ONE 2:e224.

35. Hederstedt L. 2022. Diversity of Cytochrome c Oxidase Assembly Proteins in Bacteria. Microorganisms 10:926.

36. Wu J, Pan X, Xu S, Duan Y, Luo J, Zhou Z, Wang J, Zhou M. 2019. The critical role of cytochrome c maturation (CCM) system in the tolerance of *Xanthomonas campestris* pv. *campestris* to phenazines. Pesticide Biochemistry and Physiology 156:63–71.

37. Chen L, Wang M, Huang L, Zhang Z, Liu F, Lu G. 2017. XC_0531 encodes a c-type cytochrome biogenesis protein and is required for pathogenesis in *Xanthomonas campestris* pv. *campestris*. BMC Microbiol 17:142.

38. Pena MM, Bhandari R, Bowers RM, Weis K, Newberry E, Wagner N, Pupko T, Jones JB, Woyke T, Vinatzer BA, Jacques M-A, Potnis N. 2024. Genetic and Functional Diversity Help Explain Pathogenic, Weakly Pathogenic, and Commensal Lifestyles in the Genus *Xanthomonas*. Genome Biology and Evolution 16:evae074.

39. Tamir-Ariel D, Navon N, Burdman S. 2007. Identification of Genes in *Xanthomonas campestris* pv. *vesicatoria* Induced during Its Interaction with Tomato. J Bacteriol 189:6359–6371.

40. Qian W, Jia Y, Ren S-X, He Y-Q, Feng J-X, Lu L-F, Sun Q, Ying G, Tang D-J, Tang H, Wu W, Hao P, Wang L, Jiang B-L, Zeng S, Gu W-Y, Lu G, Rong L, Tian Y, Yao Z, Fu G, Chen B, Fang R, Qiang B, Chen Z, Zhao G-P, Tang J-L, He C. 2005. Comparative and functional genomic analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. *campestris*. Genome Res 15:757–767.

41. Luneau JS, Baudin M, Quiroz Monnens T, Carrère S, Bouchez O, Jardinaud M-F, Gris C, François J, Ray J, Torralba B, Arlat M, Lewis JD, Lauber E, Deutschbauer AM, Noël LD, Boulanger A. 2022. Genome-wide identification of fitness determinants in the *Xanthomonas campestris* bacterial pathogen during early stages of plant infection. New Phytologist 236:235–248

42. Costa J, Pothier JF, Bosis E, Boch J, Kölliker R, Koebnik R. 2024. A Community-Curated DokuWiki Resource on Diagnostics, Diversity, Pathogenicity, and Genetic Control of *Xanthomonads*. MPMI 37:347–353.

43. Jiang W, Jiang BL, Xu RQ, Huang JD, Wei HY, Jiang GF, Cen WJ, Liu J, Ge YY, Li GH, Su LL, Hang XH, Tang DJ, Lu GT, Feng JX, He YQ, Tang JL. 2009. Identification of six type III effector genes with the pip box in *Xanthomonas campestris* pv. *campestris* and five of them contribute individually to full pathogenicity. Molecular Plant-Microbe Interactions 22:1401–11.

44. Popov G, Fraiture M, Brunner F, Sessa G. 2016. Multiple *Xanthomonas euvesicatoria* Type III Effectors Inhibit flg22-Triggered Immunity. MPMI 29:651–660.

45. Soni M, Mondal KK. 2018. *Xanthomonas axonopodis* pv. *punicae* uses XopL effector to suppress pomegranate immunity: XopL effector of Xap suppresses pomegranate immunity. J Integr Plant Biol 60:341–357.

46. Leong JX, Raffeiner M, Spinti D, Langin G, Franz-Wachtel M, Guzman AR, Kim J, Pandey P, Minina AE, Macek B, Hafrén A, Bozkurt TO, Mudgett MB, Börnke F, Hofius D, Üstün S. 2022. A bacterial effector counteracts host autophagy by promoting degradation of an autophagy component. The EMBO Journal 41:e110352.

47. Ortmann S, Marx J, Lampe C, Handrick V, Ehnert T-M, Zinecker S, Reimers M, Bonas U, Erickson JL. 2023. A conserved microtubule-binding region in *Xanthomonas* XopL is indispensable for induced plant cell death reactions. PLoS Pathog 19:e1011263.

48. Akimoto-Tomiyama C, Furutani A, Tsuge S, Washington EJ, Nishizawa Y, Minami E, Ochiai H. 2012. XopR, a Type III Effector Secreted by *Xanthomonas oryzae* pv. *oryzae* , Suppresses Microbe-Associated Molecular Pattern-Triggered Immunity in *Arabidopsis thaliana*. MPMI 25:505–514.

49. Wang S, Sun J, Fan F, Tan Z, Zou Y, Lu D. 2016. A *Xanthomonas oryzae* pv. *oryzae* effector, XopR, associates with receptor-like cytoplasmic kinases and suppresses PAMP-triggered stomatal closure. Sci China Life Sci 59:897–905.

50. Verma G, Mondal KK, Kulshreshtha A, Sharma M. 2019. XopR T3SS-effector of *Xanthomonas oryzae* pv. *oryzae* suppresses cell death-mediated plant defense response during bacterial blight development in rice. 3 Biotech 9:272.

51. Medina CA, Reyes PA, Trujillo CA, Gonzalez JL, Bejarano DA, Montenegro NA, Jacobs JM, Joe A, Restrepo S, Alfano JR, Bernal A. 2018. The role of type III effectors

from *Xanthomonas axonopodis* pv. *manihotis* in virulence and suppression of plant immunity: Role of type III effectors from *Xam*. Molecular Plant Pathology 19:593–606.

52. Degrassi G, Devescovi G, Bigirimana J, Venturi V. 2010. *Xanthomonas oryzae* pv. *oryzae* XKK.12 Contains an AroQ $_{\gamma}$ Chorismate Mutase That Is Involved in Rice Virulence. Phytopathology® 100:262–270.

53. Tamir-Ariel D, Rosenberg T, Navon N, Burdman S. 2012. A secreted lipolytic enzyme from *Xanthomonas campestris* pv. *vesicatoria* is expressed in planta and contributes to its virulence: *Xanthomonas campestris* pv. *vesicatoria* LipA. Molecular Plant Pathology 13:556–567.

54. Jha G, Rajeshwari R, Sonti RV. 2007. Functional Interplay Between Two *Xanthomonas oryzae* pv. *oryzae* Secretion Systems in Modulating Virulence on Rice. MPMI 20:31–40.

55. Nascimento R, Gouran H, Chakraborty S, Gillespie HW, Almeida-Souza HO, Tu A, Rao BJ, Feldstein PA, Bruening G, Goulart LR, Dandekar AM. 2016. The Type II Secreted Lipase/Esterase LesA is a Key Virulence Factor Required for *Xylella fastidiosa* Pathogenesis in Grapevines. Sci Rep 6:18598.

56. Laia ML, Moreira LM, Dezajacomo J, Brigati JB, Ferreira CB, Ferro MI, Silva AC, Ferro JA, Oliveira JC. 2009. New genes of *Xanthomonas citri* subsp. *citri* involved in pathogenesis and adaptation revealed by a transposon-based mutant library. BMC Microbiol 9:12.

57. Das A, Rangaraj N, Sonti RV. 2009. Multiple adhesin-like functions of *Xanthomonas oryzae* pv. *oryzae* are involved in promoting leaf attachment, entry, and virulence on rice. Molecular Plant-Microbe Interactions 22:73–85.

58. Pradhan BB, Ranjan M, Chatterjee S. 2012. XadM, a novel adhesin of *Xanthomonas oryzae* pv. *oryzae,* exhibits similarity to Rhs family proteins and is required for optimum attachment, biofilm formation, and virulence. Molecular Plant-Microbe Interactions 25:1157–1170.

59. Kumar Verma R, Samal B, Chatterjee S. 2018. *Xanthomonas oryzae* pv. *oryzae* chemotaxis components and chemoreceptor Mcp2 are involved in the sensing of constituents of xylem sap and contribute to the regulation of virulence-associated functions and entry into rice. Molecular Plant Pathology 19:2397–2415.

60. Li RF, Ren PD, Liu QQ, Yao JL, Wu L, Zhu GN, Xian XY, Tang JL, Lu GT. 2022. McvR, a single domain response regulator regulates motility and virulence in the plant pathogen *Xanthomonas campestris*. Molecular Plant Pathology 23:649–663.

61. Zhang Y, Callaway EM, Jones JB, Wilson M. 2009. Visualisation of *hrp* gene expression in *Xanthomonas euvesicatoria* in the tomato phyllosphere. Eur J Plant Pathol 124:379–390.

62. Mayjonade B, Gouzy J, Donnadieu C, Pouilly N, Marande W, Callot C, Langlade N, Muños S. 2016. Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. BioTechniques 61:203–205.

63. Bellenot C, Carrère S, Gris C, Noël LD, Arlat M. 2022. Genome Sequences of 17 Strains from Eight Races of *Xanthomonas campestris* pv. *campestris*. Microbiol Resour Announc 11:e00279-22.

64. Denancé N, Szurek B, Doyle EL, Lauber E, Fontaine-Bodin L, Carrère S, Guy E, Hajri A, Cerutti A, Boureau T, Poussier S, Arlat M, Bogdanove AJ, Noël LD. 2018. Two

ancestral genes shaped the *Xanthomonas campestris* TAL effector gene repertoire. New Phytol 219:391–407.

65. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546.

66. Guy E, Genissel A, Hajri A, Chabannes M, David P, Carrere S, Lautier M, Roux B, Boureau T, Arlat M, Poussier S, Noël LD. 2013. Natural Genetic Variation of *Xanthomonas campestris* pv. *campestris* Pathogenicity on *Arabidopsis* Revealed by Association and Reverse Genetics. mBio 4:e00538-12.

67. Kovach ME, Elzer PH, Steven Hill D, Robertson GT, Farris MA, Roop RM, Peterson KM. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene 166:175–176.

68. Figurski DH, Helinski DR. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc Natl Acad Sci USA 76:1648–1652.

69. Ditta G, Stanfield S, Corbin D, Helinski DR. 1980. Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc Natl Acad Sci U S A 77:7347–7351.

70. Sallet E, Roux B, Sauviac L, Jardinaud M-F, Carrere S, Faraut T, de Carvalho-Niebel F, Gouzy J, Gamas P, Capela D, Bruand C, Schiex T. 2013. Next-Generation Annotation of Prokaryotic Genomes with EuGene-P: Application to *Sinorhizobium meliloti* 2011. DNA Research 20:339–354.

71. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods 14:417–419.

72. Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048.

73. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550.

74. Blighe K. 2018. EnhancedVolcano. Bioconductor.

75. Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology 20:238.

76. Emms DM, Kelly S. 2018. STAG: Species Tree Inference from All Genes. bioRxiv 267914. Evolutionary Biology.

77. Emms DM, Kelly S. 2017. STRIDE: Species Tree Root Inference from Gene Duplication Events. Molecular Biology and Evolution 34:3267–3278.

78. Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 57:289–300.

79. Alexa A, Rahnenfuhrer J. 2010. topGO: enrichment analysis for gene ontology. R package version 2:2010.

80. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, Von Heijne G, Nielsen H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40:1023–1025.

81. Price MN, Arkin AP. 2017. PaperBLAST: Text Mining Papers for Information about Homologs. mSystems 2:e00039-17.

82. Erkes A, Grove RP, Žarković M, Krautwurst S, Koebnik R, Morgan RD, Wilson GG, Hölzer M, Marz M, Boch J, Grau J. 2023. Assembling highly repetitive *Xanthomonas* TALomes using Oxford Nanopore sequencing. BMC Genomics 24:151.

83. Zheng J, Ge Q, Yan Y, Zhang X, Huang L, Yin Y. 2023. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. Nucleic Acids Research 51:W115–W121.

84. Drula E, Garron M-L, Dogan S, Lombard V, Henrissat B, Terrapon N. 2022. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Research 50:D571–D577.

85. Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Research 46:D624–D632.

86. Yu G. 2020. Using ggtree to Visualize Data on Tree-Like Structures. Current Protocols in Bioinformatics 69:e96.

FIGURES LEGENDS

Figure 1. Gene orthology analysis of the 17 *Xanthomonas* strains used in this study. A: phylogenetic tree of all strains, each represented with a distinct coloured symbol. **B:** Lifestyle

indicates if the pathogen infects primarily leaf mesophyll (light brown), the vascular tissues (dark brown) or both (burgundy). **C:** Number of predicted protein-coding genes per genome, including those in orthogroups (dark green) and singletons (light green). **D:** Number of orthogroups per genome. **E:** Heatmap showing the number of shared orthogroups between strains. Symbols as in panel **A**. **F:** Principal component (PC) analysis plot showing the clustering of the different strains based on the presence or absence of orthogroups against the first three PCs. Shape and color of symbols identifies each of the 17 strains from the nine pathovars studied as in panel **A**.

Figure 2. Features of the HrpG* regulons in 17 *Xanthomonas* strains. **A:** Phylogenetic tree of all strains as shown in Fig. 1A. **B:** Log_2FC of *hrpX* in *hrpG** strains compared to wild-type. **C:** Proportion of differentially-expressed predicted protein-coding genes per genome (AdjPval < 0.05). **D:** Number of DEGs ($|Log_2FC| > 2$). **E:** Percentage of DEGs ($|Log_2FC| > 2$) in the core genome only considering protein-coding genes. **F:** Percentage of DEGs ($|Log_2FC| > 2$) in the accessory genome only considering protein-coding genes. **G:** Principal components analysis plot showing the first three principal components based on the average Log_2FC of genes within each orthogroups of the core genome.

Figure 3. GO term enrichment in the HrpG* regulons in 17 *Xanthomonas* strains. Only GO terms enriched amongst regulated genes ($|Log_2FC| > 2$) identified in at least three strains are depicted in the table to highlight the main commonalities across the strains (62 GO terms). A total of 197 unique GO terms were found to be enriched in at least one strain as detailed (Table S2B).

Figure 4. Regulation of orthogroups (OG) encoding putative T3Es in *Xanthomonas* strains expressing *hrpG.** Only significantly DEGs (AdjPval < 0.05) are colored. Missing orthologs are marked in gray. Orthogroup annotation is given on top of the figure, orthogroup number at bottom of the figure. Asterisks indicate protein-coding genes which are singletons.

Figure 5. Regulation of orthogroups (OG) in chemotaxis and motility in *Xanthomonas* strains expressing *hrpG.** Only significantly DEGs (AdjPval < 0.05) are colored. Missing orthologs are marked in gray. Orthogroup annotation is given on top of the figure, orthogroup number at bottom of the figure. Due to the extensive number of orthogroups potentially involved in chemotaxis, only those orthogroups that are regulated by HrpG* in at least two strains are shown, except for the secondary flagellar cluster specific to *Xcm* strains. Additionally, for both the flagellar and pilus gene clusters, only the regulation of key representative genes that encode various structural components of the flagella and pili are shown. MCPs: methyl-accepting chemotaxis proteins.

Figure 6. Regulation of orthogroups (OG) encoding carbohydrate active enzymes in *Xanthomonas* strains expressing *hrpG.** Only genes with an AdjPval < 0.05 are colored. Absent orthologs are marked in gray. Orthogroup annotations are given on top of the panel according to their CAZy family. Orthogroup numbers are given at the bottom of the panel. Due to the extensive number of orthogroups encoding carbohydrate active enzymes, only orthogroups regulated in at least two strains are shown. The cross, star and triangle symbols indicate the presence of a signal peptide within the protein, as predicted by SignalP6.

Figure 7. Regulation of orthogroups (OG) encoding proteins with proteolytic activity in *Xanthomonas* strains expressing *hrpG.** Only genes with an AdjPval < 0.05 are colored.

Absent orthologs are marked in gray. Orthogroup annotation is given on top of the figure, orthogroup number is given at the bottom of the figure. Orthogroups are annotated according to their MEROPS identifiers. Due to the extensive number of orthogroups encoding proteins with proteolytic activity, only orthogroups regulated in at least two strains are shown. The cross, star and triangle symbols indicate the presence of a signal peptide within a gene, as predicted by SignalP6.

Figure 8. Differential expression of orthogroups (OG) encoding proteins putatively involved in plant phenolic compounds metabolism. Only genes with an AdjPval < 0.05 are colored. Absent orthologs are marked in gray. Orthogroup annotation is given on top of the panel, orthogroup number is given at the bottom of the figure. Pca: protocatechuate. Hp: hypothetical protein

Figure 9. Differential expression of orthogroups (OG) encoding proteins putatively involved in cytochrome C maturation. Only genes with an AdjPval < 0.05 are colored. Orthogroup annotation is given on top of the figure. Orthogroup number is given at the bottom of the figure.

Figure 10. Schematic representation of the key processes identified to be regulated by HrpG* in this study. Red symbols annotated in bold indicate processes that are part of the *Xanthomonas* core regulon. Other red symbols indicate terms part of the group-II *Xanthomonas* core regulon. Symbols in purple highlight other regulated orthogroups which have been discussed. Dashed lines indicate hypothesized interactions, whereas continuous lines indicate interactions which are thought to be direct. T3SS: type III secretion system;

T3Es: type III effectors; T2SS: Xps type II secretion system; MCPs: methyl-accepting chemotaxis proteins.

Table 1. ***Xanthomonas*** **HrpG* core regulon members**.

Gene annotation	Orthogroup number	Average $\log_2$FC[a]	SD Average $\log_2$FC
Type III secretion system			
hpaH/hpa2	OG0000337	6.8	1.7
hrcV	OG0000338	4.7	1.8
hrcQ	OG0000339	5.6	2
hrcC	OG0001932	6.6	1.5
hrcT/hrpB8	OG0001933	5.1	1.4
hrpB7	OG0001934	5.2	1.5
hrcN/hrpB6	OG0001935	5.4	1.5
hrcL/hrpB5	OG0001936	6.1	1.6
hrpB4	OG0001937	5.6	1.5
hrpB3	OG0001938	6.2	1.7
hrpB2	OG0001939	6.7	1.4
hrpB1/hrpK	OG0001940	7	1.4
hrcU	OG0001941	6.4	1.6
hpaP/hrpC3	OG0001942	5.5	1.4
hrcR/hrpD2	OG0001943	5.6	1.6
hrcS	OG0001944	5.5	1.6
hpaA	OG0001945	5.1	1.7
hrcD	OG0001946	6	1.6
hrpD6	OG0001947	5.9	1.5
hrpE	OG0001948	6.2	1.5
hpaB	OG0001949	5.8	1.9
Type III effectors			
xopL	OG0000114	4.5	1.4
xopR	OG0000535	3.4	1.1
Others			
phoC	OG0000410	3.3	1.6
MFS transporter	OG0001668	3.5	1.5
hrpX	OG0001966	4.6	1.2

[a] Of all shown orthogroups, at least one gene copy was significantly differentially regulated in all investigated strains (AdjPval < 0.05).

Table 2. **Group-II-specific *Xanthomonas* HrpG* core regulon identified in this study**.

Gene annotation	**Orthogroup number**	**Average $\log_2 FC^a$**	**SD Average $\log_2 FC$**	**Core genome**
Putative type III secretion system translocation proteins				
xopA/hpa1	OG0003083	8.7	1.2	No
hrpF	OG0003084	5.7	1.6	No
Degradation of plant phenolic compounds				
vanK/pcaK MFS transporter	OG0000558	3.5	1.4	Yes
PcaQ-like transcription factor	OG0003126	3.8	1.4	No
Pca dioxygenase-like	OG0003127	4	1.5	No
Pca dioxygenase-like	OG0003128	4.1	1.8	No
Others				
Cytochrome P450 hydroxylase	OG0000095	4.1	1.8	Yes
Putative HTH-type MarR transcription factor	OG0000153	2.7	1.3	Yes
virK	OG0000589	4.8	1.3	Yes
lipA/*lesK* secreted lipase	OG0000935	4.6	1.3	Yes
HpaR MarR transcription factor	OG0001822	2.9	1.2	Yes
Chorismate mutase	OG0002354	3.6	1.7	Yes

[a] Of all shown orthogroups, at least one gene copy was significantly differentially regulated in all investigated group-II strains (AdjPval < 0.05).

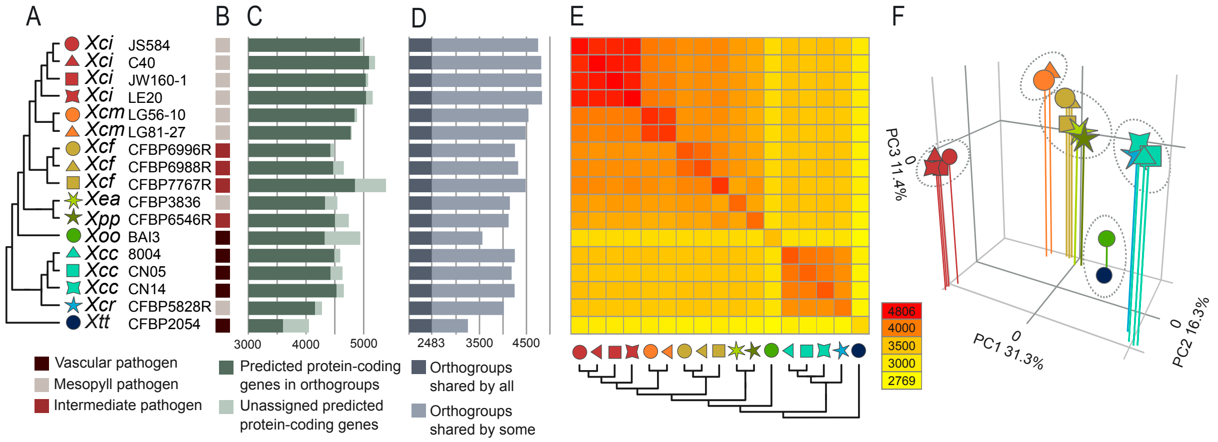


Figure 1. Gene orthology analysis of the 17 *Xanthomonas* strains used in this study. **A**: phylogenetic tree of all strains, each represented with a distinct coloured symbol. **B**: Lifestyle indicates if the pathogen infects primarily leaf mesophyll (light brown), the vascular tissues (dark brown) or both (burgundy). **C**: Number of predicted protein-coding genes per genome, including those in orthogroups (dark green) and singletons (light green). **D**: Number of orthogroups per genome. **E**: Heatmap showing the number of shared orthogroups between strains. Symbols as in panel A. **F**: Principal component (PC) analysis plot showing the clustering of the different strains based on the presence or absence of orthogroups against the first three PCs. Shape and color of symbols identifies each of the 17 strains from the nine pathovars studied as in panel A.

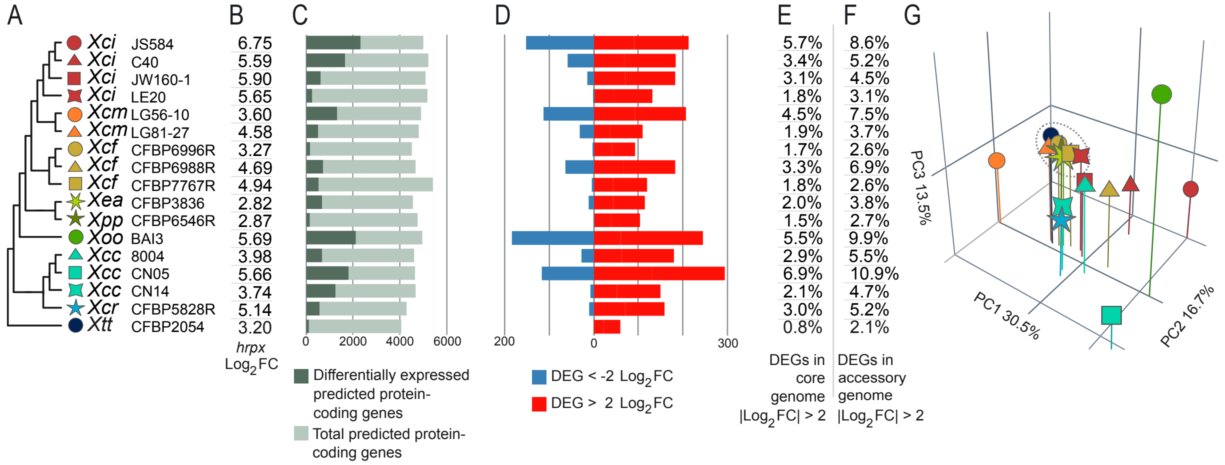
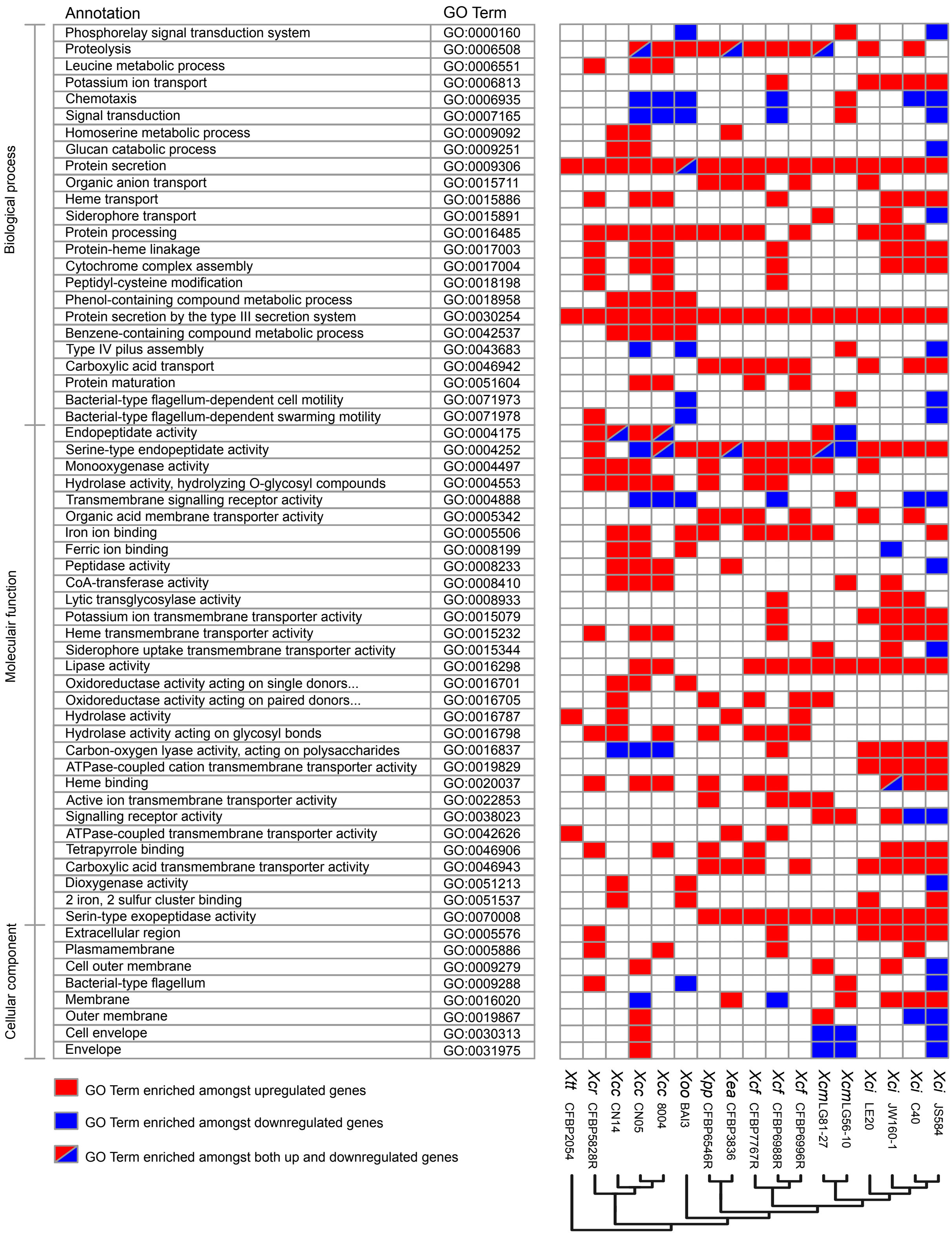


Figure 2. Features of the HrpG* regulons in 17 *Xanthomonas* strains. **A**: Phylogenetic tree of all strains as shown in Fig. 1A. **B**: Log_2FC of hrpX in hrpG* strains compared to wild-type. **C**: Proportion of differentially-expressed predicted protein-coding genes per genome (AdjPval < 0.05). **D**: Number of DEGs ($|Log_2FC| > 2$). **E**: Percentage of DEGs ($|Log_2FC| > 2$) in the core genome only considering protein-coding genes. **F**: Percentage of DEGs ($|Log_2FC| > 2$) in the accessory genome only considering protein-coding genes. **G**: Principal components analysis plot showing the first three principal components based on the average Log_2FC of genes within each orthogroups of the core genome.



	Annotation	GO Term	*Xtt* CFBP2054	*Xcr* CFBP5828R	*Xcc* CN14	*Xcc* CN05	*Xcc* 8004	*Xoo* BAI3	*Xpp* CFBP6546R	*Xea* CFBP3836	*Xcf* CFBP7767R	*Xcf* CFBP6988R	*Xcf* CFBP6996R	*Xcm* LG81-27	*Xcm* LG56-10	*Xci* LE20	*Xci* JW160-1	*Xci* C40	*Xci* JS584
Biological process	Phosphorelay signal transduction system	GO:0000160						🟦							🟥				🟦
	Proteolysis	GO:0006508				◩	🟥	🟥	🟥	◩	🟥	🟥	🟥	◩		🟥		🟥	
	Leucine metabolic process	GO:0006551		🟥		🟥	🟥												
	Potassium ion transport	GO:0006813										🟥				🟥	🟥	🟥	🟥
	Chemotaxis	GO:0006935				🟦	🟦	🟦				🟦			🟥			🟦	🟦
	Signal transduction	GO:0007165				🟦	🟦	🟦				🟦			🟥				🟦
	Homoserine metabolic process	GO:0009092			🟥	🟥				🟥									
	Glucan catabolic process	GO:0009251			🟥	🟥													🟦
	Protein secretion	GO:0009306	🟥	🟥	🟥	🟥	🟥	◩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Organic anion transport	GO:0015711							🟥	🟥	🟥		🟥			🟥			
	Heme transport	GO:0015886		🟥		🟥	🟥					🟥					🟥	🟥	🟥
	Siderophore transport	GO:0015891												🟥			🟥		🟦
	Protein processing	GO:0016485		🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥		🟥			🟥	🟥	🟥	
	Protein-heme linakage	GO:0017003		🟥		🟥	🟥					🟥					🟥	🟥	🟥
	Cytochrome complex assembly	GO:0017004		🟥		🟥	🟥					🟥					🟥	🟥	🟥
	Peptidyl-cysteine modification	GO:0018198		🟥			🟥					🟥							
	Phenol-containing compound metabolic process	GO:0018958			🟥	🟥	🟥	🟥											
	Protein secretion by the type III secretion system	GO:0030254	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Benzene-containing compound metabolic process	GO:0042537			🟥	🟥	🟥	🟥											
	Type IV pilus assembly	GO:0043683				🟦		🟦							🟥				🟦
	Carboxylic acid transport	GO:0046942							🟥	🟥	🟥	🟥	🟥			🟥		🟥	🟥
	Protein maturation	GO:0051604				🟥	🟥				🟥		🟥						
	Bacterial-type flagellum-dependent cell motility	GO:0071973						🟦							🟥				🟦
	Bacterial-type flagellum-dependent swarming motility	GO:0071978		🟥				🟦											🟦
Molecular function	Endopeptidate activity	GO:0004175		🟥	◩	🟥	◩							🟥	🟦				
	Serine-type endopeptidate activity	GO:0004252		🟥		🟦	◩	🟥	🟥	◩	🟥	🟥	🟥	◩	🟦	🟥	🟥	🟥	🟥
	Monooxygenase activity	GO:0004497		🟥	🟥	🟥			🟥		🟥	🟥	🟥	🟥		🟥			
	Hydrolase activity, hydrolyzing O-glycosyl compounds	GO:0004553		🟥	🟥	🟥	🟥		🟥		🟥	🟥							
	Transmembrane signalling receptor activity	GO:0004888				🟦	🟦	🟦				🟦			🟥			🟦	🟦
	Organic acid membrane transporter activity	GO:0005342							🟥	🟥	🟥		🟥			🟥		🟥	
	Iron ion binding	GO:0005506			🟥	🟥		🟥	🟥		🟥	🟥	🟥	🟥					🟥
	Ferric ion binding	GO:0008199			🟥	🟥		🟥									🟦		
	Peptidase activity	GO:0008233			🟥	🟥	🟥			🟥									🟦
	CoA-transferase activity	GO:0008410			🟥	🟥	🟥								🟥		🟥		
	Lytic transglycosylase activity	GO:0008933										🟥					🟥	🟥	
	Potassium ion transmembrane transporter activity	GO:0015079										🟥				🟥	🟥	🟥	🟥
	Heme transmembrane transporter activity	GO:0015232		🟥		🟥	🟥					🟥					🟥	🟥	🟥
	Siderophore uptake transmembrane transporter activity	GO:0015344												🟥			🟥		🟦
	Lipase activity	GO:0016298				🟥	🟥				🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Oxidoreductase activity acting on single donors...	GO:0016701			🟥	🟥		🟥											
	Oxidoreductase activity acting on paired donors...	GO:0016705			🟥				🟥		🟥		🟥	🟥					
	Hydrolase activity	GO:0016787	🟥		🟥					🟥			🟥						
	Hydrolase activity acting on glycosyl bonds	GO:0016798		🟥	🟥		🟥		🟥		🟥	🟥	🟥						
	Carbon-oxygen lyase activity, acting on polysaccharides	GO:0016837			🟦	🟦	🟦					🟥				🟥	🟥	🟥	🟥
	ATPase-coupled cation transmembrane transporter activity	GO:0019829														🟥	🟥	🟥	🟥
	Heme binding	GO:0020037		🟥		🟥	🟥		🟥		🟥	🟥					◩	🟥	🟥
	Active ion transmembrane transporter activity	GO:0022853							🟥			🟥	🟥	🟥					
	Signalling receptor activity	GO:0038023												🟥	🟥		🟥	🟦	🟦
	ATPase-coupled transmembrane transporter activity	GO:0042626	🟥							🟥		🟥							
	Tetrapyrrole binding	GO:0046906		🟥			🟥		🟥		🟥						🟥	🟥	🟥
	Carboxylic acid transmembrane transporter activity	GO:0046943							🟥	🟥	🟥		🟥			🟥	🟥	🟥	🟥
	Dioxygenase activity	GO:0051213			🟥			🟥											🟦
	2 iron, 2 sulfur cluster binding	GO:0051537			🟥			🟥								🟥			🟥
	Serin-type exopeptidase activity	GO:0070008							🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Cellular component	Extracellular region	GO:0005576		🟥								🟥				🟥	🟥	🟥	🟥
	Plasmamembrane	GO:0005886		🟥			🟥					🟥						🟥	
	Cell outer membrane	GO:0009279				🟥								🟥			🟥		🟦
	Bacterial-type flagellum	GO:0009288		🟥				🟦							🟥				🟦
	Membrane	GO:0016020				🟦				🟥		🟦			🟥		🟥	🟥	🟥
	Outer membrane	GO:0019867				🟥								🟥				🟦	🟦
	Cell envelope	GO:0030313				🟥								🟦	🟦				🟦
	Envelope	GO:0031975				🟥								🟦	🟦				🟦

🟥 GO Term enriched amongst upregulated genes

🟦 GO Term enriched amongst downregulated genes

◩ GO Term enriched amongst both up and downregulated genes

Figure 3. GO term enrichment in the HrpG* regulons in 17 *Xanthomonas* strains. Only GO terms enriched amongst regulated genes ($|Log_2FC| > 2$) identified in at least three strains are depicted in the table to highlight the main commonalities across the strains (62 GO terms). A total of 197 unique GO terms were found to be enriched in at least one strain as detailed (Table S2B).

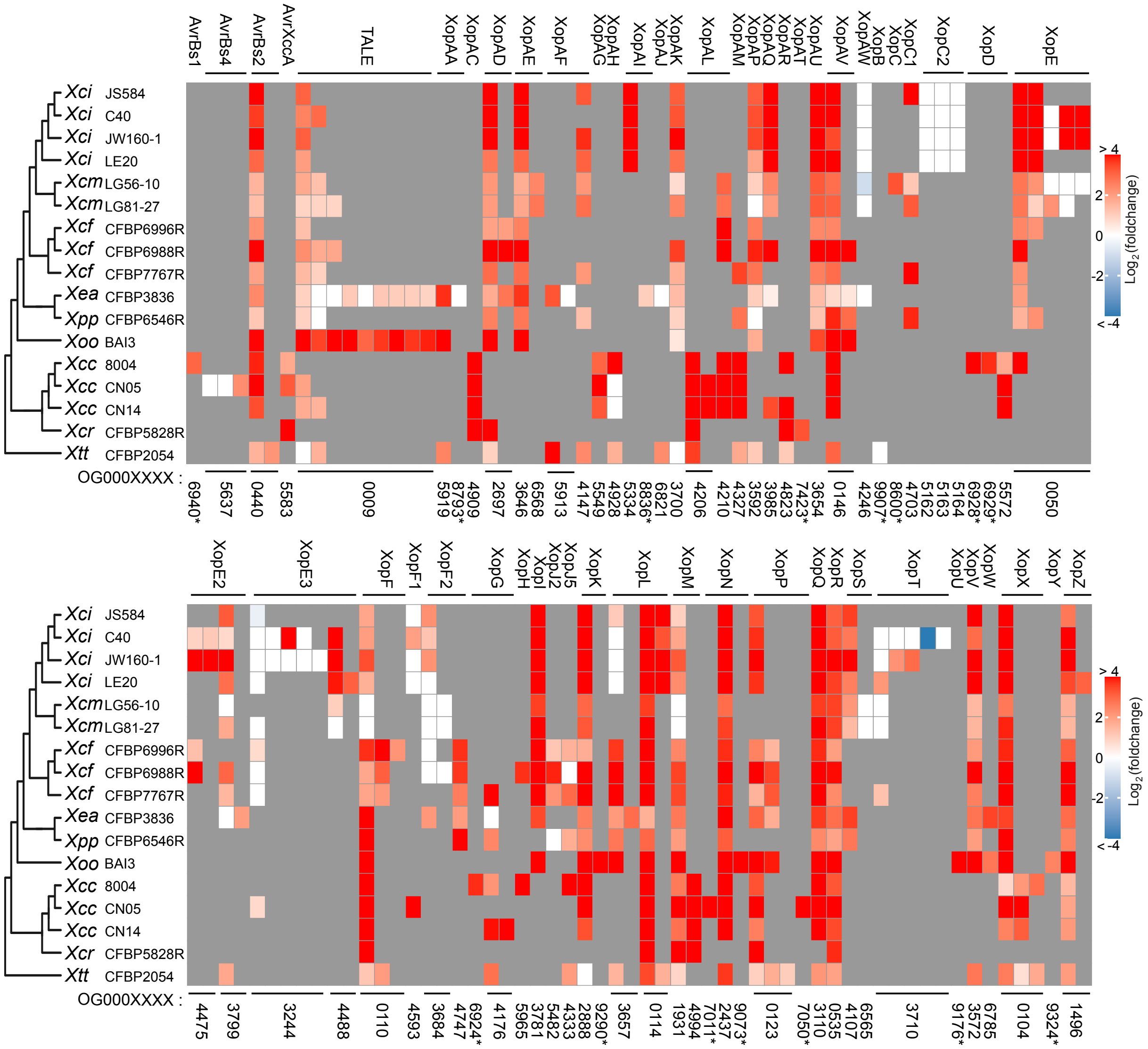


Figure 4. Regulation of orthogroups (OG) encoding putative T3Es in *Xanthomonas* strains expressing *hrpG**. Only significantly DEGs (AdjPval < 0.05) are colored. Missing orthologs are marked in gray. Orthogroup annotation is given on top of the figure, orthogroup number at bottom of the figure. Asterisks indicate protein-coding genes which are singletons.

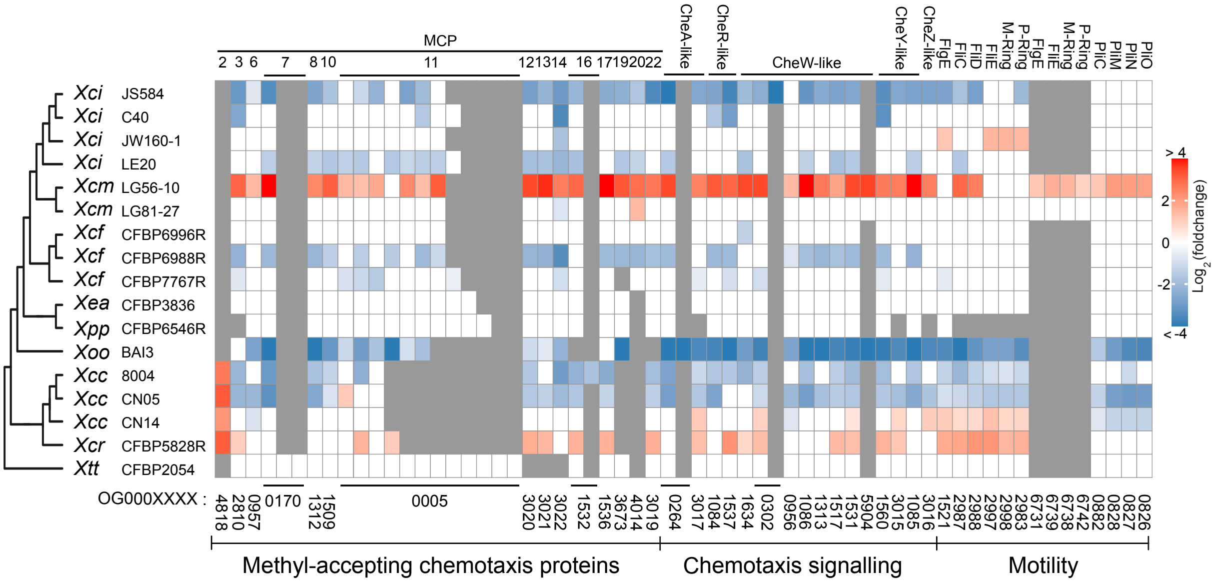


Figure 5. Regulation of orthogroups (OG) in chemotaxis and motility in *Xanthomonas* strains expressing *hrpG**. Only significantly DEGs (AdjPval < 0.05) are colored. Missing orthologs are marked in gray. Orthogroup annotation is given on top of the figure, orthogroup number at bottom of the figure. Due to the extensive number of orthogroups potentially involved in chemotaxis, only those orthogroups that are regulated by HrpG* in at least two strains are shown, except for the secondary flagellar cluster specific to

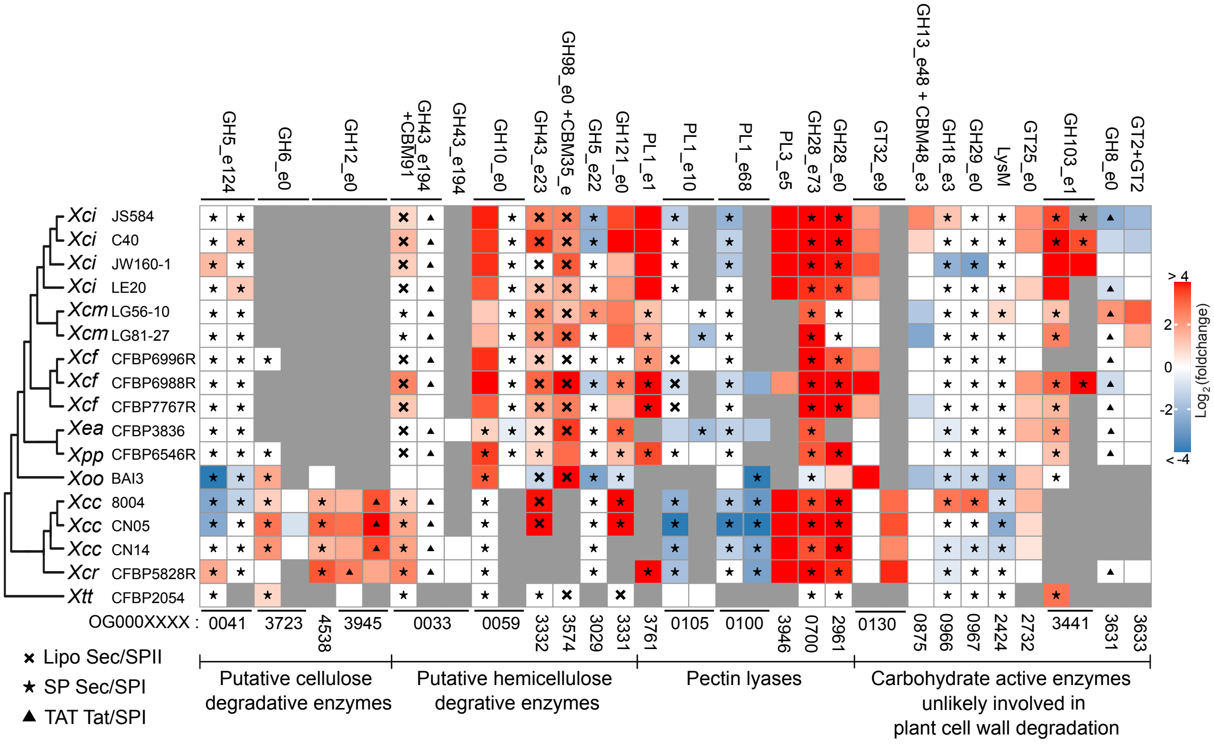


Figure 6. Regulation of orthogroups (OG) encoding carbohydrate active enzymes in *Xanthomonas* strains expressing *hrpG**. Only genes with an AdjPval < 0.05 are colored. Absent orthologs are marked in gray. Orthogroup annotations are given on top of the panel according to their CAZy family. Orthogroup numbers are given at the bottom of the panel. Due to the extensive number of orthogroups encoding carbohydrate active enzymes, only orthogroups regulated in at least two strains are shown. The cross, star and triangle symbols indicate the presence of a signal peptide within the protein, as predicted by SignalP6.

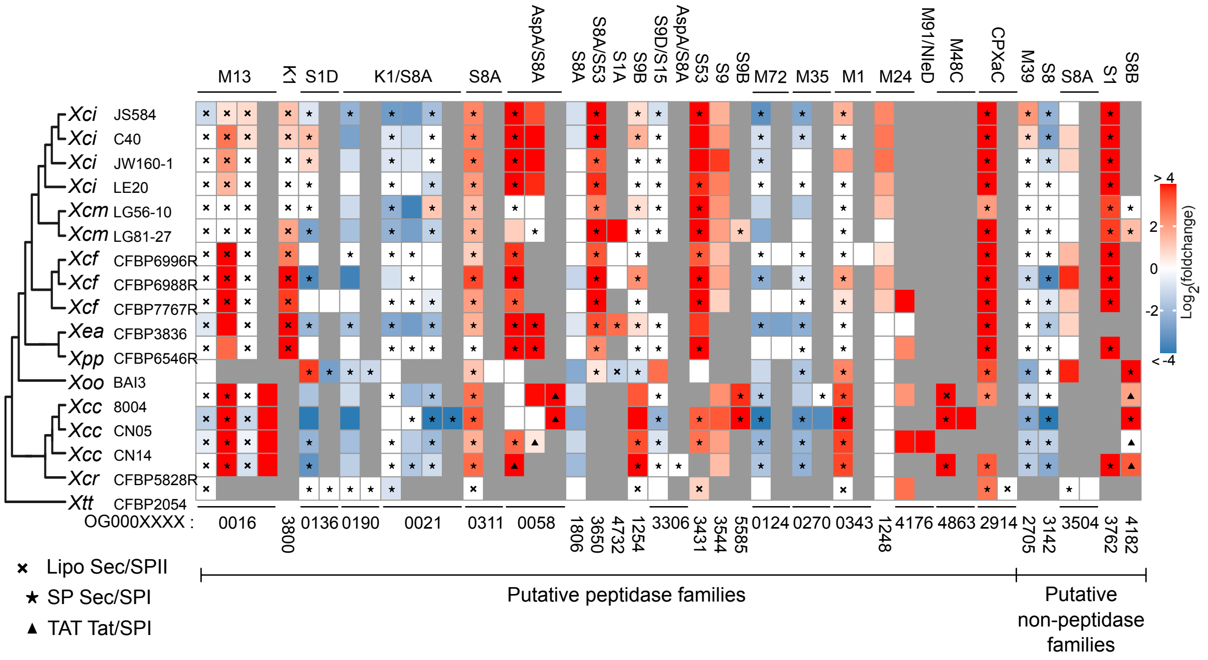


Figure 7. Regulation of orthogroups (OG) encoding proteins with proteolytic activity in *Xanthomonas* strains expressing *hrpG**. Only genes with an AdjPval < 0.05 are colored. Absent orthologs are marked in gray. Orthogroup annotation is given on top of the figure, orthogroup number is given at the bottom of the figure. Orthogroups are annotated according to their MEROPS identifiers. Due to the extensive number of orthogroups encoding proteins with proteolytic activity, only orthogroups regulated in at least two strains are shown. The cross, star and triangle symbols indicate the presence of a signal peptide within a gene, as predicted by SignalP6.

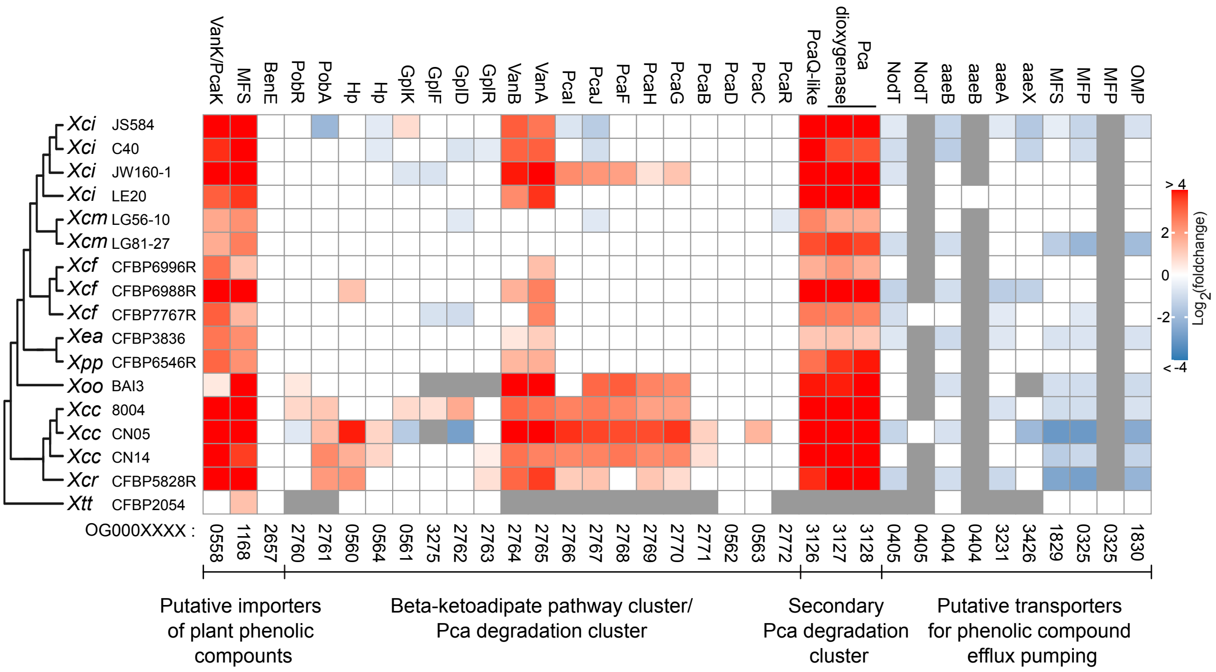


Figure 8. Differential expression of orthogroups (OG) encoding proteins putatively involved in plant phenolic compounds metabolism. Only genes with an AdjPval < 0.05 are colored. Absent orthologs are marked in gray. Orthogroup annotation is given on top of the panel, orthogroup number is given at the bottom of the figure. Pca: protocatechuate. Hp: hypothetical protein

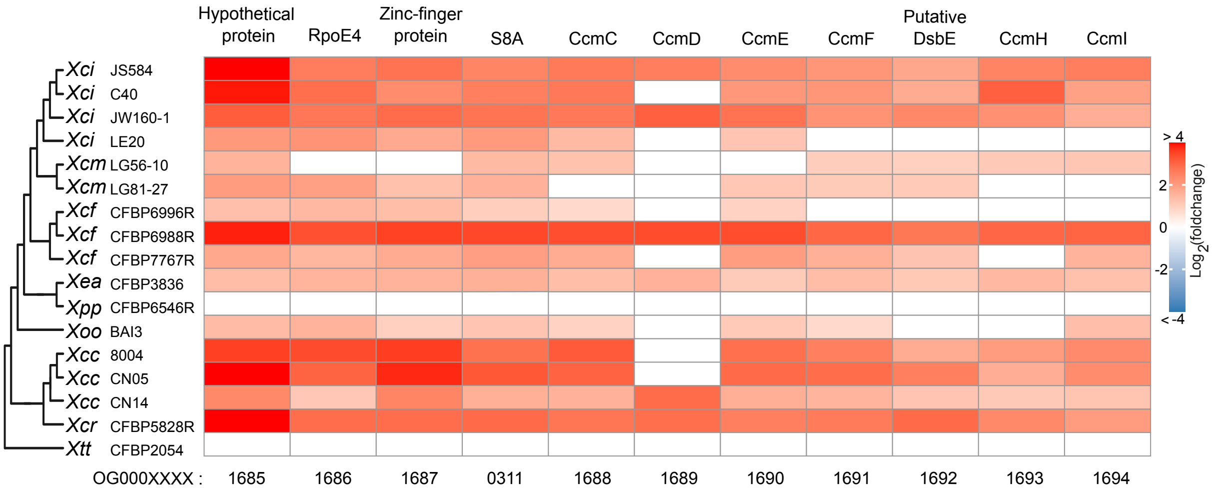


Figure 9. Differential expression of orthogroups (OG) encoding proteins putatively involved in cytochrome C maturation. Only genes with an AdjPval < 0.05 are colored. Orthogroup annotation is given on top of the figure. Orthogroup number is given at the bottom of the figure.

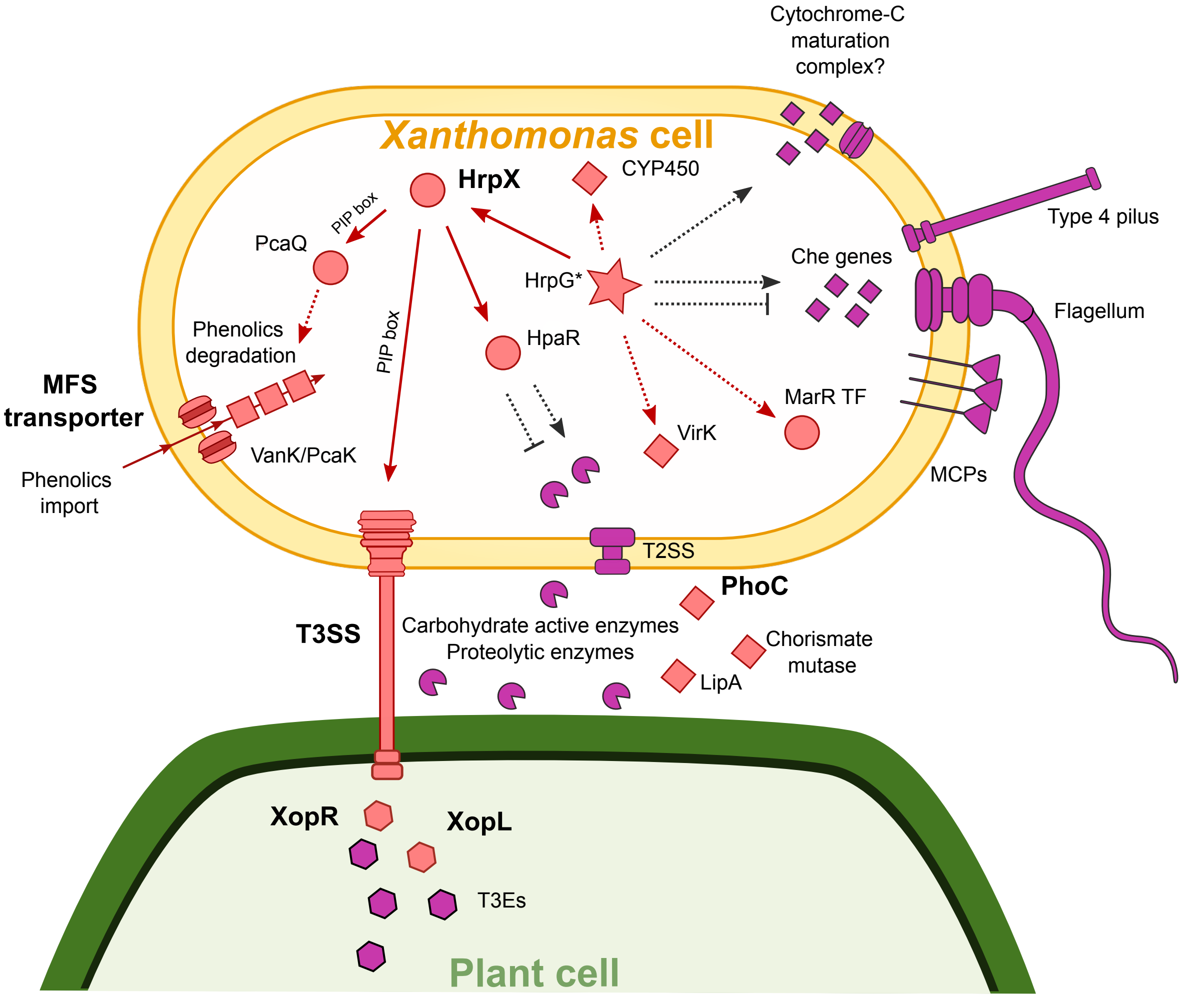


Figure 10. Schematic representation of the key processes identified to be regulated by HrpG* in this study. Red symbols annotated in bold indicate processes that are part of the *Xanthomonas* core regulon. Other red symbols indicate terms part of the group-II *Xanthomonas* core regulon. Symbols in purple highlight other regulated orthogroups which have been discussed. Dashed lines indicate hypothesized interactions, whereas continuous lines indicate interactions which are thought to be direct. T3SS: type III secretion system; T3Es: type III effectors; T2SS: Xps type II secretion system; MCPs: methyl-accepting chemotaxis proteins.