Electronic Journal of Biotechnology 42 (2019) 30–41

Contents lists available at ScienceDirect

Electronic Journal of Biotechnology

Research article

Gene expression analysis identifies hypothetical genes that may be critical during the infection process of *Xanthomonas citri* subsp. *citri*

Marcelo Luiz de Laia [a], Leandro Marcio Moreira [b,c], Janaína Fernandes Gonçalves [d], Maria Inês Tiraboschi Ferro [e], Any Caroliny Pinto Rodrigues [a], Jéssica Naiara dos Santos [a], Érica Barbosa Felestrino [c], Jesus Aparecido Ferro [e,*]

[a] *Departamento de Engenharia Florestal, Faculdade de Ciências Agrárias, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, MG, Brazil*
[b] *Departamento de Ciências Biológicas, Instituto de Ciências Exatas e Biológicas, Universidade Federal de Ouro Preto, Ouro Preto, MG, Brazil*
[c] *Núcleo de Pesquisas em Ciências Biológicas, Universidade Federal de Ouro Preto, Ouro Preto, MG, Brazil*
[d] *Instituto de Ciências Agrárias, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Unaí, MG, Brazil*
[e] *Departamento de Tecnologia, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Campus Jaboticabal, Via de acesso Prof. Paulo Donato Castellane s/n, Jaboticabal, SP, Brazil*

ARTICLE INFO

Article history:
Received 2 January 2019
Accepted 10 October 2019
Available online 23 October 2019

Keywords:
Citrus sinensis
DNA microarrays
Gene expression
Hypothetical genes
Oligonucleotide Array Sequence Analysis
RT-qPCR
Transcriptome
Type III secretion systems
Virulence
Xanthan gum

ABSTRACT

Background: Gene expression analysis via microarray is widely used in phytobacteria to validate differential gene expression associated with virulence or to compare biological profiles of wild type and mutant strains. Here, we employed DNA microarrays to study the early stages of the infection process (24, 72 and 120 h post-inoculation) of *Xanthomonas citri* subsp. citri (Xac) infecting *Citrus sinensis* to interrogate the expression profiles of hypothetical genes.

Results: Under infective conditions, 446 genes were up- and 306 downregulated. Outstanding among genes upregulated during infection were those involved in synthesizing the Type 3 Secretion System and effectors, xanthan gum and quorum-sensing induction, and flagellum synthesis and regulation. Additionally, 161 hypothetical genes were up- and 100 were downregulated, 49 of which are known to have a significant biological role. To understand hypothetical gene co-regulation or -expression, nine expression profiles including 158 genes were identified during the three infection phases. Of these, 47 hypothetical genes were identified as having expression profiles associated with at least one connected to a gene associated with adaptation and virulence.

Conclusions: Expression patterns of six differentially expressed genes were validated by quantitative reverse transcription polymerase chain reaction, thus demonstrating the effectiveness of this tool in global gene expression analysis in Xac.

How to cite: Laia ML, Moreira LM, Gonçalves F, et al. Gene expression analysis identifies hypothetical genes that may be critical during the infection process of Xanthomonas citri subsp. citri. Electron J Biotechnol 2019;42. https://doi.org/10.1016/j.ejbt.2019.10.003.

1. Introduction

Citrus canker is one of the most damaging diseases globally affecting several varieties of citrus. Among the several pathotypes of its causal agent, all bacteria belonging to *Xanthomonas* genus, found in nature (strains A, Aw, A*, B, C, and E), the A strain produces the most common and severe type A cancrosis [1,2]. The disease shows characteristic symptoms including the formation of cortical lesions that protrude on both sides of the leaves [3]. Symptoms manifest within one week to two months after inoculation, and induce the development of hyperplasic tissue, forming cancer-like lesions, a characteristic symptom very crucial in disease diagnosis [4]. Due to such injuries, fruit marketability is affected resulting in heavy losses, which are aggravated by the fact that this is not an easy bacterium to control.

Attempting to examine this phytobacterium to discover a means of control, da Silva et al. [5] performed a complete sequencing of the *X. citri* subsp. *citri* strain 306 pathotype A (Xac) in mid-2002. They identified approximately 4800 coding regions (CDSs), dozens of which were found to directly participate in inducing virulence and pathogenicity. Although most of the Xac genome has been assigned as having a probable function based on annotation, with inferences drawn from the in silico analysis, a lack of experimental research

* Corresponding author.
E-mail address: jesus.ferro@unesp.br (J.A. Ferro).
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

https://doi.org/10.1016/j.ejbt.2019.10.003

prevents accurate detection of specific genes responsible for the pathogenic and adaptation processes.

Today, considering these data, functional genomics is a priority from a research perspective and is being widely developed. Techniques such as transposon-mediated mutagenesis [6], homologous recombination [7,8], proteomic analysis [9,10,11] and RNA-seq [12] have been employed to determine the contributions of these genes/proteins to infection. Correspondingly, empirical analyses have enabled the discovery of new genes related to pathogenicity, examining potential targets to combat the pathogen.

Unfortunately, knowledge regarding the Xac genome is limited, as a probable putative function has not been assigned for nearly 40% of mapped CDSs. These CDSs have been categorized as hypothetical (HP) or conserved hypothetical (CHP), consistent with other organisms. In fact, the CDSs with uncharacterized functions should include a set of genes in Xac, as well as in other organisms, that may be critical for maintaining the connection between the pathogen and host [13,14]. However, very little interest has been shown in the functional study of these often neglected putative genes, as described earlier for other organisms [15,16,17].

From this perspective, studies involving functional genomics with global gene expression may enhance our understanding of the intricate networks of genetic relationships within the pathogen itself, as well as its relationship with its host response [18,19]. Analyzing DNA arrays (DNA microarrays or macroarrays) and identifying genes responsible for a given biological condition enable the identification of related genes and can predict new pathways, or even add regulatory networks for formerly characterized genes in those pathways [20,21]. Using this technique, complementary nucleic acid sequences are hybridized to one another, one of which is immobilized on a solid matrix [22]. This method has proven successful in genomic studies in several models, enabling inferences to be drawn regarding the functions of the hypothetical genes in other pathogenic bacteria, yeast, plants and vertebrates [23,24].

Significantly, gene expression essential for a given condition will be upregulated as redundant or undesirable, in the same situation, when its functions are negatively regulated. This principle is also applicable to virulence genes, which are also dependent on regulatory mechanisms that control appropriate gene expression in the host environment [25].

Keeping this view in mind and using a microarray platform objectively developed for this purpose [26], the present study represents the first global gene expression profile of Xac subjected to different time spans of infection in *Citrus sinensis* (L. Osbeck). The data reveal both genes classically associated with virulence and adaptation as well as a series of hypothetical genes working along with these genes. Hypothetical genes, possessing expression profiles similar to those genes known to be connected with the virulence process, enable the inference of correlated function and thus display their potential for new anti-pathogen targets.

2. Materials and methods

2.1. Bacterial isolates and growth conditions

Xac, which has a previously-sequenced genome, was grown in Petri dishes by culture in Nutrient Agar (NA) medium (3 g/L beef extract, 5 g/L peptone and 15 g/L agar) at 28°C. After 24 h, new colonies were transferred into three Erlenmeyer flasks (250 mL) each containing 50 mL of Nutrient Broth (NB) culture medium (3 g/L beef extract, 5 g/L peptone) and labeled 1, 2, and 3. After 12 h of incubation at 28°C with shaking at 200 rpm, 1 mL was aseptically drawn from each flask and transferred to another flask containing NB. Here, the concentration was adjusted to 10^8 CFU/mL (OD600 nm = 0.3) and the volume to 50 mL. Bacterial cells, after allowing 12 h of growth at 28°C with shaking at 200 rpm, were then harvested by centrifugation at 5000 × g for 5 min. The remaining cell suspension was drawn with a 1 mL syringe without a needle and infiltrated into orange leaves (*Citrus sinensis* cv. Pera). Orange plants were grown in 20 L pots.

Nine plants (three plants for each incubation period) were maintained respectively for 24, 72 and 120 h in the laboratory at 28°C with a 12/12 h photoperiod and light intensity ~2000 lx at photophase. After multiplication, inoculated leaves were collected and promptly fractionated into very thin strips. They were then kept in a beaker of sterile distilled water that was placed in an ice bath under gentle agitation. A separate beaker was used for each plant. After 5 min, leaf debris was filtered using gauze and cells were recovered by centrifugation at 5000 × g for 5 min at 4°C. Total RNA was immediately extracted. Thus, three independent biological replicates were obtained for Xac growth in vitro and *in planta*. The same procedures were performed prior to RT-qPCR analysis. Three 250 mL flasks containing 50 mL of NB and three orange plants were inoculated with cells and collected. Total RNA extraction was then performed via the same methods. The only difference was the plant incubation period, which was limited to only 72 h.

2.2. Xac total RNA extraction

An Illustra RNA Isolation Mini-RNA spin KiT (Amersham Biosciences, Little Chalfont, United Kingdom) was used to extract the RNA per the manufacturer's instructions. Total elimination of DNA was confirmed and PCR was performed using DNase I-treated RNA samples as template. PCR employed an initial denaturing step of 94°C for 3 min followed by 35 cycles of a denaturing step at 94°C for 30 s and an annealing step at 60°C for 30 s, with a final elongation step at 72°C for 2 min. After the final elongation step, samples were stored at 4°C until use. An amplification reaction was performed in the 25 μL samples with 2.5 μL 200 ng RNA, 2.5 μL PCR buffer™ (Invitrogen, Carlsbad, CA, USA), 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 300 nM of each primer and 16S rRNA, and 1 unit of Taq DNA polymerase (Invitrogen, Carlsbad, CA, USA). PCR products underwent electrophoresis on a 1% agarose gel with TAE buffer, stained with ethidium bromide and visualized under a UV transilluminator (data not shown). To verify the quality of extracted RNA, another electrophoresis was performed under the same conditions. The $A_{260/280}$ ratios of RNA samples were measured and RNA was quantified utilizing a ND-1000 spectrophotometer, NanoDrop. RNA was stored at −80°C until required.

2.3. Evaluating microarray

We used XACarray as probe [26]. To create the target, the SuperScript UPirect cDNA Labeling System (Invitrogen, Carlsbad, CA, USA) kit was used to synthesize and label cDNA with Cy3-fluorophore-dCTP™ and Cy5-fluorophore-dCTP™ (Amersham Biosciences, Little Chalfont, United Kingdom) per the manufacturer's instructions. In all experiments, cDNA from the NB cells was labeled with Cy5, while cDNA from cells grown *in planta* was labeled with Cy3. All cDNA was synthesized using 20 μg of total RNA and the fluorophore was incorporated indirectly. cDNA was quantified via absorbance detection at 550 nm for Cy3 and 650 nm for Cy5. To ensure that the analysis was accurate, equivalent quantities of fluorescent cDNA were used during microarray hybridizations. After labeling, Cy3- and Cy5-tagged samples were dried in a vacuum chamber protected from light. Cy3-tagged samples were recovered in 13.5 μL of ultra-pure water and then transferred to the Cy5-tagged sample tube. Next, 13.5 μL 4× hybridization buffer™ (Amersham Biosciences, Little Chalfont, United Kingdom) and 27 μL formamide were added to the tube. After homogenization, the target solution was heated to 90°C in the dark for 2 min, then transferred into an ice water bath for 2 min. The solutions were then placed on slides, covered and hybridized for 16 h at 42°C. Slides were then washed [27] and dried under a jet of compressed nitrogen gas. Microarrays data were then acquired using the GMS 418 Scanner Array (Affymetrix, Santa Clara, CA,

USA) to obtain images from the fluorescent channel for each of the microarray probes (Cy3 and Cy5).

Raw images were all converted to quantities using ArrayVision 8.0™ (Amersham Biosciences, Little Chalfont, United Kingdom). The foreground and background intensities were measured for each spot on the microarray representing a specific probe. Data from ArrayVision was read in R [28,29] on a Debian GNU Linux 3.2. Quality access, background correction, and print-tip loess scale normalization followed by statistical analysis were performed with the limma R package [28,29,30,31]. Differentially Expressed (DE) genes were selected using the False Discovery Rate (FDR), with a 5% p-value threshold [32].

2.4. RT-qPCR evaluation

First strand synthesis cDNA and the RT-qPCR reactions were performed using the SuperScript III First-Strand Synthesis SuperMix for RT-qPCR Kit (Invitrogen, Carlsbad, CA, USA) per the manufacturer's specifications (including random primers used in the reverse transcription reaction), except for the quantity of cDNA in each reaction, which was 20 ng. All PCR performed with SYBR Green was conducted using the 7500 Real-Time PCR instrument (Applied Biosystems, Foster City, CA, USA) using three biological and three technical replicates (for a single biological replicate). The PCR was performed using 2 min at 50°C and 10 min at 95°C followed by 40 cycles of 15 s at 95°C, and finally 1 min at 60°C. To determine the PCR efficiency, standard curves were generated using the sample cDNA at five dilutions and measured in triplicate. The reference genes used in all the experiments were *rpoB*, *atpD*, and *gyrB* [33]. We used the $2^{-\Delta\Delta CT}$ method for relative expression analysis [34,35]. The primers used are included in Table 1.

2.5. Metabolic pathways and protein complex design

Metabolic pathways were obtained from the KEGG annotation database [36,37].

2.6. K-mean clustering

K-mean clustering was used to identify pathways similarly expressed under all three conditions [38,39]. The number of clusters was assessed according to methods described by Sturn et al. [40]. Nine different clusters were identified, each of which included 2–57 different genes in triplicate. Using these parameters, gene groups with similar expression profiles were identified.

Table 1
Primer sequences used for RT-qPCR-based validation of microarray results.

Gene name	Function	GI	Forward Primer / Reverse Primer	Amplicon (bp)
hrcQ	Member of T3SS	1154474	GCAGCTGGAAGTGGACCAA GGCTGCAAACCCGACAAC	57
hrpB2	Member of T3SS	1154479	CAGCGCAGCAGATCAAGTTG CGCCGACGCTGACATTG	69
hrcS	Member of T3SS	1154472	CGACGATCTAGTGCGATTTACCT CGACCACCGGCAAGGA	68
hrpXct	Regulator of T3SS	1155337	GCCTACAGCTACATGATCACCAAT TGCGGCCACTTCGTTGA	63
XAC2613	Hp	1156684	CGAAGGGAGATGGAAGCAGTT GCTCAACAGATCGGCTGGAA	59
XAC2622	Hp	1156693	GGCGGATTCCAATATGCAGTT TGGCCCGATCAAGGTGTAG	59

3. Results

3.1. Gene expression analysis

To investigate global changes in Xac gene expression under *Citrus sinensis* infection, we used the DNA microarray platform developed by Moreira et al. [26]. It involves 2365 distinct tag sequences of unique CDSs with positive DNA × DNA hybridization signals. They were obtained by PCR from the shotgun libraries used for Xac genome sequencing, which represent 52.7% of the complete annotated genome.

Globally, of the putative 4489 annotated genes in Xac, 446 (9.94%) were found to be upregulated at least once during the infection timeline periods that were investigated. Downregulated genes numbered 306 (6.82%) and were downregulated at least once during the same time period (Table 2 and Table S1). When comparing the total number of up- and downregulated genes classified per functional annotation categories, the most up-regulated genes were found in categories I, IV, V, VII and VIII, and the most downregulated genes were found in category VII (Table 2).

The largest range of upregulated genes can be observed during the early stages of the infection process. In all, 248 upregulated genes and 130 downregulated genes were detected during the first 24 h of infection (Fig. 1A). After 72 h of infection, the total number of inducible genes decreased to 215, while the number of downregulated genes increased to 159 (Fig. 1A). At the end of the experiment, after 120 h, 232 putative genes were upregulated and 139 were downregulated (Fig. 1A). Among all genes identified, we detected 72 that were upregulated throughout the infection process but only 28 were downregulated in all conditions. When the same analyses are restricted to the hypothetical genes, it is observed that a total of 86, 72 and 82 genes are upregulated in 24, 72 and 120 h, respectively, whereas 31, 58 and 49 genes are downregulated, respectively, in these same times (Fig. 1B).

Among the genes upregulated after only 24 h of infection (Fig. 1C-a), about three-fold more were in the hypothetical category (category VIII) than any other category, as seen in Fig. 1C-d, which shows the intersection of upregulated and downregulated genes at 24 and 72 h post-infection. From this viewpoint (Fig. 1C-g), which shows the intersection of the three periods of infection, a huge induction of the expression of genes belonging to categories V, VII and VIII, which strengthens the overall expression, as described above. These values confirm the general data on the percentage of differentially expressed genes in functional annotation categories highlighted in Table 2.

3.2. Global expression and physiology of infection process

To understand the relationship between upregulated and downregulated genes, as well as their relationships with the physiology of the infection process, further analyses of the results were focused on categories V, VII and VIII to emphasize the differences in the number of the differentially expressed genes mentioned above (Fig. 1C).

A great portion of the research involving the understanding of the Xac-*Citrus* pathosystem and other related systems involves studying genes in class VII [41]. Among the 72 genes in this category (including 44 up- and 28 downregulated genes) that were differentially expressed throughout the entire infection process (Table 2) were genes involved in the synthesis of the apparatus type III secretion system (T3SS) and its effectors (T3SSe), the *rpf* genes that possess a GGDEF domain and are involved in the synthesis of xanthan gum, as well as genes involved in flagellum biosynthesis and regulation.

3.3. Type III secretion system and effectors

Regarding the T3SS, cDNA microarray data revealed that the *hrpB5* and the *hrpG* genes were upregulated within 72 h post-infection, *hpaF* within 120 h post-infection and *hrcQ* between 72 and 120 h post-

Table 2
Correlation between the Xac genome and transcriptome based on functional annotation category.

Category[a]	Genome[a]	XACarray[b]	Upregulated (%)[c]	Downregulated (%)[c]
I — Intermediary metabolism	727	424 (58.4)	77 (10.59)	60 (8.25)
II — Biosynthesis of small molecules	352	206 (58.5)	22 (6.25)	25 (7.10)
III — Macromolecule metabolism	558	311 (55.7)	44 (7.89)	35 (6.27)
IV — Cell structure	202	124 (61.4)	23 (11.39)	16 (7.92)
V — Cellular processes	391	220 (56.3)	57 (14.58)	19 (4.86)
VI — Mobile genetic elements	190	117 (61.6)	8 (4.21)	17 (8.95)
VII — Pathogenicity, virulence, and adaptation	304	194 (63.8)	44 (14.47)	28 (9.21)
VIII — Hypothetical/Conserved hypothetical genes	1658	861 (51.9)	161 (9.71)	100 (6.03)
IX — Without assigned function	107	63 (58.9)	10 (9.35)	6 (5.61)
Total	4489	2520 (56.13)	446 (9.94)	306 (6.82)

[a] According to da Silva et al. [5].
[b] Fixed individual CDSs as probes in XACarray, according to Moreira et al. [26].
[c] Values determined using the reference genome.

infection. *hrcS*, *hrcU*, *hrpB1*, *hrpD6*, *hrpB2*, *hrpF*, *hpa1*, *hpaB*, and *hrpXct* were upregulated throughout the infection process and represent 13 of the 33 genes involved in the synthesis of the apparatus; 25 of them were present on the microarray slide (Fig. 2). The sole downregulated gene, *hrcV*, was detected 24 h post-infection.

Genes reported as probable type III secretion system effectors, which include *xopAE* and x*opV*, were detected 120 h post-infection; *xopF2* and *xopE2* at 72 and 120 h, respectively, while *xopE1*, *xopN*, *xopK*, *xopE3* and *xopAI* were detected throughout the infection process (Table 3). Only those genes encoding *avrBs3* (XACb0015) and *xopI* were

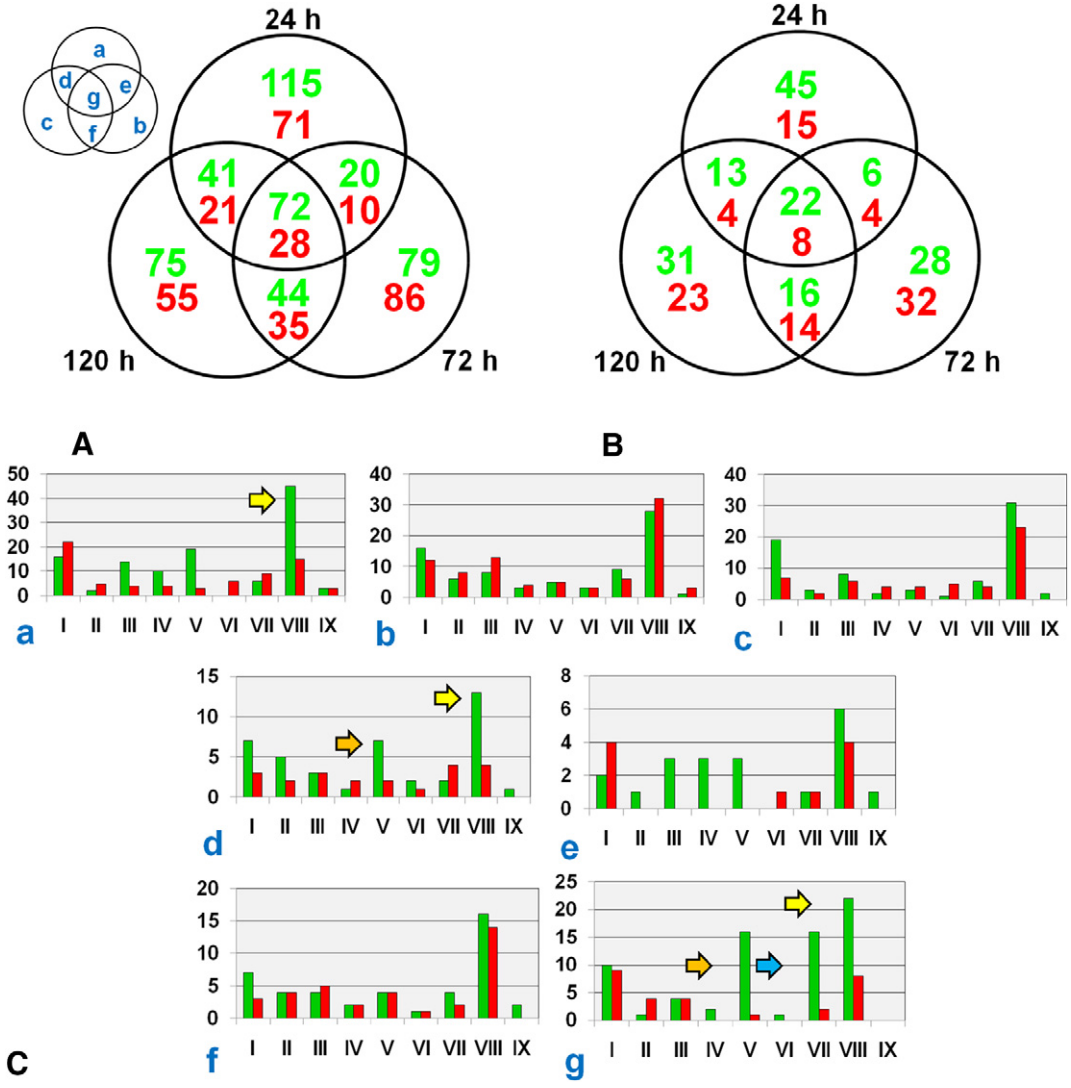

Fig. 1. Gene expression global analysis during the chronological infection. (A) Venn diagram highlighting the genes upregulated (green) and suppressed (red) over the three chronological days of the infection period (24, 72 and 120 h). (B) Venn diagram similar to that shown in A, but specific to hypothetical genes detected under the infective conditions. (C) Differential gene expression analysis based on Xac genome functional annotation categories. Each graph represents the distribution of up (green) and downregulated (red) genes previously identified in each portion of the Venn diagram (Fig. 1A). Each category description can be found in Table 2. Arrows highlight the great proportion of genes up-regulated in categories V (orange), VII (blue) and VIII (yellow) in relationship to the ones that were down-regulated. X-axis = gene category. Y-axis = number of expressed genes.

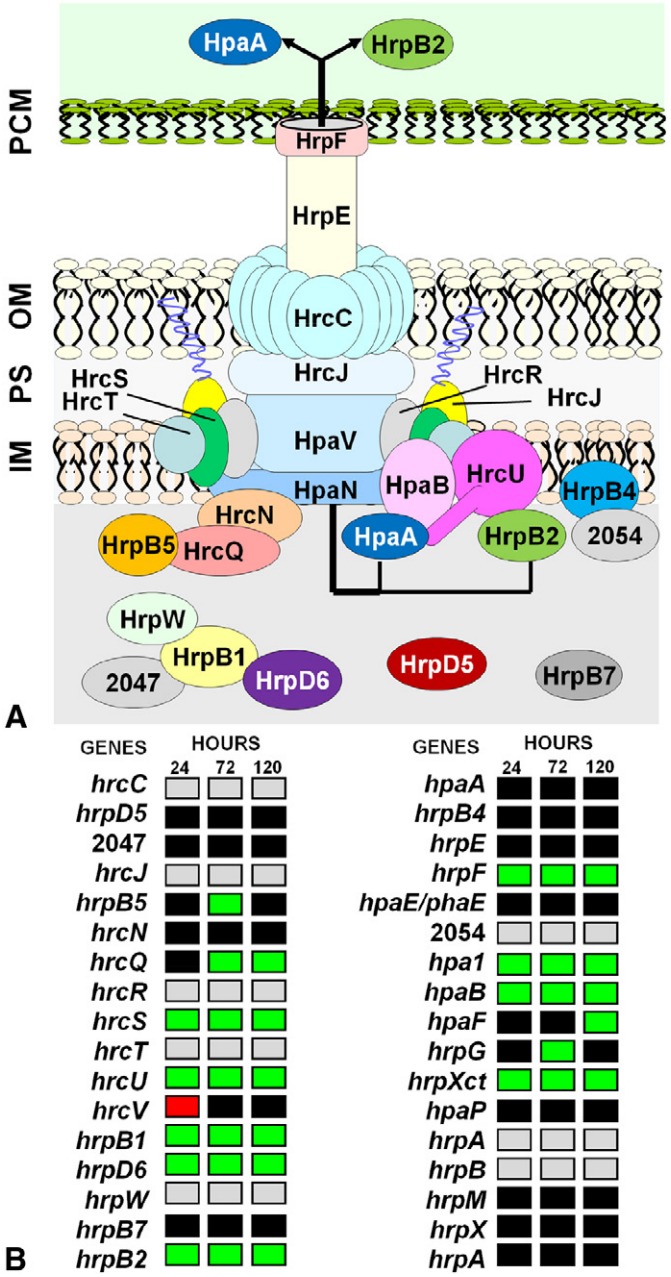

Fig. 2. Composition and determination of the expression profiles of T3SS apparatus genes. (A) Model that highlights the proteins that make up the T3SS apparatus in Xac (adapted from KEGG [36,37] and Alegria et al. [63]). Each structural or secreted protein (arrows) is encoded by the respective gene presented in B. IM — Inner membrane; OM — Outer membrane; PS — Periplasmic space; PCM — Plant-cell membrane. (B) Representation of the genes encoding proteins of the T3SS showing expression patterns during the chronological process of infection as analyzed via microarray. Green — upregulated genes; Red — downregulated genes; Gray — Not-available genes; Black — Not-detected/Not-expressed genes.

downregulated during the early stages of infection. Therefore, 24 candidate and the nine core genes of the T3SSe were upregulated, two were downregulated and only three were not found among the 14 represented in the microarray (Table 3). Notably, the *xopF2* gene was previously categorized as a putative pseudogene, but now shows an expression pattern.

Those T3SS genes upregulated during the infection process include 21 that participate in the synthesis of the apparatus; this represents about 48% of all Class VII as upregulated genes that are directly related to pathogenicity and virulence.

Table 3
Expression of genes that encode T3SSe in the Xac genome. UP — Upregulated; DOWN — Downregulated; NA — Not available; ND — Not-detected.

XAC id[a]	Gene name or family[b]	In plant		
		24 h	72 h	120 h
XAC0076	*avrBs2*	NA	NA	NA
XAC0277	*xopR*	NA	NA	NA
XAC0286	*xopE1*	UP	UP	UP
XAC0393	*xopAE*	ND	ND	UP
XAC0543	*xopX*	NA	NA	NA
XAC0601	*xopV*	ND	ND	UP
XAC0754	*xopI*	DOWN	DOWN	ND
XAC1208	*xopP*	ND	ND	ND
XAC2009	*xopZ*	ND	ND	ND
XAC2785	*xopF2*	ND	UP	UP
XAC2786	*xopN*	UP	UP	UP
XAC2922	*hrpW*	NA	NA	NA
XAC3085	*xopK*	UP	UP	UP
XAC3090	*xopL*	NA	NA	NA
XAC3224	*xopE3*	UP	UP	UP
XAC3230	*xopAI*	UP	UP	UP
XAC3666	*xopAK*	NA	NA	NA
XAC4213	*xopAD*	ND	ND	ND
XAC4333	*xopQ*	NA	NA	NA
XACa0022	*avrBs3*	NA	NA	NA
XACa0039	*avrBs3*	NA	NA	NA
XACa0065	*avrBs3*	NA	NA	NA
XACb0011	*xopE2*	ND	UP	UP
XACb0015	*avrBs3*	DOWN	ND	ND

[a] According to da Silva et al. [5].
[b] According to Moreira et al. [26] and Potnis et al. [70].

3.4. Rpf genes, involved with GGDEF and synthesis of xanthan gum

From the genes involved in quorum sensing mediated by Diffusible Signal Factor (DSF), *rpfN* was detected under infective conditions, but only within 24 h of infection; however, the *rpfC* and *rpfG* genes were observed throughout the infection process (Fig. 3). The *rpfA* and *rpfB* genes were downregulated at 24 and 120 h post-infection, respectively. Expression of genes directly regulated by *rpfG* was also noted, as was the case for the genes participating in the synthesis of the xanthan gum genes, genes that encode GGDEF domain proteins, and genes encoding cellulases and proteases. Among the genes we detected were those seen during gum infective conditions within 72 h post-infection, *gumH* within 120 h of infection, *gumF* between 24 and 120 h, and *gumL* in the late stages of infection. Of note, three (XAC1938-39 and XAC1887) of the five genes in the previously-annotated GGDEF domain protein-encoding members of the Xac genome were upregulated in the early stages of infection. Four cellulases (XAC3506, XAC1770 and XAC0028-29) and four proteases (XAC0928, XAC2831, XAC2833, XAC2853) were also identified in the infective stages (Fig. 3). The *rpf* genes, gum and cellulases that were upregulated during the infection stage constitute 27% (12 genes) of the CDSs annotated in category VII.

3.5. Flagellum biosynthesis and regulation

Our data revealed induction of the expression of genes involved in the biosynthesis and regulation of the flagellum throughout the infection process. These genes were recorded as belonging to category V (cellular process).

It is notable that 31 of the 35 genes involved in the biosynthesis of the flagellar apparatus were represented on the microarray, among which 21 (68%) were upregulated during infection (Fig. 4). Also, three of the six *che* genes present in the platform (50%) were upregulated in addition to the eight *tsr* upregulated genes among a total of 17 represented on the platform (47%). Altogether, these genes represent

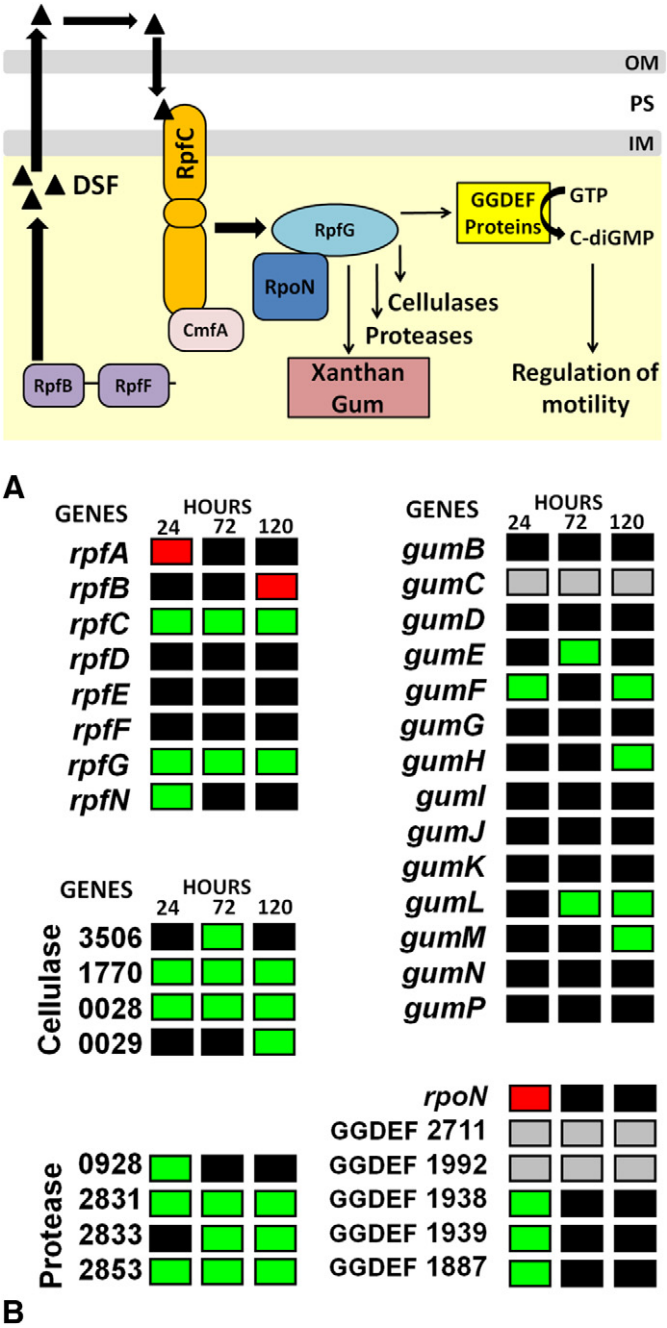

Fig. 3. Composition and determination of the expression profiles of genes involved in quorum sensing mediated by DSF. (A) Representative model of the proteins involved in the synthesis of DSF and modulation of xanthan gum, GGDEF, cellulase and protease gene expression (adapted from Vojnov et al. [69]). (B) Representation of the genes involved with quorum sensing showing temporal expression during infection as analyzed by microarray. Green — upregulated genes; Red — downregulated genes; Gray — Not-available genes; Black — Not-detected/Not-expressed genes.

32 (58%) of the 55 genes in category V and were upregulated during the infection process (Fig. 4).

3.6. Hypothetical genes

As noted above, hypothetical genes represent 51.8% (1658 genes) of the Xac genome [5]. The microarray platform included 861 unique sequences representing these genes [26], of which 261 (30.3%) showed differential expression under infective conditions.

In the initial 24 h of infection, 86 hypothetical genes were upregulated and only 31 were downregulated. At 72 h of infection, the total number of induced genes increased to 72 and 58 were suppressed. After 120 h, 82 genes were upregulated and 49 downregulated (Fig. 1B). Among all differentially expressed hypothetical genes that we identified, 22 were prominently upregulated throughout the infection and only 8 were downregulated under the same conditions (Fig. 1B).

3.7. Determining gene expression profiles

To identify relationships between all the genes differentially expressed under our experimental conditions, nine expression patterns were obtained based on the time of infection using the k-means clustering method (Fig. 5 and Table S2). Genes that were included in this cluster analysis were not categorized per the differential expression profile exhibited, as described above.

Clusters 1, 2, 3, 6, 7 and 9 are significant as they introduce at least one gene included in category VII related to virulence and pathogenicity. Therefore, the hypothetical genes included in these profiles could exhibit some similarity to the co-regulated virulent genes. Cluster 1 includes 19 genes, six of which are in category VII, one in category V, and six hypothetical genes. One fact remains clear, which is that these hypothetical genes (XAC2810, XAC3319, XACb0035, XAC1396, XAC2606 and XAC2611) are similar in their expression profiles to that of the virulence gene *pthA3* (*avrBs3* – XACb0015).

Nine genes are grouped in cluster 2. Four of these genes are involved in the synthesis and regulation of the flagellum (*motA*, *cheD* and *tsr* 2x), similar to that revealed by the hypothetical gene XAC1898 and the expression profile for the gene that encodes T3SS apparatus, *hrpXct*. The third cluster, consisting of 16 genes, includes five clustered hypothetical genes (XAC0940, XAC3214, XAC1452, XAC2637, XAC3251), four related to category V (two *tsr*, *pnuC* and *yhdG*), and one, *pmrA* (multidrug resistance efflux pump), related to category VII. Cluster 6, the largest and most representative cluster that contains 57 genes, is the one that includes four genes from category VI (Mobile Elements Genetics), two CDSs encoding plasmid proteins and two transposons. This cluster prominently includes three of the five genes directly related to virulence, *rpfB*, *vppA* (virulence plasmid protein) and *gst* (glutathione-S-transferase), as well as *virB8* and *virB9*, which are involved in the Type 4 secretion system, as well as 17 hypothetical CDSs. Clusters 7 and 9, though not-representative of many genes, are noteworthy for the expression of two virulence genes. Cluster 7 includes the expression of *gumL*, which is involved in the synthesis of the xanthan gum, along with seven hypothetical genes (XAC0899, XAC1543, XAC3720, XAC1572, XAC3287, XAC3779, XAC3984). Cluster 9 is unique, with the expression of *avrXacE3* and *egl* (cellulase) and two other hypothetical genes (XAC1806, XAC0095).

3.8. The RT-qPCR results - microarray validation

Six genes were analyzed by RT-qPCR to validate the microarray data. Four of them encode proteins that are involved in the T3SS apparatus (*hrcQ* - XAC0403; *hrcS* - XAC0401; *hrpB2* - XAC0408; *hrpXct* - XAC1266), while the other two represent hypothetical genes (XAC2613 and XAC2622). It is important to note that although the main databases (Uniprot, String, Ensembl and NCBI) still maintain the annotation of these two genes as being hypothetical, some studies have shown that these genes actually encode the VirB5 and VirB7 proteins, respectively, with the type IV secretion apparatus [42]. When performing RT-qPCR experiments, tree leaves were independently re-inoculated with Xac and cells were collected after 72 h for bacterial RNA extraction and cDNA synthesis. RT-qPCR results for all selected genes confirm the results obtained from the microarray platform and thus functionally validate the test (Fig. 6). The T3SS genes were upregulated *in planta* but downregulated in culture medium. However, the hypothetical genes were downregulated *in planta* but upregulated in culture medium.

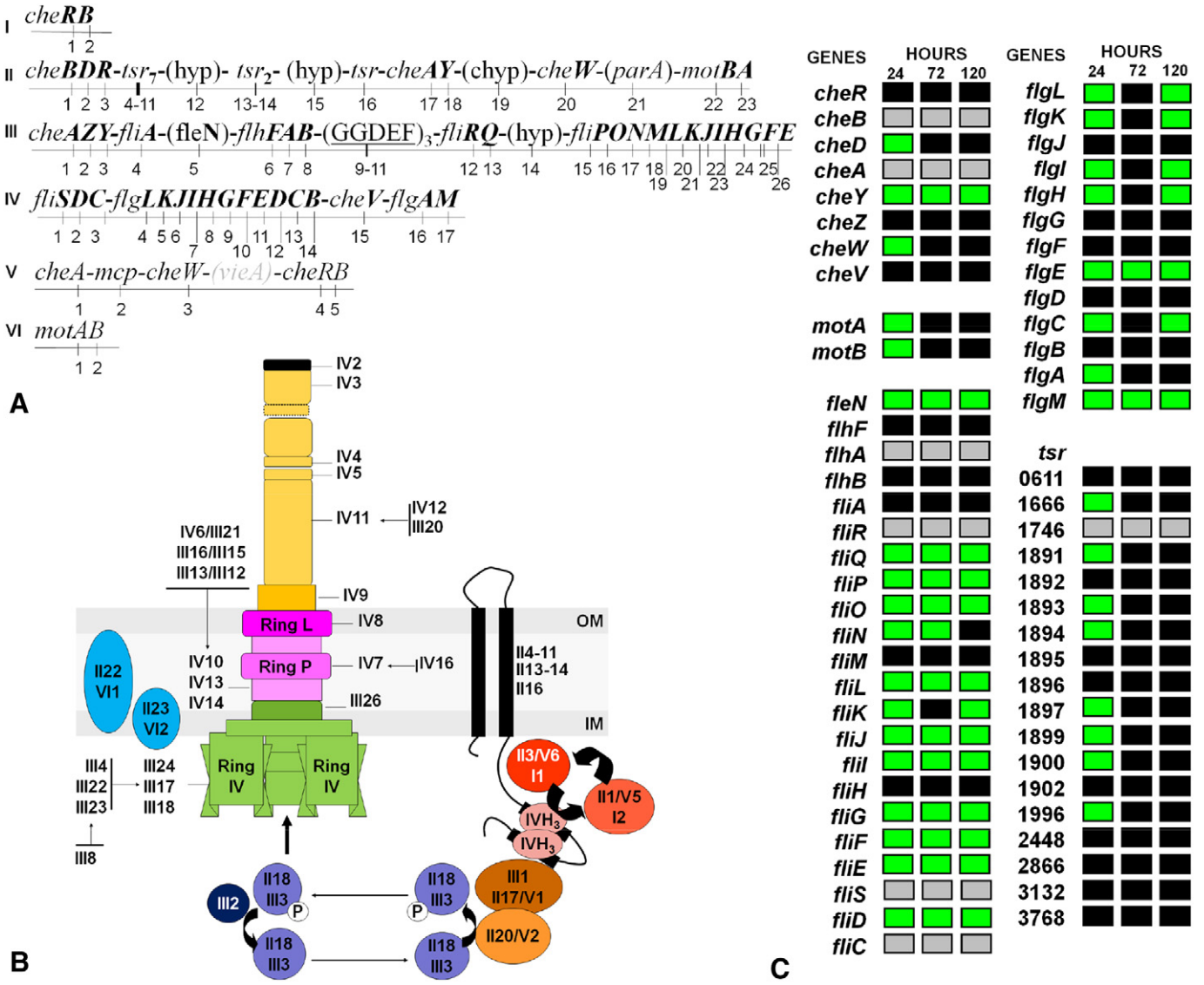

Fig. 4. Composition and determination of the expression profile of genes involved in flagellum synthesis and regulation. (A) Distribution of the genes involved in flagellum synthesis and regulation in the Xac genome. Note that these genes are categorized within five gene clusters (I to V). Each of the respective genes received a grade that can be identified in the structural Figure, prominently in B. (B) Representative model of the proteins involved in the synthesis and regulation of the flagellum. The Arabic and Roman numbers refer to the genes in the clusters shown in A. IM-Inner membrane; OM-Outer membrane. (C) Representation of the genes involved in flagellum synthesis and regulation showing temporal expression during infection as analyzed by microarray. Green — upregulated genes; Red — downregulated genes; Gray — Not-available genes; Black — Not-detected/Not-expressed genes. The arrows represent flux of displacement and interaction of these proteins in the proposed physiological scenario.

4. Discussion

Hypothetical genes were referred to as predicted protein coding regions in the genome; however, these genes become real genes only after experimental evidence that the corresponding messenger RNAs are transcribed and/or the respective proteins are synthesized. But this is just the beginning; more importantly, what are their functions? Because of this and their potential for involvement in vital biological functions or for the survival and virulence of pathogens, they have become significant targets for functional analysis [43].

Although many studies have aimed at identifying the roles of hypothetical genes in bacteria [15,16,17,44,45,46,47], the present study's main objective was to verify the possible correlation between hypothetical genes and genes co-expressed with virulence and pathogenicity, facilitating the inference of a possible relationship with these biological processes.

Based on prior knowledge regarding these genes as being associated with virulence, we attempted to identify additional candidate genes that actively play roles in the adaptation process inside the plant or even during the process virulence induction. We prioritized hypothetical genes that may be involved in these processes because they represent almost 50% of the Xac genome, and the fact that the work contributed to the discovery of the functionality of some of these hypothetical genes [6,42,48,49,50].

The time course period selected to investigate *in vivo* expression is directly related to that of physiological events previously described for the *Xanthomonas*-host pathosystem [41]. It is exactly during these stages of the infection process that classical genes related to pathogenicity and virulence are activated. The genes involved in synthesizing the T3SS apparatus and T3SSe secretion [51,52] as well as those involved in quorum sensing mediated by DSF, formation of the xanthana gum [53], and the degradation process of the plant host tissue [54] are active.

For T3SS, for example, changes in temperature, pH, oxygen tension, extracellular Ca^{2+} concentration, bile salts, and contact with the host cell (presence of specific metabolites) can regulate transcription of apparatus and secretion of effectors [55]. That is, the bacteria sense environmental and host-derived cues to determine and coordinate the stage of infection with T3SS activity, which corroborates the results

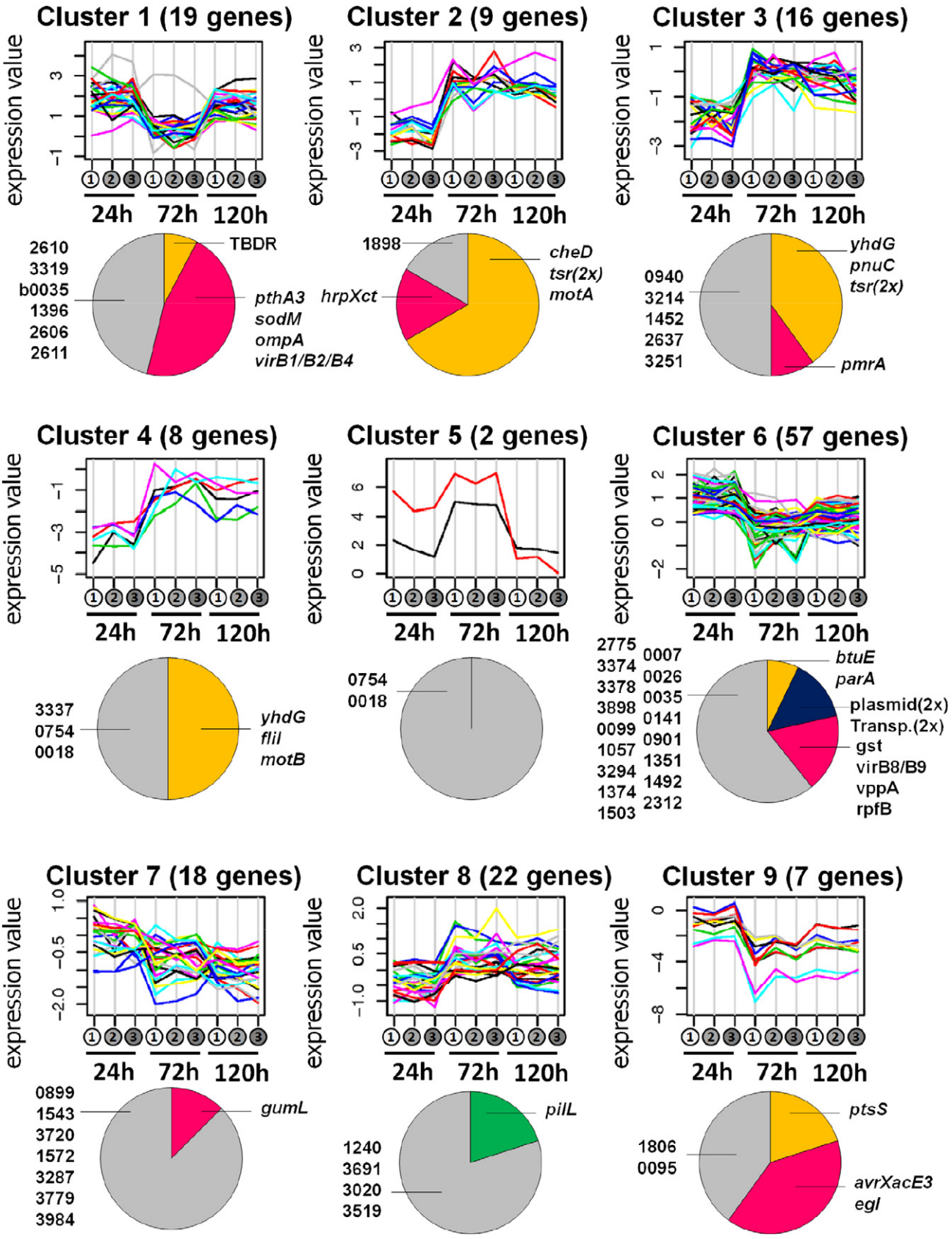

Fig. 5. Analysis of the gene expression profile during the infective process. From K-mean analysis, the genes that were differentially expressed were grouped under nine clusters according to their expression profile. The Y axis represents expression values in $\log_2$ scale. The numbers within circles 1, 2 and 3, present on the X axis, represent each experimental instance (*in planta*), at 24, 72 and 120 h, respectively. Below each cluster is a pie chart emphasizing the annotation of functional categories within these groups, highlighting their CDSs. The complete list of CDSs grouped based on these categories is shown in Supplementary Table 2. Orange — Category V; Pink — category VII; Gray — category VIII; Blue and green — categories VI and IV, respectively.

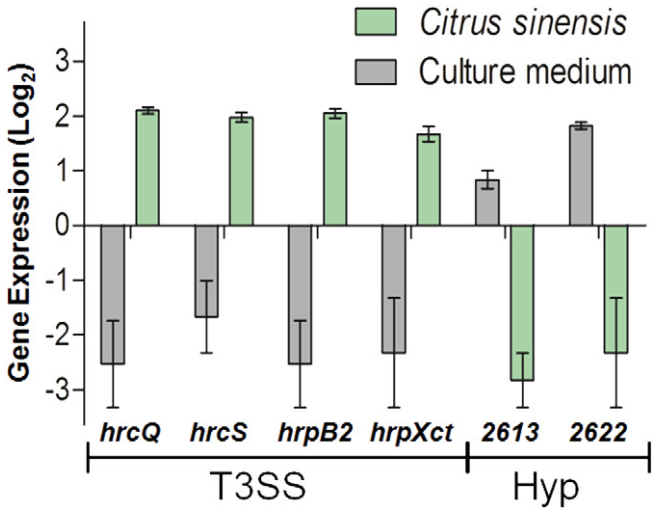

Fig. 6. Analysis of the expression of genes selected for validation using RT-qPCR. Green represents the genes expressed in the infective condition, while gray shows the non-infection condition. All genes selected are reproductive experiments regarding the overall analysis as the profile and intensity of their expression. Bars highlight the standard deviation.

found in this work. Our data validate and support the information given earlier, as genes known to be involved in these systems and biological processes were upregulated under experimental conditions (Fig. 2 and Fig. 3).

Thirteen genes encoding T3SS structural proteins apparatus were identified as upregulated under infective conditions. Furthermore, seven other genes coding for protein effector of this system were also upregulated under these conditions, five of which in the three times investigated (*xopE1*, *xopN*, *xopK*, *xopE3*, *xopAI*) and the two remaining at the later times of the infectious process (*xopE2* and *xopV*). These results not only corroborate with the described data in the literature for this pathosystem, but also point out such secreted proteins as fundamental to the pathogenesis success [56]. Curiously, *xopI* and *avrBs3*, genes, that also encode effector proteins, had a contrasting behavior, and were therefore downregulated at the initial time of infection. *xopI* has been described as an effector capable of inhibiting cell death by acting on the signal transduction blockade leading to this phenotype [57]. Considering that this function is also observed in Xac, the expression reduction of this gene in the early stages of the infectious process could contribute to the formation of canker, the latter stage of which would culminate in the necrosis of the lesion. Already for the four paralogous copies of *avrBs3* in the Xac genome, only one had its expression detected as downregulated at the initial infection times. These results may be related to the fact that it was activated at other times not investigated in this work, since its function has already been described as fundamental to Xac-induced canker formation [58].

Our data supports the conclusion that synthesizing and regulating flagellar activity are also fundamental to the adaptive processes of Xac during the early stages of infection, as described for other organisms [11]. While they are directly involved in chemotaxis, they may also be involved in protein secretion [59,60]. Of the 306 hypothetical genes detected by differential expression, either upregulated or downregulated (Fig. 1B), or by some other similar profile of functionally important genes in the aspects of adaptation and/or virulence (Fig. 5), 49 (16%) were identified as promoting a biological function previously described in the literature (Table 4). Among these, nine (18.37%) were found to be significant in a biological process involved in survival or colonization of the host plant tissue.

In their research, Yan and Wang [61] revealed that mutation of the genes XAC0007, XAC3743 and XAC0464 resulted in reduced symptoms *in planta*. However, we observed that these genes had very different expression profiles during the infection process. The XAC3743 gene was downregulated at 72 h, while XAC0464 was downregulated at 120 h after infection. XAC0007, although it does not vary in its gene expression profile, has exhibited an expression profile similar to those of genes in cluster 6, *gst*, *vppA* and *rpfB*, the latter fundamental for quorum sensing in *Xanthomonas* [62].

Mutation in the gene XAC0095 [6] produced a substantial change in water-soaking occurrence during the early stages of infection. Besides being upregulated during the three days of infection, it still exhibited an expression profile correlated to those of the nine *avrXacE3* and *egl* genes. Significantly, using two-hybrid assays, the protein encoded by the gene XAC0095 was found to interact with HrpG, which is an important regulator of T3SS apparatus synthesis and T3SSe secretion [63]. This enabled us to infer functionalities of these proteins in this same study that highlighted the phenotypic change after XAC0095 mutation. Laia et al. [6] also emphasized the mutation of XAC3294 completely eliminated symptoms when in contact with *Citrus sinensis*. XAC3294, although it did not present a gene profile of differential expression, was still placed in cluster 6 with the *vvpA*, *rpfB* and *gst* genes.

The mutant XAC1008 gene produced a massive reduction in hyperplasia and necrosis, while mutation of the XAC2344 gene increased necrosis when in contact with *Citrus sinensis* and *Citrus limonia* [64]. Both presented being upregulated within 72 h following initiation of the infection. XAC3984 mutation also reduced hyperplasia and necrosis [65]. This gene was upregulated towards the end of the investigated time (72 and 120 h), but still presented an expression profile consistent with the group that includes six other hypothetical genes and *gumL* (Cluster 7). The potential of this gene as a target for combating the disease is so highly valuable that a patent has been registered by referencing its characteristics (US 20040010131 A1).

XAC2654 was formerly annotated as a hypothetical conserved gene and was recently studied for its biological potential after being subjected to mutation by homologous recombination [7]. This gene was found to encode Plant Natriuretic Peptides (PNPs), which correspond to a class of extracellular, systematically mobile molecules that induce several responses vital to *in planta* growth and homeostasis in Xac, hence its genetic designation as XacPNP [7]. These authors contend that, although XacPNP is not expressed in nutrient-rich standard culture conditions, it is strongly upregulated under conditions that reflect the nutrient poor intercellular apoplastic environment of leaves or infected tissue, implying that XacPNP transcription can respond to the host environment. Our results suggest that this gene is likely important in the modified host response to create suitable conditions for its own survival, as it was upregulated at all the three times points in the infection process that we investigated.

Finally, two other hypothetical genes that deserve prominence are XAC2613 and XAC2622, induced in culture medium, but repressed under infectious conditions. Although still annotated as hypothetical in many databases, Alegria et al. [42] have already described and validated them as coders of VirB5 and VirB7 subunits, respectively, of the type 4 secretion apparatus (T4SS). In this way, the results observed in this work, from the differential expression levels identified in the DNA microarrays and validated by qRT-PCR (Fig. 6), corroborate data previously described by Jacob et al. [33]. These authors have demonstrated experimentally that while the type III secretion apparatus is activated under infective conditions, and repressed in culture medium, the genes involved in T4SS behave quite contrary. Thus, our work reiterates the hypothesis that T4SS is not critical to infection, but it is for other fundamental physiological functions of the cell, such as bacterial killing [66,67,68].

The results of our analysis of this collection of hypothetical genes in the present work underscore the importance of studying the hypothetical genes annotated in sequencing projects. The significance of this study will become evident as these genes are characterized as novel pathways or targets to develop methods to control the damage generated by these phytopathogens once they come into contact with the host.

Table 4
Hypothetical genes differentially expressed or with putative co-regulatory features of expression and that have been described as important in a biological function.

Locus TAG	Prod.[a]	Putative biological function (described earlier)	Differential expression[b]	Putative Co-regulation[c] (K-m cluster)	References
XAC0007	CHP	Mutant with reduction of symptoms	ND	(C6) *vppA/rpfB/gst*	[61]
XAC0077	CHP	Putative partial CUT locus, arabinose reduced TBDR	UP-120	–	[71]
XAC0095	CHP	Mutant with reduction of symptoms. Interacts with *hrpG*	UP-24/72/120	(C9) *avrXacE3/egl*	[6,63]
XAC0096	CHP	Interacts with *virD4*	UP-120		[50,72]
XAC0141	CHP	Putative partial CUT locus, CAZy associated TBDR, polygalacturonate reduced TBDR	DOWN-24/120	(C6) *vppA/rpfB/gst*	[71]
XAC0298	CHP	Putative PIP box regulatory protein	UP-120	–	This work
XAC0315	HP	Putative PIP box regulatory protein	UP-72/120	–	This work
XAC0419	CHP	Putative chaperone secretion.	DOWN-24/72/120	–	[50,72]
XAC0464	CHP	Mutant with reduction of symptoms.	DOWN-120	–	[61]
XAC0501	CHP	Homologous to *lipA*.	UP-24/72/120	–	[73]
XAC0601	HP	Putative PIP box regulatory protein. Candidate non-TAL type III effectors of Xoo.	UP-120	–	[74,75,76]
XAC0677	CHP	Protein detected on *Xanthomonas* secretome analysis.	UP-24/72/120	–	[77]
XAC0753	CHP	Protein detected on *Xanthomonas* secretome analysis.	DOWN-72	–	[77]
XAC0754	HP	Putative T3SS-e.	DOWN-24/72	(C5)	[72,78]
XAC1008	CHP	Mutant with reduction of symptoms.	UP-72	–	[64]
XAC1379	CHP	Homologous to Anti-alfa cognate (sigma factor).	UP-120	–	[79]
XAC1381	CHP	Related to copper resistance.	UP-120	–	[80]
XAC1568	CHP	Interacts with HrpG.	UP-72/120	–	[63]
XAC1971	CHP	PilZ - key receptors for cyclic di-GMP.	UP-24/72/120	–	[81]
XAC1990	CHP	Homologous to *flgN*.	UP-24/120	–	[82]
XAC2312	CHP	Homologous to Ps-TBDR/oar of *Xac*.	ND	(C6) *vppA/rpfB/gst*	[71]
XAC2344	CHP	Mutant with reduction of symptoms.	UP-24	–	[64]
XAC2530	CHP	CAZy associated TBDR.	DOWN-72	–	[71]
XAC2606	CHP	Gene reduced in infectious conditions (RT-qPCR analysis).	DOWN-120	(C1) *pthA3/sodM*	This work
XAC2610	HP	Interacts with plasmid proteins and Vir proteins.	DOWN-24/120	(C1) *pthA3/sodM*	[64]
XAC2611	CHP	Gene reduced in infectious conditions (RT-qPCR analysis).	DOWN-24/72/120	(C1) *pthA3/sodM*	This work
XAC2613	HP	Gene reduced in infectious conditions (RT-qPCR analysis).	DOWN-24/72/120	–	This work
XAC2622	CHP	Homologous to *virB7*.	DOWN-24/72/120	–	[66]
XAC2654	CHP	XacPNP protein.	UP-24-72-120	–	[7]
XAC2785	HP	Homologous to *xopF*, that flanks *xopN*. Putative pseudogene.	UP-72/120	–	[74,76,83]
XAC2786	CHP	Homologous to *xopN*.	UP-24/72/120	–	[83]
XAC3073	CHP	Putative partial CUT locus, CAZy associated TBDR, conserved locus.	DOWN-24/120	–	[71]
XAC3085	CHP	Candidate non-TAL type III effectors of Xoo.	UP-24/72/120	–	[74,76]
XAC3178	CHP	Fur regulated TBDR, conserved locus.	UP-72	–	[71]
XAC3206	CHP	CAZy associated TBDR.	UP-24/120	–	[71]
XAC3230	HP	Putative T3SS-e XAC - *xapE2*.	UP-24/72/120	–	[70,74,84,85]
XAC3294	HP	Mutant with reduction of symptoms.	ND	(C6) *vppA/rpfB/gst*	[6]
XAC3314	CHP	Protein detected on *Xanthomonas* secretome.	UP-120	–	This work
XAC3380	HP	Protein detected on *Xanthomonas* secretome.	DOWN-72/120	–	This work
XAC3408	CHP	Homologous to *ftsZ*.	UP-24/72	–	[86]
XAC3525	CHP	Putative CUT locus (unknown substrate).	UP-24/120	–	[71]
XAC3743	CHP	Biofilm production related gene.	DOWN-72	–	[87]
XAC3984	HP	Mutant with reduction of symptoms. Patented.	UP-72/120	(C7) *gumL*	[65]
XAC4131	CHP	TonB dependent transducer.	DOWN-72	–	[88]
XAC4297	CHP	Fur regulated TBDR.	DOWN-72	–	[71]
XAC4300	CHP	Fur regulated TBDR.	DOWN-72/120	–	[71]
XACb0032	CHP	Co-expressed co-purified target protein from Xac.	UP-120	–	[50,72]
XACb0033	CHP	Putative secretion chaperone.	DOWN-24	–	[50,72]
XACb0043	HP	Homologous to *virb7*.	DOWN-120	–	[1,45]

[a] HP — Hypothetical protein; CHP — Conserved hypothetical protein.
[b] UP — Upregulated; DOWN — Downregulated; 24/72/120: H after infection; ND: no difference in expression presented.
[c] (K-means clustering) followed by examples of putative co-regulated genes.

5. Conclusions

Analysis of gene expression performed during the early phases of the infection process enabled the identification of 261 differentially-expressed hypothetical genes. Among these, 49 had been formerly associated with some biological function, frequently related to events that emphasized the virulence potential of Xac. This study reveals the potential and importance of such an analysis to derive greater biological information not only for Xac, but also for other associated species. Furthermore, the appropriate verification of hypothetical gene functions facilitates more targeted investigation in the future to focus on new targets to fight diseases affecting major crops. Thus, this method can also improve the social and economic interests associated with the *Xanthomonas*-host pathosystem.

https://doi.org/10.1016/j.ejbt.2019.10.003

Acknowledgements

We sincerely thank the Fundação de Amparo a Pesquisa do Estado de São Paulo - FAPESP, the Fund for Citrus Plant Protection - FUNDECITRUS, the Coordination of Improvement of Higher Education Personnel - CAPES, National Council for Scientific and Technological Development - CNPq and Fundação de Amparo à Pesquisa do Estado de Minas Gerais for financial support.

Financial support

This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP grant # 02/13862-6), Fund for Citrus Plant Protection (FUNDECITRUS), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the biological tests and fellowship grants. Grants were also received from the

Fundação de Amparo à Pesquisa do Estado de Minas Gerais (CBB-APQ-04425-10) for Xac functional database creation. This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that they have no conflicts of interest.

References

[1] Brunings AM, Gabriel DW. *Xanthomonas citri*: Breaking the surface. Mol Plant Pathol 2003;4(3):141–57. https://doi.org/10.1046/j.1364-3703.2003.00163.x. PMid:20569374.
[2] Sun X, Stall RE, Jones JB, et al. Detection and characterization of a new strain of citrus canker bacteria from key/Mexican lime and Alemow in South Florida. Plant Dis 2004;88(11):1179–88. https://doi.org/10.1094/PDIS.2004.88.11.1179. PMid: 30795311.
[3] Gottwald TR, Graham JH. Canker. In: Timmer L, Garnsey S, Graham J, editors. Compendium of citrus diseases. APS Press; 2000. p. 5–8.
[4] Gottwald TR, Graham JH, Schubert TS. Citrus canker: The pathogen and its impact. Plant Health Progress 2002;3(1). https://doi.org/10.1094/PHP-2002-0812-01-RV.
[5] da Silva ACR, Ferro JA, Reinach FC, et al. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature 2002;417(6887): 459–63. https://doi.org/10.1038/417459a. PMid: 12024217.
[6] Laia ML, Moreira LM, Dezajacomo J, et al. New genes of *Xanthomonas citri* subsp. *citri* involved in pathogenesis and adaptation revealed by a transposon-based mutant library. BMC Microbiol 2009;9:12. https://doi.org/10.1186/1471-2180-9-12. PMid: 19149882.
[7] Gottig N, Garavaglia BS, Daurelio LD, et al. *Xanthomonas axonopodis* pv. *citri* uses a plant natriuretic peptide-like protein to modify host homeostasis. Proc Natl Acad Sci U S A 2008;105(47):18631–6. https://doi.org/10.1073/pnas.0810107105. PMid: 19015524.
[8] Kraiselburd I, Alet AI, Tondo MA, et al. A LOV protein modulates the physiological attributes of *Xanthomonas axonopodis* pv. *citri* relevant for host plant colonization. PLoS ONE 2012;7(6):e38226. https://doi.org/10.1371/journal.pone.0038226. PMid: 2267552.
[9] Soares MR, Facincani AP, Ferreira RM, et al. Proteome of the phytopathogen *Xanthomonas citri* subsp. *citri*: A global expression profile. Proteome Sci 2010;8:55. https://doi.org/10.1186/1477-5956-8-55. PMid: 21062441.
[10] Facincani AP, Moreira LM, Soares MR, et al. Comparative proteomic analysis reveals that T3SS, Tfp, and xanthan gum are key factors in initial stages of *Citrus sinensis* infection by *Xanthomonas citri* subsp. *citri*. Funct Integr Genomics 2013;14(1):205–17. https://doi.org/10.1007/s10142-013-0340-5. PMid: 24676796.
[11] Moreira LM, Facincani AP, Ferreira CB, et al. Chemotactic signal transduction and phosphate metabolism as adaptive strategies during citrus canker induction by *Xanthomonas citri*. Funct Integr Genomics 2014;15(2):197–210. https://doi.org/10.1007/s10142-014-0414-z. PMid: 25403594.
[12] Jalan N, Kumar D, Andrade MO, et al. Comparative genomic and transcriptome analyses of pathotypes of *Xanthomonas citri* subsp. *citri* provide insights into mechanisms of bacterial virulence and host range. BMC Genomics 2013;14:551. https://doi.org/10.1186/1471-2164-14-551. PMid: 23941402.
[13] Doerks T, von Mering C, Bork P. Functional clues for hypothetical proteins based on genomic context analysis in prokaryotes. Nucleic Acids Res 2004;32(21):6321–6. https://doi.org/10.1093/nar/gkh973. PMid: 15576358.
[14] Garbom S, Forsberg A, Wolf-Watz H, et al. Identification of novel virulence-associated genes via genome analysis of hypothetical genes. Infect Immun 2004;72 (3):1333–40. https://doi.org/10.1128/IAI.72.3.1333-1340.2004. PMid: 14977936.
[15] Elias DA, Mukhopadhyay A, Joachimiak MP, et al. Expression profiling of hypothetical genes in *Desulfovibrio vulgaris* leads to improved functional annotation. Nucleic Acids Res 2009;37(9):2926–39. https://doi.org/10.1093/nar/gkp164. PMid: 19293273.
[16] Kolker E, Makarova KS, Shabalina S, et al. Identification and functional analysis of 'hypothetical' genes expressed in *Haemophilus influenzae*. Nucleic Acids Res 2004; 32(8):2353–61. https://doi.org/10.1093/nar/gkh555. PMid: 15121896.
[17] Kolker E, Picone AF, Galperin MY, et al. Global profiling of *Shewanella oneidensis* MR-1: Expression of hypothetical genes and improved functional annotations. Proc Natl Acad Sci U S A 2005;102(6):2099–104. https://doi.org/10.1073/pnas.0409111102. PMid: 15684069.
[18] Cernadas RA, Camillo LR, Benedetti CE. Transcriptional analysis of the sweet orange interaction with the citrus canker pathogens *Xanthomonas axonopodis* pv. *citri* and *Xanthomonas axonopodis* pv. *aurantifolii*. Mol Plant Pathol 2008;9(5):609–31. https://doi.org/10.1111/j.1364-3703.2008.00486.x. PMid: 19018992.
[19] Fu X-Z, Gong X-Q, Zhang Y-X, et al. Different transcriptional response to *Xanthomonas citri* subsp. *citri* between kumquat and sweet orange with contrasting canker tolerance. PLoS ONE 2012;7(7):e41790. https://doi.org/10.1371/journal.pone.0041790. PMid: 22848606.
[20] Pan D, Sun N, Cheung K-H, et al. PathMAPA: A tool for displaying gene expression and performing statistical tests on metabolic pathways at multiple levels for *Arabidopsis*. BMC Bioinformatics 2003;4(56). https://doi.org/10.1186/1471-2105-4-56. PMid: 14604444.
[21] Rapaport F, Zinovyev A, Dutreix M, et al. Classification of microarray data using gene networks. BMC Bioinformatics 2007;8(35). https://doi.org/10.1186/1471-2105-8-35. PMid: 17270037.
[22] Maskos U, Southern EM. Oligonucleotide hybridizations on glass supports: A novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised *in situ*. Nucleic Acids Res 1992;20(7):1679–84. https://doi.org/10.1093/nar/20.7.1679. PMid: 1579459.
[23] Banerjee N, Zhang MX. Functional genomics as applied to mapping transcription regulatory networks. Curr Opin Microbiol 2002;5(3):313–7. https://doi.org/10.1016/S1369-5274(02)00322-3. PMid: 12057687.
[24] Lodha TD, Basak J. Plant-pathogen interactions: What microarray tells about it? Mol Biotechnol 2012;50(1):87–97. https://doi.org/10.1007/s12033-011-9418-2. PMid: 21618071.
[25] Wengelnik K, Rossier O, Bonas U. Mutations in the regulatory gene *hrp*G of *Xanthomonas campestris* pv. vesicatoria result in constitutive expression of all *hrp* genes. J Bacteriol 1999;181(21):6828–31. [PMid: 10542187].
[26] Moreira LM, de Laia ML, de Souza RF, et al. Development and validation of a *Xanthomonas axonopodis* pv. *citri* DNA microarray platform (*XACarray*) generated from the shotgun libraries previously used in the sequencing of this bacterial genome. BMC Res Notes 2010;3(150). https://doi.org/10.1186/1756-0500-3-150. PMid: 20507617.
[27] Koide T, Zaini PA, Moreira LM, et al. DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence. J Bacteriol 2004;186(16):5442–9. https://doi.org/10.1128/JB.186.16.5442-5449.2004. PMid: 15292146.
[28] Conesa A, Nueda MJ, Ferrer A, et al. maSigPro: A method to identify significantly differential expression profiles in time-course microarray experiments. Bioinformatics 2006; 22(9):1096–102. https://doi.org/10.1093/bioinformatics/btl056. PMid: 16481333.
[29] Nueda MJ, Tarazona S, Conesa A. Next maSigPro: Updating maSigPro bioconductor package for RNA-seq time series. Bioinformatics 2014;30(18): 2598–602. https://doi.org/10.1093/bioinformatics/btu333. PMid: 24894503.
[30] Smyth GK. Limma: Linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, et al, editors. Bioinformatics and computational biology solutions using r and bioconductor. New York: Springer; 2005. p. 397–420.
[31] Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics 2005;21(9): 2067–75. https://doi.org/10.1093/bioinformatics/bti270. PMid: 15657102.
[32] Aubert J, Bar-Hen A, Daudin J-J, et al. Determination of the differentially expressed genes in microarray experiments using local FDR. BMC Bioinformatics 2004;5:125. https://doi.org/10.1186/1471-2105-5-125. PMid: 15350197.
[33] Jacob TR, Laia ML, Ferro JA, et al. Selection and validation of reference genes for gene expression studies by reverse transcription quantitative PCR in *Xanthomonas citri* subsp. *citri* during infection of *Citrus sinensis*. Biotechnol Lett 2011;33(6):1177–84. https://doi.org/10.1007/s10529-011-0552-5. PMid: 21318633.
[34] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2_T^{\Delta\Delta C}$ method. Methods 2001;25(4):402–8. https://doi.org/10.1006/meth.2001.1262. PMid: 11846609.
[35] Steibel J, Poletto R, Rosa GJM. Statistical analysis of relative quantification of gene expression using real time RT-PCR data. Animal science AS of, editor. ABSTRACTS – 2005 joint annual meeting, Cincinnati – Ohio: American dairy science association – Canadian Society of Animal Science – American Society of Animal Science, vol. 83; 2005. p. 104.
[36] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28(1):27–30. https://doi.org/10.1093/nar/28.1.27. PMid: 10592173.
[37] Kanehisa M, Goto S, Sato Y, et al. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 2012;40D1:D109–14. https://doi.org/10.1093/nar/gkr988. PMid: 22080510.
[38] Lu Y, Lu S, Fotouhi F, et al. Incremental genetic K-means algorithm and its application in gene expression data analysis. BMC Bioinformatics 2004;5:172. https://doi.org/10.1186/1471-2105-5-172. PMid: 15511294.
[39] Wu F, Zhang W, Kusalik A. A genetic K-means clustering algorithm applied to gene expression data. In: Xiang Y, Chaib-draa B, editors. Advances in artificial intelligence, Canadian AI. Lecture notes in computer science (lecture notes in artificial intelligence). Berlin, Heidelberg: Springer; 2003. p. 520–6.
[40] Sturn A, Quackenbush J, Trajanoski Z. Genesis: Cluster analysis of microarray data. Bioinformatics 2002;18(1):207–8. https://doi.org/10.1093/bioinformatics/18.1.207. PMid: 11836235.
[41] Ryan RP, Vorhoelter F-J, Potnis N, et al. Pathogenomics of *Xanthomonas*: Understanding bacterium-plant interactions. Nat Rev Microbiol 2011;9:344–55. https://doi.org/10.1038/nrmicro2558. PMid: 21478901.
[42] Alegria MC, Souza DP, Andrade MO, et al. Identification of new protein-protein interactions involving the products of the chromosome- and plasmid-encoded type IV secretion loci of the phytopathogen *Xanthomonas axonopodis* pv. *citri*. J Bacteriol 2005;187(7):2315–25. https://doi.org/10.1128/jb.187.7.2315-2325.2005. PMid: 15774874.
[43] Galperin M, Koonin E. "Conserved hypothetical" proteins: Prioritization of targets for experimental study. Nucleic Acids Res 2004;32(18):5452–63. https://doi.org/10.1093/nar/gkh885. PMid: 15479782.
[44] Mazandu GK, Mulder NJ. Function prediction and analysis of mycobacterium tuberculosis hypothetical proteins. Int J Mol Sci 2012;13(6):7283–302. https://doi.org/10.3390/ijms13067283. PMid: 22837694.
[45] Nunes L, Rosato Y, Muto N, et al. Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. Genome Res 2003;13:570–8. https://doi.org/10.1101/gr.930803. PMid: 12670998.
[46] Seo Y-S, Sriariyanun M, Wang L, et al. A two-genome microarray for the rice pathogens *Xanthomonas oryzae* pv. *oryzae* and *X. oryzae* pv. *oryzicola* and its use in

the discovery of a difference in their regulation of *hrp* genes. BMC Microbiol 2008;8: 99. https://doi.org/10.1186/1471-2180-8-99. PMid: 18564427.
[47] Zhou L, Vorhoelter F-J, He Y-Q, et al. Gene discovery by genome-wide CDS re-prediction and microarray-based transcriptional analysis in phytopathogen *Xanthomonas campestris*. BMC Genomics 2011;12(359). https://doi.org/10.1186/1471-2164-12-359. PMid: 21745409.
[48] Fattori J, Prando A, Assis LHP, et al. Structural insights on two hypothetical secretion chaperones from *Xanthomonas axonopodis* pv. *citri*. Protein J 2011;30(5):324–33. https://doi.org/10.1007/s10930-011-9335-z. PMid: 21626158.
[49] Gallo M, Ferrari E, Eliseo T, et al. A new member of the ribbon-helix-helix transcription factor superfamily from the plant pathogen *Xanthomonas axonopodis* pv. *citri*. J Struct Biol 2010;170(1):21–31. https://doi.org/10.1016/j.jsb.2009.12.022. PMid: 20060909.
[50] Tasic L, Borin PF, Khater L, et al. Cloning and characterization of three hypothetical secretion chaperone proteins from *Xanthomonas axonopodis* pv. *citri*. Protein Expr Purif 2007;53(2):363–9. https://doi.org/10.1016/j.pep.2007.01.011. PMid: 17350859.
[51] Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: Discovery and function. Annu Rev Phytopathol 2010;48:419–36. https://doi.org/10.1146/annurev-phyto-080508-081936. PMid: 19400638.
[52] Socquet-Juglard D, Kamber T, Pothier JF, et al. Comparative RNA-Seq analysis of early-infected peach leaves by the invasive phytopathogen *Xanthomonas arboricola* pv. *pruni*. PLoS ONE 2013;8. https://doi.org/10.1371/journal.pone.0054196. PMid: 23342103.
[53] Guo Y, Zhang Y, Li J-L, et al. Diffusible signal factor-mediated quorum sensing plays a central role in coordinating gene expression of *Xanthomonas citri* subsp. *citri*. Mol Plant Microbe Interact 2012;25(2):165–79. https://doi.org/10.1094/MPMI-07-11-0184. PMid: 21995764.
[54] Astua-Monge G, Freitas-Astua J, Bacocina G, et al. Expression profiling of virulence and pathogenicity genes of *Xanthomonas axonopodis* pv. *citri*. J Bacteriol 2005;187 (3):1201–5. https://doi.org/10.1128/JB.187.3.1201-1205.2005. PMid: 15659697.
[55] Puhar A, Sansonetti PJ. Type III secretion system. Curr Biol 2014;24(17):R784–91. https://doi.org/10.1016/j.cub.2014.07.016. PMid: 25202865.
[56] Büttner D, He SY. Type III protein secretion in plant pathogenic bacteria. Plant Physiol 2009;150:1656–64. https://doi.org/10.1104/pp.109.139089. PMid: 19458111.
[57] Teper D, Sunitha S, Martin GB, et al. Five *Xanthomonas* type III effectors suppress cell death induced by components of immunity-associated MAP kinase cascades. Plant Signal Behav 2015;10(10):e1064573. https://doi.org/10.1080/15592324.2015.1064573. PMid: 26237448.
[58] Swarup S, Yang YN, Kingsley MT, et al. An *Xanthomonas citri* pathogenicity gene, *pth*A, pleiotropically encodes gratuitous avirulence on nonhosts. Mol Plant Microbe Interact 1992;5:204–13. https://doi.org/10.1094/MPMI-5-204. PMid: 1421509.
[59] Guttenplan SB, Shaw S, Kearns DB. The cell biology of peritrichous flagella in *Bacillus subtilis*. Mol Microbiol 2013;87(1):211–29. https://doi.org/10.1111/mmi.12103. PMid: 23190039.
[60] Kearns DB. A field guide to bacterial swarming motility. Nat Rev Microbiol 2010;8: 634–44. https://doi.org/10.1038/nrmicro2405. PMid: 20694026.
[61] Yan Q, Wang N. High-throughput screening and analysis of genes of *Xanthomonas citri* subsp. *citri* involved in citrus canker symptom development. Mol Plant Microbe Interact 2012;25(1):69–84. https://doi.org/10.1094/MPMI-05-11-0121. PMid: 21899385.
[62] Wang X-Y, Zhou L, Yang J, et al. The RpfB—Dependent quorum sensing signal turnover system is required for adaptation and virulence in rice bacterial blight pathogen Xanthomonas oryzae pv. oryzae. Mol Plant Microbe Interact 2016;29(3):220–30. https://doi.org/10.1094/MPMI-09-15-0206-R. PMid: 26667598.
[63] Alegria MC, Docena C, Khater L, et al. New protein-protein interactions identified for the regulatory and structural components and substrates of the type III secretion system of the phytopathogen *Xanthomonas axonopodis* pathovar citri. J Bacteriol 2004;186(18):6186–97. https://doi.org/10.1128/JB.186.18.6186-6197.2004. PMid: 15342589.
[64] Ferreira RM, Moreira LM, Ferro JA, et al. Unravelling potential virulence factor candidates in *Xanthomonas citri* subsp. *citri* by secretome analysis. PeerJ 2016;4: e1734. https://doi.org/10.7717/peerj.1734. PMid: 26925342.
[65] da Silva ACR, Farah SC, Quaggio RB, et al. Xanthomonas genes associated with plant pathogenicity. US 20040010131A1, 2002.
[66] Souza DP, Andrade MO, Alvarez-Martinez CE, et al. A component of the Xanthomonadaceae type IV secretion system combines a VirB7 motif with a N0 domain found in outer membrane transport proteins. PLoS Pathog 2011;7(5): e1002031. https://doi.org/10.1371/journal.ppat.1002031. PMid: 21589901.
[67] Souza DP, Oka GU, Alvarez-Martinez CE, et al. Bacterial killing via a type IV secretion system. Nat Commun 2015;6(6453). https://doi.org/10.1038/ncomms7453. PMid: 25743609.
[68] Oliveira LC, Souza DP, Oka GU, et al. VirB7 and VirB9 interactions are required for the assembly and antibacterial activity of a type IV secretion system. Structure 2016;24 (10):1707–18. https://doi.org/10.1016/j.str.2016.07.015. PMid: 27594685.
[69] Vojnov AA, do Amaral AM, Dow JM, et al. Bacteria causing important diseases of citrus utilise distinct modes of pathogenesis to attack a common host. Appl Microbiol Biotechnol 2010;87(2):467–77. https://doi.org/10.1007/s00253-010-2631-2. PMid: 20449739.
[70] Potnis N, Krasileva K, Chow V, et al. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. BMC Genomics 2011;12. https://doi.org/10.1186/1471-2164-12-146. PMid: 21396108.
[71] Blanvillain S, Meyer D, Boulanger A, et al. Plant carbohydrate scavenging through TonB-dependent receptors: A feature shared by phytopathogenic and aquatic bacteria. PLoS ONE 2007;2. https://doi.org/10.1371/journal.pone.0000224. PMid: 17311090.
[72] Lopes TP, Borin PFL, Aoki PS, et al. Comparative analysis of possible type IV secretion system chaperone (XACb0033) and its co-expressed and co-purified target protein (XACb0032) from Xanthomonas axonopodis pv. citri (xac). Resumos de trabalhos científicos - 17a RAU (Reunião Anual dos Usuários do LNLS); 2007. p. 66.
[73] Aparna G, Chatterjee A, Sonti RV, et al. A cell wall-degrading esterase of Xanthomonas oryzae requires a unique substrate recognition module for pathogenesis on rice. Plant Cell 2009;21:1860–73. https://doi.org/10.1105/tpc.109.066886. PMid: 19525415.
[74] Moreira LM, Almeida NF, Potnis N, et al. Novel insights into the genomic basis of citrus canker based on the genome sequences of two strains of Xanthomonas fuscans subsp. aurantifolii. BMC Genomics 2010;11:238. https://doi.org/10.1186/1471-2164-11-238. PMid: 20388224.
[75] Tsuge S, Terashima S, Furutani A, et al. Effects on promoter activity of base substitutions in the cis-acting regulatory element of HrpXo regulons in Xanthomonas oryzae pv. oryzae. J Bacteriol 2005;187(7):2308–14. https://doi.org/10.1128/jb.187.7.2308-2314.2005. PMid: 15774873.
[76] White FF, Yang B. Host and pathogen factors controlling the rice-Xanthomonas oryzae interaction. Plant Physiol 2009;150:1677–86. https://doi.org/10.1104/pp.109.139360. PMid: 19458115.
[77] Ferreira CB, Moreira LM, Brigati JB, et al. Identification of new genes related to virulence of Xanthomonas axonopodis pv. citri during citrus host interactions. Adv Microbiol 2017;7(1):22–46. https://doi.org/10.4236/aim.2017.71003.
[78] Jalan N, Aritua V, Kumar D, et al. Comparative genomic analysis of Xanthomonas axonopodis pv. citrumelo F1, which causes citrus bacterial spot disease, and related strains provides insights into virulence and host specificity. J Bacteriol 2011;193 (22):6342–57. https://doi.org/10.1128/JB.05777-11. PMid: 21908674.
[79] Dunger G, Garofalo CG, Gottig N, et al. Analysis of three *Xanthomonas axonopodis* pv. *citri* effector proteins in pathogenicity and their interactions with host plant proteins. Mol Plant Pathol 2012;13(8):865–76. https://doi.org/10.1111/j.1364-3703.2012.00797.x. PMid: 22435635.
[80] Behlau F, Canteros BI, Minsavage GV, et al. Molecular characterization of copper resistance genes from *Xanthomonas citri* subsp. *citri* and *Xanthomonas alfalfae* subsp. *citrumelonis*. Appl Environ Microbiol 2011;77(12):4089–96. https://doi.org/10.1128/AEM.03043-10. PMid: 21515725.
[81] Li T-N, Chin K-H, Fung K-M, et al. A novel tetrameric PilZ domain structure from xanthomonads. PLoS ONE 2011;6. https://doi.org/10.1371/journal.pone.0022036. PMid: 21760949.
[82] Khater L, Alegria MC, Borin PFL, et al. Identification of the flagellar chaperone FlgN in the phytopathogen *Xanthomonas axonopodis* pathovar *citri* by its interaction with hook-associated FlgK. Arch Microbiol 2007;188(3):243–50. https://doi.org/10.1007/s00203-007-0240-y. PMid: 17492271.
[83] Roden J, Belt B, Ross J, et al. A genetic screen to isolate type III effectors translocated into pepper cells during *Xanthomonas* infection. Proc Natl Acad Sci U S A 2004;101(47):16624–9. https://doi.org/10.1073/pnas.0407383101. PMid: 15545602.
[84] Noël L, Thieme F, Gabler J, et al. XopC and XopJ, two novel type III effector proteins from *Xanthomonas campestris* pv. vesicatoria. J Bacteriol 2003;185 (24):7092–102. https://doi.org/10.1128/JB.185.24.7092-7102.2003. PMid: 14645268.
[85] Stavrinides J, Ma W, Guttman DS. Terminal reassortment drives the quantum evolution of type III effectors in bacterial pathogens. PLoS Pathog 2006;2(10): 913–21. https://doi.org/10.1371/journal.ppat.0020104. PMid: 17040127.
[86] Martins PMM, Lau IF, Bacci M, et al. Subcellular localization of proteins labeled with GFP in *Xanthomonas citri* ssp. *citri*: Targeting the division septum. FEMS Microbiol Lett 2010;310(1):76–83. https://doi.org/10.1111/j.1574-6968.2010.02047.x. PMid: 20629754.
[87] Li J, Wang N. Genome-wide mutagenesis of *Xanthomonas axonopodis* pv. *citri* reveals novel genetic determinants and regulation mechanisms of biofilm formation. PLoS ONE 2011;6(7):e21804. https://doi.org/10.1371/journal.pone.0021804. PMid: 21750733.
[88] Aini LQ, Hirata H, Tsuyumu S. A TonB-dependent transducer is responsible for regulation of pathogenicity-related genes in *Xanthomonas axonopodis* pv. *citri*. J Gen Plant Pathol 2010;76(2):132–42. https://doi.org/10.1007/s10327-010-0227-4.