


RESEARCH ARTICLE

Open Access



Re-programming of *Pseudomonas syringae* pv. *actinidiae* gene expression during early stages of infection of kiwifruit

Peter A. McAtee¹, Lara Brian¹, Ben Curran^{1,2}, Otto van der Linden¹, Niels J. Nieuwenhuizen¹, Xiuyin Chen¹, Rebecca A. Henry-Kirk¹, Erin A. Stroud^{1,2}, Simona Nardoza¹, Jay Jayaraman^{1,3}, Erik H. A. Rikkerink¹, Cris G. Print⁴, Andrew C. Allan^{1,2} and Matthew D. Templeton^{1,2,3*} 

Abstract

Background: *Pseudomonas syringae* is a widespread bacterial species complex that includes a number of significant plant pathogens. Amongst these, *P. syringae* pv. *actinidiae* (*Psa*) initiated a worldwide pandemic in 2008 on cultivars of *Actinidia chinensis* var. *chinensis*. To gain information about the expression of genes involved in pathogenicity we have carried out transcriptome analysis of *Psa* during the early stages of kiwifruit infection.

Results: Gene expression in *Psa* was investigated during the first five days after infection of kiwifruit plantlets, using RNA-seq. Principal component and heatmap analyses showed distinct phases of gene expression during the time course of infection. The first phase was an immediate transient peak of induction around three hours post inoculation (HPI) that included genes that code for a Type VI Secretion System and nutrient acquisition (particularly phosphate). This was followed by a significant commitment, between 3 and 24 HPI, to the induction of genes encoding the Type III Secretion System (T3SS) and Type III Secreted Effectors (T3SE). Expression of these genes collectively accounted for 6.3% of the bacterial transcriptome at this stage. There was considerable variation in the expression levels of individual T3SEs but all followed the same temporal expression pattern, with the exception of *hopAS1*, which peaked later in expression at 48 HPI. As infection progressed over the time course of five days, there was an increase in the expression of genes with roles in sugar, amino acid and sulfur transport and the production of alginate and colanic acid. These are both polymers that are major constituents of extracellular polysaccharide substances (EPS) and are involved in biofilm production. Reverse transcription-quantitative PCR (RT-qPCR) on an independent infection time course experiment showed that the expression profile of selected bacterial genes at each infection phase correlated well with the RNA-seq data.

Conclusions: The results from this study indicate that there is a complex remodeling of the transcriptome during the early stages of infection, with at least three distinct phases of coordinated gene expression. These include genes induced during the immediate contact with the host, those involved in the initiation of infection, and finally those responsible for nutrient acquisition.

Keywords: RNA-seq, Type III effectors, Pathogenicity, Secreted proteins, RT-qPCR

* Correspondence: matt.templeton@plantandfood.co.nz

¹The New Zealand Institute for Plant and Food Research Limited, Auckland, New Zealand

²School of Biological Sciences, University of Auckland, Auckland, New Zealand

Full list of author information is available at the end of the article



Background

Pseudomonas syringae is a widespread bacterial species complex that comprises plant epiphytes and pathogens, as well as being found in non-plant environments such as waterways [1, 2]. Each pathovar of *P. syringae* has a relatively narrow host range related to the specific effector and secondary metabolite profile encoded by its accessory genome. Effectors are proteins that are secreted into plant cells via the Type III Secretion system (T3SS) that function to repress the host defense response [3]. The kiwifruit vine (*Actinidia* Lindl spp.) disease pathogen *P. syringae* pv. *actinidiae* (*Psa*) was first identified in Japan in 1984 [4, 5] and was subsequently found in Korea in the 1990s [6]. Both these strains caused canker symptoms, but did not spread from their country of origin. In 2008, a particularly virulent canker-causing strain of *Psa* was reported in Italy and it quickly decimated plantings of *A. chinensis* var. *chinensis* cultivars, particularly 'Hort16A', 'Hongyang' and 'Jin Tao' [7]. This strain was found in other kiwifruit-growing regions including New Zealand, Chile and China by 2010 [8].

Whole genome sequence analysis was carried out on over 25 strains of *Psa* representing isolates from all locations where *Psa* had been reported. Phylogenetic analysis of the core genome indicated that the canker-causing isolates formed three clades. The first clade comprised the initial isolates from Japan, the second those collected in the 1990s from Korea and the third the pandemic outbreak strains from Italy, New Zealand, Chile and China [9–11]. Isolates within these clades are designated as biovars [12]. The core genome of isolates from the pandemic clade (biovar 3) differed by very few Single Nucleotide Polymorphisms (SNPs) suggesting that this is a clonal population; however, the isolates from New Zealand, Italy, Chile and China each possessed a different member of a family of integrative conjugative elements [9–11]. More recently a comprehensive phylogenetic analysis of eighty *Psa* isolates has shown the origin of the pandemic strains to be China [13]. Two new biovars of *Psa* have been recently discovered in Japan [14, 15], and thus the location of the source population of *Psa* biovars has yet to be conclusively determined.

The three canker-causing biovars each had a surprisingly varied accessory genome with different complements of genes encoding effectors and toxins [11, 16]. Many of these genes are encoded on putative mobile genetic elements. While bioinformatic analysis has identified genes that might be unique to the recent outbreak clade, little is known about the expression of these and other genes that might have a crucial role in pathogenicity. Surprisingly there are few RNA-seq data on the early stages of infection of plants by pathogenic bacteria, including *P. syringae*.

Several transcriptome studies have been carried out on different *P. syringae* pathovars [17, 18]. The most

comprehensive *in planta* analysis has been of *P. syringae* pv. *syringae* (*Pss*) B728a. This pathovar is a particularly successful epiphyte as well as a pathogen of bean (*Phaseolus vulgaris* L.). Global analysis of the transcriptome as an epiphyte, pathogen, and under various stress conditions was carried out using a microarray covering > 5000 coding sequences [19, 20]. The transcript profiles indicated that success as an epiphyte is enabled by flagellar and swarming motility based on surfactant production, chemosensing, and chemotaxis. This could indicate active relocation primarily on the leaf surface. Occupation of an epiphytic niche was accompanied by high transcript levels for phenylalanine degradation, which may help to counteract phenylpropanoid-based plant defenses [19]. In contrast, intercellular or apoplastic colonization led to the high-level expression of genes for γ -aminobutyric acid (GABA) metabolism (degradation of GABA would attenuate GABA repression of virulence) and the synthesis of phytotoxins, syringolin A and two additional secondary metabolites. Perhaps surprisingly the T3SS and T3SEs were not found to be strongly induced in the apoplast [19]. Subsequent analysis of several regulatory mutants illustrated a central role for GacS, SalA, RpoN, and AlgU in global regulation in *Pss* B728a *in planta* and a high degree of plasticity in these transcriptional regulators' responses to distinct environmental signals [20].

More recently a comprehensive analysis of gene expression by *P. syringae* pv. *tomato* DC3000 (*Pto*) has been carried out on wild-type Arabidopsis and several defense gene mutants [21]. T3SS and T3SEs genes were upregulated *in planta*, as were transporter genes. A key finding was that Arabidopsis perturbs iron homeostasis in *Pto* [21].

To gain additional information about the expression of genes involved in pathogenicity, we have carried out transcriptome analysis of *Psa* grown *in vitro* on minimal media and *in planta* during the early stages of kiwifruit infection, using an RNA-seq approach. Our analysis showed that there are at least three distinct coordinated phases of gene expression and has resulted in the discovery of several uncharacterized genes that may have a role in pathogenicity.

Results

Infection assay and RNA-seq time course

Infection assays of kiwifruit tissue-cultured plantlets with *Psa* were performed in flood-inoculated tissue culture vessels under sterile conditions. This method gave consistent and reproducible infection rates, with water-soaked lesions on the underside of leaves progressively developing from day 5, necrotic lesions appearing from day 10 and plant death from four weeks (Additional file 1). A time course was carried out to assess the rate of infection and

to measure and distinguish the relative populations of apoplastic and leaf surface-colonizing (epiphytic) bacteria. Bacterial counts in the apoplast (surface sterilized samples) rose rapidly during the first six days post inoculation and reached a plateau at approximately 10^8 colony-forming units (CFU)/ cm^2 thereafter (Fig. 1). Total bacterial counts (non-surface sterilized samples), which included both apoplastic and epiphytic bacteria, rose from 10^5 to 10^8 CFU/ cm^2 during the time course. The results suggest that for the first two days of infection, the majority of live cells were located epiphytically on the surface of the plant, but that the proportion of apoplastic colonizing bacteria progressively rose from days 2 to 6 so that the apoplast became the predominant niche from that time on.

A time course of five days (120 h) was selected for the RNA-seq analysis, with a focus on very early time points to identify genes induced in the first stages of contact with the plant surface and subsequent infection. It was postulated that key genes responsible for the initiation of infection would be induced at the early stages of contact with the plant surface. Obtaining significant numbers of bacterial reads from infected plants at the early stages of infection is extremely challenging. For this reason, leaves were not surface sterilized before RNA extraction.

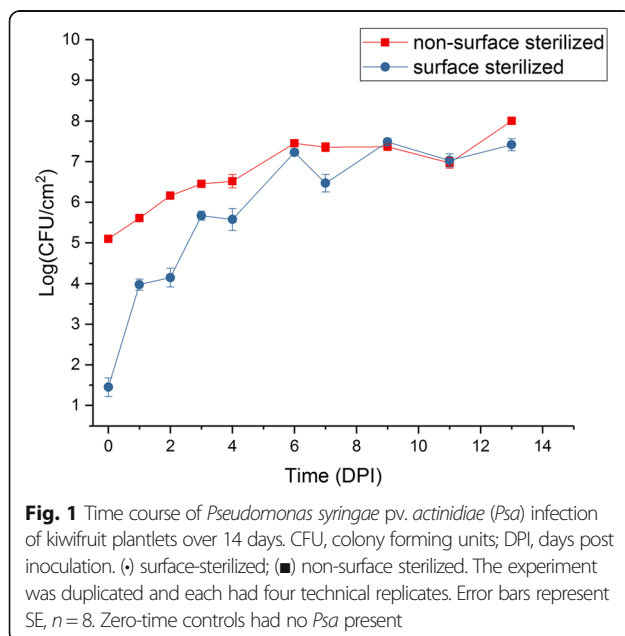
RNA-seq expression profile

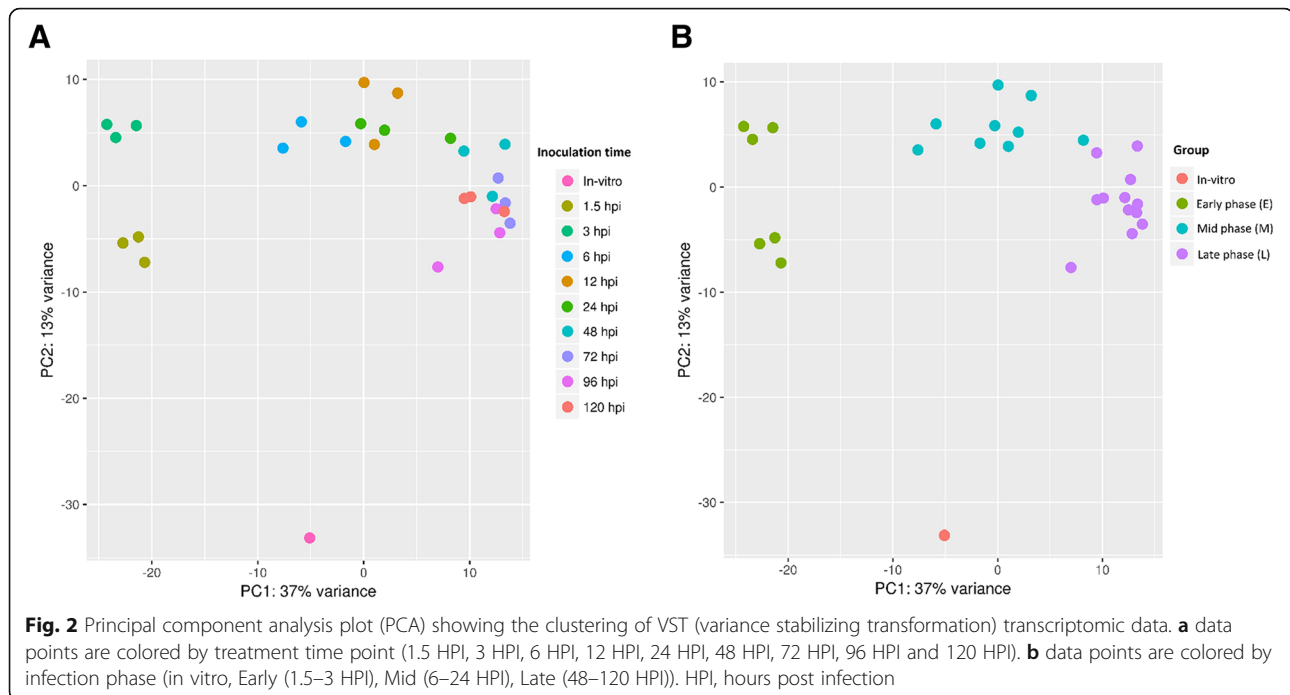
Trimmed reads were mapped onto the complete *Psa* ICMP 18884 genome (CP011972.2 and CP011973) [22]. An average of 50,000–200,000 reads mapped to the *Psa* genome for each time point. This represents between 0.2 and 0.6% of the total reads per sample. The 27 control uninfected treatments showed 50–350 reads mapping to the *Psa* genome (0.0012–0.0002% of the total reads). A

principal component analysis (PCA) was carried out on the inoculated samples to assess overall similarity, and the three biological replicates showed little variance within each time point (Fig. 2a). PCA also demonstrated that each of the *Psa*-infected tissue samples belonged to one of three major phases that closely aligned to the post inoculation period and that were distinct from the in vitro control. The component groupings included the in vitro control, an early phase of infection (1.5 and 3 h post infection, HPI), a mid-phase of infection (6, 12, and 24 HPI), and a late phase of infection (48, 72, 96, and 120 HPI) (Fig. 2b).

Heat map analysis with k-means clustering

Comparison of expression profiles is a powerful tool that can be used to identify and discover genes under the same regulatory regime. Furthermore, it was postulated that novel genes that showed similar expression profiles to known genes involved in pathogenicity might also have a role in causing disease. To identify such genes, similarities in the expression values for each gene were determined by first normalizing expression against its maximum value and then clustering by k-means analysis [23]. This analysis was restricted to those genes displaying Reads Per Kilobase per Million (RPKM) values over 50 for at least one time-point, to eliminate lowly expressed genes from the analysis [24]. Of the 5985 predicted gene models in the core and accessory genomes of *Psa*, 269 genes did not display evidence of being expressed at any sample point, 1473 had no sample point with an RPKM above 50, and 4243 genes had at least one sample point with an RPKM value above 50 (Additional file 2). Hierarchical Clustering on Principal Components (HCPC) using the remaining 4243 genes was used to partition a k-means analysis of genes into 13 clades (clusters) based on their expression profiles (Fig. 3). These were further consolidated into six groups based on their broader expression patterns (Table 1). Of these 4243 genes, 1137 were constitutively expressed, 1323 genes were down-regulated *in planta*, and a further 815 did not show significant differential expression (Table 1). The remaining 968 genes were up-regulated *in planta* compared with in vitro and thus likely to have the most direct relevance to pathogenicity. Of the upregulated genes, there were three distinct groupings that differed in their temporal patterns and level of gene expression over the time course. These groups corresponded to 107 genes induced in the early (1.5 and 3 HPI) time points (early phase), followed by a group of 311 genes highly induced between 3 and 24 HPI (mid phase). The latter group included the majority of the T3SS and T3SE genes controlled by the HrpL regulon. Finally, 550 genes increased in their expression towards the late (48–120 HPI) time points (late phase). These three phases of gene expression were similar to the groupings identified by PCA analysis (Fig. 2). Expression profiles were subsequently evaluated in more detail for





genes with known or as yet undetermined roles in pathogenicity.

Early phase of infection characterized by the induction of a type VI secretion system and nutrient adaptation

Approximately 100 genes were up-regulated immediately upon contact with the host in the early phase of infection (1.5–3.0 HPI). These were found in clade 12 from the clustering analysis (Fig. 3, Table 1). The majority of these genes were annotated as being involved in nutrient acquisition (Additional file 3). This probably reflects the adaptation to the surface of a leaf, where nutrients are scarce [25]. Genes that were particularly highly expressed included those predicted to be involved in phosphate and iron transport. In addition, some genes involved in the degradation of cell wall polymers, including a polygalacturonase (IYO_008325), were in this group. However, few genes predicted to have a direct role in pathogenicity were found. Two of the 43 annotated chemotactic response genes (chemoreceptors) found in the *Psa* genome were highly expressed during the early phase (Additional file 3), but these two are not amongst those previously functionally characterized [26]. These chemoreceptors could have a role in locating stomata or other potential sites of entry into the plant. Another set of genes that was highly induced in this phase encodes a putative Type VI Secretion System (T6SS). Effectors secreted through the T6SS have a variety of roles usually associated with killing both prokaryotic and eukaryotic cells [27]. Roles for T6SS effectors as virulence factors for animal pathogens have been well

documented; however, there is as yet no evidence for an equivalent function in plant pathogens [28, 29]. Alternatively, the T6SS induced by *Psa* may have a role under field conditions in the antagonism of competing epiphytic microbes on the leaf surface.

Mid-phase of infection characterized by expression of T3SS and T3SEs

Genes from three clades (6, 10 and 11) from the clustering analysis show increased expression 3–24 HPI (Table 1). Of these the most striking is a large transcriptional commitment to the induction of the T3SS apparatus and the expression of T3SEs, which were among the most highly upregulated genes within these time points (Fig. 3; Additional file 4). Transcripts encoding for T3SS and T3SEs rose from 1.5 HPI, peaking between 3 and 12 HPI before falling to about half maximal levels for the remainder of the time course. Between 3 and 12 HPI these genes collectively accounted for 6.3% of the total reads (Additional file 4). Expression of *hrpA1* was by far the highest of all T3SS genes, accounting for over 50% of these reads. The HrpA protein comprises the needle of the T3SS apparatus. For plant pathogens the needle is much longer than that of animal pathogens because of the need to penetrate the host cell wall, thus presumably requiring higher expression of the corresponding gene [30].

The *Psa* biovar 3 genome has 40 genes encoding T3SEs, and 35 of these are predicted to encode full-length proteins [11], including one additional T3SE (*hopBNI*) recently identified (<http://pseudomonas-syringae.org/>). The expression profile of T3SEs during the mid-phase of

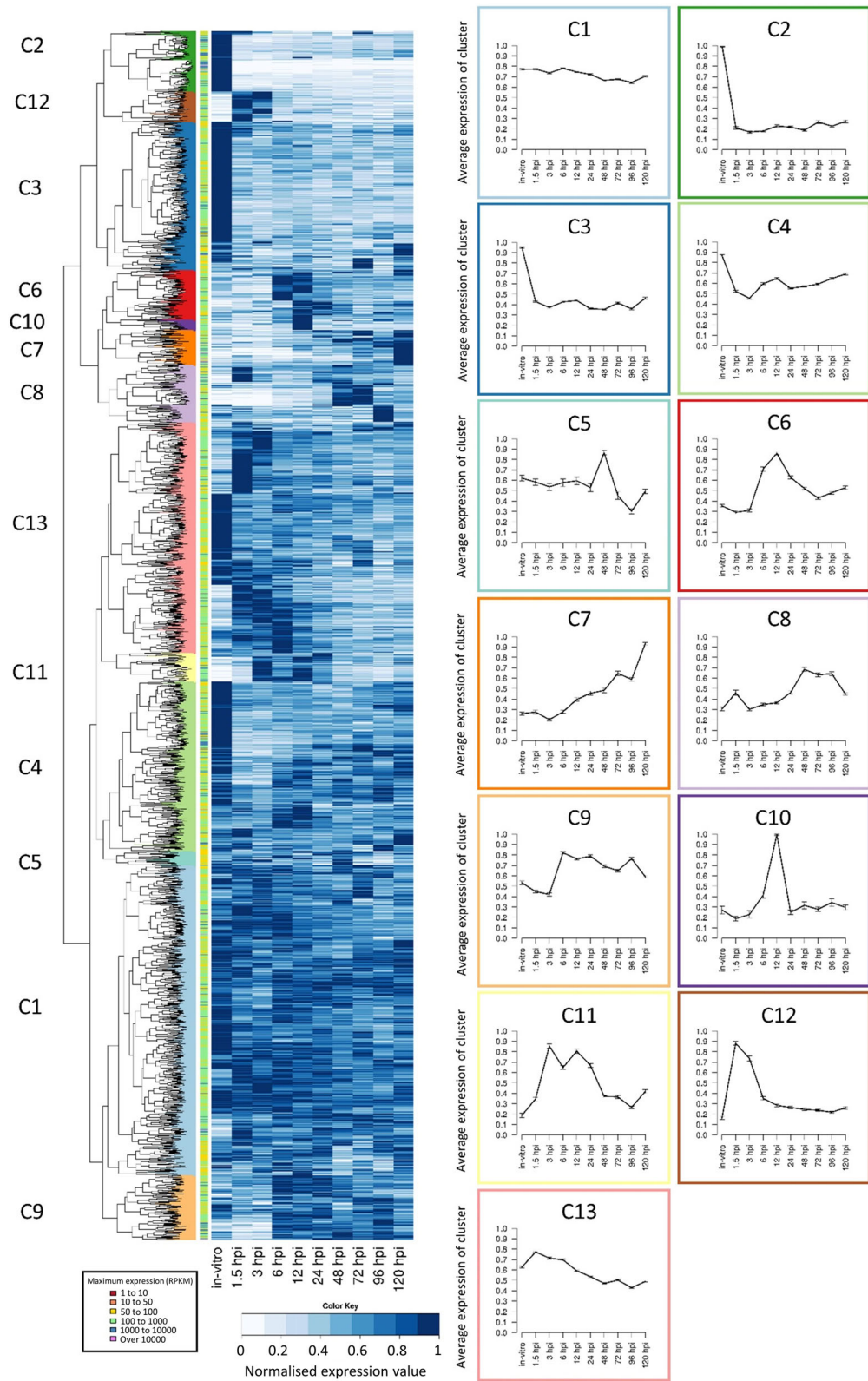


Fig. 3 Heatmap and k-means clustering showing the expression of *Pseudomonas syringae* pv. *actinidiae* (*Psa*) genes in 'Hort16A' kiwifruit plantlets post infection. Similar expression profiles were clustered into 13 distinct groups by k-means. Line graphs displaying the prototype mean expression of each cluster (**c**) are included on the right. Error bars represent standard error

Table 1 k-means clustering of 4243 genes *Psa* into 6 groups. HPI, hours post infection

Group	k-means clade	Number of genes	Description	Expression Phase
1	1 & 13	1137	Constitutively expressed genes	N/A
2	2,3, & 4	1323	Genes down-regulated <i>in planta</i>	N/A
3	5	815	Little differential expression compared to <i>in vitro</i>	N/A
4	6, 10 & 11	311	Genes upregulated (3–24 HPI) <i>in planta</i>	Mid
5	7,8, & 9	550	Genes upregulated late (48–120 HPI) <i>in planta</i>	Late
6	12	107	Early upregulated (1.5–3 HPI) genes <i>in planta</i>	Early

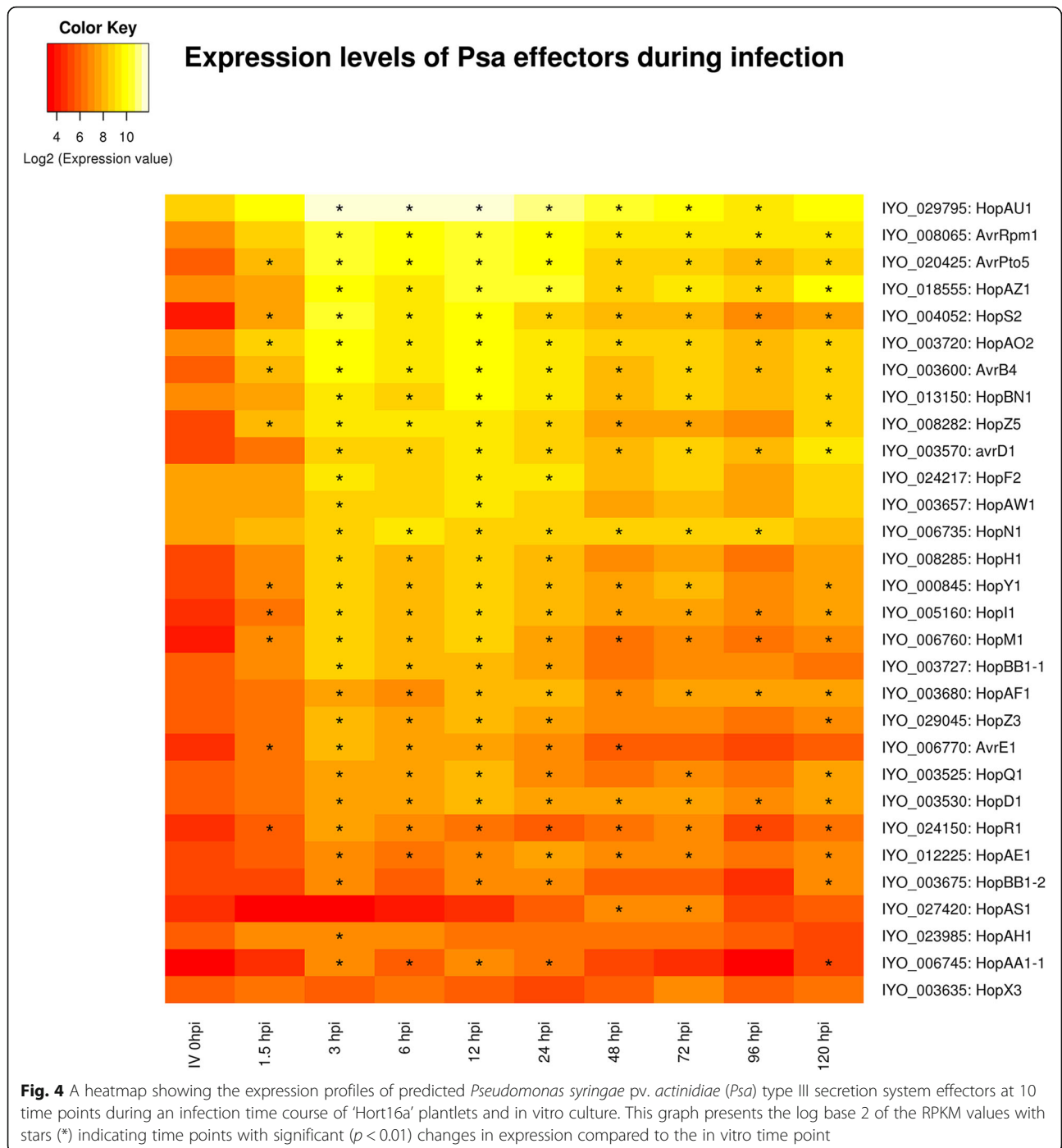
infection followed that of the T3SS transcripts, rising rapidly after 1.5 HPI, with a maximum between 3 and 12 HPI, and then falling for the rest of the time course to around 20–40% of the highest level. The expression levels of each effector varied considerably: most were relatively abundant, in particular *hopAUI*, *hopS2*, *hopAO2*, *hopAZ1*, *hopZ5* and *hopF2*, *avrRpm1*, *avrB4* and *avrPto5* peaked at over 1000 RPKM (Fig. 4, Additional file 5). However, several other effectors were weakly expressed during all the early phases of infection < 150 RPKM, such as *hopAHI* and *hopBBI-2* (Fig. 4, Additional file 5). This may be due to lack of a role for these particular genes in the infection of kiwifruit, expression at the later stages of disease development (after 120 HPI), or a role in the infection of tissues other than leaves.

The effector that displayed the most distinct temporal expression profile was *hopAS1*. This full-length (1361 residue) effector had low expression in the second phase of infection and peaked at 48 HPI. *Pto* strains that are pathogenic on *Arabidopsis thaliana* carry a C-terminal truncated version of this effector (e.g. DC3000 402 residues) and *hopAS1* is widely distributed in *P. syringae* [31]. Full-length versions caused effector-triggered immunity in almost all ecotypes of *Arabidopsis*, explaining why it effectively operates as a barrier to infection in this non-host. In contrast, deletion of the full-length version of *hopAS1* reduced virulence of *Pto* on tomato, suggesting it has a virulence function on this natural host. Both the *Pto* and *Psa* orthologs of this effector have a putative *hrp* box situated upstream of their putative start sites. In between lies a short uncharacterised potential open reading frame which could be an “unrecognised” effector chaperone. The *hopAS1* effector is also found in *P. syringae* pv. *phaseolicola*, where it was not found to be differentially expressed in response to induction of the HrpL TTSS regulatory system [32]. Unfortunately there are no strong clues about the possible biochemical function of HopAS1. It is one of the largest effectors (over 1300 residues, third largest *Psa* effector). Automated searching of the conserved domain database at NCBI identified just one tentative match (Bit score 51; E-value $5.7e^{-6}$) to

a 330 residue portion of a heterodimerization domain in the N-terminus of the chromosome maintenance protein superfamily [33, 34]. Recently HopAS1 was shown to be one of only six T3SEs from *Pto* able to bind to yeast plasma membrane, binding to several different phospho-inositol derivatives [35]. Unfortunately this research appears to have been performed with the truncated version of this gene from *Pto* DC3000. In contrast, the full-length *hopAS1* from *Psa* could not be localised when expressed in *Nicotiana benthamiana*, but this may be because it triggers cell death in that host [36].

T3SS and T3SEs are under the control of the HrpL regulon and hence their co-regulation would be expected. Several other genes that do not code for T3SEs also possess *hrp* boxes 5' to their start site. These include genes that encode a putative lytic transglycosylase (IYO_006775), M20 peptidase (IYO_027210), *apbE* involved in thiamine biosynthesis (IYO_010630), a phosphatidylserine decarboxylase (IYO_025425), and an indole acetic acid-lysine ligase (*iaal*, IYO_002060, Additional file 6). *iaal* is found adjacent to a gene encoding a multidrug and toxic compound extrusion protein (*mate*, IYO_002055) on the chromosome of many *P. syringae* and *P. savastanoi* pathovars. Some *P. savastanoi* pathovars have an additional plasmid-associated *iaal* copy linked with indole acetic acid (IAA) production and gall formation. The proteins encoded by these genes are 92% identical, and the plasmid-located copy has been expressed heterologously and functionally characterized [37]. IAAL is postulated to convert free IAA into less active conjugated forms [38]. Heterologous expression of IAAL in tobacco and potato led to abnormal developmental changes [39]. Transcript levels of *Psa iaal* were induced early in infection and, in contrast to T3SS and T3SEs, remained high throughout the infection period; however, the adjacent *mate* gene did not appear to be highly expressed during this time period (Fig. 5). In *Pto* DC3000 it has been shown that *iaal* can be both transcribed independently and co-transcribed with *mate* as an operon [40].

Other sets of genes that were strongly expressed during the mid-phase of infection (3–12 HPI) included four co-located genes on two operons that code for a diguanylate

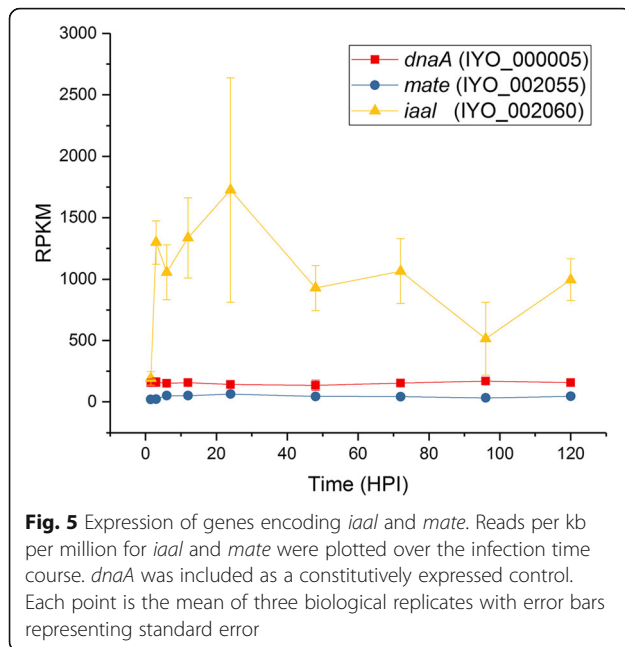


cyclase and two transcription factors, and thus may have a regulatory role (IYO_012110–25) (Additional file 7). Another set of four genes in two operons code for proteins involved in metal transport (IYO_003310–25); included in these is the highly expressed copper resistance/binding protein *copZ* (IYO_003325). Very high expression of the chemotaxis protein IYO_006420 was also observed; while not a membrane-bound chemoreceptor, it is predicted to contain a 4-helix bundle, which is a common chemoreceptor

sensor domain [26]. This protein is predicted to be structurally similar to di-iron binding proteins (Pfam 09537), suggesting an alternate role in iron acquisition as opposed to chemoreception.

Late phase of infection driven by nutrient acquisition and EPS production

A total of 550 genes were upregulated in the later phase of infection (groups 7–9). Ninety genes increased over



5-fold in expression between 1.5 and 120 HPI (Additional file 8). Of these genes, 14 were annotated to be involved in alginate and colanic acid biosynthesis and polymer export. Alginate is a hygroscopic polymer composed of D-mannuronic acid residues interspersed with L-guluronic acid residues with various degrees of acetylation [41]. This polymer has an important role in biofilm production and is well characterized in *P. aeruginosa* [42]. *P. syringae* is also known to produce alginate, but its role in pathogenicity is less well understood [43]. Recently it was shown that alginate accumulates in high amounts in the sub-stomatal spaces in *Psa*-infected leaves of kiwifruit (Sutherland et al., unpublished). A further 26 genes in this grouping were annotated as having a role in metabolite transport. This strongly suggests that as early-stage infection progresses there is a widespread induction of genes involved in metabolite transport and nutrient acquisition. These transporters are distinct from those observed in the early phase of leaf colonization.

Expression of secondary metabolite pathways during infection

Several predicted secondary metabolite biosynthesis pathways have been identified in *Psa* using either antiSMASH 4.0 [44, 45], or by similarity to known biosynthetic pathways (Additional file 9). Three of these pathways, for achromobactin, pyoverdine and yersiniabactin, are involved in iron accumulation. *Psa* biovar 3 produces fluorescent compounds, i.e. pyoverdine, when grown on King's B medium, but to a lesser extent than other *Psa* biovars [8]. The genes that code for this

pathway appear to be poorly expressed *in planta* and on minimal media (Additional file 9). Genes coding for the alternative iron siderophores yersiniabactin and achromobactin are present in *Psa* but were also expressed at low levels *in planta* (Additional file 9). It has recently been postulated that plants are able to interfere with iron homeostasis in pathogenic bacteria, which may explain the apparent lack of expression of these pathways *in planta* [21]. Alternatively, *Psa* may be using a different mechanism for acquiring iron.

Psa-infected kiwifruit leaves show a distinct chlorotic halo which is presumably the result of the diffusion of a phytotoxin [8]. In addition to the three pathways with roles in iron absorption, there were four other secondary metabolite pathways identified in the *Psa* biovar 3 genome that have potential roles in pathogenicity and might account for leaf chlorosis. *Psa* biovar 3 possesses gene clusters involved in the biosynthesis of mangotoxin, a novel non-ribosomal peptide (NRP; IYO_003775–003830), an unknown metabolite (IYO_026725–026760), and an unknown compound synthesized from chorismate; the last-named pathway is plasmid-borne (Additional file 9). The genes involved in mangotoxin biosynthesis, NRP, and the unknown metabolite did not appear to be significantly induced during the early stages of infection, although genes in the NRP pathway were constitutively expressed between 50 and 100 RPKM throughout the infection time course (Additional file 9). While BLAST/antiSMASH searches did not identify likely products of either of the unknown biosynthetic pathways, many *Pseudomonas* spp. produce surfactive molecules to wet the leaf surface to aid motility [46]. In addition, the apoplast is a relatively dry environment that pathogens often modify to increase the relative humidity. For example, *syfA* - an NRP from *Pss* - produces an extremely hygroscopic molecule that facilitates wetting of surfaces including the leaf surface and apoplast [47–49].

The uncharacterized biosynthetic pathway on the plasmid of *Psa* biovar 3 has two operons, and is adjacent to a *luxR* receptor [50]. The first operon codes for a chorismate-utilizing enzyme and a glutamine amidotransferase (annotated as anthranilate synthase I and II) [11]. The second operon codes for the biosynthesis and secretion of a putative aromatic, but uncharacterized, compound that was strongly induced *in planta* after 12 HPI and remained steady for the remainder of the time course (Fig. 6, Additional file 9). The plasmid-localized secondary metabolite pathway is not widespread in *P. syringae* but interestingly is also present in the vascular pathogen *Xylella fastidiosa*, the causal agent of Pierce's disease, and some root-associated *Pseudomonas* species [11].

Proteins secreted through the type II secretion system

In addition to the translocation of proteins through specialized structures such as the T3SS and the T6SS,

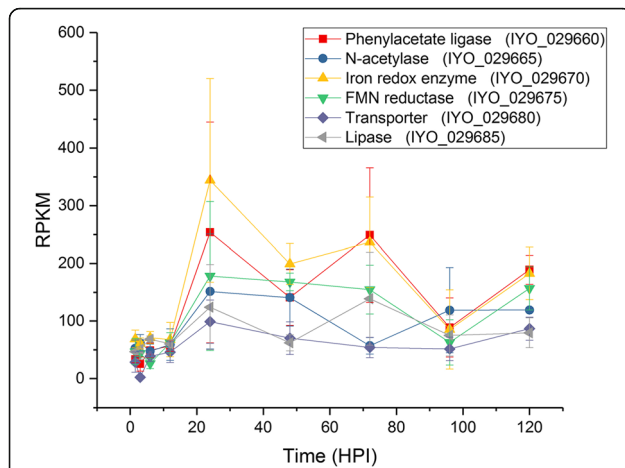


Fig. 6 Expression of genes from the plasmid-borne putative aromatic biosynthetic pathway. Reads per kb per million values for the genes in the operon coding for the biosynthetic pathway of a putative aromatic compound plotted over the infection time course. Each point is the mean of three biological replicates with error bars representing standard error

bacteria also use the Sec or Tat systems to secrete proteins into the periplasm for the Type II secretion system (T2SS) to export [51]. This system will target proteins *in planta* to the apoplast, as opposed to the cytoplasmic location of the T3SEs. This is an important function since the plant apoplast has a number of largely constitutive antimicrobial defenses such as phytoanticipins, hydrolytic enzymes and enzyme inhibitors that may need to be inactivated to facilitate colonization. Recently analysis of the *Pto* secretome identified a protease inhibitor Cip1 as playing a role in virulence of *Pto* on tomato [52]. We were therefore interested to see if there were T2SS proteins upregulated in the early and mid-phases of infection.

Type II secreted proteins (T2SP) can be identified by their canonical secretory leader sequence using SignalP [53]. Five hundred and thirty-nine proteins were predicted to be secreted. Of these proteins, 21 were induced in the early phase of infection (clade12). Of the significantly induced genes (ratio RPKM 3HPI/RPKM *in vitro* > 5), the majority are predicted subunits of membrane-bound complexes with a role in nutrient transport (Additional file 10). All of these had annotations assigned.

Twenty-six proteins with predicted leader sequences were present in the mid phase of infection (Additional file 10). Of those strongly expressed compared to *in vitro* growth, four were annotated as hypothetical proteins. However, two of these predicted gene products have similarity to enzyme inhibitors. IYO_001870 has homology to a superfamily of vertebrate lysozyme inhibitors and IYO_009660 contains a region with homology to Pfam domain 13,670, present in some putative protease inhibitors. Both these proteins may have a role in

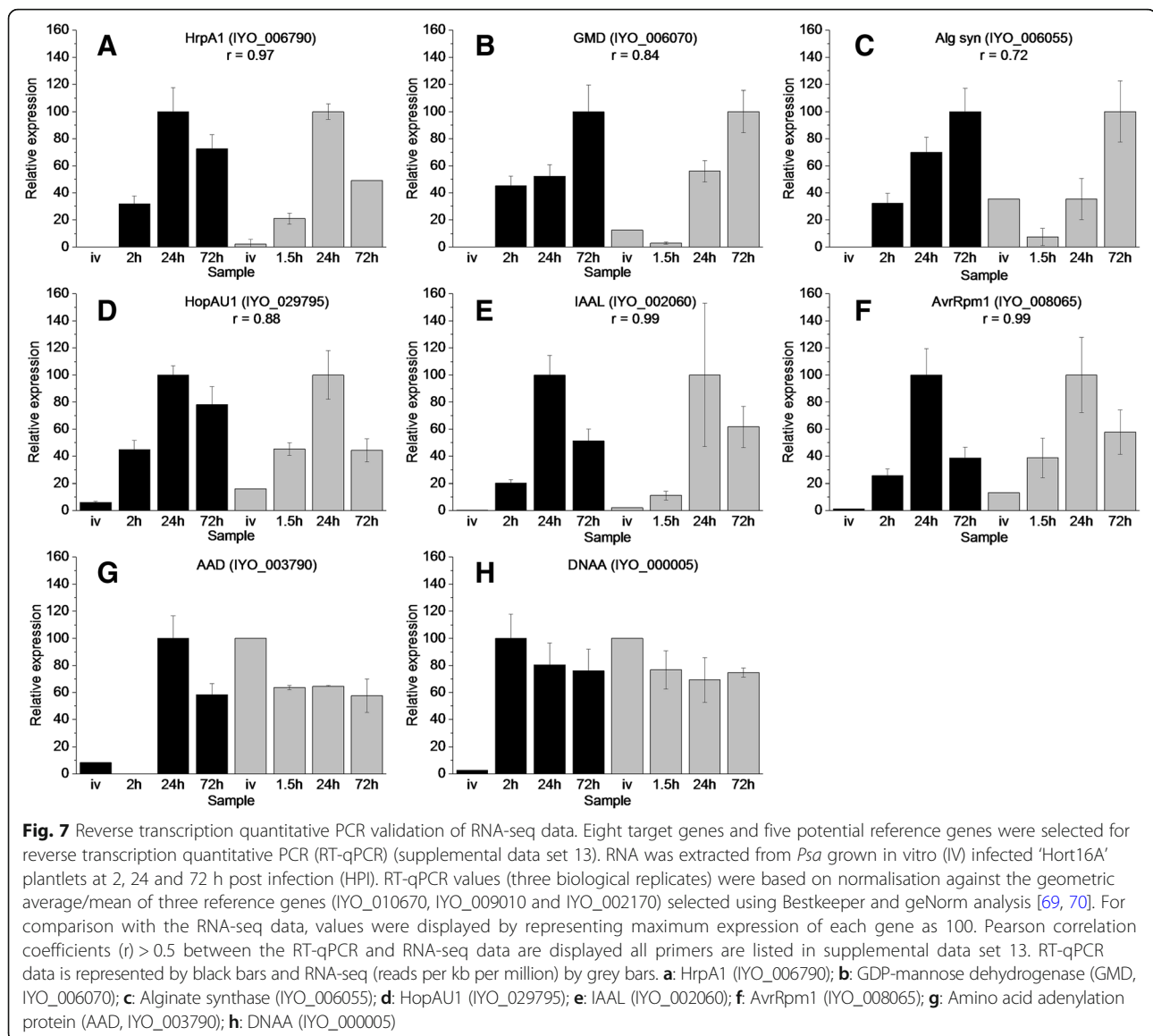
neutralizing the apoplast and are candidates for further functional analysis. Forty-two non-annotated secreted proteins identified in the *Pto* genome were screened for the ability to inhibit the tomato C14 defense-related protease and one, Cip1, was shown to be an inhibitor [52]. Of these, 37 had orthologs (95% sequence identity) in *Psa* biovar 3 but only seven were clearly differentially expressed *in planta*. Interestingly, the *Psa* ortholog of Cip1 (IYO_021465) was not differentially expressed during the time course in this study.

Validation of expression profiles using reverse transcription quantitative PCR

An independent infection time-course experiment was undertaken in order to validate the RNA-seq expression data. A total of 8 genes were selected for reverse transcription quantitative PCR (RT-qPCR) with their expression being assayed *in vitro* (IV) as well as at 2, 24 and 72 HPI (Fig. 7, Additional file 11). These time points represented the early, mid and late phases of infection previously described. When compared to the normalized expression ratio of the RNA-seq data, the normalized qPCR data from 6 of these genes displayed a regression value above 0.7 indicating a high level of consensus between the RNA-seq and qPCR datasets for these genes. Two of the selected genes (IYO 000005 and IYO 003790) displayed regression values below 0.5, however the regression of IYO 000005 increased to greater than 0.7 if the outlying IV time point was omitted. These results show that the expression profile of the bacteria at each infection phase is largely consistent with a different assay methodology and independent experimentation.

Discussion

RNA-seq was used to investigate the early stages of infection of kiwifruit plantlets by *Psa* biovar 3. This biovar is highly virulent on kiwifruit, with apoplastic CFU reaching a plateau from 6 days post inoculation in plantlet leaves. PCA and clustering analysis revealed three phases of gene expression *in planta* during early stages of colonization by *Psa*. The first was a rapid transient phase that occurred immediately upon contact with the plant. Included in these genes was a T6SS, which might have a role in pathogenesis, similarly to that in animal systems [27]. Alternatively, the T6SS may have a role in competition against epiphytic bacteria. Interestingly none of the other genes in this group had a predicted function that could have a direct role in pathogenicity. This suggests that these early expressed *Psa* genes play a role in rapid adaptation to the plant surface, since most of the bacterial counts were on the surface of the plant rather than the apoplast at this stage. Two chemotactic receptors were also highly expressed in this early phase. These are strong candidates for a role in sensing



stomata, hydathodes and other points of entry for the pathogen.

In contrast, the mid phase of infection, which occurred between 3 and 12 HPI, included the T3SS and majority of the T3SEs. These genes were the most upregulated at these time points, accounting for over 6% of the transcripts detected. Similar results were observed for *Pto* colonization of *Arabidopsis* [21]. This is in contrast to the data observed for *Pss* during the early stages of infection of bean, where a large induction of either T3SS or T3SEs was not observed. This might be due to different infection strategies of the two pathogens: *Pss* is regarded as a stronger epiphyte than other *P. syringae* pathovars, because the phylogroup it belongs to (II) has a greater focus on toxin production, and its members typically have fewer effectors than other *P. syringae*

phylogroups [1]. The difference could alternatively be attributed to the different experimental approaches employed [19]. Levels of individual *Psa* T3SE gene expression varied considerably during the time course. However, the temporal expression pattern was largely consistent between effectors, with most (25/30) fitting into the mid-phase gene expression clusters 6, 10 and 11. The most notable exception was *hopAS1*, which peaked later in expression around 48 HPI, as opposed to 3–12 HPI for most other effectors.

The roles of the T3SS and T3SE in repressing the induced host defense response are increasingly well understood. Less well understood is the repression of constitutive plant defenses in the apoplast, the inactivation of which is an essential prerequisite to the establishment of the T3SS. These defenses include phytoanticipins, cell wall degrading

enzymes, proteases and enzyme inhibitors. Furthermore, the apoplast is a relatively dry space that needs to be humidified to optimize colonization [48, 49]. Two resident proteins in the conserved effector locus (CEL), *hopM* and *avrE*, appear to be important in establishing the right humidity conditions in other *P. syringae* hosts [54], and in *Psa* these effectors both follow the mid-peak expression profile but show only average expression levels. This study has also identified two predicted proteins that may have a role in neutralizing the apoplast. One was a predicted lysozyme inhibitor (IYO_001870) and the other a predicted protease inhibitor (IYO_009660). It is likely, however, that there are further genes to be discovered that play a role in neutralizing the apoplast, including the production of potential surfactants.

The final phase comprised of genes whose expression progressively increased over the five-day (120 h) time course. Included were a raft of genes coding for proteins involved in nutrient acquisition such as transporters. Notably, these were different transporters from those induced at the very early phase of infection. There was also strong induction of genes involved in alginate and colanic acid production. These compounds are a large component of the extracellular polysaccharide substances (EPS) and known virulence factors of *Pseudomonas* [41, 43]. Their precise role in pathogenicity is not known, but they have been postulated to protect the bacteria from adversity, in this case plant defenses, and also to enhance adhesion to solid surfaces. Indeed alginate synthesis, along with ice nucleation, auxin synthesis and auxin inactivation by IAAL, is common among the canonical *P. syringae* lineages that have been traced back to a last common ancestor (LCA) 150–180 million years ago [55]. Another component predicted to be derived from the LCA is the tripartite pathogenicity island structure consisting of the *hrp/hrc* gene cluster flanked by both the CEL and an exchangeable effector locus (EEL). *Psa* shows this tripartite structure, albeit that the EEL is further away from the other two pathogenicity islands. The heat map analysis did not highlight any obvious differences in expression patterns between the effectors located on these three pathogenicity islands.

While effectors are well known to play a role in plant defense suppression, the role of many other genes expressed during infection is far less certain. This study has identified a number of non-effector genes that were strongly induced *in planta* and are likely to be having a role in establishing infection. The relative importance of these will need to be ascertained using either gene knock-outs or TraDIS (Transposon Directed Insertion-site Sequencing) [56, 57].

Conclusions

The results from this study indicate that there is a complex remodeling of the bacterial transcriptome during

the early stages of infection, with at least three distinct phases of coordinated gene expression. The first describes genes induced during the immediate contact with the host. These include the expression of a T6SS and genes annotated as involved in nutrient transport. The second phase was dominated by genes predicted to have roles in initiating infection and includes the T3SS and T3SEs. Included in this group are novel proteins that may have roles in neutralizing constitutive defenses in the apoplast. The final phase includes genes involved in nutrient transport and biofilm formation.

Methods

Infection assays

Actinidia chinensis Planch. var. *chinensis* 'Hort16A' plantlets, grown from axillary buds on Murashige and Skoog rooting medium without antibiotics in a 400-mL clear plastic tub with a sealed lid, were purchased from Multiflora (Keeling Rd., Auckland, New Zealand). Plantlets were grown at 20 °C under fluorescent lights with a 16 h on/8 h off regime and used within a month of purchase. For inoculation an overnight shake culture of *Psa* ICMP 18884 [11, 22] was grown in liquid Lysogeny Broth (LB) [58] at 20 °C and 180 rpm shaking. The cell density was determined by measuring the absorbance at 600 nm. Cells were washed in 10 mM MgSO₄ and resuspended at a cell density of 10⁷ CFU/mL (A₆₀₀ 0.01). The surfactant Silwet L-77 (Cat VIS-30, Lehle Seeds, Round Rock, TX, USA) was added to the inoculum to a concentration of 0.0025% (v/v) to facilitate leaf wetting. The inoculation method was modified from that developed for *Arabidopsis* [59]. Containers with 'Hort16A' plantlets were filled with the inoculum fully submerging the plantlets and left for three minutes. Containers were drained, the lid replaced, then incubated in a controlled climate room at 20 °C with a light/dark cycle of 16 h on/8 h off.

Growth assay

Leaf samples were taken at different times post inoculation as appropriate. Each sample consisted of four leaf discs, taken with a 1-cm diameter cork borer, from four different leaves. All four discs were taken from the same tub. To estimate CFU inside the plant (apoplast), discs were surface sterilized in 70% (v/v) ethanol for 30 s and subsequently washed in sterile Milli-Q water. Samples for estimation of total bacteria were not surface sterilized. Leaf discs were placed in Eppendorf tubes containing three stainless steel ball bearings and 300 µL 10 mM MgSO₄, and macerated in a bead crusher for 2 min at maximum speed (Storm 24 Bullet Blender, Next Advance, Averill Park, NY, USA). A dilution series of the leaf homogenates was made in sterile 10 mM MgSO₄ until a dilution of 10⁻⁸. The dilution series was plated in 5-µL droplets on LB medium supplemented with both

12.5 µg/mL nitrofurantoin and 40 µg/mL cephalexin. After 72 h of incubation at 20 °C CFU were counted for the lowest possible dilution(s), which was calculated back to the CFU per cm² of leaf area.

RNA extractions

RNA was extracted for RNA-seq analysis from *Psa* ICMP 18884 grown to late log phase at 18 °C on Hoitnik and Sinden minimal media [60]. Cells were harvested and total RNA extracted using an Ambion RNA extraction kit (Thermo Fisher, Waltham, MA, USA). RNA was extracted from 1-month-old *A. chinensis* var. *chinensis* 'Hort16A' plantlets propagated from tissue culture infected with *Psa* as described above after 1.5, 3, 6, 12, 24, 48, 72, 96 and 120 HPI. Each time point consisted of three biological replicates. Three pots were used for each time point and each biological replicate consisted of three combined plantlets sampled across each of the three pots (Additional file 12). Mock-inoculated plants (submerged in 10 mM MgSO₄ only) were used as controls for each time point. RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich, Milwaukee, WI, USA). Sequencing libraries were constructed from total RNA using the Ribo-Zero Plant procedure (Illumina, San Diego, CA, USA).

Bioinformatics and differential expression analysis

Sequencing was performed using HiSeq2000 (Illumina) by Macrogen (www.macrogen.com).

Raw RNA reads (100 bp paired end reads) were independently trimmed, quality filtered and had their adaptors removed using Trimmomatic v0.36 [61]. The trimming process involved removing the Truseq adaptor sequences and the first 15 base pairs of each read sequence. Regions of low quality calling were also removed from each read using a sliding window of one nucleotide and a quality score above 20. Reads with a length greater than 30 base pairs were selected for alignment following the trimming process.

Processed read sequences were aligned to the *Psa* ICMP 18884 genome sequence [22] using the Bowtie2 v2.25 aligner. HTSeq v0.9.1 was used in conjunction with the BAM outputs generated by Bowtie2 to count the number of alignments against the gene features defined in the corresponding *Psa* ICMP 18884 GFF3 file that had a mapping quality greater than 10 [62].

The count outputs for each gene feature acquired using HTSeq were used for differential expression analysis, principle component analysis, and for the calculation of Reads Per Kilobase per Million (RPKM) values. RPKM values were calculated by dividing the read alignment count for each gene feature in each library by the total number of reads in that library per million and then dividing this by the length of the gene in kilobases.

Analysis of mapped read counts was done using the statistical software R (version 3.4.3). Principal component and differential expression analysis was done using the R package DESeq2 (v1.18.1) [63].

K-means cluster analysis of expression data was done using the R packages ggplot2 (v2.2.1) [64], FactoMineR (v1.39) [65] and FactoExtra (v1.05) [66]. K-means analysis used the average expression of each gene across biological replicates that was normalized against the time point that displayed the highest level of expression for that gene. Hierarchical clustering on principal components was used to decide the 'optimum' minimum number of clusters (k) to partition the output of k-means analysis into. This was achieved by calculating the sum of the within-cluster variance with increasing numbers of clusters. These within cluster variances were plotted as a bar plot (Additional file 13) in order to determine the cluster number(s) that produced a notable loss of inertia (variance). A notable loss of inertia was observed at a cluster number of 2, 3, 6, 7, and 13. Partitioning of the data using a cluster number less than 13 did not adequately describe the patterns of expression in enough detail, while greater numbers generated unnecessary complexity. It was hence decided that a cluster number of 13 represented the optimal minimum number of clusters.

RT-q-PCR validation of RNA-seq data

To validate the RNA-seq dataset, total RNA was isolated from an independent infection experiment with three biological replicates using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich). RNA was DNaseI treated prior to cDNA synthesis to remove any potential genomic DNA contamination (Turbo DNA-free kit, Thermo Fisher Scientific). Relative quantification/Real-Time PCR (q-PCR) primers were designed for five reference genes and eight target genes using the Geneious software package (v10.0.3) (<https://www.geneious.com>) based on an annealing temperature of 60 °C and short (<120 bp) amplicons (supplemental file 13). As the *Psa* transcript content is low in a background of a large amount of total plant RNA, gene-specific reverse transcription (RT) primers were designed to prime the 1st strand cDNA of the *Psa* specific transcripts [67]. Each RT primer (including reference genes – see supplemental file 13) was mixed and used as a cocktail for 1st strand cDNA synthesis at a final concentration of 200 nM with 1 µg of total RNA in a 20 µl reaction according to manufacturing instructions (Superscript IV reverse transcription kit, Invitrogen). After heat denaturing of the RNA at 65 °C for 5 min, RT primers were annealed at 55 °C for 2 min and 1st strand cDNA was synthesized at 53 °C for 20 min followed by 15x dilution prior to qPCR. No RNase H step was included.

RT-qPCR was performed (four technical replicates per sample) using a LightCycler® 480 Real-Time PCR System

(Roche Diagnostics, Indianapolis, USA) using the Light-Cycler® 480 SYBR Green I Master mix and amplified according to manufacturer's instructions (60 °C annealing for 10 s followed by extension at 72 °C for 20 s for 40 cycles) [68]. None of the water (negative template) controls or plant derived samples without *Psa* infection showed any amplification with cross points (CP) < 35. Any samples with CP > 35 were considered not expressed. Primer efficiencies were determined by serial template dilution. The best reference five genes were selected based on the RNA-seq experiment that showed low coefficient of variation (standard deviation divided by mean) and after RT-qPCR were validated using the Bestkeeper [69] and geNorm [70] tools. Based on this analysis, the geometric average CP of three genes was chosen (IYO_010670, IYO_009010 and IYO_002170) to represent the reference. For visual comparison between RT-qPCR and RPKM data, values were normalized by representing maximum expression as 100. Pearson correlation coefficients (r) between RT-qPCR and RPKM data were calculated in Microsoft Excel (CORREL function).

Additional files

Additional file 1: Images illustrating the time course of symptom development of kiwifruit plantlets infected with *Psa*. (A) Abaxial side of leaf at day 0; (B) Adaxial side of leaf at day 0; (C) Pottle containing plantlets at day 0; (D) Abaxial side of leaf five days post inoculation (DPI) with water soaked lesions appearing; (E) Adaxial side of leaf 14 DPI with necrotic lesions present; (F) Plantlets 50 DPI. (PPTX 10001 kb)

Additional file 2: Reads per Kilobase per Million (RPKM) for all *Psa* genes (sheet 1) and k-means analysis for those genes with > 50 RPKM for at least one time point (sheet 2). (XLSX 2078 kb)

Additional file 3: Early induced genes ranked by the ratio of expression at three hours post infection (HPI) compared with in vitro (cutoff 5-fold). (DOCX 20 kb)

Additional file 4: Genes most highly upregulated in the mid phase of the infection time course (3–24 h post infection, (HPI)). Genes are ranked on the ratio of maximum Reads Per Kilobase per Million over that time period compared with in vitro expression (cutoff 5-fold). (DOCX 24 kb)

Additional file 5: Expression levels of individual *Psa* effectors over the infection time course. Effectors are ranked by the highest level of expression between 3 and 12 h post infection (HPI) in reads per kilobase per million (RPKM). Effectors likely to be disrupted, duplicated or pseudogenes were not included. (DOCX 22 kb)

Additional file 6: Expression of non-effector genes with upstream HrpL boxes. Genes are ranked by the ration of expression at 12 h post infection (HPI) compared with in vitro expression. (DOCX 18 kb)

Additional file 7: Genes expressed in the mid phase of infection that do not encode the T3SS or T3SEs. Effectors were ranked by the highest level of expression between 3 and 24 HPI compared with in vitro (cutoff 5-fold). (DOCX 21 kb)

Additional file 8: Genes upregulated late in the time course. Genes were ranked based on the ratio of expression at 120 h post infection (HPI) compared with 1.5 HPI. (DOCX 22 kb)

Additional file 9: Expression levels of secondary metabolite gene clusters. Means of reads per kilobase per million (RPKM) for each gene across all time points with standard deviations. (DOCX 13 kb)

Additional file 10: Reads per kilobase per million (RPKM) values of genes encoding proteins predicted to be secreted via T2SS. (DOCX 25 kb)

Additional file 11: A spreadsheet containing all the primers used for reverse transcription and reverse transcription quantitative PCR. The data includes the gene chosen, locus ID number, primer sequences, product size and PCR efficiency. (XLSX 13 kb)

Additional file 12: RNA-seq experimental design. Three pottles, each with three plantlets, were inoculated for each time point. For RNA extraction, one plantlet from each pottle, was harvested and combined for each of three biological replicates. (DOCX 229 kb)

Additional file 13: A bar plot of the inertia gain using the sum of the within-group variance with increasing cluster number (x-axis) produced using Hierarchical clustering on principal components. (JPG 742 kb)

Abbreviations

AAAD: amino acid adenylation protein; CEL: Conserved effector locus; CFU: Colony-forming units; EEL: Exchangeable effector locus; EPS: Extracellular polysaccharide substances; GABA: γ -aminobutyric acid; GMD: GDP-mannose dehydrogenase; HPCP: Hierarchical clustering on principal components; HPI: Hours post inoculation; IV: in vitro; LB: Lysogeny broth; LCA: Last common ancestor; NRP: Non-ribosomal peptide; PCA: Principal component analysis; *Psa*: *Pseudomonas syringae* pv. *actinidiae*; *Pss*: *Pseudomonas syringae* pv. *syringae*; Pto: *Pseudomonas syringae* pv. *tomato*; RPKM: Reads per kilobase per million; RT: Reverse transcription; RT-qPCR: Reverse transcription quantitative PCR; SNP: Single nucleotide polymorphism; T2SP: Type II secreted proteins; T2SS: Type II secretion system; T3SE: Type III Secreted Effectors; T3SS: Type III Secretion System; T6SS: Type VI Secretion System; TraDIS: Transposon directed insertion-site sequencing; VST: variance stabilizing transformation

Acknowledgements

We would like to thank Dr. Joanna K. Bowen and Amali Thrimawithana (Plant and Food Research) for critically reviewing this manuscript.

Funding

This work was funded by grants from the New Zealand Ministry for Business, Innovation and Employment (C11X1205) and the Bioprotection Centre for Research Excellence (NZ).

Availability of data and materials

The RNA-seq experiment is described in BioProject PRJNA472664 with separate BioSamples for each time-point (SAMN09240241–97), reads can be downloaded from the Sequence Read Archive SRP148711 [71].

Authors' contributions

BC, ACA, NN and CGP conceived of and designed the experiments. BC, RH-K, OvdL, XC, NN, ES and JJ carried out the experiments and generated the data. PM, LB, NN, EHR and MDT analyzed the data. MDT, PM, NN, SN, EHR, and ACA wrote the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate

All experiments using *Psa* were carried out with the permission of the Ministry for Primary Industries, New Zealand (CTO approval 12–05-17) and The Environmental Protection Authority, New Zealand (APP202231).

Consent for publication

N/A

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹The New Zealand Institute for Plant and Food Research Limited, Auckland, New Zealand. ²School of Biological Sciences, University of Auckland, Auckland, New Zealand. ³Bio-Protection Research Centre, Lincoln, New Zealand. ⁴Department of Molecular Medicine and Pathology, University of Auckland, Auckland, New Zealand.

Received: 5 June 2018 Accepted: 23 October 2018

Published online: 15 November 2018

References

- Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, et al. Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates. *PLoS Pathog.* 2011;7(7):e1002132.
- Monteil CL, Cai R, Liu H, Mechan Llontop ME, Leman S, Studholme DJ, et al. Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. *The New phytologist.* 2013; 199(3):800–11.
- Büttner D. Behind the lines—actions of bacterial type III effector proteins in plant cells. *FEMS Microbiol Rev.* 2016;40(6):894–937.
- Serizawa S, Ichikawa T, Takikawa Y, Tsuyumu S, Goto M. Occurrence of bacterial canker of kiwifruit in Japan description of symptoms, isolation of the pathogen and screening of bactericides. *Japanese Journal of Phytopathology.* 1989;55(4):427–36.
- Takikawa Y, Serizawa S, Ichikawa T, Tsuyumu S, Goto M. *Pseudomonas syringae* pv. *actinidiae* pv. Nov the causal bacterium of canker of kiwifruit in Japan. *Japanese Journal of Phytopathology.* 1989;55(4):437–44.
- Koh YCB, Chung H, Lee D. Outbreak and spread of bacterial canker in kiwifruit. *Korean Journal of Plant Pathology.* 1994;10:68–72.
- Balestra GM, Mazzaglia A, Quattrucci A, Renzi M, Rossetti A. Current status of bacterial canker spread on kiwifruit in Italy. *Aust Plant Dis Notes.* 2009;4:34.
- Everett KR, Taylor RK, Romberg MK, Rees-George J, Fullerton RA, Vanneste JL, et al. First report of *Pseudomonas syringae* pv. *actinidiae* causing kiwifruit bacterial canker in New Zealand. *Aust Plant Dis Notes.* 2011;6:67–71.
- Butler MJ, Stockwell PA, Black MA, Day RC, Lamont IL, Poulter RTM. *Pseudomonas syringae* pv. *actinidiae* from recent outbreaks of kiwifruit bacterial canker belong to different clones that originated in China. *PLoS One.* 2013;8(2):18.
- Mazzaglia A, Studholme DJ, Taratufolo MC, Cai RM, Almeida NF, Goodman T, et al. *Pseudomonas syringae* pv. *actinidiae* (Psa) isolates from recent bacterial canker of kiwifruit outbreaks belong to the same genetic lineage. *PLoS One.* 2012;7(5):11.
- McCann HC, Rikkerink EHA, Bertels F, Fiers M, Lu A, Rees-George J, et al. Genomic analysis of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae* provides insight into the origins of an emergent plant disease. *PLoS Pathog.* 2013;9(7):e1003503.
- Cunty A, Poliakkoff F, Rivoal C, Cesbron S, Saux ML, Lemaire C, et al. Characterization of *Pseudomonas syringae* pv. *actinidiae* (Psa) isolated from France and assignment of Psa biovar 4 to a *de novo* pathovar: *Pseudomonas syringae* pv. *actinidifoliorum* pv. Nov. *Plant Pathol.* 2015;64(3):582–96.
- McCann HC, Li L, Liu Y, Li D, Hui P, Zhong C, et al. The origin and evolution of a pandemic lineage of the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae*. *Genome Biology and Evolution.* 2017;9(4):932–44.
- Fujikawa T, Sawada H. Genome analysis of the kiwifruit canker pathogen *Pseudomonas syringae* pv. *actinidiae* biovar 5. *Sci Rep.* 2016;6:21399.
- Sawada H, Kondo K, Nakaune R. Novel biovar (biovar 6) of *Pseudomonas syringae* pv. *actinidiae* causing bacterial canker of kiwifruit (*Actinidia deliciosa*) in Japan. *Jpn J Phytopathol.* 2016;82:101–15.
- Marcelletti S, Ferrante P, Petriccione M, Firrao G, Scortichini M. *Pseudomonas syringae* pv. *actinidiae* draft genomes comparison reveal strain-specific features involved in adaptation and virulence to *Actinidia* species. *PLoS One.* 2011;6(11):17.
- Filiatrault MJ, Stodghill PV, Bronstein PA, Moll S, Lindeberg M, Grills G, et al. Transcriptome analysis of *Pseudomonas syringae* identifies new genes, noncoding RNAs, and antisense activity. *J Bacteriol.* 2010;192(9):2359–72.
- Filiatrault MJ, Stodghill PV, Wilson J, Butcher BG, Chen H, Myers CR, et al. *CrcZ* and *CrcX* regulate carbon source utilization in *Pseudomonas syringae* pathovar *tomato* strain DC3000. *RNA Biol.* 2013;10(2):245–55.
- Yu X, Lund SP, Scott RA, Greenwald JW, Records AH, Nettleton D, et al. Transcriptional responses of *Pseudomonas syringae* to growth in epiphytic versus apoplastic leaf sites. *Proc Natl Acad Sci U S A.* 2013;110(5):E425–34.
- Yu X, Lund SP, Greenwald JW, Records AH, Scott RA, Nettleton D, et al. Transcriptional analysis of the global regulatory networks active in *Pseudomonas syringae* during leaf colonization. *MBio.* 2014;5(5):e01683–14.
- Nobori T, Velásquez AC, Wu J, Kvitko BH, Kremer JM, Wang Y, et al. Transcriptome landscape of a bacterial pathogen under plant immunity. *Proc Natl Acad Sci.* 2018;115(13):E3055–E64.
- Templeton MD, Warren BA, Andersen MT, Rikkerink EHA, Fineran PC. Complete DNA sequence of *Pseudomonas syringae* pv. *actinidiae*, the causal agent of kiwifruit canker disease. *Genome Announcements.* 2015;3(5):e01054–15.
- Liu P, Si Y. Cluster analysis of RNA-sequencing data. In: Datta S, Nettleton D, editors. *Statistical analysis of next generation sequencing data.* Cham: Springer International Publishing; 2014. p. 191–217.
- Lee S, Seo CH, Lim B, Yang JO, Oh J, Lim M, et al. Accurate quantification of transcriptome from RNA-Seq data by effective length normalization. *Nucleic Acids Res.* 2011;39(2):e9.
- Mercier J, Lindow SE. Role of leaf surface sugars in colonization of plants by bacterial epiphytes. *Appl Environ Microbiol.* 2000;66(1):369–74.
- McKellar JLO, Minnell JJ, Gerth ML. A high-throughput screen for ligand binding reveals the specificities of three amino acid chemoreceptors from *Pseudomonas syringae* pv. *actinidiae*. *Mol Microbiol.* 2015;96(4):694–707.
- Kostiuk B, Unterweger D, Provenzano D, Pukatzki S. T6SS intraspecific competition orchestrates *Vibrio cholerae* genotypic diversity. *Int Microbiol.* 2017;20(3):130–7.
- Ho BT, Dong TG, Mekalanos JJ. A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe.* 2014;15(1):9–21.
- Russell AB, Peterson SB, Mougous JD. Type VI secretion system effectors: poisons with a purpose. *Nat Rev Micro.* 2014;12(2):137–48.
- Roine E, Wei W, Yuan J, Nurmiho-Lassila EL, Kalkkinen N, R M, et al. Hrp pilus: an hrp-dependent bacterial surface appendage produced by *Pseudomonas syringae* pv. *tomato* DC3000. *Proc Natl Acad Sci U S A.* 1997; 94(7):3459–64.
- Sohn KH, Saucet SB, Clarke CR, Vinatzer BA, O'Brien HE, Guttman DS, et al. HopA51 recognition significantly contributes to Arabidopsis nonhost resistance to *Pseudomonas syringae* pathogens. *The New phytologist.* 2012; 193(1):58–66.
- Mucyn TS, Yourstone S, Lind AL, Biswas S, Nishimura MT, Baltrus DA, et al. Variable suites of non-effector genes are co-regulated in the type III secretion virulence regulon across the *Pseudomonas syringae* phylogeny. *PLoS Pathog.* 2014;10(1):e1003807.
- Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, et al. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* 2017;45(Database issue):D200–D3.
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 2015; 43(Database issue):D222–6.
- Weigle BA, Orchard RC, Jimenez A, Cox GW, Alto NM. A systematic exploration of the interactions between bacterial effector proteins and host cell membranes. *Nat Commun.* 2017;8:532.
- Choi S, Jayaraman J, Segonzac C, Park H-J, Park H, Han S-W, et al. *Pseudomonas syringae* pv. *actinidiae* type III effectors localized at multiple cellular compartments activate or suppress innate immune responses in *Nicotiana benthamiana*. *Front Plant Sci.* 2017;8:2157.
- Glass NL, Kosuge T. Cloning of the gene for indoleacetic acid-lysine synthetase from *Pseudomonas syringae* subsp. *savastanoi*. *J Bacteriol.* 1986;166(2):598–603.
- Ostrowski M, Mierek-Adamska A, Porowirńska D, Goc A, Jakubowska A. Cloning and biochemical characterization of indole-3-acetic acid-amino acid synthetase PsGH3 from pea. *Plant Physiol Biochem.* 2016;107:9–20.
- Spena A, Prinsen E, Fladung M, Schulze SC, Van Onckelen H. The indoleacetic acid-lysine synthetase gene of *Pseudomonas syringae* subsp. *savastanoi* induces developmental alterations in transgenic tobacco and potato plants. *Mol Gen Genet.* 1991;227(2):205–12.
- Castillo-Lizardo MG, Aragón IM, Carvajal V, Matas IM, Pérez-Bueno ML, Gallegos M-T, et al. Contribution of the non-effector members of the HrpL regulon, *iaaL* and *matE*, to the virulence of *Pseudomonas syringae* pv. *tomato* DC3000 in tomato plants. *BMC Microbiol.* 2015;15:165.
- Franklin MJ, Nivens DE, Weadge JT, Howell PL. Biosynthesis of the *Pseudomonas aeruginosa* extracellular polysaccharides, alginate, pel, and psl. *Front Microbiol.* 2011;2:167.
- Ghafoor A, Hay ID, Rehm BHA. Role of exopolysaccharides in *Pseudomonas aeruginosa* biofilm formation and architecture. *Appl Environ Microbiol.* 2011; 77(15):5238–46.
- Markel E, Stodghill P, Bao Z, Myers CR, Swingle B. AlgU controls expression of virulence genes in *Pseudomonas syringae* pv. *tomato* DC3000. *J Bacteriol.* 2016;198(17):2330–44.
- Weber T, Blin K, Duddela S, Krug D, Kim HU, Brucoleri R et al. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 2015;43(W1):W237–W43.

45. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, et al. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* 2017;45(W1):W36–41.
46. Alsohim AS, Taylor TB, Barrett GA, Gallie J, Zhang X-X, Altamirano-Junqueira AE, et al. The biosurfactant viscosin produced by *Pseudomonas fluorescens* SBW25 aids spreading motility and plant growth promotion. *Environ Microbiol.* 2014;16(7):2267–81.
47. Hockett KL, Burch AY, Lindow SE. Thermo-regulation of genes mediating motility and plant interactions in *Pseudomonas syringae*. *PLoS One.* 2013; 8(3):e59850.
48. Burch AY, Shimada BK, Browne PJ, Lindow SE. Novel high-throughput detection method to assess bacterial surfactant production. *Appl Environ Microbiol.* 2010;76(16):5363–72.
49. Burch AY, Shimada BK, Mullin SWA, Dunlap CA, Bowman MJ, Lindow SE. *Pseudomonas syringae* coordinates production of a motility-enabling surfactant with flagellar assembly. *J Bacteriol.* 2012;194(6):1287–98.
50. Patel HK, Ferrante P, Covaceuszach S, Lamba D, Scortichini M, Venturi V. The kiwifruit emerging pathogen *Pseudomonas syringae* pv. *actinidiae* does not produce AHLs but possesses three luxR solos. *PLoS One.* 2014;9(1):e87862.
51. Green ER, Meccas J. Bacterial secretion systems – an overview. *Microbiology spectrum.* 2016;4(1). <https://doi.org/10.1128/microbiolspec.VMBF-0012-2015>.
52. Shindo T, Kaschani F, Yang F, Kovács J, Tian F, Kourelis J, et al. Screen of non-annotated small secreted proteins of *Pseudomonas syringae* reveals a virulence factor that inhibits tomato immune proteases. *PLoS Pathog.* 2016;12(9):e1005874.
53. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011;8:785.
54. Xin XF, Nomura K, Aung K, Velasquez AC, Yao J, Boutrot F, et al. Bacteria establish an aqueous living space in plants crucial for virulence. *Nature.* 2016;539(7630):524–9.
55. Xin X-F, Kvitko B, He SY. *Pseudomonas syringae*: what it takes to be a pathogen. *Nat Rev Microbiol.* 2018;16(5):316–28.
56. Mesarich CH, Rees-George J, Gardner PP, Ghomi FA, Gerth ML, Andersen MT, et al. Transposon insertion libraries for the characterization of mutants from the kiwifruit pathogen *Pseudomonas syringae* pv. *actinidiae*. *PLoS One.* 2017;12(3):e0172790.
57. Barquist L, Mayho M, C C, Cain AK, Boinett CJ, Page AJ, et al. The TraDIS toolkit: sequencing and analysis for dense transposon mutant libraries. *Bioinformatics.* 2016;32(7):1109–11.
58. Bertani G. Studies on lysogenesis. I. the mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol.* 1951;62(3):293–3000.
59. Ishiga Y, Ishiga T, Uppalapati SR, Mysore KS. Arabidopsis seedling flood-inoculation technique: a rapid and reliable assay for studying plant-bacterial interactions. *Plant Methods.* 2011;7:32.
60. Hoitink HAJ, Sinden SL. Partial purification and properties of chlorosis inducing toxins of *Pseudomonas phaseolicola* and *Pseudomonas glycinea*. *Phytopathology.* 1970;60:1236–7.
61. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20.
62. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25.
63. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
64. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Wolfgang Huber W, Liaw A, et al. Gplots: various R programming tools for plotting data. R package version. 2016;3:1.
65. Le S, Josse J, H F. FactoMineR: an R package for multivariate analysis. *J Stat Softw.* 2008;25(1):1–18.
66. Kassambara A, Mundt F. Factoextra: extract and visualize the results of multivariate data analyses. R package version. 2017;5:1.
67. Liles LC, Kumar MA, Weinschenker D. Use of gene-specific primer cocktails for first-Strand cDNA synthesis with a reverse transcriptase kit. *American Biotechnology Laboratory.* 2004;22:20–1.
68. Chen X, Yauk YK, Nieuwenhuizen NJ, Matich AJ, Wang MY, Perez RL, et al. Characterisation of an (s)-linalool synthase from kiwifruit (*Actinidia arguta*) that catalyses the first committed step in the production of floral lilac compounds. *Funct Plant Biol.* 2010;37:232–43.
69. Pfaffl MW, Tichopád A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – excel-based tool using pair-wise correlations. *Biotechnol Lett.* 2004;26:509–15.
70. Vandesompele J, Preter KD, Pattyn F, Poppe B, Roy NV, Paeppe AD, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology.* 2002;3: research0034.1.
71. Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, et al. Database resources of the National Center for biotechnology information. *Nucleic Acids Res.* 2012;40(Database issue):D13–25.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

