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

Peter A. McAtee[1], Lara Brian[1], Ben Curran[1,2], Otto van der Linden[1], Niels J. Nieuwenhuizen[1], Xiuyin Chen[1], Rebecca Henry-Kirk[1], Erin A. Stroud[1,2], Simona Nardozza[1], Jay Jayaraman[1,3], Erik H. A. Rikkerink[1], Cris G. Print[4], Andrew C. Allan[1,2], Matthew D. Templeton[1,2,3,*]

[*]*Correspondence: Matthew D. Templeton,* matt.templeton@plantandfood.co.nz,*The New Zealand Institute for Plant & Food Research Limited, Auckland, New Zealand*

Running Title: Re-programming of *Psa* gene expression during infection

Keywords: RNA-seq, type III effectors, pathogenicity, secreted proteins, RT-qPCR

Email adresses: PM : peter.mcatee@plantandfood.co.nz; LB : larabrian@hotmail.com; BC : b.curran@auckland.ac.nz; OvdL : ottovanderlinden1@hotmail.com; NN : niels.nieuwenhuizen@plantandfood.co.nz; XC : xiuyin.chen@plantandfood.co.nz; RH-K : rebecca.kirk@plantandfood.co.nz; ES : erin.stroud@plantandfood.co.nz; SN : simona.nardozza@plantandfood.co.nz; JJ : jay.jayaraman@plantandfood.co.nz; EHAR : erik.rikkerink@plantandfood.co.nz; CP : c.print@auckland.ac.nz; AA : andrew.allan@plantandfood.co.nz; MDT : matt.templeton@plantandfood.co.nz.

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.

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 (Figure 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 hours) 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

156 stages of infection is extremely challenging. For this reason, leaves were not surface
157 sterilized before RNA extraction.
158
159 **RNA-seq expression profile**
160 Trimmed reads were mapped onto the complete *Psa* ICMP 18884 genome (CP011972.2
161 and CP011973) [22]. An average of 50,000-200,000 reads mapped to the *Psa* genome for
162 each time point. This represents between 0.2 and 0.6% of the total reads per sample.
163 The 27 control uninfected treatments showed 50-350 reads mapping to the *Psa* genome
164 (0.0012-0.0002% of the total reads). A principal component analysis (PCA) was carried
165 out on the inoculated samples to assess overall similarity, and the three biological replicates
166 showed little variance within each time point (Figure 2A). PCA also demonstrated that each
167 of the *Psa*-infected tissue samples belonged to one of three major phases that closely
168 aligned to the post inoculation period and that were distinct from the *in vitro* control. The
169 component groupings included the *in vitro* control, an early phase of infection (1.5 and 3
170 hours post infection, HPI), a mid-phase of infection (6, 12, and 24 HPI), and a late phase of
171 infection (48, 72, 96, and 120 HPI) (Figure 2B).
172 **Heat map analysis with k-means clustering**
173 Comparison of expression profiles is a powerful tool that can be used to identify and discover
174 genes under the same regulatory regime. Furthermore, it was postulated that novel genes
175 that showed similar expression profiles to known genes involved in pathogenicity might also
176 have a role in causing disease. To identify such genes, similarities in the expression values
177 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 (Figure 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 (Figure 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 (Figure 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.

222

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

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

236 The *Psa* biovar 3 genome has 40 genes encoding T3SEs, and 35 of these are predicted
237 to encode full-length proteins [11], including one additional T3SE (*hopBN1*) recently
238 identified (http://pseudomonas-syringae.org/). The expression profile of T3SEs during the
239 mid-phase of infection followed that of the T3SS transcripts, rising rapidly after 1.5 HPI, with
240 a maximum between 3 and 12 HPI, and then falling for the rest of the time course to around
241 20-40% of the highest level. The expression levels of each effector varied considerably:
242 most were relatively abundant, in particular *hopAU1*, *hopS2*, *hopAO2*, *hopAZ1*, *hopZ5* and
243 *hopF2*, *avrRmp1*, *avrB4* and *avrPto5* peaked at over 1000 RPKM (Figure 4, Additional file

5). However, several other effectors were weakly expressed during all the early phases of infection < 150 RPKM, such as *hopAH1* and *hopBB1-2* (Figure 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 conjugate 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 (Figure

288 5). In *Pto* DC3000 it has been shown that *iaal* can be both transcribed independently and
289 co-transcribed with *mate* as an operon [40].

290 Other sets of genes that were strongly expressed during the mid-phase of infection (3-
291 12 HPI) included four co-located genes on two operons that code for a diguanylate cyclase
292 and two transcription factors, and thus may have a regulatory role (IYO_012110-25)
293 (Additional file 7). Another set of four genes in two operons code for proteins involved in
294 metal transport (IYO_003310-25); included in these is the highly expressed copper
295 resistance/binding protein *copZ* (IYO_003325). Very high expression of the chemotaxis
296 protein IYO_006420 was also observed; while not a membrane-bound chemoreceptor, it is
297 predicted to contain a 4-helix bundle, which is a common chemoreceptor sensor domain
298 [26]. This protein is predicted to be structurally similar to di-iron binding proteins (Pfam
299 09537), suggesting an alternate role in iron acquisition as opposed to chemoreception.

300

301 **Late phase of infection driven by nutrient acquisition and EPS production**

302 A total of 550 genes were upregulated in the later phase of infection (groups 7-9). Ninety
303 genes increased over 5-fold in expression between 1.5 and 120 HPI (Additional file 8). Of
304 these genes, 14 were annotated to be involved in alginate and colanic acid biosynthesis and
305 polymer export. Alginate is a hygroscopic polymer composed of D-mannuronic acid
306 residues interspersed with L-guluronic acid residues with various degrees of acetylation [41].
307 This polymer has an important role in biofilm production and is well characterized in *P.*
308 *aeruginosa* [42]. *P. syringae* is also known to produce alginate, but its role in pathogenicity
309 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 (Figure 6, Additional file 9). The plasmid-localized secondary metabolite pathway is not widespread in *P. syringae* but

353 interestingly is also present in the vascular pathogen *Xylella fastidiosa*, the causal agent of
354 Pierce's disease, and some root-associated *Pseudomonas* species [11].
355
356 **Proteins secreted through the type II secretion system**
357 In addition to the translocation of proteins through specialized structures such as the T3SS
358 and the T6SS, bacteria also use the Sec or Tat systems to secrete proteins into the
359 periplasm for the Type II secretion system (T2SS) to export [51]. This system will target
360 proteins *in planta* to the apoplast, as opposed to the cytoplasmic location of the T3SEs.
361 This is an important function since the plant apoplast has a number of largely constitutive
362 antimicrobial defenses such as phytoanticipins, hydrolytic enzymes and enzyme inhibitors
363 that may need to be inactivated to facilitate colonization. Recently analysis of the *Pto*
364 secretome identified a protease inhibitor Cip1 as playing a role in virulence of *Pto* on
365 tomato [52]. We were therefore interested to see if there were T2SS proteins upregulated
366 in the early and mid-phases of infection.
367 Type II secreted proteins (T2SP) can be identified by their canonical secretory leader
368 sequence using SignalP [53]. Five hundred and thirty-nine proteins were predicted to be
369 secreted. Of these proteins, 21 were induced in the early phase of infection (clade12).
370 Of the significantly induced genes (ratio RPKM 3HPI/RPKM *in vitro* >5), the majority are
371 predicted subunits of membrane-bound complexes with a role in nutrient transport
372 (Additional file 10). All of these had annotations assigned.
373 Twenty-six proteins with predicted leader sequences were present in the mid phase of
374 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 13670, 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 identify) 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 (Figure 7, Additional file 13). 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

397 increased to greater than 0.7 if the outlying IV time point was omitted. These results show
398 that the expression profile of the bacteria at each infection phase is largely consistent with
399 a different assay methodology and independent experimentation.

400

401 ## Discussion

402

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

417 In contrast, the mid phase of infection, which occurred between 3 and 12 HPI, included
418 the T3SS and majority of the T3SEs. These genes were the most upregulated at these
419 time points, accounting for over 6% of the transcripts detected. Similar results were
420 observed for *Pto* colonization of Arabidopsis [21]. This is in contrast to the data observed
421 for *Pss* during the early stages of infection of bean, where a large induction of either T3SS
422 or T3SEs was not observed. This might be due to different infection strategies of the two

423 pathogens: *Pss* is regarded as a stronger epiphyte than other *P. syringae* pathovars,
424 because the phylogroup it belongs to (II) has a greater focus on toxin production, and its
425 members typically have fewer effectors than other *P. syringae* phylogroups [1]. The
426 difference could alternatively be attributed to the different experimental approaches
427 employed [19]. Levels of individual *Psa* T3SE gene expression varied considerably
428 during the time course. However, the temporal expression pattern was largely consistent
429 between effectors, with most (25/30) fitting into the mid-phase gene expression clusters 6,
430 10 and 11. The most notable exception was *hopAS1*, which peaked later in expression
431 around 48 HPI, as opposed to 3-12 HPI for most other effectors.

432 The roles of the T3SS and T3SE in repressing the induced host defense response are
433 increasingly well understood. Less well understood is the repression of constitutive plant
434 defenses in the apoplast, the inactivation of which is an essential prerequisite to the
435 establishment of the T3SS. These defenses include phytoanticipins, cell wall degrading
436 enzymes, proteases and enzyme inhibitors. Furthermore, the apoplast is a relatively dry
437 space that needs to be humidified to optimize colonization [48, 49]. Two resident proteins
438 in the conserved effector locus (CEL), *hopM* and *avrE*, appear to be important in
439 establishing the right humidity conditions in other *P. syringae* hosts [54], and in *Psa* these
440 effectors both follow the mid-peak expression profile but show only average expression
441 levels. This study has also identified two predicted proteins that may have a role in
442 neutralizing the apoplast. One was a predicted lysozyme inhibitor (IYO_001870) and the
443 other a predicted protease inhibitor (IYO_009660). It is likely, however, that there are

444 further genes to be discovered that play a role in neutralizing the apoplast, including the
445 production of potential surfactants.

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

464 While effectors are well known to play a role in plant defense suppression, the role of
465 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 knockouts 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.

Experimental Procedures

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 (http://www.multiflora.co.nz/home.htm). 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 $MgSO_4$ and resuspended at a cell density of 10^7 CFU/mL (A_{600} 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, the 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 $MgSO_4$, 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 $MgSO_4$ 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 hours of incubation at 20°C CFU were counted for the lowest possible dilution(s), which was calculated back to the CFU per cm^2 of leaf area.

RNA extractions

RNA was extracted 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 11). Mock-inoculated plants (submerged in 10mM $MgSO_4$ 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).

526

527 **Bioinformatics and differential expression analysis**

528 Sequencing was performed using HiSeq2000 (Illumina) by Macrogen (www.macrogen.com).
529 Raw RNA reads (100 bp paired end reads) were independently trimmed, quality filtered
530 and had their adaptors removed using Trimmomatic v0.36 [61]. The trimming process
531 involved removing the Truseq adaptor sequences and the first 15 base pairs of each read
532 sequence. Regions of low quality calling were also removed from each read using a sliding
533 window of one nucleotide and a quality score above 20. Reads with a length greater than
534 30 base pairs were selected for alignment following the trimming process.
535 Processed read sequences were aligned to the *Psa* ICMP 18884 genome sequence [22]
536 using the Bowtie2 v2.25 aligner. HTSeq v0.9.1 was used in conjunction with the BAM
537 outputs generated by Bowtie2 to count the number of alignments against the gene features
538 defined in the corresponding *Psa* ICMP 18884 GFF3 file that had a mapping quality greater
539 than 10 [62].
540 The count outputs for each gene feature acquired using HTSeq were used for differential
541 expression analysis, principle component analysis, and for the calculation of Reads Per
542 Kilobase per Million (RPKM) values. RPKM values were calculated by dividing the read
543 alignment count for each gene feature in each library by the total number of reads in that
544 library per million and then dividing this by the length of the gene in kilobases. Analysis of
545 mapped read counts was done using the statistical software R (version 3.4.3). Principal
546 component and differential expression analysis was done using the R package DESeq2
547 (v1.18.1) [63].

548 K-means cluster analysis of expression data was done using the R packages ggplot2 (v2.2.1)
549 [64], FactoMineR (v1.39) [65] and FactoExtra (v1.05) [66]. K-means analysis used the
550 average expression of each gene across biological replicates that was normalized against
551 the time point that displayed the highest level of expression for that gene. Hierarchical
552 clustering on principal components was used to decide the 'optimum' minimum number of
553 clusters (k) to partition the output of k-means analysis into. This was achieved by calculating
554 the sum of the within-cluster variance with increasing numbers of clusters. These within
555 cluster variances were plotted as a bar plot (Additional file 12) in order to determine the
556 cluster number(s) that produced a notable loss of inertia (variance). A notable loss of inertia
557 was observed at a cluster number of 2, 3, 6, 7, and 13. Partitioning of the data using a cluster
558 number less than 13 did not adequately describe the patterns of expression in enough detail,
559 while greater numbers generated unnecessary complexity. It was hence decided that a
560 cluster number of 13 represented the optimal minimum number of clusters.

561

562 **RT-q-PCR validation of RNA-seq data**

563 To validate the RNA-seq dataset, total RNA was isolated from an independent infection
564 experiment with three biological replicates using the Spectrum™ Plant Total RNA Kit
565 (Sigma-Aldrich). RNA was DNaseI treated prior to cDNA synthesis to remove any
566 potential genomic DNA contamination (Turbo DNA-free kit, Thermo Fisher Scientific).
567 Relative quantification/Real-Time PCR (q-PCR) primers were designed for five reference
568 genes and eight target genes using the Geneious software package (v10.0.3)
569 (https://www.geneious.com) based on an annealing temperature of 60°C and short (< 120
570 bp) amplicons (supplemental file 13). As the *Psa* transcript content is low in a
571 background of a large amount of total plant RNA, gene-specific reverse transcription (RT)

572 primers were designed to prime the 1st strand cDNA of the *Psa* specific transcripts [67] .
573 Each RT primer (including reference genes – see supplemental file 13) was mixed and
574 used as a cocktail for 1st strand cDNA synthesis at a final concentration of 200 nM with 1
575 µg of total RNA in a 20 µl reaction according to manufacturing instructions (Superscript IV
576 reverse transcription kit, Invitrogen). After heat denaturing of the RNA at 65°C for 5 min,
577 RT primers were annealed at 55°C for 2 minutes and 1st strand cDNA was synthesized at
578 53°C for 20 minutes followed by 15x dilution prior to qPCR. No RNase H step was
579 included.
580 RT-qPCR was performed (four technical replicates per sample) using a LightCycler®
581 480 Real-Time PCR System (Roche Diagnostics, Indianapolis, USA) using the
582 LightCycler® 480 SYBR Green I Master mix and amplified according to manufacturer's
583 instructions (60°C annealing for 10 s followed by extension at 72 °C for 20 s for 40 cycles)
584 [68]. None of the water (negative template) controls or plant derived samples without *Psa*
585 infection showed any amplification with cross points (CP) <35. Any samples with CP >35
586 were considered not expressed. Primer efficiencies were determined by serial template
587 dilution. The best reference five genes were selected based on the RNA-seq experiment
588 that showed low coefficient of variation (standard deviation divided by mean) and after RT-
589 qPCR were validated using the Bestkeeper [69] and geNorm [70] tools. Based on this
590 analysis, the geometric average CP of three genes was chosen (IYO_010670,
591 IYO_009010 and IYO_002170) to represent the reference. For visual comparison between
592 RT-qPCR and RPKM data, values were normalized by representing maximum expression
593 as 100. Pearson correlation coefficients (r) between RT-qPCR and RPKM data were
594 calculated in Microsoft Excel (CORREL function).
595

Abbreviations

AAD: amino acid adenylation protein; CEL: Conserved effector locus; CFU: Colony-forming units; EEL: Exchangeable effector locus; GABA: *γ*-aminobutyric acid; GMD: GDP-mannose dehydrogenase; HPCP: Hierarchical clustering on principal components; HPI: Hours post inoculation; EPS: Extracellular polysaccharide substances; 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; T2SS: Type II secretion system; T2SP: Type II secreted proteins; T3SS: Type III Secretion System; T3SE: Type III Secreted Effectors; T6SS: Type VI Secretion System; TraDIS: Transposon directed insertion-site sequencing; VST: variance stabilizing transformation.

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 to publish

N/A

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].

Competing interests

The authors declare that they have no competing interests.

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).

Author 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.

Acknowledgements

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

640 **Author details**

641 [1]The New Zealand Institute for Plant and Food Research Limited, Auckland, New Zealand

642 [2]School of Biological Sciences, University of Auckland, Auckland, New Zealand

643 [3]Bio-Protection Research Centre, New Zealand

644 [4]Department of Molecular Medicine and Pathology, University of Auckland, New Zealand

645

References

1. 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.
2. 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.
3. Büttner D. Behind the lines–actions of bacterial type III effector proteins in plant cells. FEMS Microbiology Reviews. 2016; 40 (6):894-937.
4. 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.
5. 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.
6. Koh Y CB, Chung H, Lee D. Outbreak and spread of bacterial canker in kiwifruit. Korean Journal of Plant Pathology. 1994; 10:68-72.
7. Balestra GM, Mazzaglia A, Quattrucci A, Renzi M, Rossetti A. Current status of bacterial canker spread on kiwifruit in Italy. Australasian Plant Disease Notes. 2009; 4:34.
8. 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. Australasian Plant Disease Notes. 2011; 6:67-71.
9. Butler MI, 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.
10. 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.
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.
12. Cunty A, Poliakoff 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.
13. 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.
14. Fujikawa T, Sawada H. Genome analysis of the kiwifruit canker pathogen *Pseudomonas syringae* pv. *actinidiae* biovar 5. Scientific Reports. 2016; 6:21399.
15. 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.
16. 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.

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. Journal of Bacteriology. 2010; 192 (9):2359-72.
18. 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 Biology. 2013; 10 (2):245-55.
19. 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.
20. 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.
21. Nobori T, Velásquez AC, Wu J, Kvitko BH, Kremer JM, Wang Y et al. Transcriptome landscape of a bacterial pathogen under plant immunity. Proceedings of the National Academy of Sciences. 2018; 115 (13):E3055-E64.
22. Templeton MD, Warren BA, Andersen MT, Rikkerink EHA, Fineran PC. Complete DNA sequence of *Pseudomonas syringae* pv. *a*ctinidiae, the causal agent of Kiwifruit canker disease. Genome Announcements. 2015; 3 (5):e01054-15.
23. Liu P, Si Y: Cluster analysis of RNA-sequencing data. In: Statistical Analysis of Next Generation Sequencing Data. Edited by Datta S, Nettleton D. Cham: Springer International Publishing; 2014: 191-217.
24. 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 Research. 2011; 39 (2):e9.
25. Mercier J, Lindow SE. Role of leaf surface sugars in colonization of plants by bacterial epiphytes. Applied and Environmental Microbiology. 2000; 66 (1):369-74.
26. 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*. Molecular Microbiology. 2015; 96 (4):694-707.
27. Kostiuk B, Unterweger D, Provenzano D, Pukatzki S. T6SS intraspecific competition orchestrates *Vibrio cholerae* genotypic diversity. Int Microbiol. 2017; 20 (3):130-7.
28. 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.
29. Russell AB, Peterson SB, Mougous JD. Type VI secretion system effectors: poisons with a purpose. Nat Rev Micro. 2014; 12 (2):137-48.
30. Roine E, Wei W, Yuan J, Nurmiaho-Lassila EL, Kalkkinen N, M. R et al. Hrp pilus: an hrp-dependent bacterial surface appendage produced by *Pseudomonas syringae* pv. *tomato* DC3000. Proc Natl Acad Sci USA. 1997; 94 (7):3459-64.
31. Sohn KH, Saucet SB, Clarke CR, Vinatzer BA, O'Brien HE, Guttman DS et al. HopAS1 recognition significantly contributes to Arabidopsis nonhost resistance to *Pseudomonas syringae* pathogens. The New phytologist. 2012; 193 (1):58-66.
32. 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 pathogens. 2014; 10 (1):e1003807.
33. 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 Research. 2017; 45 (Database issue):D200-D3.
34. 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.
35. Weigele BA, Orchard RC, Jimenez A, Cox GW, Alto NM. A systematic exploration of the interactions between

bacterial effector proteins and host cell membranes. Nature Communications. 2017; 8:532.

36. 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*. Frontiers in Plant Science. 2017; 8 (2157)

37. Glass NL, Kosuge T. Cloning of the gene for indoleacetic acid-lysine synthetase from *Pseudomonas syringae* subsp. *savastanoi*. Journal of Bacteriology. 1986; 166 (2):598-603.

38. Ostrowski M, Mierek-Adamska A, Porowińska D, Goc A, Jakubowska A. Cloning and biochemical characterization of indole-3-acetic acid-amino acid synthetase PsGH3 from pea. Plant Physiology and Biochemistry. 2016; 107:9-20.

39. 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.

40. 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 Microbiology. 2015; 15:165.

41. Franklin MJ, Nivens DE, Weadge JT, Howell PL. Biosynthesis of the *Pseudomonas aeruginosa* extracellular polysaccharides, alginate, pel, and psl. Frontiers in Microbiology. 2011; 2:167.

42. Ghafoor A, Hay ID, Rehm BHA. Role of exopolysaccharides in *Pseudomonas aeruginosa* biofilm formation and architecture. Applied and Environmental Microbiology. 2011; 77 (15):5238-46.

43. Markel E, Stodghill P, Bao Z, Myers CR, Swingle B. AlgU controls expression of virulence genes in *Pseudomonas syringae* pv. *tomato* DC3000. Journal of Bacteriology. 2016; 198 (17):2330-44.

44. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R et al. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Research. 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 Research. 2017; 45 (W1):W36-W41.

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. Environmental Microbiology. 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. Applied and Environmental Microbiology. 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. Journal of Bacteriology. 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, Mecsas J. Bacterial secretion systems – an overview. Microbiology spectrum. 2016; 4 (1):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 pathogens. 2016; 12 (9):e1005874.

53. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from

transmembrane regions. Nature 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. Nature Reviews Microbiology. 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 Biology. 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 301 2016;

65. Le S, Josse J, F. H. FactoMineR: An R package for multivariate analysis. Journal of Statistical Software. 2008; 25 (1):1-18.

66. Kassambara A, Mundt F. factoextra: extract and visualize the results of multivariate data analyses. R package version 105. 2017;

67. Liles LC, Kumar MA, Weinshenker 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. Functional Plant Biology. 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. Biotechnology Letters 2004; 26:509-15.

70. Vandesompele J, Preter KD, Pattyn F, Poppe B, Roy NV, Paepe 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.

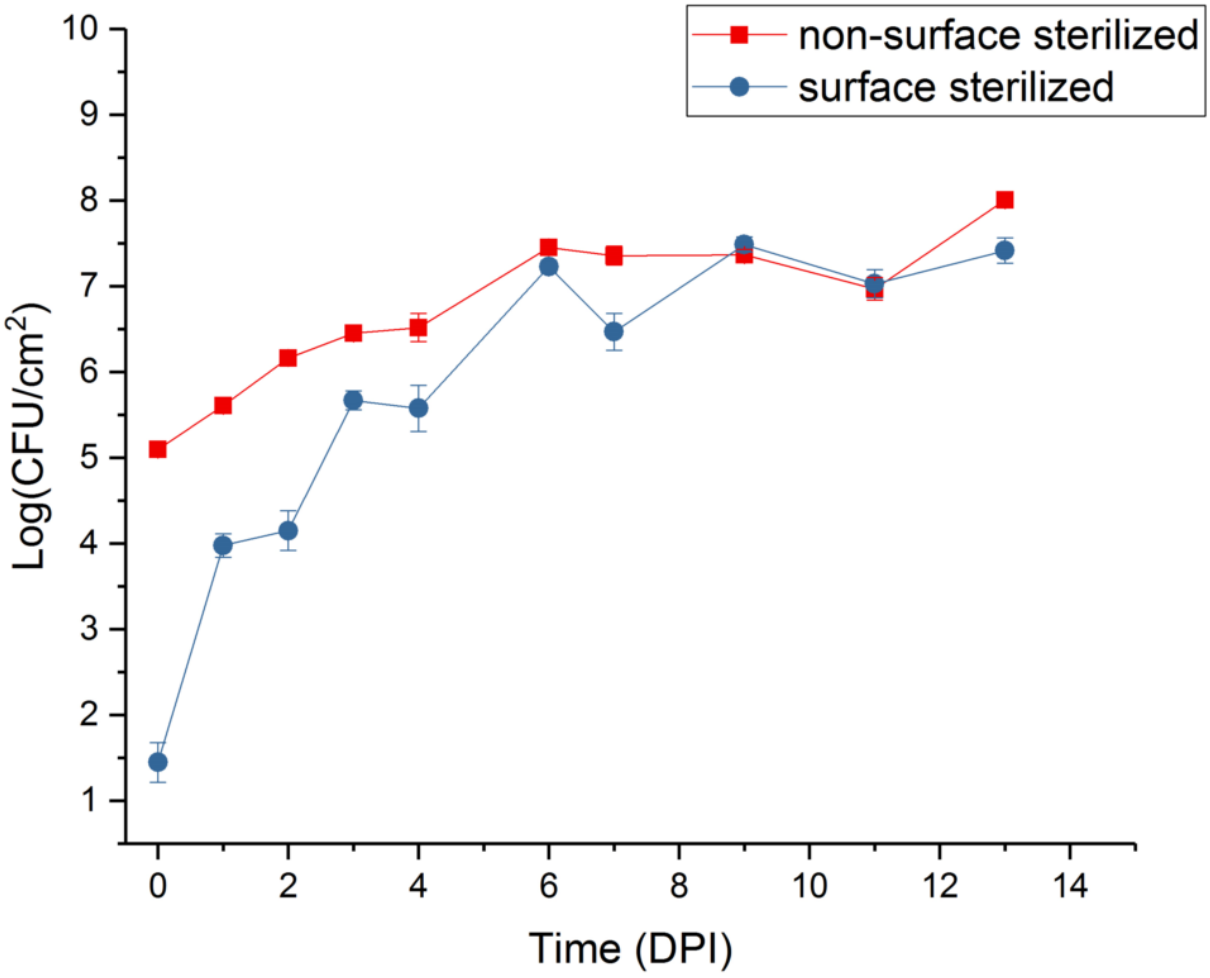
820 Table 1. k-means clustering of 4243 genes Psa into 6 groups. HPI, hours post infection.

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 *in planta*	N/A
3	5	815	Little differential expression compared to *in vitro*	N/A
4	6, 10 & 11	311	Genes upregulated (3-24 HPI) *in planta*	Mid
5	7,8, & 9	550	Genes upregulated late (48-120 HPI) *in planta*	Late
6	12	107	Early upregulated (1.5-3 HPI) genes *in planta*	Early

821

822

823 Figure 1. Time course of *Pseudomonas syringae* pv. *actinidiae* (*Psa*) infection of kiwifruit
824 plantlets over 14 days. CFU, colony forming units; DPI, days post inoculation. (●)
825 surface-sterilized; (■) non-surface sterilized. The experiment was duplicated and each
826 had four technical replicates. Error bars represent SE, n=8. Zero-time controls had no
827 *Psa* present.
828



829
830
831

Figure 2. Principal component analysis plot (PCA) showing the clustering of VST (variance stabilizing transformation) transcriptomic data. (A) data points are colored by treatment time point (1.5 HPI, 3 HPI, 6 HPI, 12 HPI, 24 HPI, 48 HPI, 72 HPI, 96 HPI and 120 HPI). (B) data points are colored by infection phase (*in vitro*, Early (1.5-3 HPI), Mid (6-24 HPI), Late (48-120 HPI)). HPI, hours post infection.

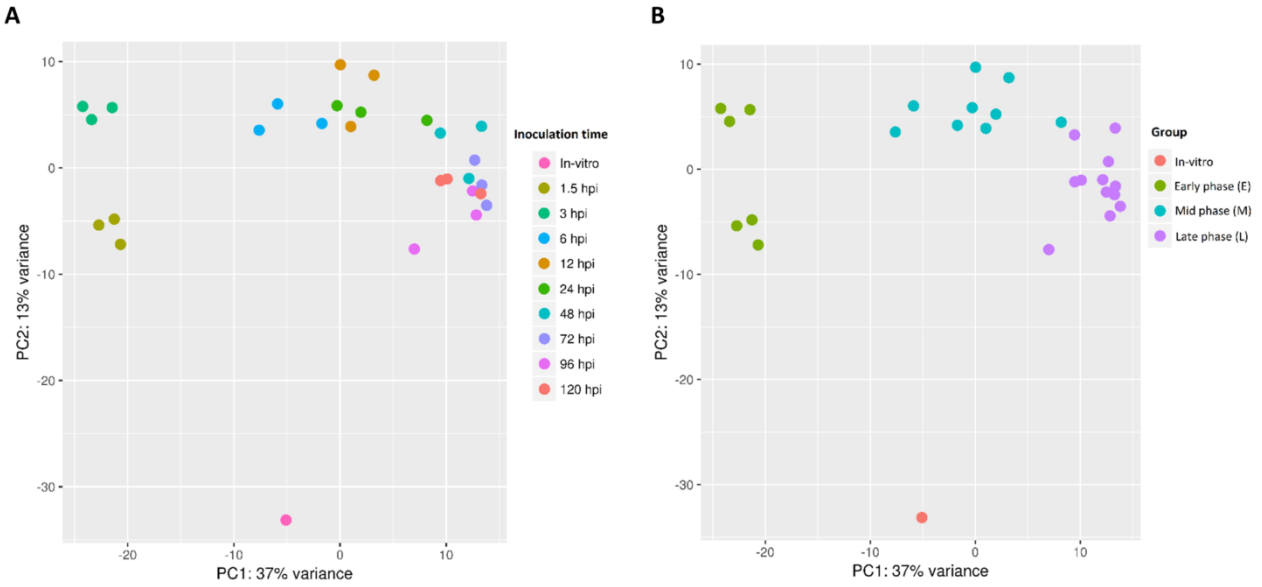


Figure 3. Heatmap and k-means clustering showing the expression of *Pseudomonas syringae* pv. *actinidiae* (*Psa*) genes in 'Hort16A' kiwifruit plantlets post infection. Similar expression profiles were clustered into 13 distinct groups by k-means. Line graphs displaying the prototype mean expression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

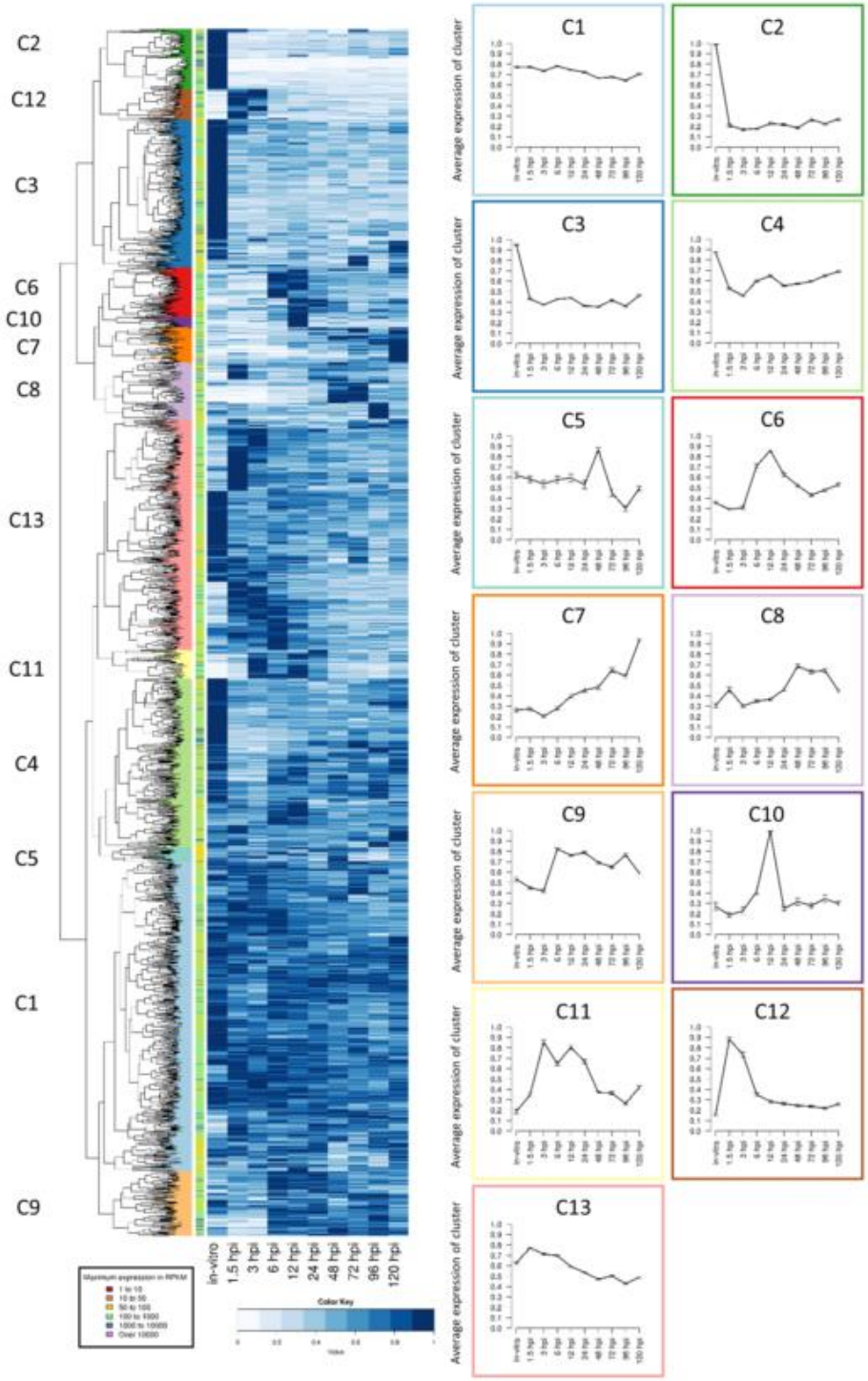
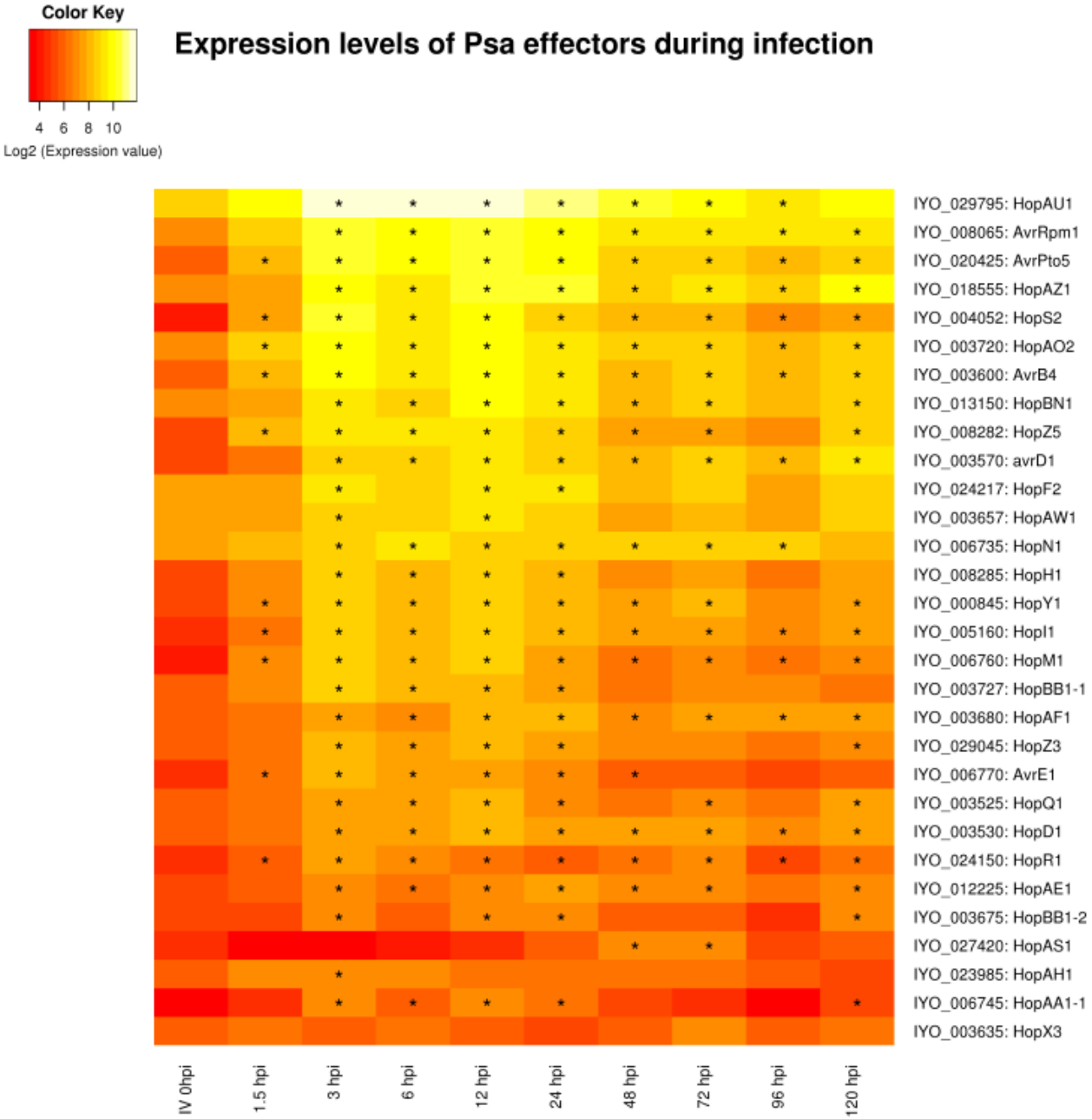
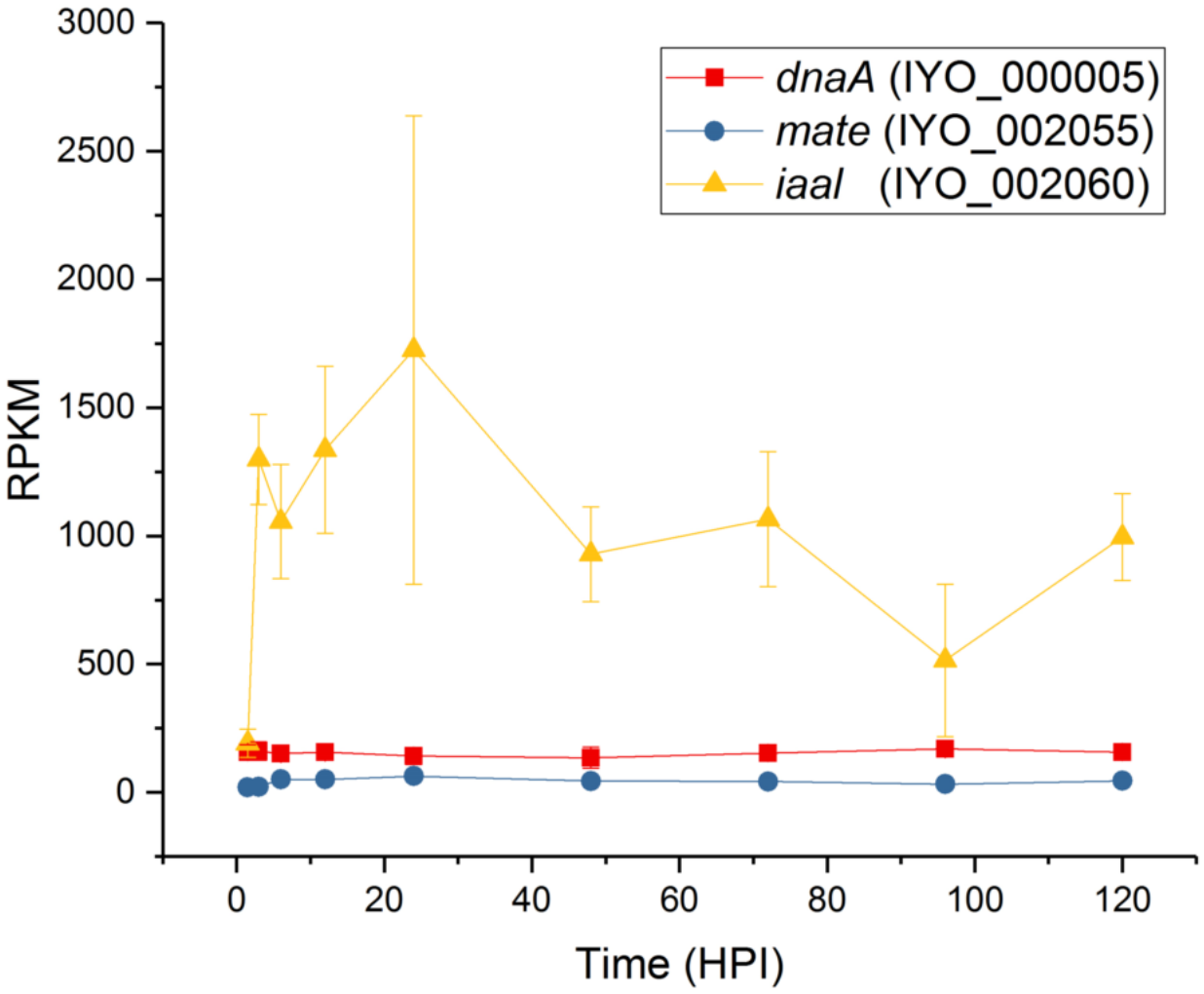


Figure 4. A heatmap showing the expression profiles of predicted *Pseudomonas syringae* pv. *actinidiae* (*Psa*) type III secretion system effectors at 10 time points during a infection time course of a Hotrt16a plantlets or in *in vitro* (IV)culture. This graph presents the log base 2 of the RPKM values with stars (*) indicating time points with significant ($p < 0.01$) changes in expression compared to the IV time point.

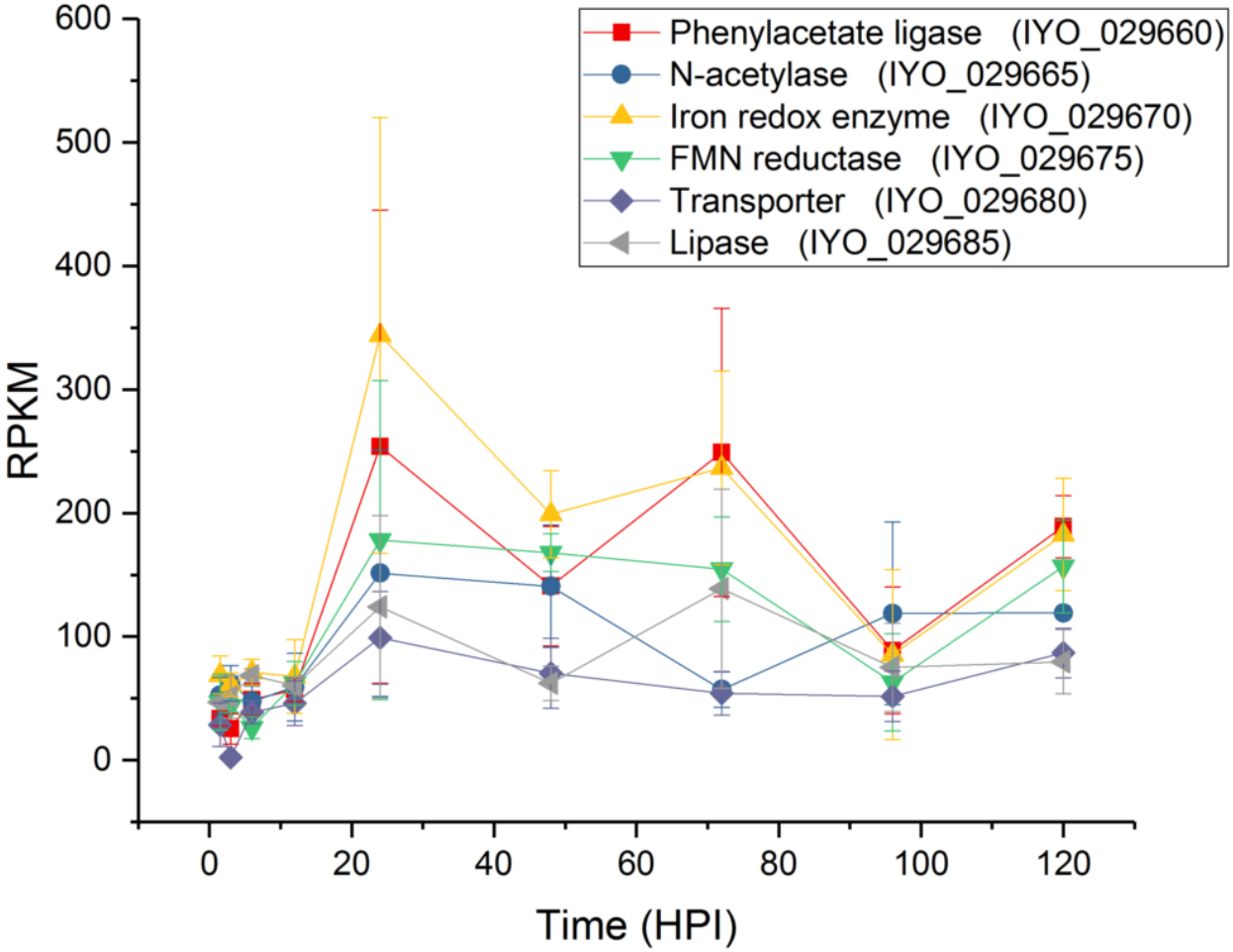


857 Figure 5. Expression of genes encoding *iaal* and *mate*. Reads per kb per million for *iaal*
858 and *mate* were plotted over the infection time course. *dnaA* was included as a
859 constitutively expressed control. Each point is the mean of three biological replicates with
860 error bars representing standard error.
861



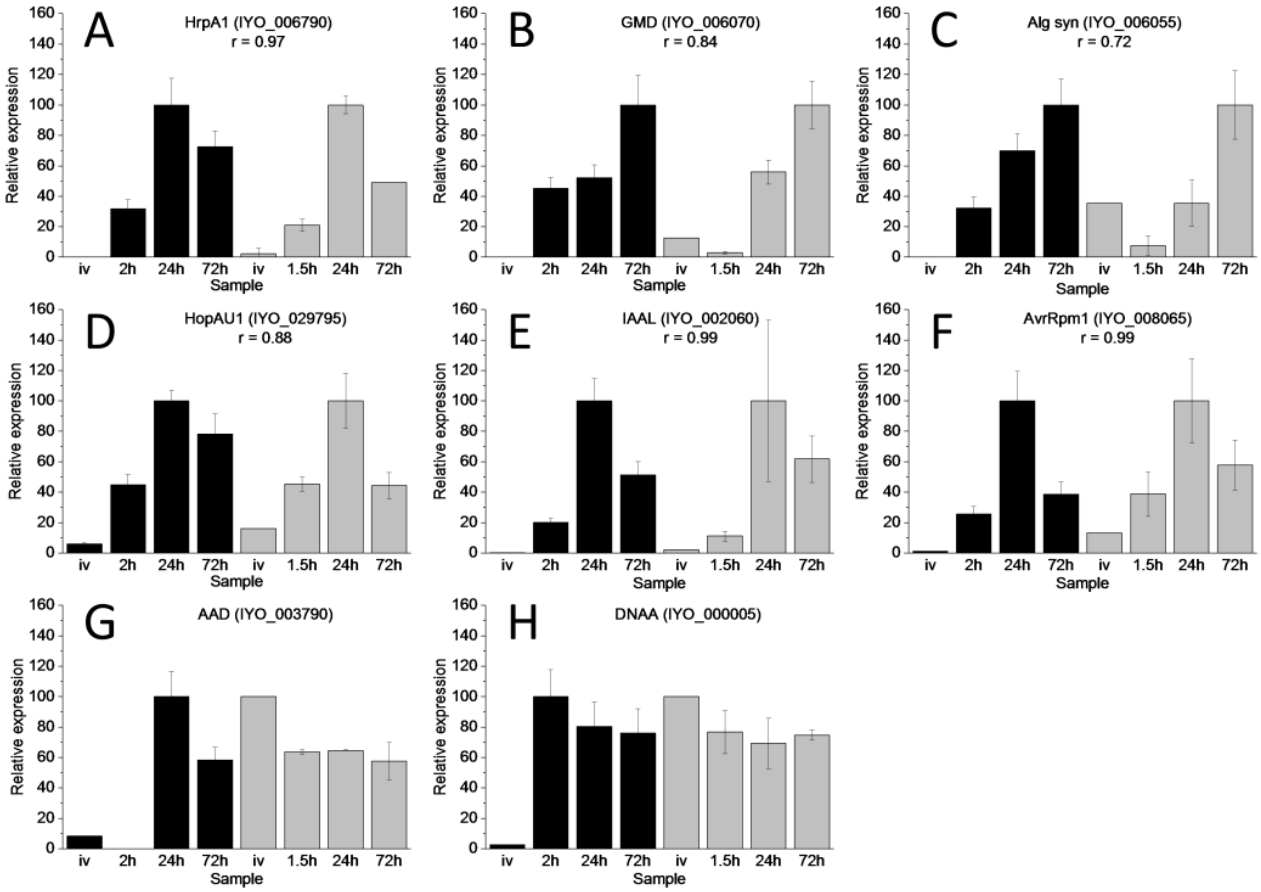
862
863
864

865 Figure 6. Expression of genes from the plasmid-borne putative aromatic biosynthetic
866 pathway. Reads per kb per million values for the genes in the operon coding for the
867 biosynthetic pathway of a putative aromatic compound plotted over the infection time
868 course. Each point is the mean of three biological replicates with error bars representing
869 standard error.
870



871
872
873

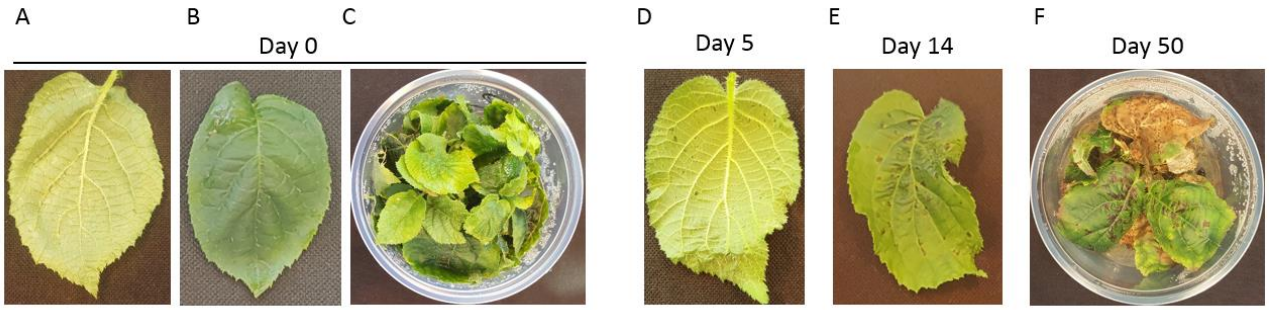
Figure 7. Reverse transcription quantitative PCR validation of RNA-seq data. Eight target genes and five potential reference genes were selected for reverse transcription quantitative PCR (RT-qPCR) (supplemental data set 13). RNA was extracted from *Psa* grown *in vitro* (IV) infected 'Hort16A' plantlets at 2, 24 and 72 hours post infection (HPI). RT-qPCR values (three biological replicates) were based on normalisation against the geometric average/mean of three reference genes (IYO_010670, IYO_009010 and IYO_002170) selected using Bestkeeper and geNorm analysis [69, 70]. For comparison with the RNA-seq data, values were displayed by representing maximum expression of each gene as 100. Pearson correlation coefficients (r) > 0.5 between the RT-qPCR and RNA-seq data are displayed all primers are listed in supplemental data set 13. RT-qPCR data is represented by black bars and RNA-seq (reads per kb per million) by grey bars. A: HrpA1 (IYO_006790); B: GDP-mannose dehydrogenase (GMD, IYO_006070); C: Alginate synthase (IYO_006055); D: HopAU1 (IYO_029795); E: IAAL (IYO_002060); F: AvrRpm1 (IYO_008065); G: Amino acid adenylation protein (AAD, IYO_003790); H: DNAA (IYO_000005).



892 **Additional file 1:**

893 Images illustrating the time course of symptom development of kiwifruit plantlets infected
894 with *Psa*. (A) Abaxial side of leaf at day 0; (B) Adaxial side of leaf at day (0); (C) Pottle
895 containing plantlets at day 0; (D) Abaxial side of leaf five days post inoculation (DPI) with
896 water soaked lesions appearing; (E) Adaxial side of leaf 14 DPI with necrotic lesions
897 present; (F) Plantlets 50 DPI (pptx).

898



899

900

901

902

903 **Additional file 2:**

904 Reads per Kilobase per Million (RPKM) for all *Psa* genes (sheet 1) and k-means analysis

905 for those genes with >50 RPKM for at least one time point (sheet 2) (xlsx).

906

907 **Additional file 3:**

908 Early induced genes ranked by the ratio of expression at three hours post infection (HPI)
909 compared with *in vitro* (cutoff 5-fold) (docx).

910

Gene ID	Gene Annotation	3 HPI/ *in vitro*	*P* value
IYO_011995	phosphate ABC transporter substrate-binding protein	148.8	4.60E-41
IYO_012000	phosphate ABC transporter permease	43.7	3.31E-29
IYO_019585	thioredoxin	31.4	6.68E-31
IYO_012010	phosphate ABC transporter ATP-binding protein	30.8	1.50E-33
IYO_018790	magnesium transporter CorA	30.8	8.40E-24
IYO_027385	ABC transporter substrate-binding protein	29.5	1.15E-19
IYO_027390	GntR family transcriptional regulator	25.0	2.15E-13
IYO_015395	hypothetical protein	23.9	1.38E-10
IYO_006115	amino acid ABC transporter substrate-binding protein	23.3	8.31E-37
IYO_027380	ABC transporter permease	21.5	1.65E-13
IYO_028665	phosphate-binding protein	20.4	5.18E-52
IYO_018545	acid phosphatase	19.5	5.16E-26
IYO_000970	ammonia channel protein	18.7	1.03E-26
IYO_006555	chemotaxis protein	17.4	3.36E-61
IYO_021410	short-chain dehydrogenase	17.2	4.98E-09
IYO_013555	type VI secretion effector protein (Hcp)	16.3	1.59E-20
IYO_013560	EvpB family type VI secretion protein	15.9	9.21E-28
IYO_028645	transcriptional regulator PhoU	14.3	7.85E-27
IYO_002185	peptidase M19	13.6	7.60E-17
IYO_010675	phosphatase	12.6	7.44E-13
IYO_021050	ABC transporter substrate-binding protein	12.5	4.36E-34
IYO_013550	Type VI secretion protein	12.2	2.07E-16
IYO_013565	type VI secretion protein	11.7	2.77E-16

IYO_001685	MFS transporter	11.6	1.44E-39
IYO_021420	polysaccharide deacetylase	10.9	4.26E-14
IYO_020035	ABC transporter substrate-binding protein	10.8	5.31E-20
IYO_027365	metallophosphatase	10.4	2.18E-08
IYO_014740	sugar ABC transporter substrate-binding protein	10.2	3.60E-29
IYO_002190	hydrocarbon binding protein	9.8	3.80E-06
IYO_020310	hypothetical protein	9.7	4.61E-06
IYO_013545	type VI secretion system protein ImpG	9.5	1.15E-16
IYO_006130	glutamine ABC transporter ATP-binding protein	9.4	4.49E-20
IYO_028615	transcriptional regulator PhoB	9.2	5.26E-17
IYO_004270	chemotaxis protein	9.1	5.83E-19
IYO_001975	nitrogen regulation protein NR(I)	9.1	2.23E-18
IYO_002205	electron transfer flavoprotein subunit alpha	8.8	4.35E-05
IYO_025185	urease accessory protein UreG	8.7	2.64E-07
IYO_021405	3-oxoacyl-ACP reductase	8.5	3.81E-04
IYO_026920	ABC transporter permease	8.5	1.68E-06
IYO_006120	amino acid ABC transporter permease	7.5	4.63E-25
IYO_002210	electron transfer flavoprotein subunit beta	7.4	1.27E-02
IYO_027370	iron ABC transporter substrate-binding protein	7.3	7.54E-10
IYO_027375	ABC transporter permease	6.8	1.39E-03
IYO_018365	MFS transporter	6.6	7.33E-05
IYO_014745	xylose isomerase	6.3	3.08E-11
IYO_018070	ATPase	6.2	9.96E-03
IYO_006260	acetyltransferase	6.0	7.07E-06
IYO_027495	sarcosine oxidase subunit alpha	5.9	5.62E-14
IYO_001120	hypothetical protein	5.8	1.04E-04
IYO_026150	acyl carrier protein	5.6	1.92E-01
IYO_004585	branched-chain amino acid ABC transporter	5.6	4.05E-16

	substrate-binding protein		
IYO_009210	quercetin 2,3-dioxygenase	5.5	4.42E-03
IYO_014735	xylose transporter	5.5	9.61E-09
IYO_025180	urease accessory protein UreF	5.5	7.13E-04
IYO_021415	MFS transporter	5.2	7.97E-05

911
912
913
914

915 **Additional file 4:**

916 Genes most highly upregulated in the mid phase of the infection time course (3-24 hours
917 post infection, (HPI)). Genes are ranked on the ratio of maximum Reads Per Kilobase
918 per Million over that time period compared with *in vitro* expression (cutoff 5-fold) (docx).
919

Gene ID	Gene Annotation	Ratio	*P*-value
IYO_006750	type III secretion protein HrpW	257.7	4.4E-64
IYO_022020	hemolysin	176.9	9.7E-44
IYO_006820	type III secretion protein	118.8	2E-88
IYO_006755	Shc Hop M1 (disrupted)	104.0	1.2E-48
IYO_006865	type III secretion system protein	77.2	2.7E-35
IYO_006790	HrpA1	77.0	1.7E-72
IYO_004052	HopS2	76.4	1E-37
IYO_006825	type III secretion protein	67.5	7.8E-44
IYO_006875	type III secretion protein	62.6	2.8E-28
IYO_006880	type III secretion protein	54.0	1.8E-49
IYO_006795	type III secretion protein HrpZ	52.4	3.5E-38
IYO_012110	Ais protein	52.2	6.6E-21
IYO_006905	RNA polymerase sigma factor HrpL	50.3	6E-38
IYO_022025	glycerol acyltransferase	47.4	4E-16
IYO_002060	IAA lysine ligase	46.4	6.2E-20
IYO_004050	type III chaperone ShcO1	44.3	8.2E-37
IYO_006830	secretin	44.1	1E-37
IYO_028770	LysR family transcriptional regulator	44.1	4.7E-25
IYO_003325	copper resistance protein CopZ	41.4	8.7E-27
IYO_006910	type III effector HrpK	39.6	3.5E-34
IYO_012005	phosphate ABC transporter	37.7	3.1E-32

	permease		
IYO_028960	sulfonate ABC transporter permease	35.7	1.4E-12
IYO_006420	chemotaxis protein	35.7	4.3E-25
IYO_014395	lytic transglycosylase	31.2	4.2E-24
IYO_014235	hypothetical protein	30.4	4.3E-21
IYO_006890	type III secretion protein	30.4	9.3E-29
IYO_006860	type III secretion system protein SsaR	30.2	5E-28
IYO_006870	type III secretion system protein	29.3	1.4E-18
IYO_006815	type III secretion protein	29.0	8.1E-30
IYO_028955	nitrate ABC transporter ATP-binding protein	28.3	4.5E-08
IYO_005735	glycoside hydrolase	25.8	1.2E-24
IYO_006900	type III secretion protein HrpJ	25.2	3.3E-29
IYO_003725	type III chaperone protein ShcF	23.8	1.1E-10
IYO_006800	type III secretion protein	23.3	9.7E-29
IYO_020425	AvrPto5	23.3	3.3E-32
IYO_006810	type III secretion protein	22.4	1.4E-30
IYO_006805	type III secretion protein	22.1	1.6E-51
IYO_006885	ATP synthase	21.3	1.6E-21
IYO_006855	type III secretory protein EscS	20.8	1.3E-06
IYO_006765	type III chaperone ShcE	18.8	5.1E-18
IYO_006760	HopM1	18.2	8.6E-32
IYO_003570	avrD1	18.0	4.8E-11
IYO_014245	membrane protein	17.9	0.01651
IYO_029290	type III chaperone protein ShcF	17.8	1.1E-11
IYO_006850	type III secretion system protein	17.5	1.3E-13

IYO_029040	type III secretion chaperone CesT	16.9	1.2E-12
IYO_008282	HopZ5	16.8	1.4E-10
IYO_009200	hypothetical protein	16.8	4.7E-16
IYO_005160	HopI1	16.6	3.7E-21
IYO_012125	diguanylate cyclase	16.1	2.4E-19
IYO_006895	type III secretion protein HrpI	15.6	2.8E-30
IYO_003600	AvrB4	15.5	4.7E-22
IYO_022140	SAM-dependent methyltransferase	14.6	3E-23
IYO_012120	AraC family transcriptional regulator	13.9	9.3E-12
IYO_002045	hypothetical protein	13.8	5.1E-05
IYO_017375	phosphonate/organophosphate ester transporter subunit	13.2	9.7E-09
IYO_018555	HopAZ1	12.9	1.3E-16
IYO_012115	XRE family transcriptional regulator	12.2	7.3E-21
IYO_006745	HopAA1-1	11.8	9.8E-12
IYO_028535	NADP transhydrogenase subunit alpha	11.6	6.2E-09
IYO_008285	HopH1	11.4	1.3E-06
IYO_008065	AvrRpm1	11.4	5E-10
IYO_010805	LuxR family transcriptional regulator	11.4	2.7E-09
IYO_029795	HopAU1	10.9	1.2E-05
IYO_006770	AvrE1	10.8	1.2E-21
IYO_000845	HopY1	10.7	1.2E-10

IYO_003720	HopAO2	10.7	7.7E-21
IYO_012140	protein tolQ	10.2	4.9E-06
IYO_022695	alkaline phosphatase	10.2	3.4E-13
IYO_006250	tail protein	10.1	2.1E-20
IYO_027360	transcriptional initiation protein Tat	9.7	1.7E-11
IYO_009265	serine/threonine protein phosphatase	9.4	4.2E-08
IYO_016255	Ais protein	9.2	3.8E-09
IYO_024150	HopR1	9.0	4.3E-25
IYO_023400	energy transducer TonB	8.9	0.00019
IYO_009660	hypothetical protein	8.9	0.00138
IYO_027435	DNA polymerase III subunit epsilon	8.4	4E-08
IYO_012610	MarR family transcriptional regulator	8.4	6.1E-05
IYO_013150	HopBN1	8.0	7.4E-12
IYO_002040	hypothetical protein	8.0	0.00434
IYO_012030	nitrite reductase	7.7	0.00011
IYO_013145	type III chaperone protein ShcF	7.7	3.8E-07
IYO_003727	HopBB1-1	7.7	1.6E-08
IYO_029288	AvrRpm2 (frameshifts)	7.5	2.3E-06
IYO_012145	biopolymer transporter TolR	7.5	0.01418
IYO_003315	metal ABC transporter ATPase	7.5	7.5E-14
IYO_028380	type III chaperone protein ShcA	7.5	3.6E-07
IYO_000385	dodecin flavoprotein	7.5	1.5E-06
IYO_006845	type III secretion system protein	7.4	4.1E-10
IYO_023505	chemotaxis protein	7.2	1.2E-07

IYO_014240	hypothetical	7.2	4.7E-09
IYO_028540	NAD(P) transhydrogenase	7.0	0.07674
IYO_003680	HopAF1	7.0	9.5E-10
IYO_001870	hypothetical protein	6.8	1.8E-12
IYO_005855	UDP-N-acetylglucosamine 2-epimerase	6.8	6.1E-13
IYO_013690	membrane protein	6.7	2.4E-07
IYO_010630	thiamine biosynthesis protein ApbE	6.6	7.4E-14
IYO_011020	chemotaxis protein	6.4	7.1E-13
IYO_023390	biopolymer transporter ExbB	6.4	0.0044
IYO_029045	HopZ3	6.3	1.6E-09
IYO_024520	voltage-gated chloride channel protein	5.9	1.1E-05
IYO_009335	Fe-S oxidoreductase	5.8	0.02136
IYO_024535	hypothetical protein	5.8	0.00527
IYO_004060	hypothetical protein	5.7	7.8E-27
IYO_021665	MFS transporter	5.7	0.00017
IYO_016185	UDP-4-amino-4-deoxy-L-arabinose-oxoglutarate aminotransferase	5.5	1.7E-05
IYO_020420	iron ABC transporter permease	5.5	1.9E-16
IYO_022135	InaA protein	5.5	3.1E-08
IYO_007455	membrane protein	5.3	1.3E-13
IYO_016195	UDP-4-amino-4-deoxy-L-arabinose formyltransferase	5.3	3.8E-08
IYO_022030	ACP phosphodiesterase	5.2	5.5E-09
IYO_006775	lytic transglycosylase	5.2	1.4E-13

IYO_012605	fusaric acid resistance protein	5.2	4.7E-05
IYO_018725	membrane protein	5.1	0.00024
IYO_004240	hypothetical protein	5.0	6E-05
IYO_014250	chemotaxis protein CheY	5.0	0.10469

920

921

922

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 hours post infection (HPI) in reads per kilobase per million (RPKM). Effectors likely to be disrupted or pseudogenes were not included (docx).

			RPKM									
Gene ID	Effector	Heat map group	*in vitro*	1.5 HPI	3 HPI	6 HPI	12 HPI	24 HPI	48 HPI	72 HPI	96 HPI	120 HPI
IYO_029795	HopAU1	11	346.4	970.6	3772.5	2876.6	3196.1	2141.7	1138.8	952.1	619.7	946.2
IYO_008065	AvrRpm1	11	123.9	365.5	1293.2	867.8	1415.6	940.2	538.8	542.9	526.2	693.6
IYO_020425	AvrPto5	11	58.8	331	1198.8	920.8	1368.6	770.7	472.4	424.3	281.7	487.2
IYO_018555	HopAZ1	11	102.1	226.8	995	616.3	1320.1	1178.3	479.8	682.5	464.8	788
IYO_004052	HopS2	11	15	181.7	1145.3	751.5	762.2	494.6	324.2	239.9	129	226.5
IYO_003720	HopAO2	11	102.7	346.6	1102.7	664.4	821.5	560.7	357.2	378.4	228.5	356.7
IYO_003600	AvrB4	11	65.6	326.8	925.5	696.7	1015.7	621.4	332.6	431.6	251.2	465.3
IYO_013150	HopBN1	11	105.3	179.4	633	468.2	843.7	622.5	327.7	440.3	232.6	342.3
IYO_008282	HopZ5	11	43.4	314.1	662.1	546.9	727.8	426.2	222	186.1	141.9	440.9
IYO_003570	avrD1	6	38.2	99.4	432.1	412.6	687.1	492.6	240.1	419.9	264.3	703.8
IYO_024217	HopF2	11	187.9	225.4	658.8	479.4	649.8	563.5	288	400.6	213.4	403
IYO_003657	HopAW1	11	195	166.1	401.6	375.5	539	402.8	208.9	254.2	187	361.9
IYO_006735	HopN1	7	200.4	301.6	380.7	510.9	478.6	467.8	383.2	394.1	347.3	314.1
IYO_008285	HopH1	11	43.1	147.6	493.5	241.2	462.3	284.1	135.7	158.6	80.9	177.6
IYO_000845	HopY1	11	42.8	121.7	426.6	279.2	459.5	317	172.9	238.8	115.9	204.8
IYO_005160	HopI1	11	22.3	96.9	370	246.3	367	324.7	169.6	217.6	143.6	181.2
IYO_006760	HopM1	11	20.3	106.4	368.9	286.3	347.5	186.1	98.1	102.9	69.2	114.3
IYO_003727	HopBB1-1	11	47.3	103.6	362.9	266.3	232.3	173.2	76.2	107	143.6	100.6
IYO_003680	HopAF1	6	46	90.6	176.3	145.4	302.4	321.1	145.9	186.5	154.3	213.1

IYO_029045	HopZ3	11	47.3	89.7	248.7	179.1	300.1	206.6	111.2	116.5	84.1	135.8
IYO_006770	AvrE1	11	27	73.9	291.1	176.3	164.6	129.1	57.6	53.6	44.9	47.8
IYO_003525	HopQ1	11	53.6	96.9	165	163.3	232.5	146.8	86.5	131.6	79.6	168.7
IYO_003530	HopD1	7	53.5	99.3	212.3	184	229.5	216.6	156.4	170.4	124.3	156.3
IYO_024150	HopR1	11	24.1	58.5	217.6	103	94	68.1	77.2	112.8	45.6	94.9
IYO_012225	HopAE1	11	31.2	53.1	133.8	95.1	142.7	174.9	115.4	111.7	71.9	121.5
IYO_003675	HopBB1-2	6	32.1	40.7	129.6	67.4	148.5	104.5	55.6	50.1	21.9	136.4
IYO_027420	HopAS1	9	24.9	11.4	12.3	17.1	21.7	60.3	129.9	109.4	38.4	53.4
IYO_023985	HopAH1	1	53.5	111.4	111.3	110.1	81.9	93.5	70.5	82.1	57.6	42.3
IYO_006745	HopAA1-1	11	9.2	27.7	107.2	60.6	109.3	80.5	32.5	27.7	12.8	31
IYO_003635	HopX3*	5	53.1	93.4	67.9	70.6	65.6	33.7	63.6	107.1	59.7	88.2
IYO_023205	HopAM1-1	ND										
IYO_008385	HopAM1-2	ND										

Additional file 6:

Expression of non-effector genes with upstream HrpL boxes. Genes are ranked by the ration of expression at 12 hours post infection (HPI) compared with *in vitro* expression (docx).

Gene ID	Gene Annotation	Cluster	12 HPI/*in vitro*	*P*-value
IYO_002060	IAA lysine ligase	11	36.0	6.20E-17
IYO_006775	lytic transglycosylase	11	5.2	5.56E-11
IYO_010630	thiamine biosynthesis protein ApbE	11	5	5.79E-14
IYO_027210	peptidase M20	11	3.5	1.16E-05
IYO_025425	phosphatidylserine decarboxylase	11	3.4	1.27E-01
IYO_002055	multidrug transporter Mate	7	2.5	1.42E-02
IYO_000225	AraC transcription factor	8	1.5	2.89E-01
IYO_008215	Transporter	1	1.2	3.33E-01

943 **Additional file 7:**

944 Genes expressed in the mid phase of infection that do not encode the T3SS or T3SEs.
945 Effectors were ranked by the highest level of expression between 3 and 24 HPI compared
946 with *in vitro* (cutoff 5-fold) (docx).

947

Gene ID	Gene Annotation	Max 3-24hr/*in vitro*	*P*-value
IYO_022020	hemolysin	176.8965134	9.7E-44
IYO_012110	Ais protein	52.15177309	6.6E-21
IYO_022025	glycerol acyltransferase	47.41414785	4E-16
IYO_002060	IAA lysine ligase	46.43964929	6.2E-20
IYO_006830	secretin	44.12981547	1E-37
IYO_028770	LysR family transcriptional regulator	44.1000272	4.7E-25
IYO_003325	copper resistance protein CopZ	41.36623134	8.7E-27
IYO_012005	phosphate ABC transporter permease	37.70442846	3.1E-32
IYO_028960	sulfonate ABC transporter permease	35.71674044	1.4E-12
IYO_006420	chemotaxis protein	35.6663794	4.3E-25
IYO_014395	lytic transglycosylase	31.15779288	4.2E-24
IYO_014235	hypothetical protein	30.42819287	4.3E-21
IYO_028955	nitrate ABC transporter ATP-binding protein	28.30708753	4.5E-08
IYO_005735	glycoside hydrolase	25.83644908	1.2E-24
IYO_006885	ATP synthase	21.27236361	1.6E-21
IYO_014245	membrane protein	17.92814971	0.01651
IYO_009200	hypothetical protein	16.75685698	4.7E-16
IYO_012125	diguanylate cyclase	16.10227356	2.4E-19
IYO_022140	SAM-dependent methyltransferase	14.56526101	3E-23
IYO_012120	AraC family transcriptional regulator	13.86039793	9.3E-12
IYO_002045	hypothetical protein	13.78948606	5.1E-05
IYO_017375	phosphonate/organophosphate ester transporter	13.19762412	9.7E-09

	subunit		
IYO_012115	XRE family transcriptional regulator	12.15821469	7.3E-21
IYO_028535	NADP transhydrogenase subunit alpha	11.6258924	6.2E-09
IYO_010805	LuxR family transcriptional regulator	11.35613537	2.7E-09
IYO_012140	protein tolQ	10.24591846	2.9E-05
IYO_022695	alkaline phosphatase	10.1825476	3.4E-13
IYO_006250	tail protein	10.12427557	2.1E-20
IYO_027360	transcriptional initiation protein Tat	9.73508113	1.7E-11
IYO_009265	serine/threonine protein phosphatase	9.394839603	4.2E-08
IYO_016255	Ais protein	9.226696251	3.8E-09
IYO_023400	energy transducer TonB	8.931602714	0.00019
IYO_009660	hypothetical protein	8.858717065	0.00138
IYO_027435	DNA polymerase III subunit epsilon	8.414824567	4E-08
IYO_012610	MarR family transcriptional regulator	8.394245239	6.1E-05
IYO_002040	hypothetical protein	7.97028053	0.00434
IYO_012030	nitrite reductase	7.725691653	0.00011
IYO_012145	biopolymer transporter TolR	7.522804764	0.01832
IYO_003315	metal ABC transporter ATPase	7.519255209	7.5E-14
IYO_000385	dodecin flavoprotein	7.463415413	1.5E-06
IYO_023505	chemotaxis protein	7.205132932	1.2E-07
IYO_014240	hypothetical	7.163653846	4.7E-09
IYO_028540	NAD(P) transhydrogenase	6.987308302	0.07674
IYO_001870	hypothetical protein	6.828648569	1.8E-12
IYO_005855	UDP-N-acetylglucosamine 2-epimerase	6.828589164	6.1E-13
IYO_013690	membrane protein	6.663626577	2.4E-07
IYO_010630	thiamine biosynthesis protein ApbE	6.594972584	7.4E-14
IYO_011020	chemotaxis protein	6.419561085	7.1E-13
IYO_023390	biopolymer transporter ExbB	6.37203283	0.0044

IYO_024520	voltage-gated chloride channel protein	5.859449978	1.1E-05
IYO_009335	Fe-S oxidoreductase	5.832304804	0.01462
IYO_024535	hypothetical protein	5.789314032	0.00527
IYO_004060	hypothetical protein	5.695280181	7.8E-27
IYO_021665	MFS transporter	5.684151414	0.00017
IYO_016185	UDP-4-amino-4-deoxy-L-arabinose-oxoglutarate aminotransferase	5.505631007	1.7E-05
IYO_020420	iron ABC transporter permease	5.488234589	1.9E-16
IYO_022135	InaA protein	5.460963656	3.1E-08
IYO_007455	membrane protein	5.347108089	1.3E-13
IYO_016195	UDP-4-amino-4-deoxy-L-arabinose formyltransferase	5.310883424	3.8E-08
IYO_022030	ACP phosphodiesterase	5.183953282	5.5E-09
IYO_006775	lytic transglycosylase	5.161048509	1.4E-13
IYO_012605	fusaric acid resistance protein	5.150295806	4.7E-05
IYO_018725	membrane protein	5.138633214	0.00024
IYO_004240	hypothetical protein	4.992962804	6E-05
IYO_014250	chemotaxis protein CheY	4.96231316	0.10469

948
949
950

951 **Additional file 8:**

952 Genes upregulated late in the time course. Genes were ranked based on the ratio of

953 expression at 120 hours post infection (HPI) compared with 1.5 HPI (docx).

954

Gene ID	Gene Annotation	120HPI RPKM/1.5 HPI RPKM	*P*-value
IYO_016130	ABC transporter	119.9	6.92E-05
IYO_016135	acyl-CoA synthetase	93.1	7.11E-06
IYO_006065	glycosyl transferase	75.6	4.34E-08
IYO_013825	alkanesulfonate monooxygenase	51.2	4.06E-15
IYO_006070	GDP-mannose dehydrogenase	51.1	7.92E-40
IYO_009605	YqcI/YcgG family protein	39.1	8.72E-07
IYO_023405	biopolymer transporter ExbD	34.3	6.02E-05
IYO_026605	monooxygenase	33.2	2.77E-11
IYO_017250	energy transducer TonB	28.3	2.71E-08
IYO_017245	biopolymer transporter ExbB	28.0	5.43E-07
IYO_027905	transporter	26.9	3.73E-10
IYO_016105	acyl-CoA dehydrogenase	24.8	1.23E-07
IYO_005580	hypothetical protein	24.2	2.93E-12
IYO_027910	aliphatic sulfonates transport ATP-binding subunit	23.7	5.36E-06
IYO_006055	alginate biosynthesis protein	22.7	2.51E-06
IYO_006050	alginate regulatory protein	20.0	6.29E-08
IYO_009255	sulfonate ABC transporter ATP-binding protein	19.9	8.75E-03
IYO_006015	mannose-1-phosphate guanylyltransferase	19.3	1.91E-22
IYO_018210	calcium-binding protein	19.2	2.34E-16
IYO_008315	lipoprotein	18.3	8.74E-18
IYO_016100	acyl-CoA dehydrogenase	18.2	1.37E-03
IYO_016120	ABC transporter permease	16.7	7.26E-03

IYO_026615	N5,N10-methylene tetrahydromethanopterin reductase	16.5	3.86E-07
IYO_015950	polar amino acid ABC transporter permease	15.6	3.90E-02
IYO_006060	hemolysin D	15.3	5.79E-11
IYO_006040	alginate O-acetyltransferase	15.1	1.94E-09
IYO_026620	methionine ABC transporter substrate-binding protein	15.0	1.95E-03
IYO_026625	ABC transporter	15.0	6.50E-07
IYO_027920	ABC transporter substrate-binding protein	14.1	2.03E-09
IYO_015300	ABC transporter permease	13.1	3.01E-03
IYO_026675	sulfonate ABC transporter ATP-binding protein	12.7	2.93E-04
IYO_006030	poly(beta-D-mannuronate) O-acetylase	12.3	2.14E-09
IYO_011010	catalase	12.0	4.25E-14
IYO_016095	5,10-methylene tetrahydromethanopterin reductase	11.7	2.57E-07
IYO_014785	sugar ABC transporter	11.4	1.52E-06
IYO_026610	acyl-CoA dehydrogenase	11.0	1.56E-05
IYO_006025	alginate O-acetyltransferase	11.0	1.43E-07
IYO_012440	hypothetical protein	10.7	1.66E-02
IYO_020560	peptidase M4	10.4	3.46E-09
IYO_011310	NAD(P)H-dependent FMN reductase	10.2	9.59E-06
IYO_024090	porin	10.1	1.15E-08
IYO_003290	hypothetical protein	10.0	9.64E-02
IYO_010385	lipoprotein	10.0	1.85E-02
IYO_027985	hypothetical protein	9.8	1.90E-07
IYO_006045	poly(beta-D-mannuronate) C5 epimerase	9.6	1.11E-09
IYO_006035	poly(beta-D-mannuronate) lyase	9.6	4.76E-11
IYO_014780	sugar ABC transporter substrate-binding protein	9.5	1.68E-06
IYO_016140	monooxygenase	9.1	1.75E-06
IYO_017240	biopolymer transporter ExbD	9.1	1.69E-05

IYO_011305	lysine transporter LysE	8.7	8.38E-05
IYO_001790	taurine transporter ATP-binding subunit	8.7	6.15E-02
IYO_026775	alpha/beta hydrolase	8.6	1.81E-07
IYO_024095	ABC transporter substrate-binding protein	8.5	6.24E-08
IYO_027980	ABC transporter permease	8.1	6.65E-12
IYO_026630	ABC transporter permease	7.8	2.58E-02
IYO_001460	prophage PssSM-01	7.7	2.51E-03
IYO_011840	hemolysin D	7.7	2.28E-03
IYO_013055	aldolase	7.4	1.26E-04
IYO_026680	taurine dioxygenase	7.2	9.94E-08
IYO_014465	hypothetical protein	6.9	1.11E-09
IYO_014710	lipoprotein	6.8	7.90E-09
IYO_017510	lipoprotein	6.8	1.75E-01
IYO_001820	hypothetical protein	6.8	4.01E-02
IYO_027915	alkanesulfonate transporter permease subunit	6.7	2.33E-04
IYO_016110	branched-chain amino acid ABC transporter ATP-binding protein	6.6	4.90E-03
IYO_009290	LTXXQ domain-containing protein	6.6	1.85E-15
IYO_006935	hypothetical protein	6.5	2.86E-08
IYO_016115	ABC transporter permease	6.3	8.18E-05
IYO_009230	sulfurtransferase	6.1	2.96E-04
IYO_013050	nitrate ABC transporter substrate-binding protein	6.0	2.86E-02
IYO_020620	hypothetical protein	5.8	2.46E-05
IYO_011220	Fis family transcriptional regulator	5.8	2.93E-06
IYO_029660	coenzyme F390 synthetase (plasmid)	5.6	1.54E-01
IYO_027965	sulfate ABC transporter ATP-binding protein	5.4	7.96E-04
IYO_005875	hypothetical protein	5.4	7.26E-14
IYO_016125	ABC transporter permease	5.4	7.52E-02

IYO_001465	prophage PssSM-01	5.3	1.02E-04
IYO_011375	class V aminotransferase	5.3	4.08E-06
IYO_009250	ABC transporter permease	5.3	1.92E-01
IYO_001810	ribonucleotide reductase	5.2	1.16E-14
IYO_004495	hypothetical protein	5.1	9.88E-03
IYO_001805	transposase	5.1	3.42E-06
IYO_028000	diguanylate cyclase	5.0	3.01E-06
IYO_009615	serine dehydratase	5.0	1.85E-07

955

956

957 **Additional file 9:**

958 Expression levels of secondary metabolite gene clusters. Means of reads per kilobase
959 per million (RPKM) for each gene across all time points with standard deviations (docx).

960

Secondary metabolite pathway	Function	Gene members	Induction levels *in planta*
Novel Non-ribosomal peptide synthetase	unknown	IYO_003775-003830	Constitutive expression *in planta* (average RPKM 67 +/- 43)
Pyoverdine	Iron chelation	IYO_010820-010910	Constitutive expression *in planta* (average RPKM 41 +/- 43)
Achromobactin	Iron chelation	IYO_013460-013515	Constitutive expression *in planta* (average RPKM 17 +/- 9)
Yersiniabactin	Iron chelation	IYO_013840-013910	Constitutive expression *in planta* (average RPKM 15 +/-19)
Unknown	unknown	IYO_026725-026760	Weak constitutive *in planta* (expression 48 +/-16 RPKM)
Mangotoxin	Inhibitor of ornithine deacetylase	IYO_028470-028715	Weak expressed *in planta* (average RPKM 13 +/- 9)
Plasmid-borne pathway	unknown	IYO_029645-029685	Induced late *in planta* (see figure 5)

961

962 **Additional file 10:**

963 Reads per kilobase per million (RPKM) values of genes encoding proteins predicted to be secreted via T2SS (docx).

964

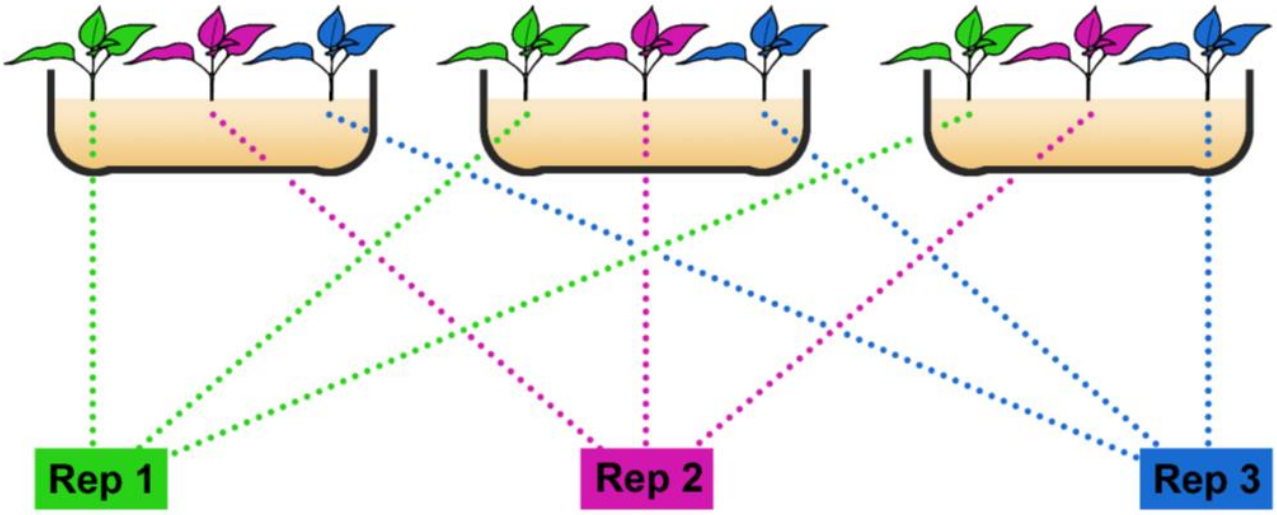
# ID	Function	Phase	*in vitro*	1.5 HPI	3 HPI	12 HPI	24 HPI	8 HPI	72 HPI	96 HPI	120 HPI
IYO_011995	phosphate ABC transporter substrate-binding protein	Early	39.5	3366.6	5876.1	3269.3	2774.8	1220.1	1143.4	1114.5	1774.3
IYO_019585	thioredoxin	Early	59.4	1525.9	1863.2	983.2	433.9	374.4	329.6	350.7	469.3
IYO_027385	ABC transporter substrate-binding protein	Early	10.4	222.5	306.4	105.9	117.6	54.4	68.2	7.9	58.9
IYO_006115	amino acid ABC transporter substrate-binding protein	Early	117.7	3408.6	2742.5	843.8	829.8	596.2	750.2	622.4	722.3
IYO_028665	phosphate-binding protein	Early	33.6	759.5	683.8	278.3	194.3	145.9	138.5	166.4	210.9
IYO_000970	ammonia channel protein	Early	43.5	1336.6	811.4	274.2	268.1	231.2	409.6	259.8	426.5
IYO_021410	short-chain dehydrogenase	Early	6.6	142.6	113.3	26.3	33.4	19.3	4.4	19.6	12.1
IYO_010675	phosphatase	Early	32.6	228.2	410.1	158.8	97.9	151.1	56.9	80.3	89.7
IYO_021050	ABC transporter substrate-binding protein	Early	55.0	945.8	685.8	281.4	281.5	212.0	245.3	245.3	224.0
IYO_020035	ABC transporter substrate-binding protein	Early	131.3	1169.0	1415.7	583.5	416.4	253.4	327.7	249.7	292.0
IYO_014740	sugar ABC transporter substrate-binding protein	Early	59.2	490.2	604.3	243.3	163.4	143.6	89.5	131.6	96.6
IYO_020310	hypothetical protein	Early	24.9	380.2	242.5	139.5	139.0	187.8	112.7	43.1	75.4
IYO_004585	branched-chain amino acid ABC transporter substrate-binding protein	Early	87.8	956.2	489.1	174.2	232.2	130.6	255.0	172.8	338.3
IYO_025190	protein hupE	Early	25.0	198.5	123.6	36.9	61.1	38.8	90.6	89.1	123.2

IYO_006385	porin	Early	1383.7	7387.9	6833.1	1618.0	1721.0	1130.5	1055.6	984.5	945.0
IYO_020485	glycine/betaine ABC transporter substrate-binding protein	Early	33.5	134.6	129.0	35.6	46.6	43.4	44.4	37.1	32.5
IYO_008325	polygalacturonase	Early	164.0	369.3	564.7	150.2	140.3	107.1	97.9	83.8	107.4
IYO_021455	Methylamine utilization protein MauL	Early	94.4	582.6	317.5	319.1	399.1	422.6	247.9	279.5	147.7
IYO_004580	urea ABC transporter permease	Early	14.9	74.3	43.4	20.5	16.3	35.5	30.8	19.4	22.0
IYO_026915	amino acid ABC transporter substrate-binding protein	Early	50.4	426.4	139.6	38.5	65.7	44.0	87.4	35.9	98.4
IYO_006365	sugar ABC transporter substrate-binding protein	Early	1316.7	2310.1	3559.7	671.4	762.4	398.4	474.9	375.7	335.4
IYO_024670	hypothetical protein part of ICE	Early	34.2	23.9	18.9	59.0	31.6	29.2	22.7	8.5	38.9
IYO_006805	type III secretion protein	Mid	57.1	309.6	1263.6	1201.0	751.4	414.5	501.0	276.3	450.1
IYO_002045	hypothetical protein	Mid	73.0	63.8	30.4	1006.5	102.8	120.8	106.7	104.0	50.3
IYO_009660	hypothetical protein	Mid	31.6	11.5	0.0	280.3	84.6	121.0	73.3	66.4	81.0
IYO_002040	hypothetical protein	Mid	60.1	51.1	27.5	479.1	77.9	74.5	104.8	106.2	117.1
IYO_001870	hypothetical protein	Mid	510.6	318.1	478.1	2793.5	2242.8	1524.5	1518.2	1974.0	1162.6
IYO_006020	alginate O-acetyltransferase	Mid	88.2	49.0	27.9	423.9	197.5	278.0	383.9	259.6	417.2
IYO_023395	TonB-dependent receptor	Mid	25.2	19.7	20.4	98.7	101.4	46.8	37.8	47.0	83.2
IYO_008760	sorbosone dehydrogenase	Mid	105.4	49.9	40.3	391.5	269.6	218.2	157.1	176.8	125.4
IYO_018720	sorbosone dehydrogenase	Mid	60.3	37.9	25.8	224.0	111.0	97.4	102.4	122.0	93.4
IYO_027210	peptidase M20	Mid	53.5	35.4	114.4	189.8	132.1	98.8	88.0	68.1	119.3
IYO_022715	phospholipid-binding protein	Mid	8230.8	2550.5	2568.7	29098.2	18992.7	15528.5	13209.7	16287.6	10178.4

IYO_004060	hypothetical protein	Mid	1959.4	7898.0	5063.3	6887.1	5478.9	4687.6	4501.3	3573.6	3898.1
IYO_019200	BNR/Asp-box repeat-containing protein	Mid	56.0	117.1	157.2	196.2	133.8	88.1	59.6	80.2	66.9
IYO_004055	membrane protein	Mid	34.6	39.5	114.7	86.9	62.9	37.4	39.6	39.0	33.7
IYO_006835	type III secretion protein	Mid	312.2	275.5	532.0	847.3	672.8	399.1	381.4	330.7	383.1
IYO_020600	ABC transporter substrate-binding protein	Mid	76.1	47.3	60.0	198.5	143.9	138.9	98.6	106.3	119.6
IYO_022515	toluene tolerance protein	Mid	116.7	73.1	109.9	251.8	129.2	148.1	174.9	121.3	128.6
IYO_006575	superoxide dismutase	Mid	178.1	46.6	35.6	375.8	202.9	114.6	173.8	176.6	223.3
IYO_022005	hypothetical protein	Mid	45.1	33.6	34.6	88.2	69.7	55.2	0.0	17.9	66.5
IYO_027595	phosphorylcholine phosphatase	Mid	50.4	33.8	36.3	82.4	97.7	53.8	62.8	63.3	64.2
IYO_011990	hypothetical protein	Mid	1900.4	539.1	489.4	3491.9	1831.7	1693.3	1220.1	1023.3	836.5
IYO_027840	ABC transporter substrate-binding protein	Mid	70.8	93.0	65.3	56.8	99.4	36.2	24.1	37.9	42.6
IYO_017485	hypothetical protein	Mid	300.2	114.9	115.4	411.4	188.4	259.0	298.3	216.1	202.3
IYO_010560	cytochrome C	Mid	86.6	54.4	56.6	92.6	76.9	66.5	80.4	56.4	49.9
IYO_014770	hypothetical protein	Mid	208.1	169.6	157.6	200.6	190.5	206.4	197.7	172.8	410.9
IYO_011885	type III effector	Mid	25.2	6.4	3.6	7.2	7.8	8.5	6.0	3.4	2.9

969 **Additional file 11:**

970 RNA-seq experimental design. Three pottles, each with three plantlets, were inoculated
971 for each time point. For RNA extraction, one plantlet from each pottle, was harvested and
972 combined for each of three biological replicates (docx).

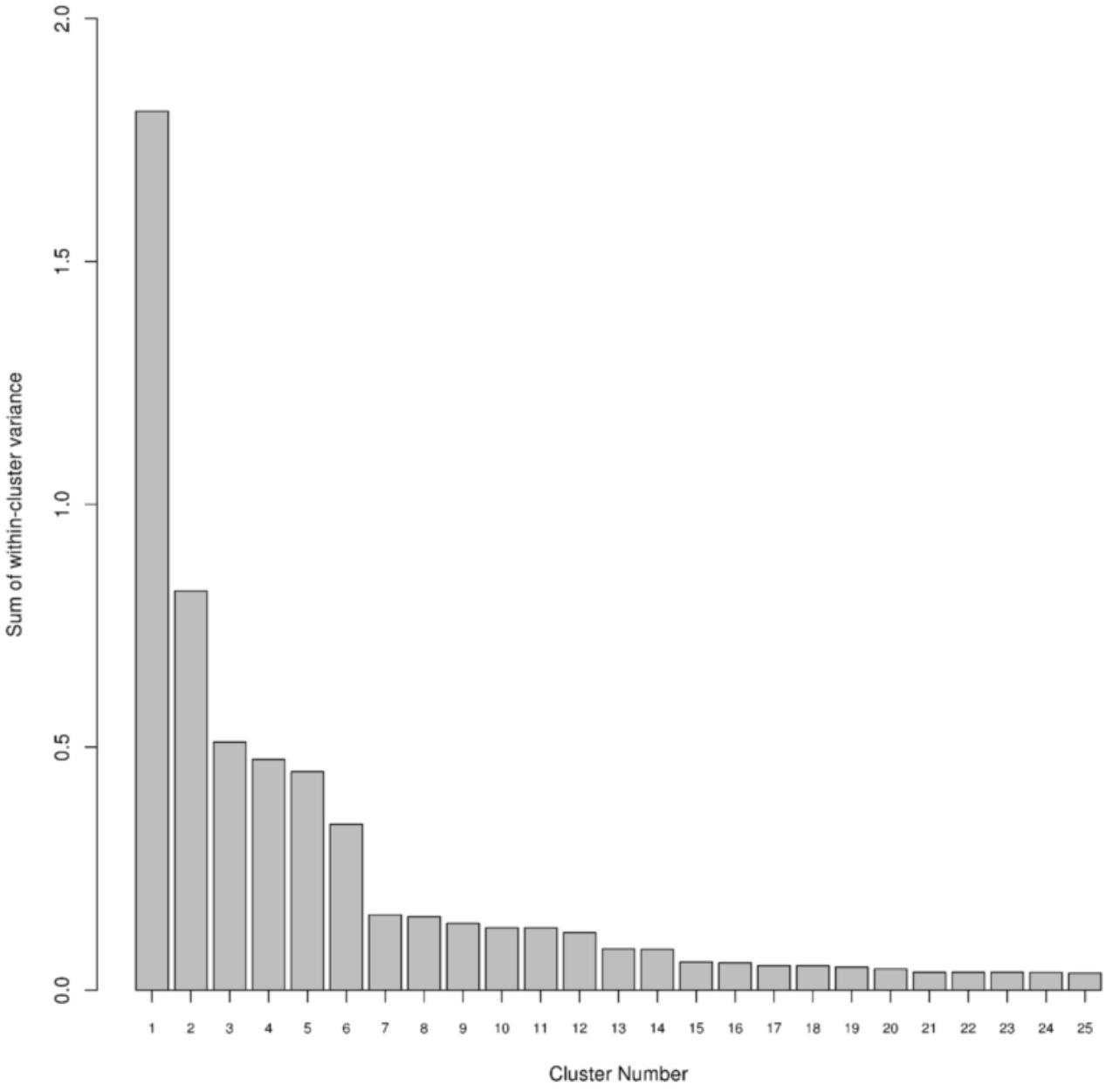


973

974 **Additional file 12:**

975 A bar plot of the inertia gain using the sum of the within-group variance with increasing
976 cluster number (x-axis) produced using Hierarchical clustering on principal components
977 (tiff).

978



979

980

981 **Additional file 13:**

982 A spreadsheet containing all the primers used for reverse transcription and reverse
983 transcription quantitative PCR. The data includes the gene chosen, locus ID number,
984 primer sequences, product size and PCR efficiency (xlsx).