A high-throughput, near-saturating screen for type III effector genes from *Pseudomonas syringae*

Jeff H. Chang*, Jonathan M. Urbach†, Terry F. Law*, Larry W. Arnold‡, An Hu§, Saurabh Gombar*, Sarah R. Grant*¶, Frederick M. Ausubel†, and Jeffery L. Dangl*‡¶‖

Departments of *Biology CB#3280 and ‡Microbiology and Immunology CB#7290, and ¶Curriculum in Genetics, Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC 27599 ; †Department of Molecular Biology, Massachusetts General Hospital, Cambridge, MA 02139; and §Syngenta Biotechnology, Inc., P.O. Box 12257, 3054 Cornwallis Road, Research Triangle Park, NC 27709

Communicated by Brian J. Staskawicz, University of California, Berkeley, CA, December 22, 2004 (received for review November 4, 2004)

Pseudomonas syringae **strains deliver variable numbers of type III effector proteins into plant cells during infection. These proteins are required for virulence, because strains incapable of delivering them are nonpathogenic. We implemented a whole-genome, high-throughput screen for identifying *P. syringae* type III effector genes. The screen relied on FACS and an arabinose-inducible *hrpL* σ factor to automate the identification and cloning of HrpL-regulated genes. We determined whether candidate genes encode type III effector proteins by creating and testing full-length protein fusions to a reporter called Δ79AvrRpt2 that, when fused to known type III effector proteins, is translocated and elicits a hypersensitive response in leaves of *Arabidopsis thaliana* expressing the RPS2 plant disease resistance protein. Δ79AvrRpt2 is thus a marker for type III secretion system-dependent translocation, the most critical criterion for defining type III effector proteins. We describe our screen and the collection of type III effector proteins from two pathovars of *P. syringae*. This stringent functional criteria defined 29 type III proteins from *P. syringae* pv. *tomato*, and 19 from *P. syringae* pv. *phaseolicola* race 6. Our data provide full functional annotation of the *hrpL*-dependent type III effector suites from two sequenced *P. syringae* pathovars and show that type III effector protein suites are highly variable in this pathogen, presumably reflecting the evolutionary selection imposed by the various host plants.**

host–microbe interaction | plant pathogenesis | *Arabidopsis* | FACS

The commonly studied plant bacterial pathogen, *Pseudomonas syringae*, is subdivided into pathovars based on their ability to cause disease on one or more distinct host species. During infection, *P. syringae* and other Gram-negative pathogens deliver type III effector proteins via a type III secretion system (TTSS) (*hrp*/*hrc* genes in *P. syringae*; ref. 1) from the bacterium into host cells (2). *P. syringae* strains incapable of delivering type III effectors are nonpathogenic (3). Thus, the type III effectors each strain delivers are required for pathogenicity. In contrast, if just one of the type III effectors is "recognized" by the plant immune system's surveillance machinery [disease resistance (R) proteins], a battery of host responses is triggered, including localized programmed cell death, termed the hypersensitive response (HR) (4). In this case, the pathogen is rendered avirulent; its multiplication is limited, and it does not cause disease. As a consequence, some type III effector genes have been functionally defined as avirulence (*avr*) genes. Recognition of type III effector proteins by corresponding R proteins may therefore limit the particular host range of individual *P. syringae* pathovars.

P. syringae type III effector genes share several characteristics. Their expression is coordinately regulated with the TTSS-encoding genes by the alternative σ factor, HrpL (5). Genes encoding both the TTSS and type III effector proteins also share a cis-element (*hrp*-box) in their promoters (6). Finally, delivery of type III effector proteins into the host cell depends on the TTSS and a loosely defined N-terminal signal in the type III effector protein that includes >10% serine or proline in the first 50 aa, an aliphatic amino acid or proline at position 3 or 4, and the absence of negatively charged amino acids in the first 12 residues (7–9).

Nonsaturating genetic screens have relied on these characteristics. Guttman *et al.* (9) identified 15 protein fusions defining genes whose N termini were sufficient for TTSS-dependent translocation. These included 12 previously uncharacterized type III effector proteins. Boch *et al.* (10) identified genes that were induced by *in planta* conditions and regulated by HrpL *in vitro*. Six of these encoded previously identified type III effector proteins. Bioinformatic approaches relying on the predicted shared characteristics of type III effector genes have also been used to identify candidate type III effector genes. The genomes of three *P. syringae* pathovars, *tomato* race DC3000 (*Pto*) (www.pseudomonas-syringae.org), *phaseolicola* race 6–1448a (*Pph6*) (www.pseudomonas-syringae.org), and *syringae* race B728a (*Psy*) (www.jgi.doe.gov/index.html) have been sequenced (11). Genes encoding putative type III effector proteins and helpers, referred to as *hop* (hrp-dependent outer protein) genes were identified in these genomes by the presence of putative *hrp*-boxes (12, 13), and ORFs with N-terminal amino acid compositions consistent with known Hops (8, 9, 14, 15). Helper proteins are secreted from the bacterium via the TTSS but likely do not function within the host cell. Their presumptive roles are to assist the TTSS in delivery of type III effector proteins (15). From these efforts, *Pto* was estimated to deliver ≈40 different Hops (7–10, 12–15). Far fewer Hops were predicted in *Psy* B728a (14). Thirty Hops were predicted from *Pph6* based on homology to known and predicted Hops from other strains (16). The exact number of type III effector proteins based on these predictions has not been experimentally validated.

We modified a technique termed differential fluorescence induction (DFI) (17) for high-throughput discovery of HrpL-regulated genes from pathovars of *P. syringae*. Genomic libraries of *P. syringae* were screened for clones expressing GFP in a HrpL-dependent manner by using FACS. Identified clones were sequenced and assembled. Resulting contiguous DNA sequences (contigs) were examined for the characteristics of type III effector genes mentioned above. We amplified and cloned full-length candidate genes in-frame to the coding region of the C-terminal 177 aa of the type III effector protein AvrRpt2, from which the N-terminal 79 aa are deleted (Δ79AvrRpt2; refs. 9, 18, and 19). Cells expressing these fusion proteins were infiltrated into plants expressing RPS2, the R protein that "recognizes" AvrRpt2, to determine which of the candidate genes encode proteins that can translocate Δ79AvrRpt2, and are hence type III effector proteins.

Abbreviations: TTSS, type III secretion system; R, disease resistance; HR, hypersensitive response; *Pto*, *P. syringae* pv. *tomato* race DC3000; *Pph6*, *P. syringae* pv. *phaseolicola* race 6-1448a; *Psy*, *P. syringae* pv. *syringae*; DFI, differential fluorescence induction; Hop, Hrp-dependent outer protein.

Data deposition: Sequences corresponding to genes identified in *Pph6* have been deposited in the GenBank database [accession nos. AY803993 (*avrB3*), AY803994 (*hopAT1*), AY803995 (*hopAU1*), AY803996 (*hopAV1*), AY803997 (*hopAW1*), and AY803998 (HrpL-regulated gene)].

‖To whom correspondence should be addressed. E-mail: dangl@email.unc.edu.

We present our method and analyses of the type III effector suites from *Pto* and *Pph6*. Our data provide full functional annotation for the HrpL-regulated type III effector proteins from these two strains, and provide a tool with which to characterize the type III effector suites from the entire range of *P. syringae* pathovars. The two strains carry different sets of type III effectors. Essentially all of the type III effector genes are located in nonorthologous positions in the respective genomes. These findings suggest that type III effector suites are evolutionarily very fluid, and that host range can be limited by "recognition" of type III effector protein by host R proteins, as well as by the lack of positive virulence functions in the pathogen.

Methods

Bacterial Strains and Growth Conditions. *P. syringae* strains *Pto*, *PtoΔhrpL* and *Pph6* (all rifampicin resistant) were grown in King's B (KB) media at 28°C with shaking or on KB media agar plates at 28°C. Before FACS sorting, *PtoΔhrpL* was grown in minimal media, modified from ref. 20 to include 1% glycerol, 0.5% dextrose, 10 mM L-glutamate with or without 200 mM or 75 mM L-(+)-arabinose (*Supporting Text*, which is published as supporting information on the PNAS web site), at 28°C with shaking. *Escherichia coli* DH5α was used in all cloning procedures and grown on LB agar plates or in SOB media at 37°C with shaking. The following antibiotics were used for both *P. syringae* and *E. coli*: 50 μg/ml rifampicin, 30 μg/ml kanamycin, 5 μg/ml tetracycline (10 μg/ml for agar plates), and 25 μg/ml cyclohexamide. *PtoΔhrpL* was provided by B. Staskawicz (University of California, Berkeley).

Construction of Plasmids. pCFS40 (21) was modified to carry an idealized Shine–Delgarno sequence from *gene10* of phage T7, upstream of $hrpL_{Pto}$ (*Supporting Text*). Another vector, pBBR1-MCS2 (22), was modified to carry in three different reading frames, an operon fusion of Δ*79avrRpt2* followed by GFP3 (23). The operons are flanked by RNA terminators (*Supporting Text*). These are referred to as DFI vectors. The vector in frame 1 was converted into a Gateway-ready vector by digesting with *Sma*I and ligating in conversion cassette C.1 (Invitrogen).

Library Construction. DNA was extracted from *P. syringae* (24), purified by using CTAB (25), and either partially digested with *Tsp*509I, *Alu*I, *Bst*UI, *Hae*III, or *Rsa*I or physically sheared by using a double stroke shearing device (Fiore Automation, Salt Lake City, UT). Fragments from 0.5–0.8 kb, 0.8–1.4 kb, and 1.4–2.0 kb were extracted and cloned into either *Eco*RI- or *Sma*I-digested and shrimp alkaline phosphatase (SAP)-treated DFI vectors. *E. coli* colonies carrying clones of similarly sized ligation products were pooled and mated *en masse* by modified triparental mating (P. Ronald, personal communication), with *PtoΔhrpL* + pBAD::*hrpL* and pRK2013.

Construction of Full-Length Gene Fusions. Candidate HrpL-induced gene fragments were amplified from cells harboring the DFI plasmids (*Supporting Text*). Contigs were assembled and layered onto the genome sequences of *Pto* or *Pph6* (ref. 11 and www.pseudomonas-syringae.org). Full-length genes and operons were cloned by two-step PCR (*Supporting Text*). PCR products were recombined by BP reaction into pDONOR207 following the manufacturer's recommendations (Invitrogen). All products were sequenced and then recombined into DFI vector 1 by using LR (Invitrogen). Candidate genes were individually mated into *Pto* + pBAD::*hrpL* by triparental mating (P. Ronald, personal communication) with pRK2013.

FACS. FACS was performed on a MoFlo (Cytomation). Analysis by FACS was performed on a FACscan from Becton Dickinson. One microliter of overnight culture cells was diluted into 400 μl of 1× PBS. Events shown in histograms (GFP, FL1) were gated on both side scatter (SSC) and forward scatter. Both were detected in log, and events were triggered on SSC. A total of 10,000 events were collected for each analysis (see *Supporting Text*).

Plant Infiltrations. Infiltrations of leaves of Arabidopsis *rpm1–3* were done as described (26). The HR was scored blind no more than 26 h after inoculation. Results were compared to leaves infiltrated with *Pto* carrying a HrpL-regulated full-length *avrRpt2* or an empty vector.

Results

Vectors and Library Construction for DFI. We created three DFI vectors for "trapping" HrpL-regulated promoters (Fig. 1*A* and *Supporting Text*). These have cloning sites, in all three reading frames upstream of a promoterless avirulence gene lacking the N-terminal coding 79 aa (Δ*79avrRpt2*; refs. 9, 18, and 19). Following Δ*79avrRpt2* in these DFI vectors is a promoterless *GFP3* gene (23). We also created pBAD::*hrpL*, a separate vector with an arabinose-inducible $hrpL_{Pto}$ (*hrpL*; *Supporting Text*). DNA fragments from *Pto* or *Pph6* were cloned into each of the three DFI vectors and mobilized into *PtoΔhrpL* conditionally complemented with pBAD::*hrpL*. The library representation was at least 10-fold excess for each genome in each reading frame.

The DFI Screen for *Pto* HrpL-Regulated Genes. *Pto* was screened for HrpL-regulated genes by using FACS as described in Fig. 1 and the *Supporting Text*. We identified 43 HrpL-regulated genes/operons (*Supporting Text*, Table 2, and Fig. 3, which are published as supporting information on the PNAS web site). Five conserved HrpL-regulated TTSS promoters, *hrpA*, *hrpF*, *hrpJ*, *hrpP*, and *hrpK* (27), were present in contigs represented by many different clones in *Pto* as well as *Pph6*, supporting the contention that the screens were near saturating (Table 2). Fifty-nine *hop* genes were previously proposed in *Pto* and their protein products were classified (www.pseudomonas-syringae.org) as type III effector proteins or helper proteins, based on biological tests (15), or as candidate Hops (14, 15). We found only 29 of the 39 proposed effectors and helpers. Eight of the 10 genes we missed either lacked *hrp*-boxes, are potentially in operons, and/or are located on an endogenous plasmid (Table 2 and www.pseudomonas-syringae.org) that our lab isolate of *Pto* lacks (data not shown). Therefore, only two of the previously identified *hop* genes were actually missed in our screen: *hopAA1–2* and *schV-hopV1*. We also found eight of the 20 genes previously listed as candidate *hop* genes (www.pseudomonas-syringae.org) and seven that were previously suggested to be HrpL-regulated or previously unidentified (Table 2 and ref. 10).

We created 61 full-length gene fusions to Δ*79avrRpt2*::*GFP3*, accounting for 50 of the 59 predicted *hop* genes (from the plasmid-bearing strain of *Pto*; ref. 11 and www.pseudomonas-syringae.org). We included 10 categorized *hop* genes (two that escaped our screen, four that lack *hrp*-boxes, and four that are located on the endogenous plasmid), and 13 additional candidate *hop* genes (eight of which were identified in our screen, five of which were not). Forty-seven of these 61 gene fusions exhibited HrpL-dependent induction (Table 1 and Table 3, which is published as supporting information on the PNAS web site). We observed a strong correlation between HrpL-induction and the presence of a *hrp*-box. Nine of 11 predicted genes without predictable *hrp*-boxes were not induced by HrpL, but seven of these were induced when cloned as operon fusions to their respective upstream genes that do contain *hrp* boxes (Table 3, not bold). Only two possible operons with clear *hrp* boxes were not induced by HrpL in our assays (Table 3).

The DFI Screen for *Pph6* HrpL-Regulated Genes. We also screened *Pph6* and identified 41 HrpL-induced genes or operons (Table 2), including 23 of the 30 predicted *hop* genes in *Pph6* (16). Forty-five genes/operons were fused to Δ*79avrRpt2*::*GFP3* (Table 2). We elected not to retest seven other predicted candidate *hop* genes (16);

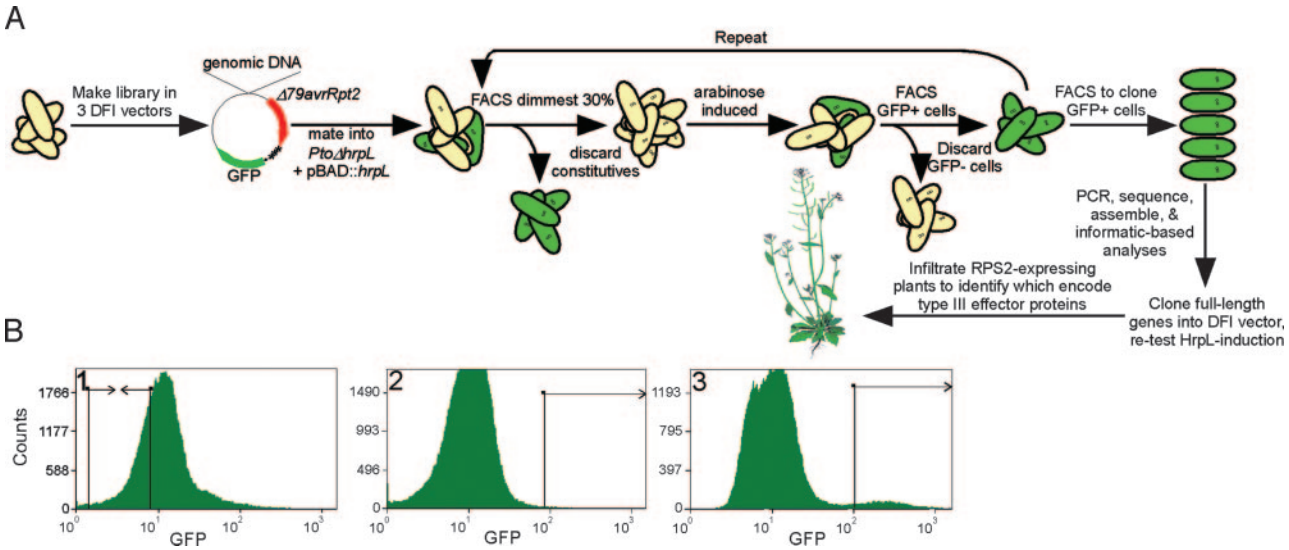


Fig. 1. A high-throughput, FACS-based screen for *P. syringae* type III effector genes. (*A*) Diagram depicting the flow of our screen. Libraries were constructed in DFI vectors 1–3 upstream of Δ*79avrRpt2*::*GFP3* and mobilized into *Pto* carrying pBAD::*hrpL*. Clones carrying HrpL-regulated inserts were isolated in a four-step process using FACS. Cells were first grown in modified minimal medium lacking arabinose, and the least fluorescent ≈30% of cells were collected to eliminate those constitutively expressing GFP. These ≈200,000 cells were subsequently grown in minimal medium with arabinose, and a small population of cells with the highest level of GFP expression, compared to those of a negative control population grown in the absence of arabinose, were collected by FACS (see *B*). These two steps were repeated, except that the GFP-positive cells were individually cloned by FACS after the second arabinose induction. DNA inserts were amplified, sequenced, and assembled. Contigs were analyzed, and full-length genes and operons were identified, cloned upstream of Δ*79avrRpt2*::*GFP3*, and tested for HrpL-induction as well as translocation of Δ79AvrRpt2 into leaves of RPS2-expressing plants. (*B*) Three FACS histograms from the screen for HrpL-induced genes from *Pto*. Histogram 1 shows distribution of GFP-fluorescence of the original library before any enrichment. The boxed area represents the region that was FACS cloned (least fluorescent 28.82% with a mean of 5.73. Histogram 2 shows distribution of GFP-fluorescence after growth in arabinose. A fluorescent population of the brightest 0.44% (mean of 227.66) of cells was sorted (compared to 0.34% and 147.24, respectively, in the same fluorescence range in uninduced controls of the same cell population). Histogram 3 shows distribution of GFP-fluorescence after all four FACS enrichment steps. The boxed area represents the region from which individual cells were FACS cloned; 4.66% with a mean of 325.50 versus 0.42% and 231.13, respectively, for the uninduced negative control. Each histogram represents at least 40,000 cells. *x* axis shows GFP fluorescence in log. *y* axis shows number of cells in each channel.

two corresponded to *hrpZ* and *hrpK* of the *hrp*/*hrc* cluster, the other five were not identified in our screen and have no functional characteristic of a type III effector gene ascribed to them. We confirmed that 36 of 45 fusions were HrpL-induced (Tables 1 and 3). These defined 35 genes, including 21 of the 23 predicted *hop* genes of *Pph6* identified in our screen.

Characterization of HrpL-Induced Proteins for Delivery into Plants. TTSS-dependent translocation of full-length proteins into plant cells is the most direct and most stringent test for defining type III effector proteins (14). All full-length fusions to Δ79AvrRpt2 derived from *Pto* and *Pph6* described above (in wild-type *Pto*), were infiltrated into leaves of RPS2-expressing plants, relying on induction of the endogenous HrpL by *in planta* conditions. This experiment determines which candidate genes encode type III effector proteins, but it will not identify helper proteins. We also analyzed the translated sequences of all of the genes for putative TTSS-dependent sequences (Tables 1 and 3 and refs. 8, 9, 14, and 15). These characteristics include >10% serine or proline in the first 50 aa, an aliphatic amino acid or proline at position 3 or 4, and the absence of negatively charged amino acids in the first 12 residues. We found a high correlation between experimentally translocated proteins and those having amino acids consistent with at least two of the proposed secretion signal rules (Table 1). Proteins that were unable to deliver Δ79AvrRpt2 generally had discordance with these rules or were not HrpL-regulated (Table 3).

Twenty-five proteins from *Pto* reliably elicited RPS2-dependent HR in multiple experiments (Table 1). Twenty of these were identified in our screen; five were missed for reasons defined above. Thirty-two of the 61 fusion proteins consistently failed to elicit the HR (Table 3). AvrPphE$_{Pto}$ (HopX1) and HopG1 elicited weak, inconsistent HR. Nevertheless, we classified them as type III effector proteins. We were unable to clone full-length genes of *hopR1* or *avrE1*$_{Pto}$, but we cautiously list them as type III effectors based on orthology to known type III effector proteins. The sum of our HrpL-dependent expression data and Δ79avrRpt2 translocation data (Tables 1 and 3) suggests that the previous informatics-based predictions significantly overestimated the number of type III effector genes in *Pto* (www.pseudomonas-syringae.org and refs. 7, 14, and 15). We conclude that *Pto* has at least 24, and possibly 28, HrpL-regulated, translocated type III effector proteins, plus *hopO1-2*, which is apparently translocated but HrpL-independent.

Seventeen fusion proteins from *Pph6* reproducibly elicited an HR (Table 1). Twelve of these proteins had been predicted based on homology to other Hops (16). We additionally found a member of the AvrB family (28) not previously identified in *Pph6* (16); we named this protein AvrB3$_{Pph6}$. HopM1$_{Pph6}$ and AvrE$_{Pph6}$ fusions to Δ79avrRpt2 did not elicit reliable HR phenotypes, but were cautiously classified as type III effectors based on orthology to known proteins that traverse the TTSS (Table 1 and ref. 29). Twenty-seven *Pph6* fusion proteins consistently failed to elicit an RPS2-dependent HR (Table 3). Twenty corresponded to genes/operons identified in our screen, and are therefore HrpL-regulated, but not delivered to host cells. One of them, HopV1$_{Pph6}$, did not give a reliable HR, and its translated sequence is not similar enough to the translocated HopV1$_{Pto}$ to use orthology in classifying HopV1$_{Pph6}$ as a type III effector protein. The N-terminal regions do not align and HopV1$_{Pph6}$ is ≈244 residues longer than HopV1$_{Pto}$. The remaining seven proteins correspond to predicted *Pph6* homologs of *Pto* candidate type III effector genes (Table 3). We conclude that *Pph6* encodes at least 17, and possibly 19, HrpL-regulated and translocated type III effector proteins.

Identification of Four Previously Undescribed Type III Effectors in *Pph6*. We identified four previously undescribed type III effector proteins in *Pph6* (Table 1 and Table 4, which is published as

Table 1. Full-length proteins of *P. syringae* capable of delivering Δ*79AvrRpt2* into plant cells

Protein	Score†	Ind‡	Translocation signal*		
			% S	Aliph. AA	Lack-AA
Pto					
Group 1					
HopK1		4.00	18	4(I)	Y
HopY1	7.3	19.33	16	3(I)	Y
HopH1	17.2	45.33	10	4(P)	Y
HopC1	9.4	43.67	16	3(I) 4(V)	Y
HopD1	10.0	7.50	14	3(P) 4(L)	Y
HopQ1	13.6	14.67	24	4(P)	Y
HopAM1-1	5.6	23.33	16	N	Y
HopAA1-1	11.4	7.25	18	3(I)	Y
HopAF1	5.9	5.50	26	3(L)	Y
HopP1	10.4	19.25	20	N	Y
HopAB2	15.2	19.43	18	4(I)	Y
AvrPto1	19.3	36.00	16	4(I)	Y
HopE1	16.0	46.00	14	4(V)	Y
HopAA1-2	6.8	1.79	20	3(I)	Y
HopAR1	7.7	6.50	20	3(P) 4(L)	Y
HopI1	7.9	17.83	14	4(L)	Y
HopAM1-2	5.6	25.50	16	N	Y
Group 2					
HopO1-1	5.1	14.50	24	4(I)	Y
HopV1	5.4	10.00	6	N	N
HopO1-2§	11.4	1.00	28	3(I)	Y
HopF2	12.3	1.83	20	3(I)	Y
HopN1	6.0	8.67	14	3(I)	Y
HopM1	3.0	23.00	10	N	Y
HopB1§	8.0	25.50	12	3(P) 4(V)	Y
HopA1	6.7	18.00	10	3(P) 4(I)	Y
Group 3					
HopR1¶	12.8	3.13	18	4(V)	Y
AvrE1¶	11.2	NT	16	4(P)	Y
HopX1‖	4.2	1.67	14	3(I)	Y
HopG1‖	10.4	19.33	10	3(I)	Y
HopB1§		1.00	12	3(P) 4(V)	Y
HopO1-2§		1.00	28	3(I)	Y
Pph6					
Group 1					
AvrB3	8.7	20.83	20	4(I)	Y
HopAB2	0.5	3.00	16	4(I)	Y
HopAU1	9.1	11.60	10	3(P) 4(V)	Y
HopG1	8.1	9.67	14	3(I)	Y
HopK1		16.78	24	4(I)	Y
HopAW1	7.7	22.40	10	3(V)	Y
AvrB2-3	9.1	18.33	26	4(V)	Y
HopQ1	13.6	11.00	22	4(P)	Y
HopW1-1	6.4	16.00	18	3(P)	Y
HopAV1	7.8	2.25	14	N	Y
HopAF1	7.0	17.67	24	3(L)	Y
HopD1	10.0	21.25	14	3(P) 4(L)	N
HopI1	7.4	23.00	16	4(L)	Y
HopAT1	1.2	13.20	8	4(I)	Y
HopR1	9.9	4.73	16	4(V)	Y
Group 2					
HopX1	9.2	23.22	20	3(I)	Y
HopF2	9.6	17.50	22	3(I)	Y
Group 3					
HopM1‖	3.6	4.67	4	4(P)	Y
AvrE1‖	10.0	1.67	10	3(L)	Y

Proteins in bold are encoded by genes identified in our screen. Genes from each strain are divided into three groups: group 1, expressed as single genes; group 2, expressed in an operon downstream from putative chaperone-encoding genes, except HopB1 and HopX1(*Pph6*), which were expressed downstream from *hrpK*; group 3, inconsistent translocation results.

*Criteria as defined in refs. 8, 9, 14, and 15. Criteria that did not correlate to those of a type III effector protein are underlined.

†Bit score for the *hrp*-box; scores >2.0 are considered reliable; <2.0 are underlined. Scores <0 are not listed.

‡Fold Induction = mean FACS-derived GFP fluorescence of the positive peak divided by the mean fluorescence of uninduced cells. Negative control, empty vector; 1-fold induction. Positive control, AvrRpm1 upstream of GFP; 122-fold induction. NT, not tested.

§Elicited strong HR in 2/3 experiments as both gene and operon fusions, and therefore included in groups 2 and 3.

¶Could not construct full-length fusions to Δ79AvrRpt2.

‖Caused tissue collapse in only one of three experiments.

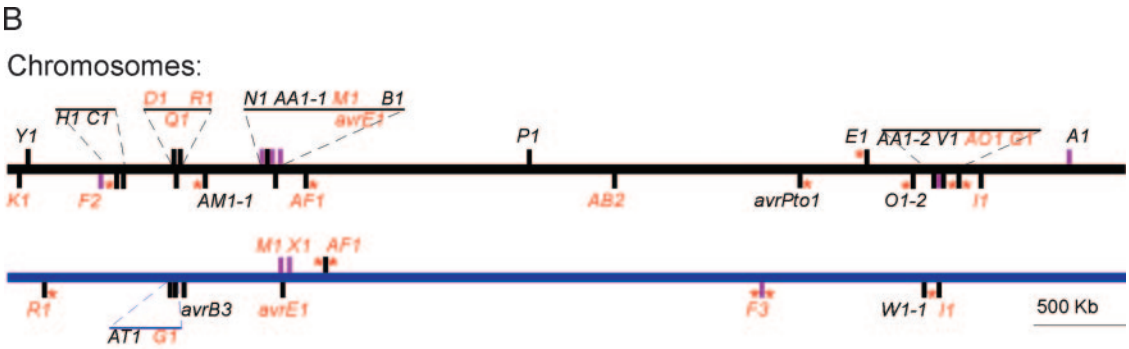
supporting information on the PNAS web site). HopAT1 is predicted to be only 88 aa long. HopAU1 and HopAV1 have homology to predicted proteins from two pathovars of *Xanthomonas* and *Ralstonia*, respectively. A HopAV1 ortholog was identified in a screen for HrpB-induced genes of *Ralstonia* (30). HopAW1 has homology to the translated product of the so-called ORF3 from the pAV511 plasmid of *Pph* race 7 (31), except that *ORF3* from *Pph* race 7 has a 10-bp deletion relative to that of *Pph6*, potentially truncating the protein to 90 residues.

***Pto* and *Pph6* Share 13 Type III Effector Proteins.** These common proteins are, largely, highly related, and homology in most cases spans the length of the deduced protein (Fig. 2*A*). However, when we compared the genomic context of all of the type III effector genes (Fig. 2*B* and *Supporting Text*) only the *avrE1* and *hopM1* genes from each strain are immediately flanked by orthologous ORFs. Both of these genes are adjacent to the *hrp/hrc* cluster, within the conserved effector locus of each strain (27). Two other orthologs share one common flanking ORF, but have rearrangements relative to one another on their other side. In this case, the five predicted ORFs 5′ to $hopI1_{Pto}$ and $hopI1_{Pph6}$ are orthologous and colinear (Fig. 2*C*). In contrast, although only 2.5 kb of DNA separates $hopI1_{Pph6}$ from its 3′ neighboring ORF, 40.2 kb and six mobile element-like genes separate $hopI1_{Pto}$ from the same 3′ orthologous ORF. Interestingly, the chromosomal $virPphA_{Pto}$ ($hopAB2_{Pto}$) is flanked by ORFs that are orthologous to those flanking the nontranslocated $virPphA_{Pph6}$ ($hopAB3_{Pph6}$) gene located on the *Pph6* chromosome (Fig. 2*C*), but not to those that flank the translocated $virPphA_{Pph6}$ ($hopAB2_{Pph6}$) that resides on the *Pph6* plasmid.

$AvrPphE_{Pph6}$ (*hopX1*) is part of the *hrpK* operon. By contrast, $avrPphE_{Pto}$ (*hopX1*) is located on an endogenous plasmid (15). Likewise, *avrPphD* (*hopD1*), *hopQ1*, and *avrRps4* (*hopK1*) are located on the plasmid of *Pph6*, but orthologs are found on the *Pto* chromosome (Fig. 2*B*). *AvrPphD* (*hopD1*) and *hopQ1* are separated by 116 bp in both *Pto* and *Pph6*. The 116-bp spacer regions differ by only 3 bp between the two strains. However, there is no other apparent conservation of flanking orthology between these genes as they have shuttled between chromosome and plasmid. The remaining common type III effector genes were found scattered throughout the chromosomes of their respective strains and do not share common flanking ORFs with their orthologs.

Discussion

P. syringae pathovars rely on their suites of type III effector proteins to colonize evolutionarily diverse plant hosts spanning ≈200 million years of evolution (32). These proteins manipulate the eukaryotic host cell to suppress host defense responses, facilitate nutrient acquisition, and contribute to colony size and spread (33, 34). A detailed understanding of the sets of type III effector genes in strains across this bacterial species will help unravel how host range is determined. For example, *Pto* is an aggressive *Arabidopsis* pathogen, but no strain of *Pph* is able to cause disease or trigger hypersensitive cell death on this host. *Pto* and *Pph* share 13 type III effector proteins. We reason that this set in *Pph6* is insufficient for pathogenicity on *Arabidopsis*. This finding implies that one or more of the 16 extra-*Pto* type III effectors, perhaps in combination with the shared set, would render *Pph6* pathogenic on *Arabidopsis*. Conversely, *Pph6* is a pathogen of bean, and its set of 19 type III effector proteins is sufficient for pathogenicity, apparently in combination with the right set of host target alleles. However, *Pto* triggers an HR on bean. This finding suggests either that one or more of the 16 extra-*Pto* type III effectors is recognized by bean plants or allelic differences in one or that more of the 13 shared type III effector proteins in *Pto* allow for its recognition by bean. Only through definition of the entire type III effector gene suite can these hypotheses be evaluated carefully.



A

Pph 6	*Pto* DC3000	%ID	%Aln	Region
AvrRps4	HopK1	73.5	44.5	N
HopF3	HopF2	39.4	100.0	FL
HopD1	HopD1	94.2	100.0	FL
HopD1	HopAO1	63.1	42.2	N
HopQ1	HopQ1	98.7	100.0	FL
HopR1	HopR1	86.0	100.0	FL
HopM1	HopM1	85.0	76.4	N
AvrE1	AvrE1	83.0	100.0	FL
HopAF1	HopAF1	85.9	100.0	FL
HopAB1	HopAB2	51.7	100.0	FL
HopG1	HopG1	53.8	100.0	FL
HopI1	HopI1	82.7	39.7	N
		79.4	20.7	C
HopX1	HopX1	67.0	100.0	FL

Fig. 2. Genome context of type III effector genes in *Pto* and *Pph6*. (*A*) Orthologous type III effector proteins were compared by using BLASTX analysis of deduced *Pph6* proteins against the translated sequences from *Pto*. Significant homologies are given as % identity (%ID), % of the proteins that aligned (% Aln) and regions of alignment; *n*, amino terminal, FL, full-length entire protein; C, carboxy terminal. (*B*) Type III effector genes are depicted along horizontal bars representing the bacterial chromosome or plasmid (black, *Pto*; blue, *Pph6*). Marks above and below the bar are the genes encoded on the top and bottom strand, respectively. Black and purple marks represent genes encoding type III effector genes and operons, respectively. Letter abbreviations correspond to the unified naming nomenclature of type III effector genes. Orthologous type III effector genes are highlighted in red. Predicted mobile elements are denoted with a red asterisks. (*C*) The orthologous copies of *HopI1* as well as $hopAB2_{Pto}$ and *hopAB3* occupy similar regions in their respective genomes. Flanking ORFs had >80% identity along the length of the sequence. Orthologous genes are connected by dotted red lines. Gray boxes, orthologous flanking ORFs; black bars or boxes, *Pto*; blue bars or boxes, *Pph6*; red asterisks, mobile elements.

Horizontal transmission is likely a mechanism by which *P. syringae* pathovars acquire type III effector genes. All but three of the 13 sets of orthologs common to *Pto* and *Pph6* are found scattered throughout the genomes and plasmids, with almost no flanking colinearity. Consequently, few extant type III effector genes were likely present before the divergence of these two strains. In contrast, two helper genes not linked to the *hrp/hrc* cluster, *hopAJ1* and *hopAK1*, are flanked by orthologous ORFs, suggesting that they may have been present before the divergence of the two pathovars (data not shown). These data support a mix-and-match strategy used by *P. syringae* to acquire type III effector genes. The observation that several type III effector genes are associated with remnants of mobile elements (Fig. 2*B*, red asterisks) supports the notion that horizontal transmission is a key mechanism for acquisition of type III effector genes (35, 36).

Alternatively, or in addition, intra- and intergenic changes can also contribute to the evolution of type III effector suites of *P. syringae* strains. The plasmid of *Pph6* (Fig. 2*B*) carries five type III effector genes and *ORF4* in common with, and colinear to, orthologs on pAV511 of *Pph* race 7 (*Pph7*) (31). However, the *avrPphF* ($hopF3_{Pph6}$ and $hopF1_{Pph7}$) operons are found in different locations between the two strains: on the chromosome of *Pph6* (Fig. 2*B*) and between *avrD1* and *avrPphC* (*avrB2*) on the plasmid in *Pph* 7 (31). The respective ORF1 and ORF2 proteins encoded from the *avrPphF* operon in *Pph6* are quite different from their counterparts in *Pph* 7 (27% and 39% identity, respectively), hence the subclassification of $avrPphF_{Pph6}$ as *hopF3* (16). Interestingly, another copy of *ORF4* is located on the chromosome, just 3′ to $hopF3_{Pph6}$. This finding suggests that a pathogenicity island located on a plasmid was common to an ancestral pathovar of *Pph6* and -*7*, and that subsequent rearrangements have occurred since their divergence. *Pph6* also carries two identical HrpL-regulated genes of 433 bp ($hopW1\text{–}2_{Pph6}$ and $\text{-}3_{Pph6}$) that are homologous to *hopW1–1* (BLASTN = 1.0×10^{-83}, Table 3 and ref. 16). The two fragments are oriented in the same direction and flank a total of five type III effector genes and a HrpL-regulated shikimate kinase-encoding gene (Fig. 2*B*, red asterisks). This structure may indicate a recombination mechanism by which the plasmid could integrate/excise from the genome, like the PAI plasmid pFKN from *P. syringae* pv. *maculicola* M6 (37).

The recognition of any single type III effector protein by a plant R protein will render that individual bacterium avirulent on an otherwise normally susceptible host plant. One method to avoid "recognition" is inactivating the relevant type III effector genes. The genomes of these two *P. syringae* are graveyards of inactive type III effector genes. Both pathovars carry HrpL-expressed pseudogenes interrupted by transposable elements. When fused to Δ79AvrRpt2, none of these fusion proteins elicited an RPS2-dependent HR (Table 3). *Pph6* carries a nearly identical copy of AvrB3 elsewhere in the genome, but it too has a transposon insertion. We did not identify this gene in our screen and conse-

quently did not test it for HrpL-regulation or its protein for translocation. *HopAB3* encodes a functional gene, but not a functional type III effector protein, likely because of the loss of the secretion signal (Table 3 and Fig. 4, which are published as supporting information on the PNAS web site). Finally, *Pph6* carries an inactivated ortholog of *hopAA1*. Although it has an identifiable *hrp*-box, it is not expressed in a HrpL-dependent manner, has a premature stop codon, and is not delivered into plants. Strains may also carry type III effectors that mask the perception of other type III effector proteins (38–40).

We confirmed fewer type III effectors from *Pto* than previously predicted both experimentally and by informatic-based analyses (www.pseudomonas-syringae.org and refs. 7 and 15). This variance is likely due to different assays and criteria for defining type III effectors. Our definition is stringent, based on HrpL-dependent expression from native promoters and the ability to translocate full-length fusion to Δ79AvrRpt2 directly into plant cells. Previous analyses used, for example, secretion of candidate type III effectors into media after expression from the strong, constitutive *npt II* promoter (8). These approaches may resurrect transcriptionally silent or HrpL-independent ORFs as Hops (14) or lead to misclassification of helper proteins as type III effector proteins.

In some cases, protein fusions expressed from the promoter of an endogenous type III effector gene may be insufficient to produce enough protein for delivery of an RPS2-dependent HR. This may be the case for $avrPphE_{Pto}$ and $avrE_{Pph6}$, which had some of the lowest fold-induction ratios and maximum mean GFP fluorescence values (1.67/25 and 1.67/50, respectively) of the HrpL-induced genes in our experiments (ranges of fold-induction and maximum GFP-fluorescence were 1.14–46.67 and 25–2,160, respectively). However, very little protein needs to be expressed and delivered to elicit the RPS2-dependent HR (14). These weakly HrpL-induced genes will nevertheless be identified in our screen because of their HrpL-dependent gene expression. These genes may require more sensitive assays than fusions to Δ79AvrRpt2, which rely on their endogenous promoters, for proper classification as type III effector proteins.

Nearly all Δ79AvrRpt2 protein fusions gave reproducible positive or negative RPS2-dependent HR results. However, a few results were inconclusive. Some fusions to Δ79AvrRpt2 might be unstable, unable to pass through the TTSS, or misfolded to preclude triggering of RPS2-dependent HR. Arguing against this is our finding that $HopR1_{Pph6}$ (>200 kDa before the addition of Δ79AvrRpt2) did elicit RPS2-dependent HR. Therefore, we believe that few, if any, of the higher expressed proteins were incorrectly categorized.

In summary, we created a high-throughput, near-saturating screen for type III effector genes of *P. syringae* and corroborated our biological results with bioinformatics analyses. The expeditious nature of our screen enabled thorough screening for type III effector genes, and our isolation of all five *hrp* promoters from each strain suggests that the screen is probably saturating. Our results with *P. syringae* pv. *tomato* and *phaseolicola* indicate that this method will be useful for identifying entire suites of type III effector genes from diverse pathovars of *P. syringae* whose genomes have not been sequenced. We have targeted 13 additional strains for analyses based on their phylogenetic relationships (32), maximizing both evolutionary distance among strains and their respective hosts. We envision that comparisons between pathovars will be useful in understanding the function and evolution of type III effectors throughout this species. These findings should be generalized to all pathogens that rely on type III systems to colonize a wide variety of host species.

We thank Ms. Brittain Fish and Dr. Han-Suk Kim for technical assistance, Drs. Jurg Schmid and Steve Goff and Syengenta Biotechnology for support, and the members of the J.L.D., S.R.G., and F.M.A. laboratories for their insightful comments. We thank The Institute for Genome Research (TIGR) for providing *Pto* and *Pph6* genomic sequences before publication. This research was supported by National Institutes of Health Grants RO1GM066025 (to J.L.D. and S.R.G.) and RO1GM48707 (to F.M.A.). J.H.C. was funded by National Institutes of Health/National Research Service Award Fellowship F32-GM20296-02S1.

1. Bogdanove, A. J., Beer, S. V., Bonas, U., Boucher, C. A., Collmer, A., Coplin, D. L., Cornelis, G. R., Huang, H. C., Hutcheson, S. W., Panopoulos, N. J., *et al*. (1996) *Mol. Microbiol.* **20,** 681–683.
2. Hueck, C. J. (1998) *Microbiol. Mol. Biol. Rev.* **62,** 379–433.
3. Lindgren, P. B., Peet, R. C. & Panapoulos, N. J. (1986) *J. Bacteriol.* **168,** 512–522.
4. Dangl, J. L. & Jones, J. D. G. (2001) *Nature* **411,** 826–833.
5. Xiao, Y., Heu, S., Yi, J., Lu, Y. & Hutcheson, S. W. (1994) *J. Bacteriol.* **176,** 1025–1036.
6. Innes, R. W., Bent, A. F., Kunkel, B. N., Bisgrove, S. R. & Staskawicz, B. J. (1993) *J. Bacteriol.* **175,** 4859–4869.
7. Schechter, L. M., Roberts, K. A., Jamir, Y., Alfano, J. R. & Collmer, A. (2004) *J. Bacteriol.* **186,** 543–555.
8. Petnicki-Ocwieja, T., Schneider, D. J., Tam, V. C., Chancey, S. T., Shan, L., Jamir, Y., Schechter, L. M., Janes, M. D., Buell, R., Tang, X., *et al*. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 7652–7657.
9. Guttman, D. S., Vinatzer, B. A., Sarkar, S. F., Ranall, M. V., Kettler, G. & Greenberg, J. T. (2002) *Science* **295,** 1722–1726.
10. Boch, J., Joardar, V., Gao, L., Robertson, T. L., Lim, M. & Kunkel, B. N. (2002) *Mol. Microbiol.* **44,** 73–88.
11. Buell, C. R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I. T., Gwinn, M. L., Dodson, R. J., Deboy, R. T., Durkin, A. S., Kolonay, J. F., *et al*. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 10181–10186.
12. Zwiesler-Vollick, J., Plovanich-Jones, A. E., Nomura, K., Bandyopadhyay, S., Joardar, V., Kunkel, B. N. & He, S. Y. (2002) *Mol. Microbiol.* **45,** 1207–1218.
13. Fouts, D. E., Abramovitch, R. B., Alfano, J. R., Baldo, A. M., Buell, C. R., Cartinhour, S., Chatterjee, A. K., D'Ascenzo, M., Gwinn, M. L., Lazarowitz, S. G., *et al*. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 2275–2280.
14. Greenberg, J. T. & Vinatzer, B. A. (2003) *Curr. Opin. Microbiol.* **6,** 20–28.
15. Collmer, A., Lindeberg, M., Petnicki-Ocwieja, T., Schnieder, D. J. & Alfano, J. R. (2002) *Trends Microbiol.* **10,** 462–469.
16. Lindeberg, M., Stavrinides, J., Chang, J. H., Alfano, J. R., Collmer, A., Dangl, J. L., Greenberg, J. T., Mansfield, J. W. & Guttman, D. S. (2005) *Mol. Plant–Microbe Interact.*, in press.
17. Valdivia, R. H. & Falkow, S. (1997) *Science* **277,** 2007–2011.
18. Mudgett, M. & Staskawicz, B. (1999) *Mol. Microbiol.* **32,** 927–941.
19. Guttman, D. S. & Greenberg, J. T. (2001) *Mol. Plant–Microbe Interact.* **14,** 145–155.
20. Huynh, T. V., Dahlbeck, D. & Staskawicz, B. J. (1989) *Science* **245,** 1374–1377.
21. Newman, J. R. & Fuqua, C. (1999) *Gene* **227,** 197–203.
22. Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., Jr., & Peterson, K. M. (1995) *Gene* **166,** 175–176.
23. Cormack, B. P., Valdivia, R. H. & Falkow, S. (1996) *Gene* **173,** 33–38.
24. Syn, C. K. & Swarup, S. (2000) *Anal. Biochem.* **278,** 86–90.
25. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1989) *Molecular Cloning*: *A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
26. Mackey, D., Holt, B. F., III, Wiig, A. & Dangl, J. L. (2002) *Cell* **108,** 743–754.
27. Alfano, J. R., Charkowski, A. O., Deng, W. L., Badel, J. L., Petnicki-Ocwieja, T., van Dijk, K. & Collmer, A. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 4856–4861.
28. Tamaki, S., Dahlbeck, D., Staskawicz, B. J. & Keen, N. T. (1988) *J. Bacteriol.* **170,** 4846–4854.
29. DebRoy, S., Thilmony, R., Kwack, Y.-B., Nomura, K. & He, S. Y. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 9927–9932.
30. Cunnac, S., Occhialini, A., Barberis, P., Boucher, C. & Genin, S. (2004) *Mol. Microbiol.* **53,** 115–128.
31. Jackson, R. W., Athanassopoulos, E., Tsiamis, G., Mansfield, J. W., Sesma, A., Arnold, D. L., Gibbon, M. J., Murillo, J., Taylor, J. D. & Vivian, A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 10875–10880.
32. Sawada, H., Suzuki, F., Matsuda, I. & Saitou, N. (1999) *J. Mol. Evol.* **49,** 627–644.
33. Chang, J. H., Goel, A. K., Grant, S. R. & Dangl, J. L. (2004) *Curr. Opin. Microbiol.* **7,** 11–18.
34. Espinosa, A. & Alfano, J. R. (2004) *Cell Microbiol.* **6,** 1027–1040.
35. Rohmer, L., Guttman, D. S. & Dangl, J. L. (2004) *Genetics* **167,** 1341–1360.
36. Kim, J. F., Charkowski, A. O., Alfano, J. R., Collmer, A. & Beer, S. V. (1998) *Mol. Plant–Microbe Interact.* **11,** 1247–1252.
37. Rohmer, L., Kjemtrup, S., Marchesini, P. & Dangl, J. L. (2003) *Mol. Microbiol.* **47,** 1545–1562.
38. Mackey, D., Belkhadir, Y., Alonso, J. M., Ecker, J. R. & Dangl, J. L. (2003) *Cell* **112,** 379–389.
39. Ritter, C. & Dangl, J. L. (1996) *Plant Cell* **8,** 251–257.
40. Tsiamis, G., Mansfield, J. W., Hockenhull, R., Jackson, R. W., Sesma, A., Athanassopoulos, E., Bennett, M. A., Stevens, C., Vivian, A., Taylor, J. D., *et al*. (2000) *EMBO J.* **19,** 3204–3214.

Supporting Text

Construction of Plasmids. All oligonucleotide sequences used are listed in Table 5. pBAD::*hrpL* construction: $hrpL_{Pto}$ (*hrpL*) was PCR amplified by using *Pfu* (Stratagene) and the oligonucleotides HYZ046 and HYZ045, then cloned into pCR2.1 (Invitrogen). The pCR2.1::*hrpL* was sequenced, digested with *Nde*I and *Bam*HI, and cloned into *Nde*I-*Bam*HI-digested pCR2.1::*GFP2* (L. Rohmer and J.L.D., unpublished data), replacing the *GFP* gene. This plasmid carries a hybrid *npt II-lacZ* promoter followed by an idealized Shine-Delgarno from *gene10* of phage T7 and a mutant version of GFP (1). The *hrpL* gene was then digested with *Xba*I and *Hin*dIII to capture the Shine-Delgarno box, and cloned into *Nhe*I-*Hind*III digested pCFS40 (2). pBAD::*hrpL* exhibited tight regulation in the absence of arabinose and up to ~90% of a positive control population exhibited high expression in the presence of arabinose (data not shown). Expression was most significantly induced in modified minimal media. Expression of the endogenous *hrpL* gene was also strongly repressed in this media (data not shown).

Differential Fluorescence Induction (DFI) Vectors. pBBR1-MCS2 (3) was modified to carry the RNA terminators from rrnBT_1T_2 and phage T4 to prevent read-through from genes on the vector (J.H.C., unpublished data). Mutant *GFP3* (1) was cloned into the modified vector digested with *Eco*RI and *Eco*RV as an *Eco*RI-*Sph*I (blunt) fragment. GFP3 has an excitation peak corresponding to the argon laser (488 nM) of the FACS (1). Oligos HYZ017 and HYZ018, adding three stop codons in each frame, were annealed and cloned upstream from the Shine-Delgarno sequence from *gene10* of phage T7 and *GFP3* as an *Eco*RI-*Xba*I fragment, yielding vector 125.1. *AvrRpt2* was modified to change a valine codon at position 95. This mutation did not alter the amino acid sequence of AvrRpt2 in frame 1, but did introduce a stop codon closer to the N terminus in frame 2. Strains carrying full-length *avrRpt2* with this codon change still elicit an RPS2-dependent HR. The *avrRpt2* fragment encoding the 177 C-terminal amino acids (D *79avrRpt2*) was PCR amplified in six separate reactions using Pfu and HYZ122, HYZ124, or HYZ126 with HYZ129 and HYZ123, HYZ125, or HYZ127 with HYZ128. PCR products were directly cloned into *Eco*RI-*Bam*HI digested 125.1 vector following the sticky-end PCR protocol (4). Translational fusions providing *hrpL*-regulated promoters and N-terminal sequences sufficient for type III-dependent translocation can deliver D 79AvrRpt2 fusion proteins into plant cells to elicit RPS2-dependent HR (5, 6). The resulting DFI vectors are depicted in Fig. 1A. DFI vector 1 was converted into a Gateway-ready vector by digesting with *Sma*I and ligating in conversion cassette C.1 (Invitrogen). Plasmids pCFS40 and pBBR1-MCS2 were provided by C. Fuqua and Kovach *et al.*, respectively (2, 3).

Candidate HrpL-induced gene fragments were amplified from cells harboring the DFI plasmids using *Taq* polymerase and primers, HYZ163 and HYZ166 (Table 5). PCR products were treated with 5 units of exonuclease I and 0.5 unit of SAP at 37°C for 40 min and heat terminated at 80°C for 30 min. Products were directly sequenced by using the BigDye Terminator Cycle Sequencing Reaction method and vector-specific primers HYZ042 and HYZ152 (Table 5) on an ABI 3700 or 3730.

Full-length genes were amplified by using AccuPrime Pfx and gene-specific primers that contained a portion of the B1 and B2 universal Gateway primers followed by another round of amplification using B1 and B2 that overlapped each set of gene-specific primers (Table 5). We included ~100 bp of upstream sequence genes with identifiable *hrp*-boxes. When no *hrp*-box could be detected, we included ~750 bp of upstream sequence from the predicted ATG. The largest distance we observed between a *hrp*-box and a start codon for a confirmed HrpL-dependent gene is 533 bp in *hopAD1*. No stop codons were included. We also created operon fusions that carried all of the genes potentially regulated by a predicted upstream *hrp*-box-containing gene.

Bioinformatic Analyses. Perl was used to automate the analyses of sequence data that were stored in a MySQL relational database. Base calling and sequence quality assessment of raw ABI

sequence files was performed using phred (7, 8). Reads for each genome were assembled by using phrap (www.phrap.org) . Contigs were trimmed to remove sequences of phrap <20 averaged over a nine-base sliding window. Trimmed contigs were analyzed by using blast against GBBCT (ftp://ftp.ncbi.nih.gov/ncbi-asn1). blastx results were used for first pass gene identification. Second pass gene identification was conducted by using glimmer (9). *Hrp*-boxes were identified by using hmmer (http://hmmer.wustl.edu) trained from previously identified *hrp*-boxes (J.M.U. and F.M.A., unpublished).

The nearest predicted ORFs flanking each type III effector gene were identified from a table of predicted ORFs for *Pto* (ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/Pseudomonas_syringae/NC_004578.ptt) and by blastx of flanking sequences for *Pph6*. For comparing the ORFs flanking type III effector genes, no putative ORFs of less than 125 bp in length were used, and all regions homologous to mobile elements were ignored.

The DFI Screen. We screened each library under two inducing conditions for each strain, once to isolate more weakly expressed *hrpL*-dependent promoters (200 mM arabinose for 24 h in the first round of enrichment; 200 mM arabinose for 6 h before FACS cloning) and once to clone the *hrpL*-dependent genes whose high levels of expression might lead to GFP toxicity (200 mM arabinose for 6 h in the first round of enrichment; 75 mM for 4 h before FACS cloning). Growth in arabinose alternated with growth in the absence of arabinose. Cells were sorted after each step.

A total of 2,590 sequencing reads from *Pto* assembled into 124 contigs (Fig. 3). Of these, 84 were reconfirmed as HrpL-inducible by FACS analysis; 96.6% of the reads assembled into these contigs (Fig. 3). We did not retest the remaining 40, accounting for only 87 reads, for HrpL induction, and considered them to be false positives. We used a HMM-based search to determine that 65 of 124 contigs carried, or were associated with contigs that carried, identifiable *hrp*-boxes (Fig. 3). Searches against the genome sequence of *Pto* and the GBBCT database were used to identify contigs related to known HrpL-induced genes. Sixty-six of 84 contigs carried known or predicted HrpL-induced genes. For *Pph6*, 2,347 reads assembled into 51 contigs of greater than three reads. HrpL-dependent induction of representative clones, HMM-based searches and homology-based searches were performed as for *Pto* (data not shown).

Modifications to the Hop Protein Database (www.pseudomonas-syringae.org) as Reflected in Table 2. Changes reflect the reclassification of two candidate Hops, HopS1 and HopAG, as type III effectors proteins (7) and include the *hrp* operons and other HrpL-induced genes. HrpA1, HrpZ1, and HrpK were moved from "Identified, confirmed type III effector or helper genes" to "*hrp* genes, operons; measure of saturation". We also indicate the newly proposed *hop* names for each gene, where applicable (11). A standard nomenclature for type III secretion system (TTSS)-delivered proteins has been proposed, using the Hop (for Hrp outer protein) name, followed by an assigned letter, number, and subscript identifying the pathovar and strain from which they were isolated (11). For simplicity, we include here the subscript only when it is unclear which of the two strains in our study is being referenced, or when the type III effector protein is clearly related to a previously described *avr* gene. Families of Hops are further divided into subfamilies based on phylogenetic analyses and their relationships are denoted by the common assigned letter but followed by a different number. Paralogs share the same letter and number but are further defined with an additional hyphenated number.

1. Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., Jr., & Peterson, K. M. (1995) *Gene* **166,** 175-176.

2. Huynh, T. V., Dahlbeck, D. & Staskawicz, B. J. (1989) *Science* **245,** 1374-1377.

3. Newman, J. R. & Fuqua, C. (1999) *Gene* **227,** 197-203.

4. Zeng, G. (1998) *BioTechniques* **25,** 206-208.

5. Guttman, D. S., Vinatzer, B. A., Sarkar, S. F., Ranall, M. V., Kettler, G. & Greenberg, J. T. (2002) *Science* **295,** 1722-1726.

6. Mudgett, M. & Staskawicz, B. (1999) *Mol. Microbiol.* **32,** 927-941.

7. Ewing, B. & Green, P. (1998) *Genome Res.* **8,** 186-194.

8. Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998) *Genome Res.* **8,** 175-185.

9. Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. (1999) *Nucleic Acids Res.* **27,** 4636-4641.

10. Schechter, L. M., Roberts, K. A., Jamir, Y., Alfano, J. R. & Collmer, A. (2004) *J. Bacteriol.* **186,** 543-555.

11. Lindeberg, M., Stavrinides, J., Chang, J. H., Alfano, J. R., Collmer, A., Dangl, J. L., Greenberg, J. T., Mansfield, J. W. & Guttman, D. S. (2005) *Mol. Plant--Microbe Interact.*, in press.

Supplemental Figure 3

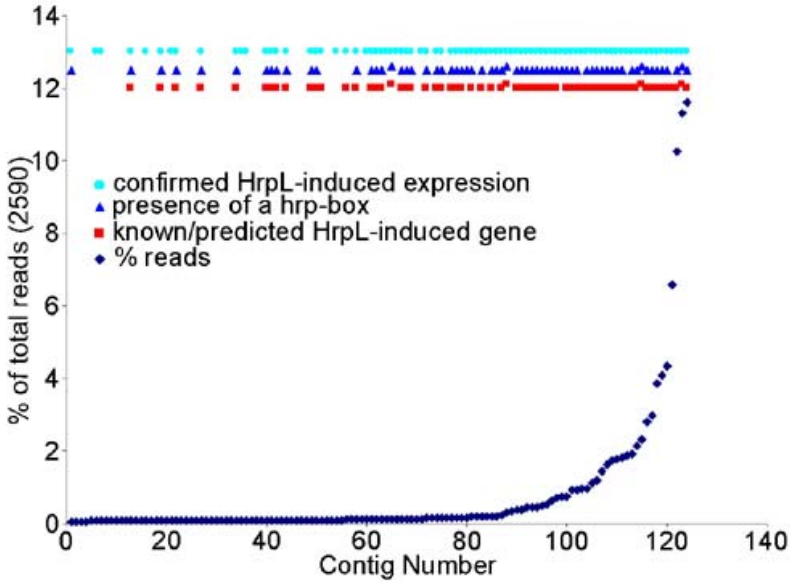


Fig. 3. Distribution of assembled contigs containing HrpL-induced *Pto* clones. The X-axis represents contig number (to 124). The Y-axis shows percent representation [reads in a contig/total number of assembled reads (2,590); blue diamonds]. Also shown are contigs with identifiable *hrp*-boxes (65 total; blue triangles), whether representative clones were confirmed regulated by HrpL (84 total; turquoise circles), and whether a contig carried a known HrpL-induced gene from *Pto* (66 total; red squares). Symbols elevated relative to others depict contigs with more than one *hrp*-box or HrpL-induced gene. Contig distribution in the *Pph*6 was similar to *Pto* (data not shown). Fifty-one HrpL-induced contigs of greater than 3 reads were assembled from a total of 2,347 reads. All *hrp* genes were identified (Table 2).

Supplemental Figure 4

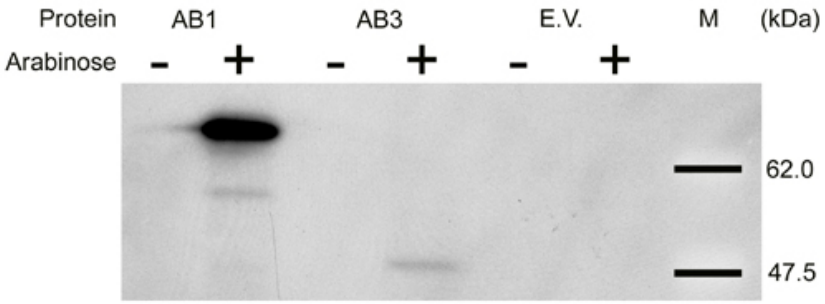


Fig. 4. Western blot showing that *hopAB3* is regulated by HrpL. Full-length genes (promoter, ORF, and no stop codon) were recombined by LR reaction into a broad-host range vector carrying a carboxy-terminal hemagglutinin (HA) epitope tag sequence (not shown). These were transferred to *Pto* + pBAD::*hrpL* by three-way mating. Cells carrying *hopAB1* (AB1) and *hopAB3* (AB3) fused to the HA sequence or the empty HA vector (E.V.) were grown in modified minimal media with (+) or without (-) 100 mM arabinose for 24 h. Equivalent concentrations of cells were spun down, resuspended in 1´SDS loading buffer, and resolved by SDS/PAGE on a 10% acrylamide gel. Blots were probed with anti-HA and then with horseradish peroxidase-conjugated anti-rat antibodies. The predicted sizes of the protein products were 59.6 kDa and 43.3 kDa for HopAB1 and HopAB3, respectively. The translated recombination and HA epitope sequences and HA add an additional ~2.7 kDa. HopAB1 is a functional type III effector homolog of HopAB3 and was thus our positive control for HrpL-dependent expression.

Table 2. Definition of the HrpL regulon from *Pto* and *Pph* 6

PSPTO *	Hop†	Alt name	*Pto* ‡	*Pph* ‡	PSPTO *	Hop†	Alt name	*Pto* ‡	*Pph* ‡
Identified, confirmed type III effector or helper genes					***hrp* genes, operons; measure of saturation**				
852	*hopAJ1*	*hopPmaG*	52	185	1381, 1382	*hrpA1-Z1*	*hrpA-Z*	6	30
4101	*hopAK1*	*hopPmaH*	29	9	1387	NA	*hrpF*	5	12
2678	*hopP1*	*hopPtoP*	105		1403	NA	*hrpJ*	63	41
1373	*hrpW1*	*hrpW*	8	29	1398	NA	*hrpP*	34	13
1377	*avrE1*	*avrE*	3	4	1405	*hrpK1*	*hrpK*	77	52
1022	*hopAM1-1*	*avrPpiB1*	48		**Candidate Hop genes identified**				
4001	*avrPto1*	*avrPto*	85		4703	*hopAQ1*	*holPtoAA*	51	
3087	*hopAB2*	*virPphA*	12	12	883	*hopR1*	*holPtoR*	51	235
4776	*hopI1*	*hopPmaI*	17	14	4593	*hopT1-2*	*holPtoU2*		
5354	*hopA1*	*hopPsyA*	18(O)		4993, 4996	NA	*hopPtoB2 (Tn)*		
1372	*hopAA-1*	*hopPtoA1*	6		875	NA	*ipx47*		
4718	*hopAA1-2*	*hopPtoA2*			5061	*hopAN1*	*ipx53*		
1406	*hopB1*	*hopPtoB1*	77(O)		61	*hopY1*	*holPtoY*	14	
589	*hopC1*	*avrPpiC2*	312		4588	*hopS2*	*holPtoZ2*	32(O)	
876	*hopD1*	*avrPphD1*	64	131	877	*hopQ1-1*	*hopPtoQ*	25	67
4722	*hopAO1*	*avrPphD2*	279		869	NA	ORF20		
4331	*hopE1*	*hopPtoE*	320		835	NA	ORF23	44(O)	46(O)
502	*schF-hopF2*	*avrPphF*	118(O)	40 (O)	836	NA	ORF24	44(O)	46(O)
4727	*hopG1*	*hopPtoG*	47	28	474	*hopAS1*	ORF01152		
588	*hopH1*	*hopPtoH*	308		905	*hopAH2*	*holPsyAH*		
44	*hopK1*	*avrRps4*	26	32	906	*hopAI1*	*holPsyAI*		
1375	*schM-hopM1*	*hopPtoM*	8(O)	6 (O)	4732	*hopQ1-2*	*holPtoQ2*		
1370	*schN-hopN1*	*hopPtoN*	51(O)		4590	*hopT2*	*holPtoU3*		
4597, 4595	*hopS1*	*hopPtoS4*	81		4724, 4726	*hopD::IS52*	*hopPtoD1-related(Tn)*	279	
	hopS1' ¶	*hopPtoS4'*	81		4592	*hopO1-2*	*hopPtoO2*		
901	*hopAG*	*hopPtoAG*	10(O)		4188	NA	ORF36		
	hopAG' ¶	*hopPtoAG'*	10		**Other HrpL-induced genes§**				
4691	*hopAD1*	*hopPtoI*	6		1378	NA	CEL ORF1	3	13
1568	*hopAF1*	*avrXv3*	29	100	2105	NA	ORF07856	7	
4720	*schV-hopV1*	*hopPtoV*		4(O)	1407	Not predicted		17	
A0005	*hopAM1-2*	*avrPpiB2*	P		4680	NA	*cfl*	179	
A0018	*hopO1-1*	*hopPtoO*	P		371	NA	*iaaL*	111	
A0019	*hopT1-1*	*hopPtoT1*	P(NHO)		4705	NA	*corS*	20	
A0012	*hopX1*	*avrPphE*	P	52(O)	834	NA	ORF22	44(O)	46(O)
1179	*hopJ1*	*hopPmaJ*	NH		4589	NA	ORF14	32(O)	
2872	*hopL1*	*hopPtoL*	NH		NA	*avrD1*			31
501	*hopU1*	*hopPtoS2*	(NHO)		NA	*hopAB3*			3
4594	*hopO1-2*	*hopPtoS3*	(NHO)		NA	NA	glycosyl hydrolase, family 25		47
NA	*hopW1-1*			73	NA	NA	polygalacturonase		30
NA	*hopW1-2 -3*			30	NA	NA	shikimate kinase 2		80
NA	*avrB3*			26	NA	NA	AY803998		235
NA	*avrB2*			58	NA	NA	orphan hrp-box		7
NA	*hopAT1*			185					
NA	*hopAU1*			*33*					
NA	*hopAV1*			*79*					
NA	*hopAW1*			*46*					

*www.pseudomonas-syringae.org/. PSPTO= *P. syringae* pv *tomato* DC3000 predicted ORF #. †Hop = *hrp* outer protein. NA= not a *hop* gene or not previously assigned a *hop* name. ‡ # of sequencing reads in contig; (O) = Expressed from an operon. (NHO)= may be expressed from an operon. The number of reads associated with the contig are listed regardless of whether the ORF was present in the contig sequence. Blank = not identified; P = located on plasmid; NH = did not identify a *hrp*-box. § PSPTO1378, PSPTO2105, and PSPTO1407 were re-confirmed. *Pto "o*thers" represent genes located on the contig that may not be the HrpL-induced gene. Genes highlighted in gray are present in *Pph* 6 and

not in *Pto*. ¶ Data are from ref. 1)

1. Schechter, L. M., Roberts, K. A., Jamir, Y., Alfano, J. R. & Collmer, A. (2004) *J. Bacteriol.* **186,** 543-555.

Table 3. Full-length proteins of *P. syringae* that did not deliver Δ79AvrRpt2 into plant cells

					Translocation signal ¶				
Upstream Protein†	Common Name	Unified Name or PSPTO#	Score‡	Fold-Ind.§	% S	% P	Aliphatic AA	Lack of - AA	Comments regarding HR assay
				***Pto* DC3000**					
		PSPTO0834	7.5	29.00	4	2	4(P)	N	
		HopAJ1	10.3	15.67	18	4	N	Y	
		HopAG'	3.1	2.50	12	8	3(P) 4(I)	Y	
		HrpW1	10	46.00	12	12	3(I)	Y	
	CEL ORF1	**PSPTO1378**	6.6	16.00	10	14	4(V)	Y	
		PSPTO01407	9.4	26.88	12	2	3(L)	N	
	ORF07856	**PSPTO2105**	13.3	8.70	6	2	N	N	
		HopAK1	10.5	24.80	6	8	4(I)	Y	
		PSPTO4589	19.1	22.50	6	2	N	N	
		HopS1'	11.4	46.67	24	2	3(I)	Y	Spotty, weak and infrequent tissue collapse
		HopAD1	1.8	1.14	16	6	3(I)	Y	
		HopAQ1		10.00	16	8	4(I)	Y	
		SchO	5.1	15.67	10	2	N	N	
		HopU1		1.00	10	4	3(I)	Y	
		PSPTO0836		1.00	10	0	N	N	
		HopJ1		Con	4	4	4(I)	N	
	AvrF	PSPTO1376		1.00	6	4	N	N	
		PSPTO01409		1.00	6	6	N	Y	
		HopL1		1.00	4	0	4(L)	Y	
		HopS2		1.00	14	2	N	Y	
		HopT1-2		1.00	12	6	N	N	
		HopT1-1		1.00	14	4	4(V)	Y	
SchF-HopF2		HopU1	12.3	2.67	10	4	3(I)	Y	
SchM1-HopM1	AvrF	PSPTO1376	3.0	8.55	6	4	N	N	
PSPTO1407		**PSPTO1408**	9.4	4.50	8	6	3(P)	Y	
PSPTO1407-1408		PSPTO1409	9.4	3.00	6	6	N	Y	
PSPTO4589		HopS2	19.1	14.67	14	2	N	Y	Caused tissue collapse in 1 of 4 experiments
HopS1'-HopO1-2		HopT1-2	11.4	1.00	12	6	N	N	
SchO-HopO1-1		HopT1-1	5.1	10.00	14	4	4(V)	Y	
PSPTO0834		PSPTO0835	7.5	20.00	8	2	3(V)	N	
PSPTO0834-PSPTO0835		PSPTO0836	7.5	26.60	10	0	N	N	
HopS1'		HopS1	11.4	1.00	24	2	3(I)	Y	
				***Pph* 6**					
		HopAK1	9.8	20.33	8	16	4(I)	Y	
	AvrD	**AvrD1**	15.1	26.50	14	2	4(L)	N	
		HrpW1	8.4	32.67	12	10	3(I)	Y	
		HopAJ1	10.3	11.57	10	6	N	Y	
	AvrPphE	**HopX1 w/o HrpK**		1.00	20	10	3(I)	Y	
	Orphan hrp-box		7.3	22.67	Orphan *hrp*-box				
	CEL ORF1		6.6	8.40	12	8	4(V)	Y	
		HopAB3	11.8	6.67	8	6	3(L)	N	
	glycosyl hydrolase		0.9	13.33	6	0	N	N	
		PSPTO0834	9.3	22.77	4	4	N	N	

		HopW1-2, -3	4.3	4.50	20	10	N	N	
	polygalacturonase		5.2	4.29	8	0	3(V)	Y	
	205555-205079		5.1	4.50	4	6	3(I)	Y	
	AvrF			1.00	4	6	4(P)	N	
		HopJ1		Con	6	4	4(L)	N	
		HopAA1-1 part 1	10.6	1.00	20	2	N	N	
		PSPTO0835		1.00	12	2	N	N	
		PSPTO0836		1.00	12	2	3(V) 4(I)	N	
		PSPTO2105	12.2	2.55	14	6	N	Y	
PSPTO0834		PSPTO0835	9.3	17.33	12	2	N	N	
PSPTO0834-PSPTO0835		PSPTO0836	9.3	20.38	12	2	3(V) 4(I)	N	
HopAA1-1 part 1		HopAA1-1 part 2	10.6	1.00	20	2	N	N	
Orphan hrp-box	transposases		7.3	2.00	Downstream of Orphan *hrp*-box				
SchV1		HopV1	8.6	12.25	10	4	N	Y	
SchM-HopM1		AvrF	3.6	2.25	4	6	4(P)	N	

Proteins in **bold** encoded from genes identified in our screen. *Proteins are presented in three groups, separated by horizontal bold lines. Group 1: identified in our screen. Group 2: not identified in our screen but previously predicted to be type III effectors. Group 3: expressed from operons. †For operon fusions. ‡Bit score for the *hrp*-box; scores above 2.0 are reliable; <2.0 are highlighted in gray. Scores less than 0 are not listed. §Mean fluoresence of the most positive peak divided by the mean fluoresence of uninduced cells. Negative control: empty vector; 1.0-fold induction. Positive control: AvrRpm1 upstream of GFP; 122-fold induction. "con" indicates GFP expression was positive (>100) in the absence of induction, and was not induced by arabinose. ¶ Data are from refs. 1-4. Criteria that did not correlate to those of a type III effector protein are highlighted in gray.

1. Petnicki-Ocwieja, T., Schneider, D. J., Tam, V. C., Chancey, S. T., Shan, L., Jamir, Y., Schechter, L. M., Janes, M. D., Buell, R., Tang, X., *et al*. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 7652-7657.

2.Guttman, D. S., Vinatzer, B. A., Sarkar, S. F., Ranall, M. V., Kettler, G. & Greenberg, J. T. (2002) *Science* **295,** 1722-1726.

3. Greenberg, J. T. & Vinatzer, B. A. (2003) *Curr. Opin. Microbiol.* **6,** 20-28.

4. Collmer, A., Lindeberg, M., Petnicki-Ocwieja, T., Schnieder, D. J. & Alfano, J. R. (2002) *Trends Microbiol*. **10,** 462-469.

Table 4. BLASTP results of type III effectors identified from *Pph* 6

Type III effector*	Genbank accession no.	Homolog	Genbank accession number	E-value, BLASTP
HopAT1	AY803994	None		
HopAU1	AY803995	conserved hypothetical protein of *Xanthomonas campestris* pv. *campestris* str. ATCC 33913	gi|21229709|ref|NP_635626.1|	2.0 X 10^{-29}
		conserved hypothetical protein of *X. axonopodis pv. citri str. 306*	gi|21243513|ref|NP_643095.1|	4.0 X 10^{-27}
HopAV1	AY803996	putative transmembrane protein of *Ralstonia solanacearum* GMI1000	gi|17548953|ref|NP_522293.1|	0.0
HopAW1	AY803997	ORF3 from the plasmid pAV511 from *Pph* race 7 strain 1449b	gi|5702218|gb|AAD47205.1|	2.0 x 10^{-37}

*Proteins were named according to the unified naming nomenclature for type III effectors

Table 5. Oligonucleotide sequences

Primer Name	Gene	strand	Sequence (5' to 3')
HYZ042	*T4 RNAP terminator*	Top	GCTTGCTCAATCAATCACC
HYZ045	*hrpL*	Bottom	AAGCTTTCAGGCGAACGGGTCG
HYZ046	*hrpL*	Top	CATATGTTTCAGAAGATTCTCATC
HYZ017	*3 stop codons*	Top	AATTCTGAGTAGGTGAT
HYZ018	*3 stop codons*	Bottom	CTAGATCACCTACTCAG
HYZ122	*delta79avrRpt2 frame 1*	Top	GATCCCGGGAATTCAAGCACGAGACGGGCGG
HYZ123	*delta79avrRpt2 frame 1*	Top	CCGGGAATTCAAGCACGAGACGGGCGG
HYZ124	*delta79avrRpt2 frame 2*	Top	GATCCCGGGAATTCCAAGCACGAGACGGGCGG
HYZ125	*delta79avrRpt2 frame 2*	Top	CCGGGAATTCCCAAGCACGAGACGGGCGG
HYZ126	*delta79avrRpt2 frame 3*	Top	GATCCCGGGAATTCCCAAGCACGAGACGGGCGG
HYZ127	*delta79avrRpt2 frame 3*	Top	CCGGGAATTCCCAAGCACGAGACGGGCGG
HYZ128	*delta79avrRpt2*	Bottom	AATTTTAGCGGTAGAGCATTGCGTG
HYZ129	*delta79avrRpt2*	Bottom	TTAGCGGTAGAGCATTGCGTG
HYZ152	*GFP*	Bottom	GCTCGAACTATCTGCGTTGGC
HYZ163	*T4 RNAP terminator*	Top	GCATAAAGCTTGCTCAATCAATCACC
HYZ166	*GFP*	Bottom	CGTAAGGAACGTGCGTGAGACGGAAC
B1universal	*B1 site*	Top	GGGGACAAGTTTGTACAAAAAAGCAGGCT
B2 universal	*B2 site*	Bottom	AGATTGGGGACCACTTTGTACAAGAAAGCTGGGT

Supplemental Type III effector score sheet

AvrRps4	= HopPtoK	=HopK1	
HopF3	=HopF2	=AvrPphF	
HopD1	=AvrPphD	=HopAO1	(see note below.)
HopQ1	=ORF19	=HolPtoQ	
HopR1	=HolPtoR		
HopM1	=ORF3 of CEL	= HopPtoM	
AvrE1	=AvrE		
HopAF1	=AvrXv3		
HopAB1	=HopAB2	=VirPphA	
HopG1	=HopPtoG		
HopI1	=HopPmaI		
HopX1	=AvrPphE		

These are the two copies HopD1 and HopD2, hence 13 total effectors shared.