

ABSTRACT

Title of Thesis: NOVEL EXCHANGEABLE EFFECTOR LOCI
ASSOCIATED WITH THE *PSEUDOMONAS SYRINGAE*
HRP PATHOGENICITY ISLAND: EVIDENCE FOR
INTEGRON-LIKE ASSEMBLY FROM TRANSPOSED
GENE CASSETTES

Degree candidate: James Carl Charity

Degree and year: Master of Science, 2003

Thesis directed by: Dr. Steven W. Hutcheson
Department of Cell Biology and Molecular Genetics

Pseudomonas syringae strains use a type III secretion system (TTSS) to translocate effector proteins that assist in the parasitism of host plant cells. Some genes encoding effector proteins are clustered in the exchangeable effector locus (EEL) associated with the *hrp* pathogenicity island. A PCR-based screen was developed to amplify the EEL from *P. syringae* strains. Of the 86 strains screened, the EEL was successfully amplified from 29 predominately North American *P. syringae* pv. *syringae* strains using *hrpK* and *queA*-derived primers. Among the amplified EEL, ten distinct types of EEL were identified that could be classified into six families distinguishable by genetic composition. Six alleles of known effectors were found that sufficiently differed from their orthologs to expect distinct activity. One of these apparently novel effector proteins, HopPsyB, was 35% identical to its only known ortholog, the effector HopPsyA. HopPsyB was able to induce the hypersensitive

response in tobacco, which is consistent with effector activity. Additionally, *P. syringae* pv. tomato DC3000 expressing a HopPsyB':AvrRpt2 fusion elicited the HR in RPS2⁺ *Arabidopsis thaliana*, which suggested HopPsyB was secreted in a type III-dependent manner. Moreover, *P. syringae* pv. *syringae* DC3000 carrying HopPsyB exhibited slightly enhanced virulence in several *Brassica* spp. These results are consistent with the hypotheses that the EEL is a source of disparate effectors involved with the pathogenicity of *P. syringae* strains and that the EEL evolved independently of the central conserved region of the *hrp* pathogenicity island, possibly by integron-like assembly of transposed gene cassettes.

NOVEL EXCHANGEABLE EFFECTOR LOCI ASSOCIATED WITH THE
PSEUDOMONAS SYRINGAE HRP PATHOGENICITY ISLAND: EVIDENCE FOR
INTEGRON-LIKE ASSEMBLY FROM TRANSPOSED GENE CASSETTES

by

James Carl Charity

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2003

Advisory Committee:

Dr. Steven W. Hutcheson, Chair
Dr. Charles F. Delwiche
Dr. Richard C. Stewart

©Copyright by

James C. Charity

2003

ACKNOWLEDGEMENTS

Thanks to everyone who made this possible—Dr. Steve Hutcheson for giving a fledgling undergraduate the opportunity to work in his lab; Jamie Bretz, Nate Ekborg, Mike Howard, and Lili Losada for making the day-to-day grind much less tedious; my parents for being there when it counted; and my sister, Christina, who insisted I mention her by name (as if I would have forgotten you!).

TABLE OF CONTENTS

Acknowledgements.....	ii
List of Tables	iv
List of Figures.....	v
List of Abbreviations	viii
Introduction.....	1
Specific Aims.....	9
Materials and Methods.....	10
Results	16
Discussion.....	82
References.....	88

LIST OF TABLES

1.	Strains.....	13
2.	Plasmids	14
3.	Primers	15
4.	Properties of the strains whose EELs were amplified.....	19
5.	GC content of the genes in the EELs of the representative strains.....	73

LIST OF FIGURES

1.	The exchangeable effector locus (EEL)	8
2.	Amplification of the EELs from <i>Psy</i> B728a, <i>Pto</i> DC3000, and <i>Psy</i> 61	18
3.	Sequence alignment of ShcA orthologs	24
4.	Sequence alignment of HopPsyA orthologs.....	25
5.	The EELs of Family I.....	26
6.	Sequence alignment of the intergenic region between <i>queA</i> and <i>hopPsyB1</i> in <i>Psy</i> B5 and the intergenic region between <i>eelA</i> and <i>hopPsyB2</i> in <i>Psy</i> 5D4198	27
7.	Sequence alignment of AvrPphE, HopPphE1, and the deduced <i>Pph</i> BK378 AvrPphE8 gene product.....	30
8.	Sequence alignment of EelF1 and the deduced <i>Pph</i> BK378 EelF4 gene product.....	31
9.	Sequence alignment of the intergenic region between <i>hopPsyE1</i> and <i>hopPsyC1</i> in <i>Psy</i> B728a and the intergenic region between <i>avrPphE8</i> and <i>hrpK</i> in <i>Pph</i> BK378	32
10.	The EELs of Families I and II.....	33
11.	Amplification of diagnostic fragments in Family III EELs	36
12.	Properties of <i>Psy</i> B452 ' <i>hopPsyC</i> '	37
13.	<i>Psy</i> B452 contained alleles of <i>shcV</i> and <i>hopPsyV</i>	38
14.	Sequence alignment of a site-specific recombinase in <i>Pto</i> DC3000 with short, and likely not expressed, ORFs in <i>Psy</i> B728a and <i>Psy</i> B452.....	39
15.	Sequence alignment of the intergenic region between <i>hopPsyC1</i> and <i>hrpK</i> in <i>Psy</i> B728a and the intergenic region between ' <i>hopPsyC</i> ' and <i>hrpK</i> in <i>Psy</i> B452	40
16.	Sequence alignment of the intergenic region between <i>hopPsyV1</i> and <i>hopPsyC1</i> in <i>Psy</i> B728a and the intergenic region between <i>hopPsyV2</i> and ' <i>hopPsyC</i> ' in <i>Psy</i> B452	41

17.	The EELs of Families I, II, and III.....	42
18.	Sequence alignment of HopPsyC1 and the deduced HopPsyC2 gene product.....	44
19.	Sequence alignment of HopPsyE1, AvrPphE8, and the deduced HopPsyE2 gene product.....	45
20.	Sequence alignments with EelF3 and EelG3	46
21.	The EELs of Families I, II, III, and IV.....	47
22.	Amplification of diagnostic fragments in Family V EELs.....	50
23.	Sequence alignments with HopPtoB5 and EelG2	51
24.	Transposable elements in the <i>Ppe</i> 5846 EEL.....	52
25.	Sequence alignment of the intergenic region between <i>hopPtoB1</i> and <i>hrpK</i> in <i>Pto</i> DC3000 and the intergenic region between <i>hopPtoB5</i> and <i>hrpK</i> in <i>Ppe</i> 5846	53
26.	Sequence alignment of the intergenic region between the tRNA _{leu} gene and <i>eelG1</i> in <i>Pto</i> DC3000 and the intergenic region between the tRNA _{leu} gene and <i>eelG2</i> in <i>Ppe</i> 5846	54
27.	The EELs of Families I, II, III, IV, and V	55
28.	Sequence alignment of HopPtoB1, HopPtoB5, and HopPsyG	57
29.	Sequence alignments with EelF and EelG orthologs from the <i>Psy</i> DH015 EEL.....	58
30.	Sequence alignment of transposase ISPsy5 in <i>Pto</i> DC3000 with a short, likely not expressed, ORF in <i>Psy</i> DH015	59
31.	Sequence alignment of the intergenic region downstream of <i>hrpK</i> in <i>Pto</i> DC3000, <i>Ppe</i> 5846, and <i>Psy</i> DH015.....	60
32.	Sequence alignment of the intergenic region between <i>tnpA</i> ' and <i>hopPtoB1</i> in <i>Pto</i> DC30000 and the intergenic region between <i>eelF5</i> and <i>hopPsyG</i> in <i>Psy</i> DH015	61
33.	The EELs of Families I, II, III, IV, V, and VI.....	62

34.	Sequence alignment of the intergenic region between <i>hopPsyC1</i> and <i>hrpK</i> in <i>Psy</i> B728a and the intergenic region between <i>hopPsyC2</i> and <i>hrpK</i> in <i>Psy</i> W4N15.....	66
35.	Sequence alignment of the intergenic region between <i>hopPsyE1</i> and <i>hopPsyC1</i> in <i>Psy</i> B728a and the intergenic region between <i>hopPsyE2</i> and <i>hopPsyC2</i> in <i>Psy</i> W4N15.....	67
36.	Sequence alignment of the intergenic region downstream of <i>hopPsyE</i> from <i>Psy</i> B728a and <i>Psy</i> W4N15	68
37.	Sequence alignment of the intergenic region downstream of <i>eelF</i> from <i>Pph</i> BK378, <i>Psy</i> W4N15, <i>Psy</i> DH015, and <i>Pto</i> DC3000.....	69
38.	Sequence alignment of the intergenic region upstream of <i>eelF</i> from <i>Pph</i> BK378, <i>Psy</i> W4N15, <i>Psy</i> DH015, and <i>Pto</i> DC3000.....	70
39.	Sequence alignment of the intergenic region downstream of the tRNA _{leu} gene from <i>Pph</i> BK378, <i>Psy</i> W4N15, <i>Psy</i> DH015, and <i>Pto</i> DC3000	71
40.	Sequence alignment of the intergenic region downstream of the tRNA _{leu} gene from the representative stains of Families IA, IB, II, IV, VA, VB, and VI.....	72
41.	Two putative gene cassettes identified in the EELs of Families IIIA, IV, and VA.....	74
42.	Activity of HopPsyB and HopPsyA in tobacco	78
43.	Evidence for type III secretion-dependent translocation of HopPsyB.....	79
44.	Multiplication of <i>Pto</i> DC3000 expressing and not expressing the <i>Psy</i> B5 <i>shcB-hopPsyB</i> operon in <i>A. thaliana</i> leaves	80
45.	Ectopic expression of HopPsyB1 increased virulence of <i>Pto</i> DC3000 in some <i>Brassica oleracea</i> varieties	81

LIST OF ABBREVIATIONS

<i>avr</i> gene	avirulence gene
bp	base pair
CCR	central conserved region
CEL	conserved effector locus
cfu	colony forming unit
EEL	exchangeable effector locus
HR	hypersensitive response
GC	guanidine/cytosine
I	identity
kb	kilobase
kDa	kilodalton
ORF	open reading frame
PCR	polymerase chain reaction
pv.	pathovar
<i>R</i> gene	resistance gene
S	similarity
TTSS	type III secretion system

INTRODUCTION

Phytopathogenic bacteria cause many diseases, including aberrant growth, wilts, scorches, blights, soft rots, and leaf spots, and affect almost every cultivated plant. Although it is difficult to estimate the cost of bacterial pathogenesis, microbial diseases cause the United States farming industry to lose approximately \$9.1 billion per year (Agrios, 1997). Worldwide, annual crop loss due to bacterial soft rot, a single type of bacterial disease, amounts to about \$100 million (Perombelon and Salmond, 1995). Characterizing phytopathogenic bacteria inhabiting the phyllosphere will allow researchers to develop methods for combating plant disease and maximizing crop yield.

Some bacteria, such as *Pseudomonas syringae*, *Xanthomonas campestris*, *Ralstonia solanacearum*, and *Erwinia amylovora*, are facultatively parasitic. These pathogens often arrive on the surface of leaves by rainsplash and grow epiphytically, enduring the stresses of temperature, ultraviolet light, and competition with other members of the phyllosphere, until a wound on the leaf provides the opportunity for entry into the apoplast (Reviewed in Hirano and Upper, 2000). *P. syringae* and *X. campestris* are members of a select group of phytopathogens that encode ice nucleation proteins to assist the formation of frost-induced wounds (Reviewed in Hirano et al., 1995; Lindow, 1983). Once inside the apoplast, these bacteria secrete proteins that promote their reproduction and colonization (Reviewed in Lindgren, 1997).

Plant cells, however, do not rely on the benevolence of potential parasites; they have co-evolved with their pathogens to have an active biochemical defense system

(Keen, 1992; Keen, 1981). In general, there are two types of interactions between plants and their bacterial pathogens: compatible and incompatible. A compatible interaction is associated with bacterial colonization, establishment of disease, and development of late necrosis. An incompatible interaction is associated with reduced or absent colonization. A plant's lack of susceptibility to a particular pathogen is often due to a recognition event that induces localized cell death, known as the hypersensitive response (HR), in host cells near the site of inoculation (Klement, 1982). During the HR, a plant cell dissolves its intracellular membranes (Lamb et al., 1989), strengthens its cell wall (Lamb et al., 1989), and increases production of active oxygen species (H_2O_2 , O_2^- , and OH^\cdot) (Baker et al., 1991; Slusarenko et al., 1991) and antimicrobial substances such as phytoalexins (Dixon and Lamb, 1990). Each of these responses is an effort to create an extracellular environment that is inhospitable for phytopathogens and hinders the progression of disease.

Some bacterial pathogens use a type III secretion system (TTSS) to attempt to parasitize plant cells. Notably, TTSSs are not limited to plant pathogens; animal-pathogenic *Salmonella typhimurium* (Groisman and Ochman, 1993), *Shigella flexneri* (Venkatesan et al., 1992), enteropathogenic *Escherichia coli* (Jarvis et al., 1995), and *Yersinia* species (Cornelis, 1994) also contain TTSSs. A TTSS employs a flagellum-like structure to transport virulence proteins, called effectors, from the cytosol of a bacterium directly into host cells (Reviewed in He, 1998). Once inside a plant cell, effectors promote the leakage of nutrients into the apoplast (Dangl and Jones, 2001). For most effectors, the mechanism of action has not been determined (Jin et al., 1997).

Initiation of the HR is dependent on a set of plant genes known as resistance (*R*) genes. In resistant plants, at least one of the effectors interacts with the products of *R* genes to elicit the HR (Dangl and Jones, 2001). Bacterial genes encoding effectors that induce the HR are sometimes referred to as avirulence (*avr*) genes. There is a cognate, or “gene-for-gene,” relationship between *avr* and *R* genes (Staskawicz et al., 1984). That is, a plant must express a specific *R* gene to recognize a specific translocated Avr protein and initiate the HR. For example, *P. syringae* pv. *glycinea* race 4 strains normally have a compatible relationship with specific soybean cultivars, and *P. syringae* pv. *glycinea* race 6 strains normally have an incompatible relationship with these soybean cultivars. However, if a race 4 strain expresses the *avr* gene *avrA* isolated from race 6, it becomes incompatible with the soybean cultivars (Staskawicz et al., 1984). In this case, the soybean cultivars contain *R* genes whose products are capable of recognizing AvrA translocated by the TTSS.

The genes encoding structural and regulatory components of TTSSs usually cluster in a pathogen’s genome (Lindgren, 1997). The cluster of genes encoding the TTSS for *P. syringae* (Alfano et al., 2000; Hutcheson 1999), *X. campestris* (Bonas et al., 1991; Daniels et al., 1988), *E. amylovora* (Steinberger and Beer, 1988), and *R. solanacearum* (Boucher et al., 1986) has been termed the *hrp/hrc* cluster (*hrp* for *HR* and *pathogenicity*, and *hrc* for *HR* and *conserved*). The *P. syringae* *hrp* pathogenicity island is one of the most studied *hrp/hrc* clusters.

P. syringae is a rod shaped, gram-negative, usually fluorescent strict aerobe that is currently classified based on rRNA sequence as a member of the gamma subgroup of the Proteobacteria. It is oxidase and arginine dihydrolase negative, which

distinguishes it from most other fluorescent pseudomonads (Doudoroff and Pallaroni, 1974; Hirano and Upper, 2000). *P. syringae* was isolated from diseased lilac (*Syringa vulgaris*) in 1899, and was characterized and named after its host in 1902 by C. J. J. van Hall (Young, 1991; Hirano and Upper, 2000). Subsequently isolated strains of *P. syringae*, as well as the type strain, were found to be facultative pathogens of a variety of plant species. They were eventually assigned an additional infrasubspecific designation, pathovar (pv.), according to their host ranges (Dye et al., 1980; Pallaroni, 1984), which are often narrow relative to other phytopathogenic bacteria (Bradbury, 1986). There are currently more than 40 recognized pathovars (Young et al, 1992; Hirano and Upper, 2000).

The *P. syringae hrp/hrc* cluster is a 25-kilobase chromosomal tripartite pathogenicity island composed of an exchangeable effector locus (EEL), a conserved effector locus (CEL), and a central conserved region (CCR) that expresses the TTSS (Alfano et al., 2000). There is a well-studied regulatory system that controls expression of the *hrp* TTSS. One of the key features is an alternative sigma factor called HrpL, which activates expression of effectors in addition to other structural and regulatory components of the TTSS (Xiao et al., 1994). Notably, a cosmid clone, pHIR11, with a 31 kilobase (kb) genomic insert containing the *hrp* cluster from *P. syringae* pv. *syringae* 61 is able to confer type III secretion-dependent translocation of effectors when transformed into normally saprophytic *E. coli* (Hutcheson et al, 1989). *P. syringae* genes encoding known effectors are preceded by a HrpL-dependent promoter (Innis et al., 1993). Effectors also appear to contain a rather cryptic amphipathic type III secretion signal in their first 50 amino acid residues (Petnicki-

Ocwieja et al, 2002). In susceptible plants, the translocated effectors, known as Avr, Vir, or Hop proteins, suppress or modify the cellular defense responses of the host (Chen et al., 2000) or alter the physiology of the host cell to favor growth of the pathogen (Dangl and Jones, 2001; Hutcheson, 2001). Translocated effectors are essential for the virulence of *P. syringae* strains, but also control the host range by initiating host defenses, including the HR, in resistant plants. Strains with similar host ranges likely encode comparable sets of effectors that are necessary for virulence in their hosts; strains with distinct host ranges likely differ in the effectors they translocate into plant cells. Consequently, to understand the pathogenicity and host range of a given *P. syringae* strain, it is important to identify the effectors that are expressed and translocated by that strain.

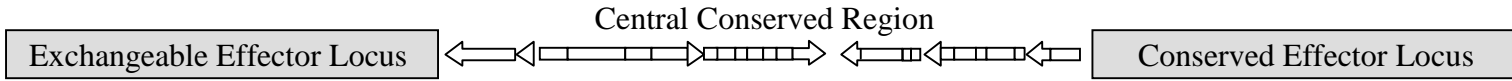
Several methods have been used to survey *P. syringae* strains for translocated effectors. Most effector genes have been identified by screening a genomic library of one strain for genes that affect the host range of another strain (Leach and White 1996; Staskawicz et al., 1984; Vivian and Gibbon, 1997). Bioinformatic analysis of the *P. syringae* pathovar (pv.) tomato DC3000 genome identified 51 putative or verified effectors that appear to be expressed from HrpL-dependent promoters (Fouts et al., 20002) or carry type III secretion signals (Guttman and Greenberg, 2001; Guttman et al., 2002; Petnicki-Ocwieja et al., 2002). Some translocated proteins have been identified by using randomly generated genomic fusions to an '*avrRpt2* cassette that requires TTSS-dependent translocation for activity in *RPS2* lines of *Arabidopsis thaliana* (*RPS2* is the cognate *R* gene for *avrRpt2*) (Guttman et al. 2002). In some cases, culture filtrates have been used to identify secreted proteins (Mudgett and

Staskawicz, 1999; van Dijk et al. 1999; Yuan and He, 1996). Unfortunately, most *P. syringae* strains secrete effectors into liquid media below the threshold of detection (Alfano and Collmer, 1997; Hutcheson, 1999; Li et al., 1992). The results of these screens indicate that effector genes can be dispersed throughout the genome (Kim et al., 1998), clustered in plasmid-borne pathogenicity islands (Jackson et al., 1999), or associated with the *hrp/hrc* cluster in either the EEL or the CEL (Alfano et al., 2000; Hutcheson, 1999).

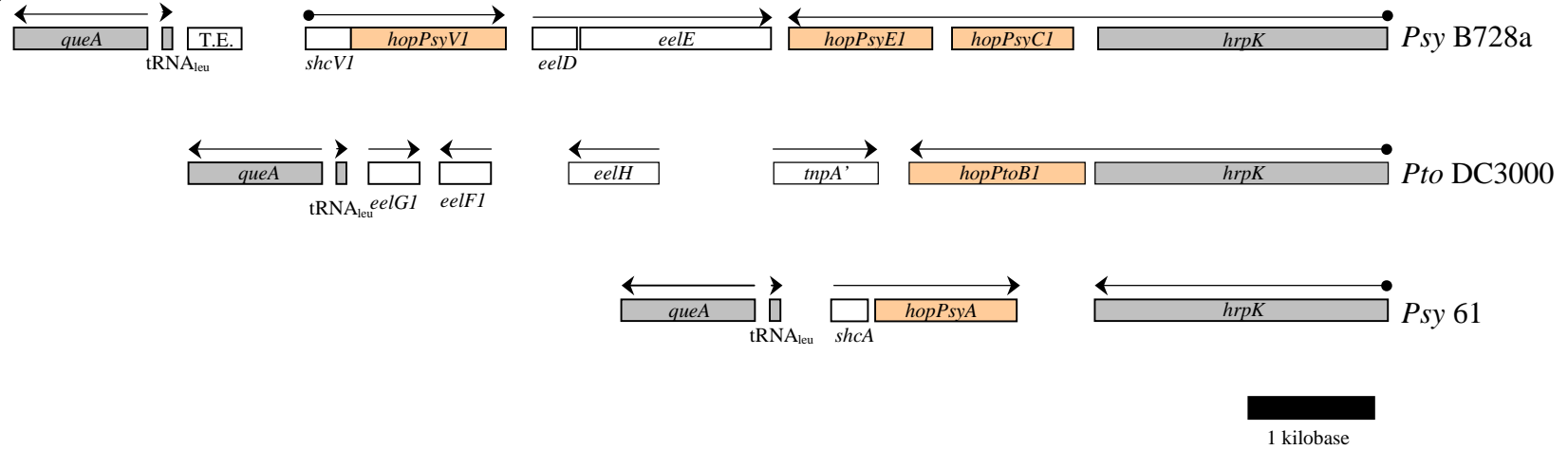
Although effectors located in the EEL are unlikely to be solely responsible for a strain's host range, the EEL is remarkable because it appears to be a highly variable source for effectors (Alfano et al., 2000). Each strain characterized thus far contains a distinct EEL with at least one demonstrated effector (Figure 1) (Alfano et al. 2000; Heu and Hutcheson, 1993; Mansfield et al., 1994; Petnicki-Ocweija et al., 2002). In *P. syringae* pv. *syringae* (*Psy*) 61, a weak bean pathogen, the EEL consists of an operon containing *hopPsyA* (also known as *hrmA*), which encodes a TTSS-dependent effector (Alfano et al., 1997; Heu and Hutcheson et al., 1993), and *shcA*, which encodes a chaperone for HopPsyA (van Dijk et al., 2002). The EEL of *Psy* B728a, a more virulent bean pathogen, includes six genes, three of which—*hopPsyCI*, *hopPsyE1*, and *hopPsyVI*—encode translocated effectors (Alfano et al., 2000). In *P. syringae* pv. *tomato* (*Pto*) DC3000, a pathogen of *Arabidopsis* and tomato, the EEL encodes the effector HopPtoB1 and contains four other genes that are entirely dissimilar from the genes in the *Psy* 61 and *Psy* B728a EELs (Alfano et al., 2000; Fouts et al., 2002; Petnicki-Ocweija et al., 2002).

The EELs in *Psy* 61, *Pto* DC3000, and *Psy* B728a are bordered on one side by *queA* and a tRNA_{leu} gene, and on the other side by *hrpK* (Alfano et al., 2000). *queA* is a housekeeping gene involved in the modification of tRNA nucleoside 34 (Reuter et al., 1991) and is not functionally linked to the TTSS; *hrpK*, which is located in the CCR of the *hrp/hrc* cluster (Alfano et al. 2000), has an unknown function. The presence of these conserved genes adjacent to the EEL introduced the possibility of using polymerase chain reaction (PCR) to amplify the EEL from other *P. syringae* strains. Such a study may facilitate the identification of novel effector genes, alleles of known effector genes, and an association between a strain's host range and the genetic composition of its EEL. This report describes the properties of the EELs and effector genes amplified from 29 *P. syringae* strains isolated mostly from North American sources.

A.



B.



∞

Figure 1. The exchangeable effector locus (EEL). (A) Effectors associated with the *hrp* pathogenicity island flank the cluster of operons that make up the central conserved region and are located in either the EEL or the conserved effector locus (Alfano et al., 2000). (B) The EEL has been characterized for three strains: *Psy* B728a, *Pto* DC3000, and *Psy* 61 (Alfano et al., 2000). Arrows indicate the direction of transcription. A circle at the tail of an arrow indicates the transcriptional unit is controlled by a HrpL-dependent promoter. Tan shading indicates the gene encodes an effector. Elements that are not part of the EELs are gray. An open reading frame that is likely not expressed but has similarity to transposable elements is labeled "T.E."

SPECIFIC AIMS

The exchangeable effector locus (EEL) borders one side of the central conserved region of the *hrp/hrc* cluster in *P. syringae*, and is a source for effectors that are translocated by the type III secretion system. Variation within the EELs of *Psy* 61, *Pto* DC3000, and *Psy* B728a indicates other *P. syringae* strains might contain unique effectors in the EEL that may affect their host ranges. A conserved gene (*queA* or *hrpK*) borders each side of the EEL. The purpose of this project was as follows:

1. Amplify the EEL from *P. syringae* strains.

This may be accomplished by designing and using *hrpK*- and *queA*-specific primers to PCR screen *P. syringae* strains. At least one EEL must be amplified before addressing the other specific aims.

2. Determine the sequences of the EELs.

After an EEL has been amplified, it may be partially sequenced. If a particular EEL appears to be divergent from the previously characterized EELs, it may be completely sequenced.

3. Identify potentially novel effectors.

Sequence data generated from specific aim (2) may be used to identify novel effectors that could function in *P. syringae* pathogenicity.

MATERIALS AND METHODS

Bacterial strains

Pseudomonas syringae and *Escherichia coli* strains were grown in King's B (KB) medium at 25°C and 37°C, respectively (King et al., 1954). Rifampin (200 µg/ml) and spectinomycin (100 µg/ml) were added to the media when appropriate.

Previously described strains, plasmids, and primers that were used are listed in Tables 1, 2, and 3, respectively.

PCR amplification of the EEL

Colony PCR reactions employed a Hybaid PCR Sprint™ thermocycler using a 10 minute extension time and either ProofSprinter™ (Hybaid) or ProofPro™ (Continental Lab Products) DNA polymerase enzyme mixtures. The primers K2688 and Q920 were chosen based on their ability to amplify the previously characterized EELs of *Psy* 61, *Psy* B728a, and *Pto* DC3000. All other variables were set according to standard protocol or manufacturer's instructions for long-range PCR. PCR products were gel purified and extracted using the Bio-Rad (Hercules, CA) Prep-A-Gene™ Master Kit prior to sequencing. DNA was sequenced at the University of Maryland Biotechnology Institute using an ABI Model 3100 Automated Sequencer.

Diagnostic PCR to determine conservation of *Pto* DC3000- and B728-like EELs

Published sequences of *Pto* DC3000 and *Psy* B728a were used to create primers internal to selected ORFs. Primers were appropriately paired and utilized in *Taq* polymerase (Invitrogen)-based PCR amplification with a 5-minute extension time. For

the *Psy* B728a-like EEL, the following primer sets were used: B1F and BER amplified *hopPsyC1* to *hopPsyE1*; BEF and B5R amplified *hopPsyE1* to *hopPsyV1*; B5F and Q920 amplified *hopPsyV1* to *queA*. For the *Pto* DC3000-like EEL, the following primer sets were used: D1F and D2R amplified *hopPtoB1* to *eelH*; D2F and D3R amplified *eelH* to *eelF1*; and D3F and Q920 amplified *eelF1* to *queA*.

Construction of '*avrRpt2* fusions to *Psy* B5 *hopPsyB1*

A 259 base pair (bp) fragment carrying the 5'-portion of *hopPsyB1* (includes the native ribosome binding site and the first 86 codons), was PCR amplified from *Psy* B5 using the primers B5EEL1950-X and B5EEL1565-RI. A 528 bp fragment carrying the 3'-portion of *avrRpt2* (including the last 176 codons) was amplified using the primers *avrRpt2*-412-R1 and *avrRpt2*-1028-H. Both fragments were gel-purified using the Bio-Rad Prep-A-Gene™ Master Kit, digested with *EcoRI*, and ligated. The desired fusion was amplified from the ligation mixture using primers B5EEL1950-X and *avrRpt2*-1028-H, gel-purified, and ligated into pDSK519 as a *XbaI*, *HindIII* fragment to create pJCB5EEL2-AR2.

Plant assays

For pathogenicity trials, overnight cultures grown at 25° C were diluted to an OD₆₀₀ of 0.1, diluted 10³ fold into water and infiltrated into leaves of the indicated plants. Plant responses were scored daily and bacterial populations were monitored using the leaf disk assay of Bertoni and Mills (Bertoni and Mills, 1987). For assessment of the ability to elicit the HR, overnight cultures were harvested, washed in water, and

resuspended to 10^8 colony forming units (cfu)/mL for *Pto* DC3000 derivatives and 10^9 cfu/mL for *E. coli* MC4100 derivatives. Leaf tissue was inoculated using a syringe and scored for responses at 18 or 24 hours, as indicated.

Table 1. Strains.

Strain	Properties	Reference
<i>P. syringae</i> strains <i>Pto</i> DC3000	tomato / <i>Arabidopsis</i> pathogen; Rif ^r derivative of NCBPP1106 isolated in the United Kingdom; carries a Family VA EEL	(Cuppels, 1986)
<i>Psy</i> 61	weak bean pathogen; carries a Family IA EEL	(Baker et al., 1987)
<i>Psy</i> B728a	Bean pathogen; carries a Family IIIA EEL	(Hirano et al., 1999)
<i>Pph</i> 1032A	Bean pathogen; carries a Family II EEL	(Mansfield et al., 1994)
<i>E. coli</i> strains DH5 α	(r _k ⁻ m _k ⁻) <i>recA1 relA1</i> Δ (<i>argF-lacZYA</i>)U169 ϕ 80 <i>dlacZDM15</i>	Invitrogen
MC4100	F' Δ (<i>argF-lacZYA</i>)U169	(Casadaban, 1976)

Table 2. Plasmids.

Plasmid	Properties	Reference
pDSK519	incQ, Kn ^r	(Keen et al., 1988)
pDSK600	incQ, Sp ^r , triple <i>lacUV5</i> promoter, mcs	(Murillo et al., 1994)
pHIR11	<i>hrp</i> PAI from <i>Psy</i> 61 cloned into pLAFR3	(Huang et al., 1988)
pHIR11-2070	pHIR11 derivative carrying a <i>hopPsyA::TnphoA</i> insertion	(Huang et al., 1991)
pJC' <i>avrRpt2</i> -600	617 bp <i>EcoR1-HindIII</i> fragment carrying C-terminal effector domain of <i>avrRpt2</i> ligated into pDSK600	This report
pJCB5EEL2-AR2	259 bp <i>XbaI-HindIII</i> fragment carrying <i>hopPsyB</i> RBS and N-terminal 86 codons ligated into pDSK519	This report
pLAFR3	incP-1 Tc ^r cosmid vector	(Staskawicz et al., 1987)
pSHB5EEL1-600	The 3.5 kb EEL from <i>P. syringae</i> B5 cloned as a <i>XbaI-EcoR1</i> fragment into pDSK600	(Charity et al., 2003)
pYX1L	the <i>shcA-hopPsyA</i> operon cloned as a <i>BamH1-EcoR1</i> fragment in pLAFR3	(Pirhonen et al., 1996)
pYXL2B	expresses the <i>hrpL</i> locus from vector promoter	(Xiao et al., 1994)

Table 3. Primers

Primer	Properties	Reference
K2688	CTGGGCGGACAGATGATC	This report
Q920	AACGCCGAAACCAGCATCAA	This report
D1F	GCAGAGTCAGGGTCATCAG	This report
D2R	TTCCAGCACAGCATCCAGTT	This report
D2F	GGTGCGGTAAGTGTGAGAAA	This report
D3R	CGGTGTGGTGGTCATTGTC	This report
D3F	TCATTGGCACTTCGCTACCT	This report
B1F	CGATGTGGGTGAGCCTAATG	This report
BER	GGAGCGGGATTCAGGATGT	This report
BEF	GTGGTCGCTATCGGCAAGAA	This report
B5F	TACTTTCAGCATTAGGCAACG	This report
B5R	GTCTGCCCTCGCCACGCTGTA	This report
B5EEL1950-X	GCTCTAGAGCTGTATATCGGCTTTA	This report (Charity et al., 2003)
B5EEL1565-RI	CGGAATTCCGCGCGCTTGTCATTTCAGTA	(Charity et al., 2003)
<i>avrRpt2</i> -412-RI	CGGAATTCCACGAGACGGGCGGTTCAAG	This report
<i>avrRpt2</i> -1028-H	CCCAAGCTTTAGGGACCAAAAAGCCAGAC	This report
P197-R1	CCGGAATTCGGACGGTCTATATAAGGAGG	This report
PB5ORF1-XbaI	GCTCTAGACAGTTCGGGATTGACAGG	This report

RESULTS

Development of a PCR screen for the *P. syringae* EEL

Conserved regions within alleles of *queA* and *hrpK* from *Psy* 61, *Pto* DC3000, and *Psy* B728a were used to design oligonucleotide primers, K2688 and Q920, for amplifying the EEL (Figure 2A,B,C). The K2688/Q920 primer pair was used with a commercial long-range PCR amplification kit to amplify the EEL from *Psy* 61, *Pto* DC3000, and *Psy* B728a. The expected fragment sizes were generated (Figure 2D) and were identified as EEL by sequencing approximately 400 nucleotides downstream of *hrpK*.

Characterization of *hrpK*-associated EELs amplified from *P. syringae* strains

Using the K2688/Q920 primer pair and a commercial long range PCR amplification kit, 86 *P. syringae* strains were screened for the EEL. The EEL was successfully amplified from 29 predominantly North American *P. syringae* pv. *syringae* strains that were originally isolated from a variety of plants (Table 4). The amplified EELs were initially differentiated by size, nucleotide sequence adjacent to *hrpK*, and in some cases, the amplification of diagnostic internal fragments. The complete nucleotide sequence was obtained for a representative of each EEL identified as novel (Table 4).

Twelve strains had an EEL equivalent to the EEL of *Psy* 61, *Pto* DC3000, or *Psy* B728a. Seventeen strains contained an EEL that was different from the previously characterized EELs. The EELs were classified into six families based on their constituent genes. A representative for each novel EEL was fully sequenced (Table 4). The absence of an accession number in Table 4 indicates that a strain did not have

its EEL completely sequenced. When this applies, the strain's EEL was classified based on its size, its nucleotide sequence immediately downstream of *hrpK* (approximately 400 nucleotides), and possibly other diagnostic methods, as discussed below. Strains classified as having the same type of EEL retained >98% identity to the representative EEL across sequenced portions of the region.

Several of the strains used were recalcitrant to EEL amplification using *hrpK*- and *queA*-derived primers. The attempt was made to amplify the EEL from a subset of these strains using Q920 and a primer complementary to a conserved region of *hrpL*, which lies directly upstream of *hrpK*. The EELs of three strains were amplified in this manner. However, each of these EELs was equivalent to one of the EELs amplified in the previous survey. The reason for the inability to amplify the EEL from some strains has not been determined, but it may be that these strains have especially long EELs, sequence divergence at the primer sites preventing primer binding or elongation, or the absence of closely linked *queA* and *hrpK* or *hrpL* alleles.

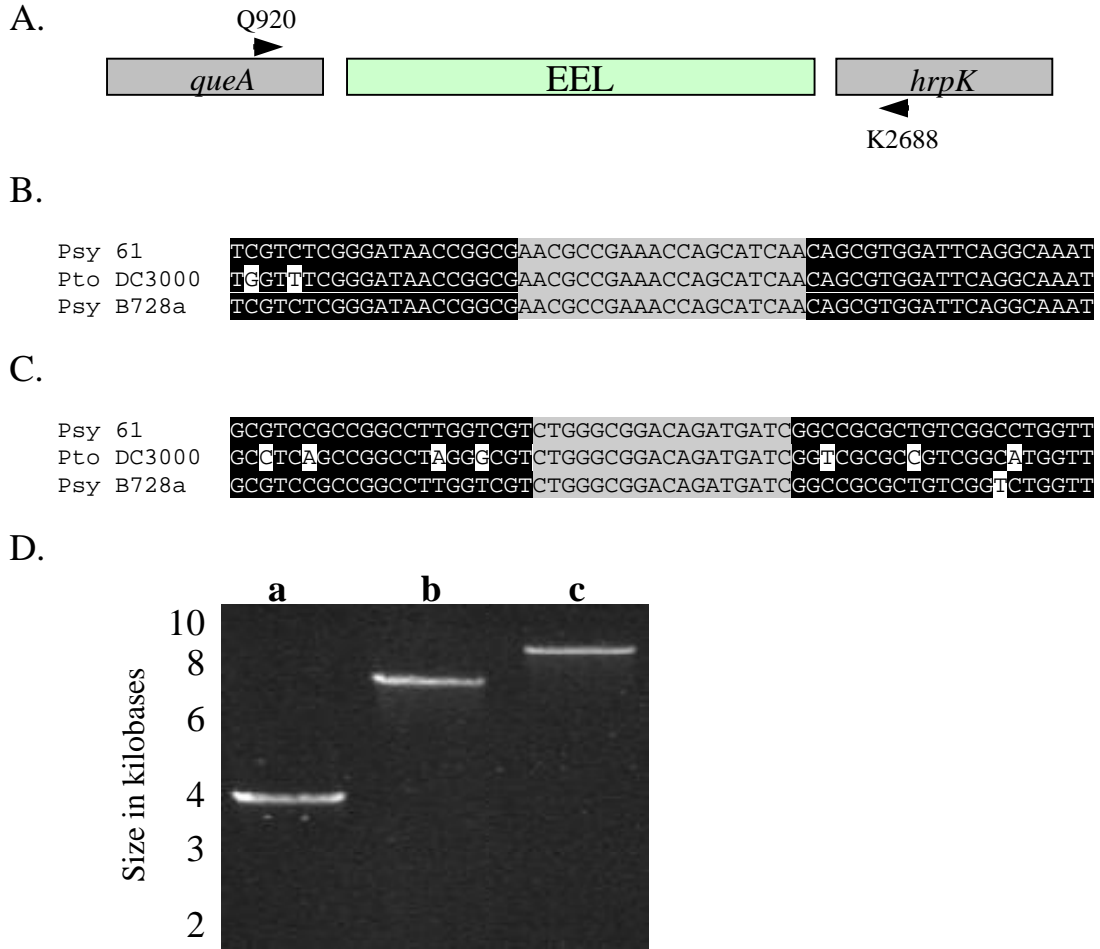


Figure 2. Amplification of the EELs from *Psy* B728a, *Pto* DC3000, and *Psy* 61. (A) Oligonucleotide primers Q920 and K2688 were designed to be complementary to internal regions of *queA* and *hrpK*, as depicted by the arrows. (B) The sequence of the primer Q920 is shown (gray) along with flanking sequence from the *queA* alleles in *Psy* 61, *Pto* DC3000, and *Psy* B728a. Non-conserved nucleotides have a white background. (C) The sequence of the primer K2688 is shown (gray) along with flanking sequence from the *hrpK* alleles in *Psy* 61, *Pto* DC3000, and *Psy* B728a. Non-conserved nucleotides have a white background. (D) Primers Q920 and K2688 and ProofPro™ polymerase (Continental Lab Products, San Diego, CA) were used to PCR amplify the EELs from whole cells of (a) *Psy* 61, (b) *Pto* DC3000, and (c) *Psy* B728. The amplified fragments were gel-purified and sequenced from *hrpK* to verify they contained EELs

Table 4. Properties of the strains whose EELs were amplified..

Strain ^{a,f}	Plant Source ^b	Geographical Location ^c	Amplified PCR Product ^d	EEL Type ^e	Accession Number
<i>Psy</i> B76	Tomato	GA	4.0	IA	
<i>Psy</i> B458	Orange	CA	4.0	IA	
<i>Psy</i> PDDCC3523	Tomato	AU	4.0	IA	
<i>Psy</i> PS-1	Peach	NC	4.0	IA	
<i>Psy</i> PS-14	Apricot	NC	4.0	IA	
<i>Psy</i> Ps6-SB	Soybean	NE	4.0	IA	
<i>Psy</i> 84-15	Tomato	GA	4.0	IA	
<i>Psy</i> 85-274	Tomato	GA	4.0	IA	
<i>Psy</i> B5	Soybean	MC	3.5	IB	AY147022
<i>Psy</i> B6	Soybean	MC	3.5	IB	AY147023
<i>Psy</i> 464	Corn	SD	3.5	IB	AY147017
<i>Psy</i> B301D	Pear	UK	3.5	IB	AY147020
<i>Psy</i> BK034	Wheat	MN	3.5	IB	AY147024
<i>Psy</i> PDDCC3907	Pear	GR	3.5	IB	AY147027
<i>Psm</i> 567	Cherry	UK	3.5	IB	
<i>Psy</i> 5D4198	Plum	CA	8.0	IC	AY147019
<i>Pph</i> BK378	Bean	MN	4.0	II	AY147025
<i>Ppa</i> B138	Apple	GA	4.0	II	
<i>Psy</i> B362	Bean	CA	8.5	IIIA	
<i>Psy</i> Ps1Bean	Bean	WA	8.5	IIIA	
<i>Psy</i> B452	Orange	CA	5.5	IIIB	AY147021
<i>Psy</i> B460	Lemon	CA	5.5	IIIB	
<i>Psy</i> BK035	Wild Rice	MN	5.5	IIIB	

Table 4 continued.

Strain ^{a,f}	Plant Source ^b	Geographical Location ^c	Amplified PCR Product ^d	EEL Type ^e	Accession Number
<i>Psy</i> W4N15	Apple	WA	5.5	IV	AY147028
<i>Psy</i> W4N27	Pear	OR	5.5	IV	
<i>Psy</i> NCPFB1053	Millet	ET	7.5	VA	
<i>Pto</i> 2844	Tomato	UK	7.5	VA	
<i>Ppe</i> 5846	Peach	FR	6.0	VB	AY147018
<i>Psy</i> DH015	Bean	WI	5.0	VI	AY147026

Superscripts:

^a See (Denny et al., 1988) for strain references. Abbreviations: *Pmo*, *P. syringae* pv. morsprunorum; *Ppa*, *P. syringae* pv. papulans; *Ppe*, *P. syringae* pv. persicae; *Pph*, *P. syringae* pv. phaseolicola; *Psy*, *P. syringae* pv. syringae; *Pto*, *P. syringae* pv. tomato. Representative strains used for each type of EEL are shown in bold. *Psy* 61, *Psy* B728A and *Pto* DC3000 were used as the representative strains for type IA, IIIA, and VA EEL.

^b Type of plant the strain was isolated from or reported to cause symptoms in.

^c Geographical region the strain was reported to be isolated from. Abbreviations: AU, Australia; CA, California, USA; ET, Ethiopia; FR, France; GA, Georgia, USA; GR, Greece; MC, Manitoba, Canada; MN, Minnesota, USA; NC, North Carolina, USA; NE, Nebraska, USA; OR, Oregon, USA; SD, South Dakota, USA; UK, United Kingdom; WA, Washington, USA; WI, Wisconsin, USA.

^d The indicated strain was used to initiate PCR amplification of the EEL using primers Q920 and K2468, as described in Materials and Methods. The indicated size was determined by agarose gel electrophoresis and represents the amplified EEL plus approximately 1.5 kb of sequence internal to *queA* and *hrpK*.

^e Genetic organizations of EELs are shown in Figure 33. Placement into the indicated families was based on sequence similarities.

^f The following strains were negative for PCR amplification of the EEL using Q920 and K2688: *Psy* 480, *P. syringae* pv. delphinii 529, *Pto* 2424, *P. syringae* pv. berberidis 4116, *P. syringae* pv. antirrhini 4303, *P. syringae* pv. savastanoi 4352, *Pmo* 5795, *P. syringae* pv. maculicola (*Pma*) #1, *Pma* #5, *Pma* #10, *Pma* 1083-3, *Psy* 5D417, *Pto* 832F, *Pto* B117, *Pto* B118, *Pto* B120, *Pto* B121, *Pto* B122, *Pto* B125, *Pph* B130, *Psy* B15, *Pto* B19, *Psy* B359, *Psy* B368, *Psy* B382, *Psy* B3A, *Psy* B407, *Psy* B427, *Psy* B455, *Pto* B67, *Pto* B88, *Psy* BK036, *Pto* CNBP1323, *Pto* DAR26742, *Pto* 30555, *Pto* DAR31861, *Psy* HS191, *Pto* JL1053, *Pto* JL1060, *Pto* JL1075, *Pto* JL1105, *Pto* JL1120, *Pto* NCPFB880, *Pph* NK343, *Pto* PDDCC 3357, *Pto* PDDCC 4355, *Pto* PT14, *Pto* PT21, *Pto* PT30, *Pto* RG-4, *Psy* S-4B-1, *Psy* SD19, *Pto* T1, *Pto* T4B1, *Psy* W4N108, *Psy* W4N43.

Family I EEL

Sixteen strains carried an EEL related to the EEL in *Psy* 61. In eight strains—*Psy* B76, *Psy* B458, *Psy* PDDCC3523, *Psy* PS-1, *Psy* PS-14, *Psy* Ps6-SB, *Psy* 84-15, and *Psy* 85-274—the locus was equivalent to the EEL in *Psy* 61. The fragments amplified from these strains were about the same size as the fragment amplified from *Psy* 61 and initial sequence (approximately 400 nucleotides) downstream of *hrpK* had >98% identity with the *Psy* 61 EEL. This type of EEL was classified into Family IA (representative strain: *Psy* 61).

The EELs in seven additional strains—*Psy* B5, *Psy* B6, *Psy* 464, *Psy* B301D, *Psy* BK034, *Psy* PDDCC3907, and *Psy* 567—contained two genes. The first gene downstream of *hrpK* was similar to *Psy* 61 *shcA* and the second gene downstream of *hrpK* was similar to *Psy* 61 *hopPsyA*. The orientation of the genes was reversed relative to the EEL of *Psy* 61, and the two genes appeared to be expressed as part of the *hrpK* operon. The apparent translational start of the *shcA* ortholog, designated *shcB1*, overlapped the stop codon of *hrpK*, and the stop codon of *shcB1* was immediately followed by the translational start of the *hopPsyA* ortholog, designated *hopPsyB1*. The deduced *Psy* B5 ShcB1 product was 175 amino acid residues long and had a predicted mass of 19.5 kDa. Four minor sequence variants of ShcB1 were detected. ShcB1 shared 56% identity (I) and 69% similarity (S) with *Psy* 61 ShcA, but was 65 residues longer than ShcA at the N-terminus (Figure 3). *Psy* B5 *hopPsyB1* encoded a 382 amino acid residue polypeptide with a predicted mass of 41.9 kDa. The amino-terminal 44 residues shared 68% I and 82% S with the amino terminus of HopPsyA (Figure 4). This domain, which likely carries the type III secretion signal

(Mudgett and Staskawicz, 1999), was serine rich, as observed in other *P. syringae* TTSS-dependent effectors (Guttman et al., 2002). It did not, however, fit many of the criteria proposed by Petnicki-Ocweija and associates for TTSS translocated effectors. For example, the deduced HopPsyB1 product has hydrophobic residues at positions three and four, a glutamic acid residue at position 12, and multiple cysteine residues. However, other translocated effectors, including HopPsyA, also display these characteristics. Over its whole length, HopPsyB1 shared only 35% I and 45% S with HopPsyA and required several gaps to align with HopPsyA. This suggested that the effector activity of the two proteins might be different. Since the EEL in these strains appeared to be an inversion of the *Psy* 61 EEL, this type of EEL was classified into Family IB (representative strain: *Psy* B5) (Figure 5).

A third variant of the Family I EEL was identified in *Psy* 5D4198. This EEL encoded orthologs of ShcB and HopPsyB (>97% amino acid residue identity; Figures 3 and 4) and carried three additional genes, designated *eelA*, *eelB*, and *eelC*, downstream of the *hopPsyB2* allele. *eelA* encoded a 256 amino acid residue polypeptide with a predicted mass of 30 kDa; *eelB* encoded a 243 amino acid residue polypeptide with a predicted mass of 27.2 kDa; *eelC* encoded a 545 amino acid residue polypeptide with a predicted mass of 62 kDa. *eelB* and *eelC* likely formed an operon since they were separated by only 97 bp, were oriented in the same direction, and were not separated by an obvious Rho-independent transcriptional terminator. Neither orthologs nor conserved functional domains could be located in online databases for any of the three gene products. Additionally, no obvious secretion signals were found. A partial duplication of the tRNA_{leu} gene was identified between

eelA and *hopPsyB2* (Figure 6), indicating that *eelA* and *eelBC* might have been part of a 4 kb insertion into the *tRNA_{leu}* gene. Consistent with this interpretation, the region between the truncated *tRNA_{leu}* gene and *hopPsyB2* shared 95% identity with the intergenic region separating the *tRNA_{leu}* gene and *hopPsyB1* in the *Psy* B5 EEL (Figure 6). No obvious HrpL-dependent promoters were detected upstream of *eelA* or *eelBC*, suggesting neither transcriptional unit is part of the HrpL-dependent regulon. This type of EEL was classified into Family IC (representative strain: *Psy* 5D4198) (Figure 5).

```

B5      1  MQWSRAATLSSVLFRECVVLQTSMINTPPGEVIYRPRSFVQC SGHRGVLMTDRLYKTVLA
464    1  MQWSRAALSSILFRECVVLQTSMINTPPGEVIYRPRPFVQC SGHPCVLM TDRLYKTVLA
5D4198 1  MQWSRAAALSSILFRECVVLQTSMINTPPGEVIYRPRPFVQC SGHRGVLMTDRLYKTVLA
ShcA   1  -----

B5      61  DLSTSLTMQPLMFDDTGACDVVVDEEIALKVVVDP-VFQ RLLLI GLMDISPDLPLQRLLS
464    60  DLSTSLTMQPLMFDDTGACDVVVDEEIALKVVVDP-VFQ RLLLI GLMDISPDLPLQRLLS
5D4198 61  DLSTSLTMQPLMFDDTGACDVVVDEEIALKVVVDP-VFQ RLLLI GLMDISPDLPLQRLLS
ShcA   1  -----MEMPALAFDDK GACNMIIDKAFALTLR DDDTHQRLLLI GLLEPHEDLPLQRLLA

B5      120  GALNPLFNDGPGLGWHAGSELYIGFKAIPRENVSVVTLKQAIAELVEWIKRWRDAH-
464    119  GALNPLFNDGPGLGWHAGSELYIGFKAIPRENVSVVTLKQAIAELVEWIKTWRDAH-
5D4198 120  GALNPLFNDGPGLGWHAGSELYIGFKAIPRENVSVVTLKQAIAELVEWIKTWRDAH-
ShcA   56  GALNPLVNA GPGIGWDEQSGLYHAYQSIPREKVSVEMLKLEIAGLVEWMMKWRERART

```

Figure 3. Sequence alignment of ShcA orthologs. The deduced sequences of ShcB from the EELs of *Psy* B5, *Psy* 464, and *Psy* 5D4198 were aligned to ShcA from the EEL of *Psy* 61 using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background.

```

B5      1  MNPIQTRFSNVEALRHSEVDVQELKAHQIEVGGKCYDIRAAANNDLTVQRSDKQMAMSK
464     1  MNPIQTRFSNVEALRHSEVDVQELKAHQIEVGGKCYDIRAAANNDLTVQRSDKQMAMSK
5D4198 1  MNPIQTRFSNVEALRHSEVDVQELKAHQIEVGGKCYDIRAAANNDLTVQRSDKQMAMSK
HopPsyA 1  MNPIHARFSSVEALRHSNVDIQAIKSEGCQLEVNCKRYEIRAAADGSIIVLRLPDCQSKADK

B5      61  FFKKAGLSGSSGSQSDQIAQVLNDRKRGSSVSRILIRQGQTHLGRMQFNIEEGQGSSAATSV
464     61  FFKKAGLSGSSGSQSDQIAQVLNDRKRGSSVPRLMRQGQTHLGRMQFNIEEGQGSSAATSV
5D4198 61  FFKKAGLSGSSGSQSDQIAQVLNDRKRGSSVPRILIRQGQTHLGRMQFNIEEGQGSSAATSV
HopPsyA 61  FFKGAAHLIGGQSQRQAQIAQVLN-EKAAAVPRLDR----MLGRRFDLKGGSSAVCAAIK

B5      121  QNSRLPNGRLVNSSILQWAEKAKANGSTSSSALYQIYAKELPRVELLPRTTEHRACLAHYM
464     121  QNSRLPNGRLVNSSILQWAEKAKANGSTSSSALYQIYAKELPRVELLPRTTEHRACLAHYM
5D4198 121  QNSRLPNGRLVNSSILQWAEKAKANGSTSTLSALYQIYAKELPRVELLPRTTEHRACLAHYM
HopPsyA 116 AADSRLTSKQTFASFCQWAEKAEALCGRYRNRYLHDLQEGHARHNAYECGRVKNTTWKRYR

B5      181  KLNGKDGISIWPFQFDGVRGLQLKHDTKVFMMNPKAADEFYKIERSGTQFPDEAVKARL
464     181  KLNGKDGISIWPFQFDGVRGLQLKHDTKVFMMNPKAADEFYKIERSGTQFPDEAVKARL
5D4198 181  KLNGKDGISIWPFQFDGVRGLQLKHDTKVFMMNPKAADEFYKIERSGTQFPDEAVKARL
HopPsyA 176 LSITRKTISYAPQIHDRE--EELDLGRVIAEDRNARTGFFRMVPK-DQRAPEINSGRL

B5      241  TINVKPQFQKAMVDAAVRLTAERHDIITAKVAGPAKIGTITDAAVFVSGDFSAAQTLAK
464     241  TINVKPQFQKAMVDAAVRLTAERHDIITAKVAGPAKIGTITDAAVFVSGDFSAAQTLAK
5D4198 241  TINVKPQFQKAMVDAAVRLTAERHDIITAKVAGPAKIGTITDAAVFVSGDFSAAQTLVK
HopPsyA 233 TIGVEPKYGAQLALAMATLMDKHKSVTQCKVVGPAKVCQQTDSALTYINCDLAKAVKLGFE

B5      301  ELQAL--LPDDAFINHTPAGMQSMGKGLCYAERTPQDR TSHGMSRASIIESALADTSRSS
464     301  ELQAL--LPDDAFINHTPAGMQSMGKGLCYAERTPQDR TSHGMSRASIIESALADTSRSS
5D4198 301  ELQAL--LPDDAFINHTPAGMQSMGKGLCYAERTPQDR TSHGMSRASIIESALADTSRSS
HopPsyA 293 KLKKLSGIPPEGFVEHTPLSMOSTICLGLSYAESVEGQPS SHGQARTHVIMDALKGG--P

B5      359  LEKKLRNAFKSAGYNPDNPAFRLE-
464     359  LEKKLRNAFKSAGYNPDNPAFRLE-
5D4198 359  LEKKLRNAFKSAGYNPDNPAFRLE-
HopPsyA 351 MENRLKMALAERGYDPEINPALRARN

```

Figure 4. Sequence alignment of HopPsyA orthologs. The deduced sequences of HopPsyB from the EELs of *Psy* B5, *Psy* 464, and *Psy* 5D4198 were aligned with HopPsyA from the EEL of *Psy* 61 using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background.

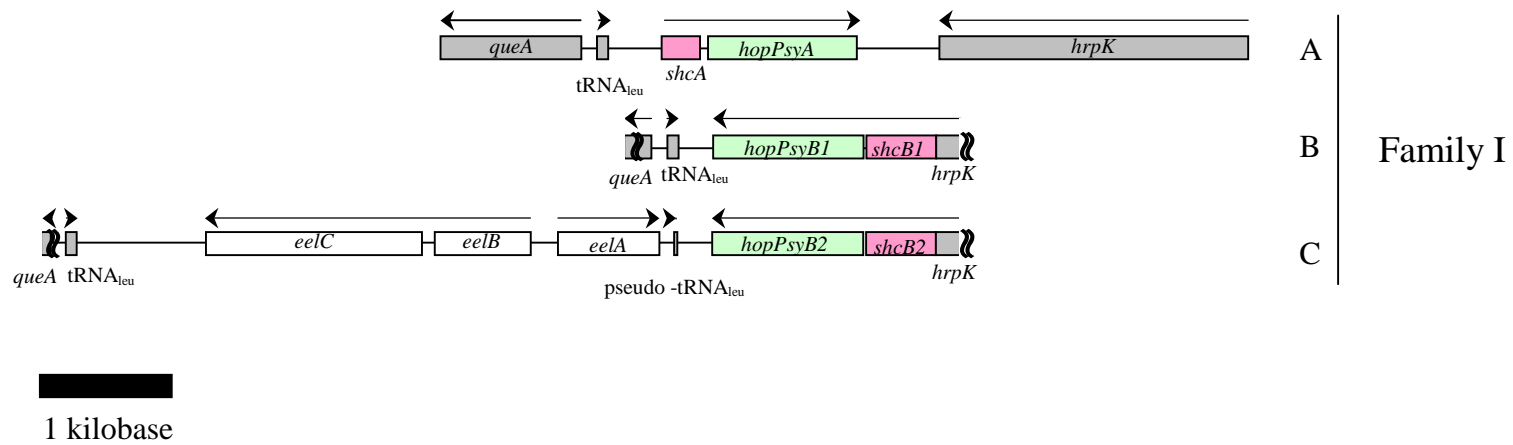
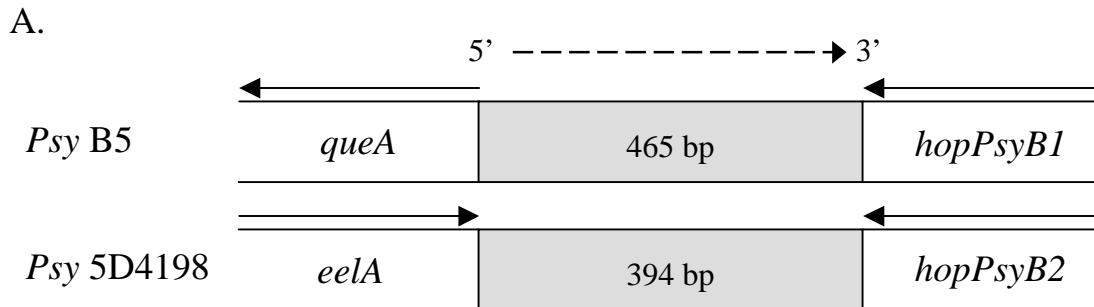


Figure 5. The EELs of Family I. The EELs are aligned by their *hrpK* sequences. *hrpK* and *queA* were not fully sequenced for the representatives of Family IB or IC. Arrows indicate the direction of transcription. Conserved regions surrounding the EELs are gray. Shared colors denote similar genes. Open boxes identify genes dissimilar from other genes in the EELs of Family I. Family representatives: IA, *Psy* 61; IB, *Psy* B5; IC, *Psy* 5D4198.



B.

```

B5      1  GATCGGGTTCGTTTAGCAGGGCCGGAAAGTTTATCCGGTTTGACGCCATTAGTAAAAAC
5D4198  1  -----TGGATTGGATGAACT

B5      61  CTGCCTAAATCCCTGTTGACCAACAGAAAACATCCTTATACTTCGCCGCCATTGAGCC
5D4198  17  CTCAGCAAGCGTCTC-----CCCTCATCAACTGTGAGTAGGACT--C-----GAGCC

B5      121  CTGATGGCGGAATTGGTAGACGCGCGGATTCAAAATCCGTTTTTCGAAAAGAAGTGGGAGT
5D4198  63  -TCAAGCCGACAT-----ACCC--CTG--TAAAAGTTCGCCTCACAGGTCAAT--AGT

B5      181  TCGATTCTCCCTCGGGGCACCACCTTTAAAAAAGACCTTGAAATTCAAGGTCTTTTTTTT
5D4198  109  TCGATTCTCCCTCGGGGCACCACCTTTAAAAAAGACCTTGAAATTCAAGGTCTTTTTTTT

B5      241  CGTCTGGTGAAAGTGCCTTGTGTGTCGTTTACCGTATGCATCTCCTGAAAAAATCTGCC
5D4198  169  CGTCTGGTGAAAGTGTCTTGTGTGTCGTTTACCGCATGCATCTCCTGAAAAAATCTGCC

B5      301  TGGCTCAATCTCTAAAAACGCCTTTCTCTGTCTCAGGGCAGGCTGCCTGAAACATCAAACAG
5D4198  229  TGGCTCAATCTCTAAAAACGCCTTTCTCTGTCTCAGGGCAGGCTGCCTGAAACATAAAACAG

B5      361  ATTCACGGCAAGCCTGTGTTTCGAGCAAGCGAAGGATACGCCGGAGAATGGCCCCCTCAGGT
5D4198  289  ATTCACGGCAAGCCTGTGTTTCGAGCAAGGACGGATACGCCCTGAGAATGGCCCCCTCAGGT

B5      421  CCGTCTTCCTGGGCCCGGACTTTGGGATATTCAGCAGTTCATAC
5D4198  349  CCGTCTTCCTGGGCCCGGACTTTGGGATGTTTCAGCAATTCATAC

```

Figure 6. Sequence alignment of the intergenic region between *queA* and *hopPsyB1* in *Psy B5* and the intergenic region between *eelA* and *hopPsyB2* in *Psy 5D4198*.

(A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the IGRs displayed in (A). The full-length tRNA_{leu} of *Psy B5* has a gray background. Non-conserved nucleotides have a white background.

Family II EEL

Initial sequencing of the 4 kb fragments amplified from *P. syringae* pv. phaseolicola (*Pph*) BK378 and *Pph* B130 identified an open reading frame (ORF) downstream of *hrpK* that had similarity to *avrPphE*, which was originally discovered in *Pph* 1032A (race 4) (Mansfield et al., 1994). After completely sequencing the fragment from *Pph* BK378, it was found that the *Pph* BK378 EEL consisted of two ORFs. The product of the first ORF downstream of *hrpK* was similar to AvrPphE from *Pph* 1032A and HopPsyE1 from the *Psy* B728a EEL. In the product of the BK378 *avrPphE* allele, residues G191, W235, and E310, which have been reported as essential for R2-specific avirulence activity (Stevens et al., 1998), were conserved. However, the *Pph* BK378 AvrPphE was 60 residues longer than the *Pph* 1032A AvrPphE. A similar 3'-extension is present in an *avrPphE8* allele found in *P. syringae* pv. phaseolicola race 8 strains (Stevens et al., 1998), including *P. syringae* pv. phaseolicola 2656A, so the *Pph* BK378 allele was designated *avrPphE8* (Figure 7). This extension attenuates the activity of AvrPphE (Stevens et al., 1998). The second ORF was located 154 bp downstream of *avrPphE8* and exhibited 83% I and 87% S with *eelF1* from the *Pto* DC3000 EEL. In recognition of its similarity to *eelF1*, this ORF was named *eelF4*. Notably, EelF4 was missing the fourth 16 residue repeat found in EelF1 (Figure 8). The functions of EelF1 and the repeats are unknown.

The intergenic region between *hrpK* and *avrPphE8* in *Pph* BK378 was less than 41% identical to the intergenic region between *hopPsyC1* and *hopPsyE1* in *Psy* B728a (Figure 9). This indicated that the genetic arrangement of the *Pph* BK378 EEL was likely not the result of a deletion in the *Psy* B728a EEL; conversely, the *Psy* B728a

EEL likely did not arise by an insertion into the *Pph* BK378 EEL. Accordingly, the *Pph* BK378 EEL was classified into Family II (representative strain: *Pph* BK378), a separate group from the *Psy* B728a EEL (Figure 10).

Alleles of *queA* and a *tRNA_{leu}* gene bordered the *Pph* BK378 EEL (Figure 10). Although a nearly identical *avrPphE* allele is downstream of a *hrpK* allele in *Pph* 1032A (Mansfield et al., 1994; Stevens et al., 1998), the chromosomal location of the *hrp* PAI in this strain was not established. The detection of *queA* and the *tRNA_{leu}* gene in *Pph* BK378 indicated the location of the *hrp* PAI in strains carrying a Family II EEL was likely identical to the location of the *hrp* PAI in other strains with characterized EELs.


```

AvrPphE      1  MR IHSAGHS L P A P G P S V E T T E K A V O S S - S A Q N P A S C S S Q T E R P E A G S T Q V R P N Y P Y S S V K
HopPsyE1    1  MR I H S S C H G I S G P V S S A E T V E K A V O S S A Q A Q N E A S H S G P S E H P E S R S C Q A R P N Y P Y S S V K
AvrPphE8    1  MR I H S A G H S L P A P G P S V E T T E K A V O S S - S A Q N P A S C S S Q T E R P E A G S T Q V R P N Y P Y S S V K
IS2656A
-----

AvrPphE     60  T R L P P V S S T G Q A I S D T P S S L P G Y L L R R L D R R P L D E D S I K A L V P A D E A L R E A R R A L P F G R
HopPsyE1    61  T R L P P V A S A G Q S L S E T P S S L P G Y L L R R L D R R P L D Q D A I K G L I P A D E A V G E A R R A L P F G R
AvrPphE8    60  T R L P P V S S T G Q A I S D T P S S L P G Y L L R R L D R R P L D E D S I K A L V P A D E A L R E A R R A L P F G R
IS2656A
-----

AvrPphE    120  G N I D V D A Q R T H L Q S G A R A V A A K R L R K D A E R A G H E P M P E N D E M N W H V L V A M S G Q V F G A G N C
HopPsyE1    121  G N I D V D A Q R S N L E S G A R T L A A R L R K D A E T A G H E P M P E N E D M N W H V L V A M S G Q V F G A G N C
AvrPphE8    120  G N I D V D A Q R T H L Q S G A R A V A A K R L R K D A E R A G H E P M P E N D E M N W H V L V A M S G Q V F G A G N C
IS2656A
-----

AvrPphE    180  G E H A R I A S F A Y G A L A Q E S G R S P R E K I H L A E Q P C K D H V W A E T D N S S A G S S P I V M D P W S N G A
HopPsyE1    181  G E H A R I A S F A Y G A S A Q E K G R A G D E N I H L A A Q S C E D H V W A E T D D S S A G S S P I V M D P W S N G P
AvrPphE8    180  G E H A R I A S F A Y G A L A Q E S G R S P R E K I H L A E Q P C K D H V W A E T D N S S A G S S P I V M D P W S N G A
IS2656A
-----

AvrPphE    240  A T L A E D S R F A K D R S A V E R T Y S F T L A M A A E A G K V A R E T A E N V L T H T T S R L Q K R L A D Q L P N V
HopPsyE1    241  A V F A E D S R F A K D R R A V E R T D S F T L S T A A K A G K I T R E T A E K A L T Q A T S R L Q Q R L A D Q Q A Q V
AvrPphE8    240  A T L A E D S R F A K D R S A V E R T Y S F T L A M A A E A G K V A R E T A E N V L T H T T S R L Q K R L A D Q L P N V
IS2656A
-----

AvrPphE    300  S P L E G G R Y Q P E K S V L D E A F A R R V S D K L N S D D P R R A L Q M E I E A V G V A M S L G A E G V K T V A R Q
HopPsyE1    301  S P V E G G R Y R Q E N S V L D D A F A R R V S D M L N N A D P R R A L Q V E I E A S G V A M S L G A Q G V K T V V R Q
AvrPphE8    300  S P L E G G R Y Q P E K S V L D E A F A R R V S D K L N S D D P R R A L Q M E I E A V G V A M S L G A E G V K T V A R Q
IS2656A
-----

AvrPphE    360  A P K V V R Q A R S V A S S K G M P P R R -----
HopPsyE1    361  A P K V V R Q A R G V A S A K G M S P R A T -----
AvrPphE8    360  A P K V V R Q A R S V A S S K E G L K K P F L O K S K P V N I G A P A R F L S K K M G F F R G Y L K A C L H E D N V S I
IS2656A    1  ----- K E G L K K P F L O K S K P V N I G A P A R F L S K K M G F F R G Y -----

AvrPphE
HopPsyE1
AvrPphE8    420  R T L S N K K L I A L P M I N F Q E N R S
IS2656A
-----

```

Figure 7. Sequence alignment of AvrPphE, HopPsyE1, and the deduced *Pph* BK378 AvrPphE8 gene product. AvrPphE from *Psy* 1302A (race 4), HopPsyE1 from the *Psy* B728a EEL, AvrPphE8 from the *Pph* BK378 EEL, and the product of an insertion sequence observed in a *Pph* 2656A (race 8) *avrPphE* allele were aligned using ClustalW version 1.8. Conserved residues have a black background and similar residues have a gray background.

```

EelF1 1 MRAYKNLTAKIGGFLLALTIIGTSLPAFAVNDCDLDNDNSTGATCGGNDKDLNDNVTDAL
EelF4 1 MRVYNALTAKIGGLLVLTMVGTSLPAFAVNDCDMDNDNSTDARCGGNDKDLNDNVTDAL

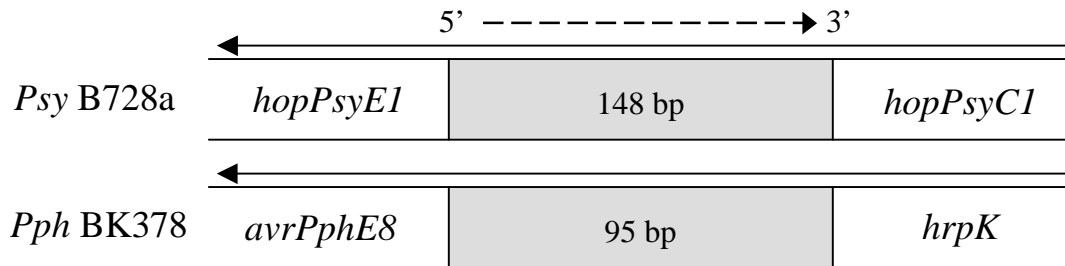
--Repeat 1--> <---Repeat 2---> <---Repeat 3---> <---Repeat 4-
EelF1 61 AFGGNDKMDNDHHTDAAFGGNDKDLNDHHTDAAFGGNDKDLNDNKTDAAFGGNDRDL
EelF4 61 AFGGNDKMDNDHHTDAAFGGTDKDLNDNHHTDAAFGGTDRDLNDNNTDKYHG-----

----->
EelF1 121 DNDNNTDNYNGTPSAAKK
EelF4 115 -----SVPSAAKK

```

Figure 8. Sequence alignment of EelF1 and the deduced *Pph* BK378 EelF4 gene product. EelF1 from the *Pto* DC3000 EEL was aligned with EelF4 from the *Pph* BK378 EEL using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background. Repeated domains are overlined and numbered. Note that one of the domains in EelF4 was disrupted.

A.



B.

```

B728a 1  -TGTTACCTCGCTG-----CACTCGTCGTACCTCGAAAAGCTTACTCAAACCAATCA
BK378 1  ATGTT-CCTCACTGATTGCGGGCGCTCATCGCACCCGAAAAGCTCATTACGGCCAA-CA

B728a 52 CTGCCAACGTTCCACACAAATGGAACGTTGGCCCGTCTCATTAAAAGCAGCGCAGTACGT
BK378 59 CAGG-AACC-----AC-CGATCTCAACGCT--CTCGTCACTGTCAA-----

B728a 112 TACTGGCAGCGTTCAAACGCCAGTAACGCTGCAACGA
BK378 -----

```

Figure 9. Sequence alignment of the intergenic region between *hopPsyE1* and *hopPsyC1* in *Psy* B728a and the intergenic region between *avrPphE8* and *hrpK* in *Pph* BK378 (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.

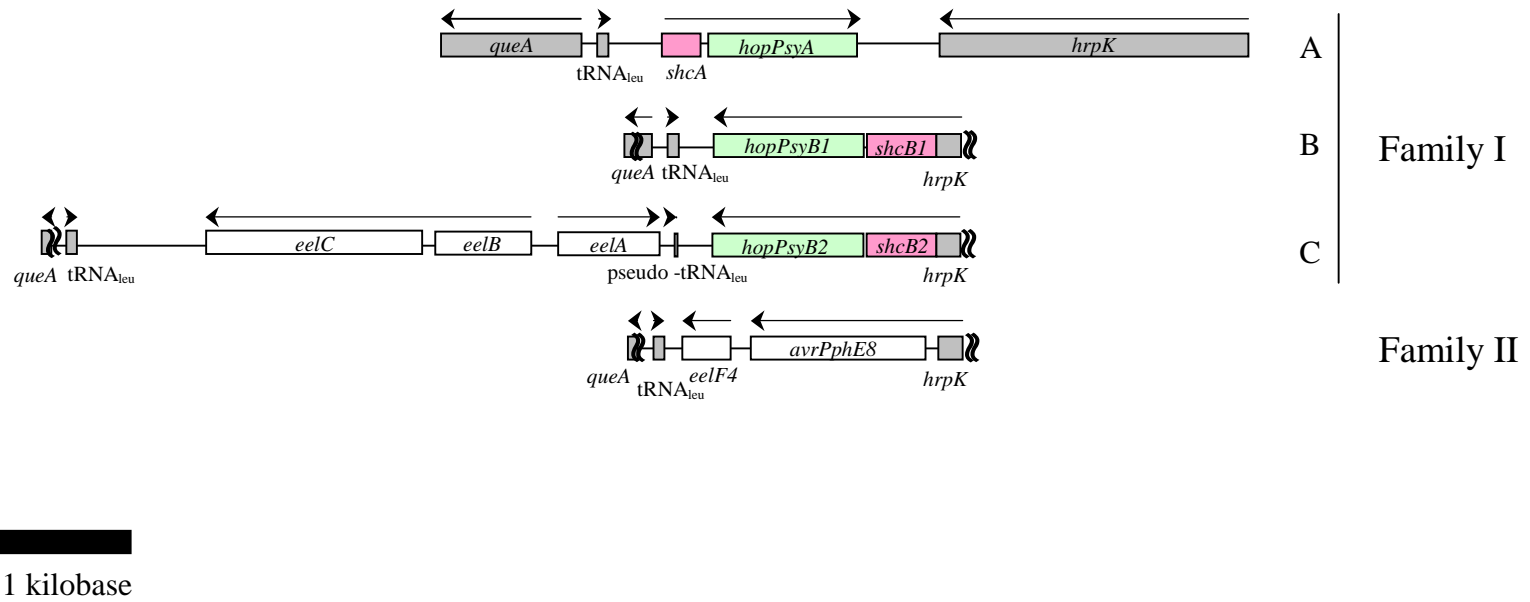


Figure 10. The EELs of Families I and II. The EELs are aligned by their *hrpK* sequences. *hrpK* and *queA* were not fully sequenced for the representatives of Families IB, IC, or II. Arrows indicate the direction of transcription. Conserved regions surrounding the EELs are gray. Shared colors denote similar genes. Open boxes identify genes dissimilar from other genes in the EELs of Families I and II. Family representatives: IA, *Psy* 61; IB, *Psy* B5, IC, *Psy* 5D4198; II, *Pph* BK378.

Family III EEL

Two strains—*Psy* B362 and *Psy* Ps1-Bean—appeared to carry an EEL equivalent to the *Psy* B728a EEL. Their amplified fragments were about the same size as the fragment amplified from *Psy* B728a, and their sequence (approximately 400 nucleotides) downstream of *hrpK* was >98% identical to the region downstream of *hrpK* in the *Psy* B728a EEL. To test if the *Psy* B362 and *Psy* Ps1-Bean EELs were likely equivalent to the *Psy* B728a EEL, the published nucleotide sequence of the *Psy* B728a EEL (locus AF232005) was used to choose pairs of primers for amplifying internal regions of *Psy* B728a. A diagnostic PCR screen showed that all fragments amplified from *Psy* B728a could also be amplified from *Psy* B362 and *Psy* Ps1-Bean (Figure 11), indicating the genetic organization of the EELs in these two strains was likely the same as the genetic organization of the *Psy* B728a EEL. No attempt was made to identify strain-specific sequence polymorphisms in these EELs. This type of EEL was classified into Family IIIA (representative strain: *Psy* B728a).

Three additional strains—*Psy* B452, *Psy* B460, and *Psy* BK035—appeared to have an ortholog of *Psy* B728a *hopPsyC1* downstream of *hrpK*, but had EELs 3 kb smaller than *Psy* B728a. The diagnostic PCR screen was repeated, and it was observed that these strains might lack an *avrPphE* allele and the transposition-associated *eeIDE* operon (Figure 11).

Complete sequencing of the *Psy* B452 EEL revealed that there was nearly a complete deletion of the *hopPsyC* allele. Only the coding sequence for the N-terminal 3 to 71 amino acid residues were conserved, and the truncated ORF lacked an

identifiable ribosome binding site and start codon. This ORF was designated 'hopPsyC' (Figure 12). Alleles of *shcV*, which encodes a putative chaperone, and *hopPsyV*, which encodes an effector with similarity to *avrXct*, were present, as in *Psy* B728 (Figure 13). These alleles were designated *shcV* and *hopPsyV2*, respectively. In *Psy* B452, the *hopPsyV2* allele has a frameshift mutation in codon 208 that likely inactivates effector activity. In *Psy* B452, as in *Psy* B728a, there was a short ORF between the tRNA_{leu} gene and *shcV2* that had sequence similarity to a site-specific recombinase in *Pto* DC3000 (Figure 14).

The intergenic region between *hrpK* and 'hopPsyC' in *Psy* B452 was nearly identical to the intergenic region between *hrpK* and *hopPsyC1* in *Psy* B728a (Figure 15), and the intergenic region downstream of *hopPsyV2* (the intergenic region downstream of where similarity with the *hopPsyV1* allele ends, which is 613 bp past the stop codon of *hopPsyV2*) was nearly identical to the intergenic region downstream of *Psy* B728a *hopPsyV1* (Figure 16), indicating the type of EEL in these strains (classified into Family IIIB [representative strain: *Psy* B452]) was likely a deletion derivative of a Family IIIA EEL (Figure 17).

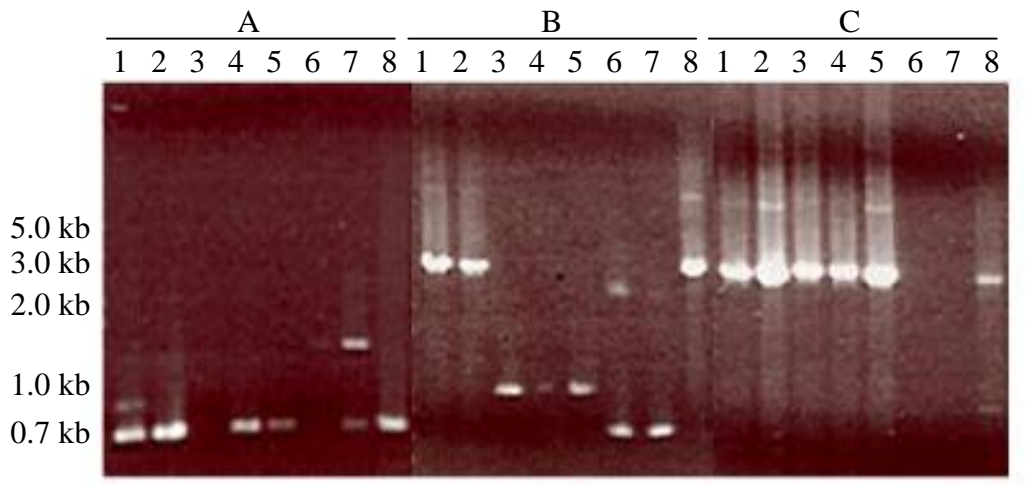


Figure 11. Amplification of diagnostic fragments in Family III EELs. The sizes of marker bands are labeled on the left. Primer pairs and strains: A, B1F and BER, which amplify a 630 bp fragment from *Psy* B728a extending from *hopPsyCI* to *hopPsyE1*; B, BEF and B5R, which amplify a 2898 bp fragment from *Psy* B728a extending from *hopPsyE1* to *hopPsyVI*; C, B5F and Q920, which amplify a 2651 bp fragment extending from *hopPsyVI* to *queA*; 1, *Psy* B362; 2, *Psy* Ps1-bean; 3, *Psy* B452; 4, *Psy* B460; 5, *Psy* BK035; 6, *Psy* W4N15; 7, *Psy* W4N27; 8, *Psy* B728a. The fragments amplified from *Psy* B362 and *Psy* Ps1-bean were about the same size as the fragments amplified from *Psy* B728a. The remaining strains clustered by banding pattern into two groups: Group 1—*Psy* B452, *Psy* B460, and *Psy* BK035; Group 2—*Psy* W4N15 and *Psy* W4N27.

A.

```

HopPsyC1    1  ---MGCVSSKASVISSDSFRASYTNSPEASSVHQARTPRCGELQGPQVSRLMPYQQALV
'HopPsyC'   1  MSLKSCVSSKASVISSDSFRASYTNSPEASSVHQARTPRCGELQGPQVSRLMPCQQALV

HopPsyC1    58  GVARWPNPHFNRRDDAPHQMEYGESFYHKSRELGASVANGEIETFQELWSEARDWRASRAG
'HopPsyC'   61  GVARWEDPHFN-----

HopPsyC2    118  QDARLFSSSRDPNSSRAFVTPITGPIYEFKDRFANRKDGEKHKMDFLPHSNTFRFHGKI
'HopPsyC'   72  -----SLLN-----RGTVVPLVFGEEKVKG-----LSAHEI STSIPGKNNRFMY---

HopPsyC1    178  DGERLPLTWISISSDRRADRTKDPYQRLRDQGMNDVGEPNVMLHTQAEYVPKIMQHVEHL
'HopPsyC'   -----

HopPsyC1    238  YKAATDAALSDANALKKLAEIHWWTVQAVPDFRGSAAKAELCVRSIAQARGMDLPPMRLG
'HopPsyC'   -----

HopPsyC1    298  IVPDLEALTMPLKDFVKSIEGFFEHN
'HopPsyC'   -----

```

B.

Shine-Delgarno sequence and start codon for HopPsyC1:

G T T A A A A A G G T G A G C C C G A T G G

Psy B452 '*hopPsyC*' was missing the start codon of *hopPsyC1*, but might still be expressed upstream from a GTG start codon. Nonetheless, expression of 'HopPsyC' would likely be low since the possible GTG start codon was immediately downstream of the Shine-Delgarno sequence:

G T T A A A A A G G T G A G C C T G A A G A

Figure 12. Properties of *Psy* B452 '*hopPsyC*'. (A) Sequence alignment of HopPsyC1 and the deduced *Psy* B452 'HopPsyC' gene product. HopPsyC1 from the *Psy* B728a EEL and 'HopPsyC' from the *Psy* B452 EEL were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background. (B) 'HopPsyB' was unlikely to be highly expressed due to a weak Shine-Delgarno site.

A.

B728a ShcV	1	MTLERIEQQNTLFVYLCVGTLPASSTLLSDILAANLFHYGSSDGAAGLDEKNNVLL
B452 ShcV	1	MTLERIEQQNTLFVYLCVGTLPASSTLLSDILAANLFHYGSSDGAAGLDEKNNVLL
B728a ShcV	61	FQRFDPRLRIDEDHFVSACVQMI EVAKIWRAKLLHGHSAPLASSTRLT KAGLMLTMAGTIR
B452 ShcV	61	FQRFDPRLRIDEDHFVSACVQMI EVAKIWRAKLLHGHSAPLASSTRLT KAGLMLTMAGTIR

B.

HopPsyV1	1	MNISGPNRRQGTQAENTESASSSSVTNPPLQRGEGRRLRRQDALPTDIRYNANQTATSPQ
HopPsyV2	1	MNISGPNRRQGTQAENTESASSSSVTNPPLQRGEGRRLRRQDALPTDIRYNANQTATSPQ
HopPsyV1	61	NARAAGRYESGASSSGANDTPQAE G SMPSSSAFLQFRLAGGRNHSELENFHTMMLNSPKA
HopPsyV2	61	NARAAGRYESGASSSGANDTPQAE G SMPSSSAFLQFRLPGRTHSELED FHTMMLNSPKA
HopPsyV1	121	SRGDAIPEKPEAIPKRLLLEKMEP INLAQLALRDKDLHEYAVMVCNQVKKGEGPNSNITQG
HopPsyV2	121	SRGDAIPEKPEAIPKRLLLEKMEP INLAQLALRDKDLHEYAVMVCNQVKKGEGPNSNITQG
HopPsyV1	181	DIKLLPLFAKAENTRNPGLNLHTFKSHKDCYQAIKEQNRDIQKNKQSLSMRVVYPPFKKM
HopPsyV2	181	DIKLLPLFAKAENTRNPGLNLHTFKMS-----
HopPsyV1	241	PDHHIALDIQLRYGHRPSIVGFESAPGNIIDAAEREILSALGNV KIKMVGNFLOYSKTDCC
HopPsyV2		-----
HopPsyV1	301	TMFALNNALKAFKHHEEYTARLHNGEKQVPIPATFLKHAQSKSLVENHPEKDTTVTKDQG
HopPsyV2		-----
HopPsyV1	361	GLHMETLLHRNRAYRAQRSAGQHVTSIEGFRMQEIKRAGDFLAANRVRAKEP
HopPsyV2		-----

Figure 13. *Psy* B452 contained alleles of *shcV* and *hopPsyV*. (A) Sequence alignment of ShcV and the deduced *Psy* B452 ShcV gene product. ShcV from the *Psy* B728a EEL and ShcV from the *Psy* B452 EEL were aligned using ClustalW version 1.8. Identical residues have a black background. (B) Sequence alignment of HopPsyV1 and the deduced HopPsyV2 gene product. HopPsyV1 from the *Psy* B728a EEL and HopPsyV2 from the *Psy* B452 EEL were aligned as in (A).

```

B728a ORF 1 I E F G P D F H S H S D G K I P R Q R Q R Q P R G F V L A L T D T A A R T A K P R E K L Y C L T A T A G L C L E V T R G
B452 ORF 1 I E F G P D F Y S H S D G K I P R Q R Q -- P R G F V L A L T D T A V R T A K P R E K L Y C L T D T A G L C L E I T R G
integrase 1 ----- M P L S D T T I R T A K P K D K L Y R L T D A N G L C L E I A P S

B728a ORF 61 G S K L Q R F R Y R S A R T A N M S I T K P E ----- T L P K P I Q R A L N Q I A H
B452 ORF 59 G S K L W R F R Y R F G K A K M M E L G -----
integrase 34 G S K L W R Y R Y R F N G K A K M L A L G A Y P A V T L L K A R Q L R D S A R Q L L V E G N D P G E H K K T A Q Q A Q K

B728a ORF 99 S R P L L Y Q A A C R D Q I C K ----- E I D T L L A R G M S H Q D A T E P L R A C E P T L D -- P D Y -----
B452 ORF 80 A Y P T V T L A K A R E R R ----- E D A R Q R H R A C G V R R -----
integrase 94 V E G L I F E T L A R E W F A Y N S P R W A E S T T Y K A K L Y M E N D L I P G I G A R P V K A L T R P D L V D L V R K

B728a ORF -----
B452 ORF -----
integrase 154 V E A R G T L N A A G K I R Q W L H Q I F R Y G L A K G V V E S N P A T D L D V V A A P Q K A P R H H P H V P F S E L P

B728a ORF -----
B452 ORF -----
integrase 214 E L L E I C D G S K I N T L T R C A I H L L V L T A V R P G E L R N A P W S E F D L D A A T W A I P K E R M K A R R P H

B728a ORF -----
B452 ORF -----
integrase 274 V V P L P T Q A V D I L R Q L Q P I T G R Y P L V F A G Q H N P S R P M S E N T I N K A L R L L G Y E N R Q T G H G F R

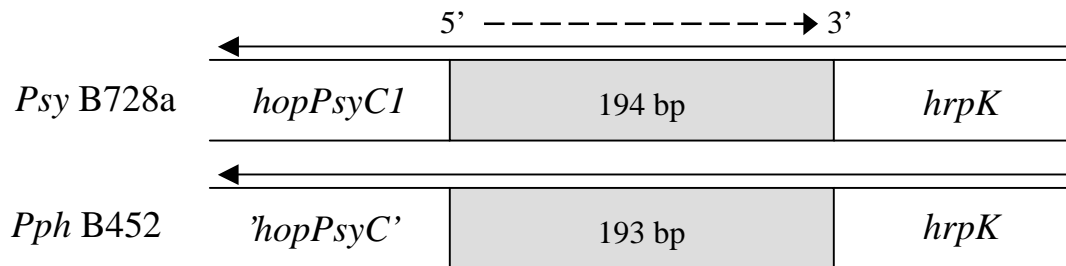
B728a ORF -----
B452 ORF -----
integrase 334 H L L S T E L N G R G Y N K D W I E R Q L A H G D A D G I R D T Y N H A S Y L E Q R R G M M Q A W A D S I D A L S A G S

B728a ORF -----
B452 ORF -----
integrase 394 N V V S I K R Q A

```

Figure 14. Sequence alignment of a site-specific recombinase in *Pto* DC3000 with short, and likely not expressed, ORFs in *Psy* B728a and *Psy* B452. The site-specific recombinase in *Pto* DC3000 (locus tag PSPTO5344, accession NP_795075) and the short ORFs located between the *tRNA_{leu}* gene and *shcV* in *Psy* B728a and *Psy* B452 were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background.

A.

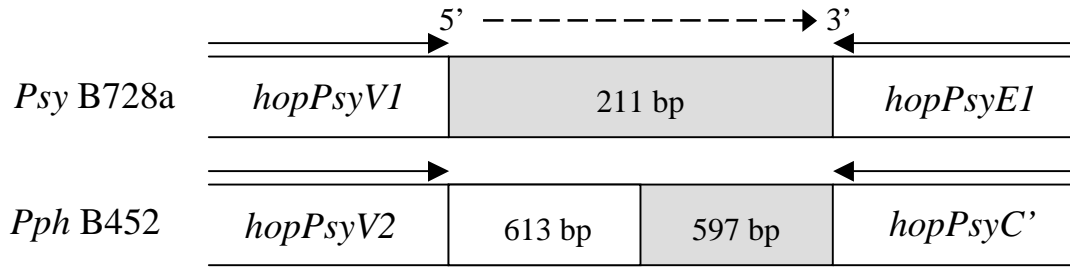


B.

B728a	1	CGGGCTCACCTTTTTTAACTATTTAATTGAATGTCAGGCCTGACAAAAGTCCGATATTTG
B452	1	CAGGCTCACCTTTTTTAACTGTTTAATTGAATGTCAGGCCTGACAAAAGCCGATATTTG
B728a	61	TCAAATAGCCTCGATTTTTTAGCGTCTAGCGCTAACGCTTCCATAAACTTTATGCGGC
B452	61	TCAAATAGCCTCCGTTTTTAGCGTCTAGCGCTAACGCTTCCATAAACTTTATGCGTC
B728a	121	CGATTAAACAGCTCACACGGGATCTGTAAAATGAAGGCTCAGCATTCAAGGCGTCTGAG
B452	121	CGATTAAACAGCTCACACGG-ATCTGTAGAAATGAAGGCTCAGCATTCAAGGCGTCTGAG
B728a	181	CCGACTCAATTCAA
B452	180	CCGACTCAATTCAA

Figure 15. Sequence alignment of the intergenic region between *hopPsyC1* and *hrpK* in *Psy* B728a and the intergenic region between *'hopPsyC'* and *hrpK* in *Psy* B452. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.

A.



B.

Sequence alignment of the intergenic region between *hopPsyV1* and *hopPsyE1* in *Psyllid* B728a and the intergenic region between *hopPsyV2* and *hopPsyC'* in *Psyllid* B452. For *Psyllid* B452, the aligned intergenic sequence is downstream of where similarity with the *hopPsyV1* allele ends, which is 613 bp past the stop codon of *hopPsyV2*. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the gray IGRs displayed in (A). Conserved nucleotides have a black background.

```

B728a 1 CTCACGTCCTCTGAAAAACGCGCCTTACGGTTGGCGCGTTTTGTCCGAGGACAGGTGTT
B452 1 CTCACGTCCTCTGAAAAACGCGCCTTACGGTTGGCGCGTTTTGTCCGAGGACAGGTGTT

B728a 61 CGAGAAGAATAAAGGTTGTCACCTCAAGAGGCAGGAGGAACGAGTCGAGGCTCGGCTTGC
B452 61 CGAGAAGAATAAAGGTTGTCACCTCAAGAGGCAGGAGGAACGAGTCGAGGCTCGGCTTGC

B728a 121 TCAGTAAGCGC----CTCATGAACCCACATTCCAAACCGGCC-ATTGAAGCTTGATGC-
B452 121 TCAGCAAACGAAAAGCTGGCGTAACGGGTTTGGCCAGCCCCCTATTGGAGCCGGGCGC

B728a 175 TGT-----GCTCCCGATTCTTTTCAGGAGCCAGCACCTCATG-----
B452 181 TGTATCGCGGCCTCAGCTCCACTGCAAGACGAATCAGGTTATAGAAAGACGTGCGTCACT

B728a -----
B452 241 TTGCTGAAAGAGCGCCAACGCTGTGACTGCGAAGAAAAGTCCGAGATCACGCACA

B728a -----
B452 301 CGCAACTCAGCGTTTATATCCTGTAGCGCCTCAAGCATCAGCGTATTCAAGGGCGCACCA

B728a -----
B452 361 TTGGACATTTCCGAGACCTGCTCAAGGAGATTGTTGGGAAACCTATCAACAGCTGACGCA

B728a -----
B452 421 TCCAGATACTCGGCTTCTGTTTCTTGCTTGTTTTTCAGCCTTACAAGCAGATCTTCAGGC

B728a -----
B452 481 AGCTTTATATTCTTGAGCTCACATCCCAGAGACAGCGTTACAAGATCGCCAAGGCTTCTG

B728a -----
B452 541 ATCTCCACTTCTCGATTGCCACCGTCTGCCTCCCCGGCCAGGCGATGCTCATTTGC

```

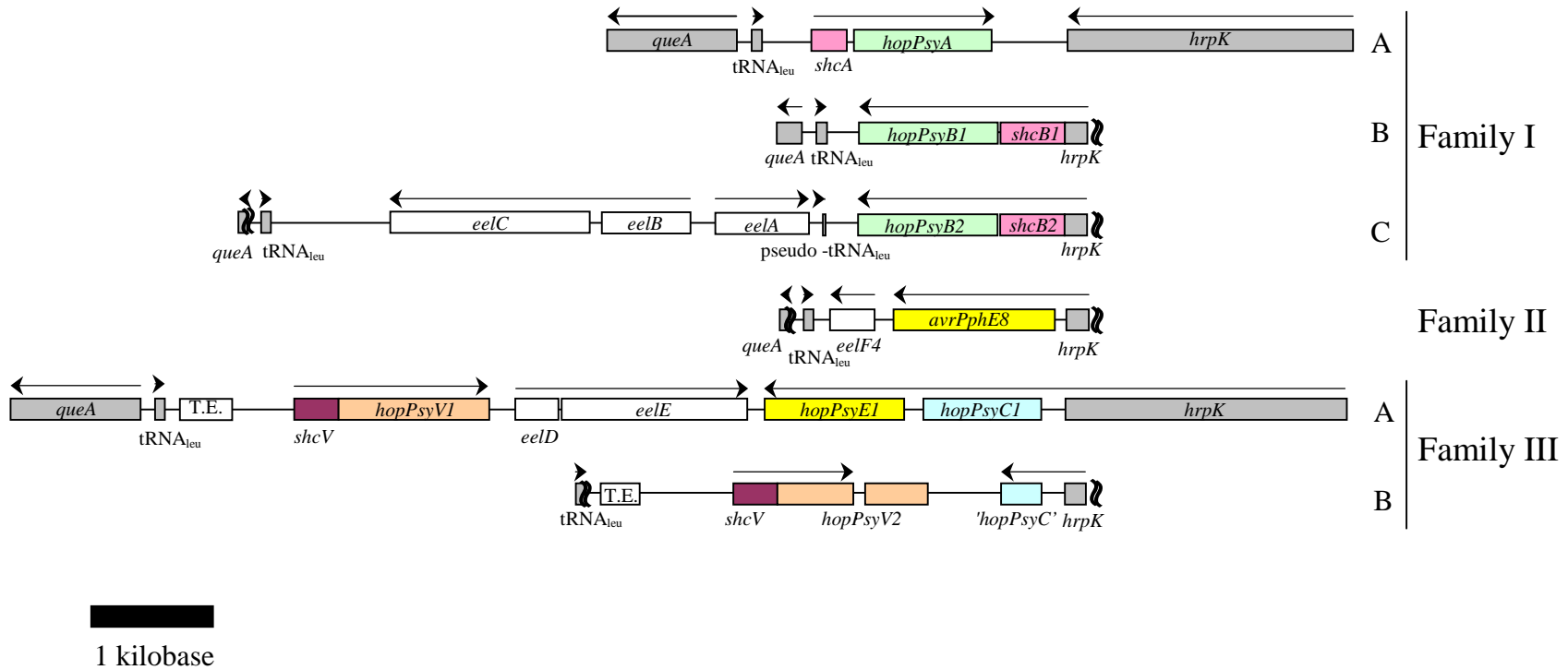


Figure 17. The EELs of Families I, II, and III. The EELs are aligned by their *hrpK* sequences. *hrpK* and *queA* were not fully sequenced for the representatives of Families IB, IC, II, or IIIB. Arrows indicate the direction of transcription. Conserved regions bordering the EELs are gray. Shared colors denote similar genes. Open boxes identify genes dissimilar from other genes in the EELs of Families I, II, and III. Open reading frames that have similarity to transposable elements are labeled "T.E." Family representatives: IA, *Psy* 61; IB, *Psy* B5; IC, *Psy* 5D4198; II, *Pph* BK378; IIIA, *Psy* B728a; IIIB, *Psy* B452.

Family IV EEL

The EEL of *Psy* W4N15 and *Psy* W4N27 amplified as a 5.5 kb fragment and had an ORF immediately downstream of *hrpK* that was similar to *Psy* B728a *hopPsyC1*. The diagnostic PCR experiment (Figure 11) displayed that these two strains might have an EEL different from the Family IIIA and IIIB EELs. Complete sequencing of the *Psy* W4N15 EEL revealed that this locus contained a *hrpK-hopPsyC-hopPsyE* operon, as in the Family IIIA EEL, and alleles of *eelF* and *eelG*, as in the *Pto* DC3000 EEL. In *Psy* W4N15, the deduced product of the first gene downstream of *hrpK* was 74% I and 80% S to HopPsyC1, and was designated HopPsyC2 (Figure 18). The deduced product of the next gene downstream of *hopPsyC2* was 97% identical to HopPsyE1 from the *Psy* B728a EEL, and was named HopPsyE2 (Figure 19). The deduced gene products of the *eelF* and *eelG* alleles were 85% identical and 89% identical, respectively, to *Pto* DC3000 EelF1 and EelG1, and were appropriately named EelF3 and EelG3 (Figure 20). This type of EEL was classified into Family IV (representative strain: *Psy* W4N15) (Figure 21).

```

HopPsyC1      1  MGCVSSKASVISSDSFRASYTNSP-----EASSVHORARTPRCGELQGPOVSRLMPYQQA
HopPsyC2      1  MGCVSSKASVISSDSFRASYTNSPGAFARES SSVHNRARTPRY GELQGQA SRLMPYQQA

HopPsyC1     56  LVGVARWPNPHFNRDDAPHQMEYGESFYHKSRELGASVANGELTIFQELWSEARDWRASR
HopPsyC2     61  LVGVARWPDSHFNRDDAPHQMEYGCQSFYNKSRALGGSVANGELTGSFQELWDEARDWRASR

HopPsyC1    116  AGQDARLFSSSRDPNSSRAFTVPIITGPTYEFLKDRFANRKDGEKHKMMDFLPHSNTFRFHG
HopPsyC2    121  AGADADVFLSPRDPNSYREYATPLAEQYSYIKDRFANRKDGEVVGSPADFLPKSKTFRISG

HopPsyC1    176  KIDGERLPLTWISISSDRRADRTKDPYQRLRDQGMNDVGEPNVMLHTQA EYVPKIMQHVE
HopPsyC2    181  KIDGEQIPLTRITVSKDRHADRMADPYPRLRNQGHGDLGEPNLSHTSAEYVQIMSHVE

HopPsyC1    236  HLYKAATDAALSDANALKKLAETHWWTVQAVPDERGSAAKAELCVR SIAQARGMDLPPMR
HopPsyC2    241  SLHQSATDPSVSDSHALKTLAEMHWWMAHAMPDKRGSAAKTELCVR SIAQARGMDLPPMR

HopPsyC1    296  LGIVPDLEALTMPLKDFVKS YEGFF EHN
HopPsyC2    301  LGIVPDLEALTMPLKDFVKS YQGF SRK-

```

Figure 18. Sequence alignment of HopPsyC1 and the deduced HopPsyC2 gene product. HopPsyC1 from the *Psy* B728a EEL and HopPsyC2 from the *Psy* W4N15 EEL were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background.

```

HopPsyE1      1  MR IHSSGHGTS CPVSSAETVEKAVQSSAQANEASHSGPSEHPESRSCQARPNYPYSSVK
HopPsyE8      1  MRIHSAAGHSLPAPGPSVETTEKAVQS-SSAONPASCSSTQTEPEAGSTQVRPNYPYSSVK
HopPsyE2      1  MR IHSSGHGTCAPVSSAETVEKAVQSSAQANEASHSGPSEHPESRSCQARPNYPYSSVK

HopPsyE1     61  TRLPPVASAGQSLSETPSSLPGYLLRRLDRRPLDQDAIKGLIPADEAVGEARRALPFGR
HopPsyE8     60  TRLPPVSSITGOAISDTPSSLPGYLLRRLDRRPLDEDSIKLVPADAEALREARRALPFGR
HopPsyE2     61  TRLPPVASAGQSLSETPSSLPGYLLRRLDRRPLDQDAIKGLIPADEAVGEARRALPFGR

HopPsyE1    121  GNIDVDAQRSNLESGARTLAARRLRKDAETAGHEPMPENEDMNWHVLVAMSGQVFGAGNC
HopPsyE8    120  GNIDVDAQRTEHQSGARAVAAKRLRKAERAGHEPMPENEDMNWHVLVAMSGQVFGAGNC
HopPsyE2    121  GNIDVDAQRSNLESGARTLAARRLRKDAETAGHEPMPES-EDMNWHVLVAMSGQVFGAGNC

HopPsyE1    181  GEHAR IASFAYGAS AQEKGRAGDENIHLAAQSGEDHVWAE TDDSSAGSSP IVM DPWSNGP
HopPsyE8    180  GEHAR IASFAYGALAQE SGRSPREK IHLAEQEGKDHVWAE TDN SSAGSSP IVM DPWSNGA
HopPsyE2    181  GEHAR IASFAYGALAQEKGR TGDENIHL SAQSGEDHVWAE TDDSSAGSSP IVM DPWSNGP

HopPsyE1    241  AVFAEDSRFAKDRRAVERTDSFTLSTAAKAGKITRETAEKAL TQATSRLQORLADQQAQV
HopPsyE8    240  AILAEEDSRFAKDRSAVERTYSFTLAMAAEAGK VARETAENVLTH TTSRLQORLADQLPNV
HopPsyE2    241  AVFAEDSRFAKDRSAVERTDSFTLSTAAEAGKITRETAEKAL TQATSRLQORLADQQAQV

HopPsyE1    301  SPVEGGRYRQENSVLDDAFARRVSDMLNNADPRRALQVEIEASGVAMSLGAQGQVKTVVVQ
HopPsyE8    300  SPLEGGRYRQEPKSVLDEAFARRVSDKLNSDPRRALQMEIEAVGVAMSLGAGGVKTVARQ
HopPsyE2    301  SPVEGSR YRQENSVLDDAFARRVSDMLNNADPRRALQVEIEASGVAMSLGAQGQVKTVVVQ

HopPsyE1    361  APKVVRQARGVASAKGMSPRAT-----
HopPsyE8    360  APKVVRQARSVASSKEGLKPPFLQKSKPVNIGAPARFLSKKMGFFRGYKACLHEDNVSI
HopPsyE2    361  APKVVRQARGVASAKGMSQRAT-----

HopPsyE1    -----
HopPsyE8    420  RTLSNKKLIALPMINFQENRS
HopPsyE2    -----

```

Figure 19. Sequence alignment of HopPsyE1, AvrPphE8, and the deduced HopPsyE2 gene product. HopPsyE1 from the *Psy* B728a EEL, AvrPphE8 from the *Psy* BK378 EEL, and HopPsyE2 from the *Psy* W4N15 EEL were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background.

A.

```

EelF1      1  MRAYKNTAKIGGFLLALTIIGTSLPAFAVNDCDLDNDNSTGATCGGNDKDLDNDNVTDA
EelF4      1  MRVYNALTAKIGGLLVLTMVGTSLPAYAVNDCDMDNDNSTDARCGGNDKDLDNDNVTDA
EelF3      1  MRAYKTLTAKIGGLLLALTIVGTSLETYAVNDCDMDNDNSTDATCGGNDKDLDNDNVTDA

EelF1      61  AFGGNDKMDNDHHTDAAFGGNDKDLDNDHHTDAAFGGNDKDLDNDNKTDAAFGGNDRDL
EelF4      61  AFGGNDKMDNDHHTDAAFGGTDKDLDNDNHTDAAFGG-----TDRDL
EelF3      61  AFGGNDKMDNDNHTDAAFGGTDKDLDNDNHTDASFGGNDKMDNDHHTDAAFGGNDRDL

EelF1      121  DNDNNTDNYNGTPSAAKK-
EelF4      104  DNDNNTDKYHGSVPSAAKK
EelF3      121  DNDNNTDKYDGASSAAKK

```

B.

```

EelG1      1  MNKIVYVKAYFKPIGEEVSVKVPTGEIKKGFFGDKELMKKETQWQQTGWSDCQIDGERLS
EelG3      1  MNKIVYVKAYFKPVGEEVIVKVPTGEIKKGFFGKEVMRKETRWQQTGWSDCQIDGERLS

EelG1      61  KDVEDAVAQLNADGYEIQTVLPILSGAYDYALKYRYEIRHNRTELSPGDQSYVFGYGYSF
EelG3      61  KDVEDAVAQLNADGYEIQTVLPILSGAYDYALKYRYEMRHDRTDLDSRDLSYVFGYGYSF

EelG1      121  TEGVTLVAKKFQSSAS
EelG3      121  TEGVTLVAKKFQSSAR

```

Figure 20. Sequence alignments with EelF3 and EelG3. (A) Sequence alignment of EelF1, EelF4, and the deduced EelF3 gene product. EelF1 from the *Pto* DC3000 EEL, EelF4 from the *Pph* BK378 EEL, and EelF3 from the *Psy* W4N15 EEL were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background. (B) Sequence alignment of EelG1 and the deduced EelG3 gene product. EelG1 from the *Pto* DC3000 EEL and EelG3 from the *Psy* W4N15 EEL were aligned as in (A).

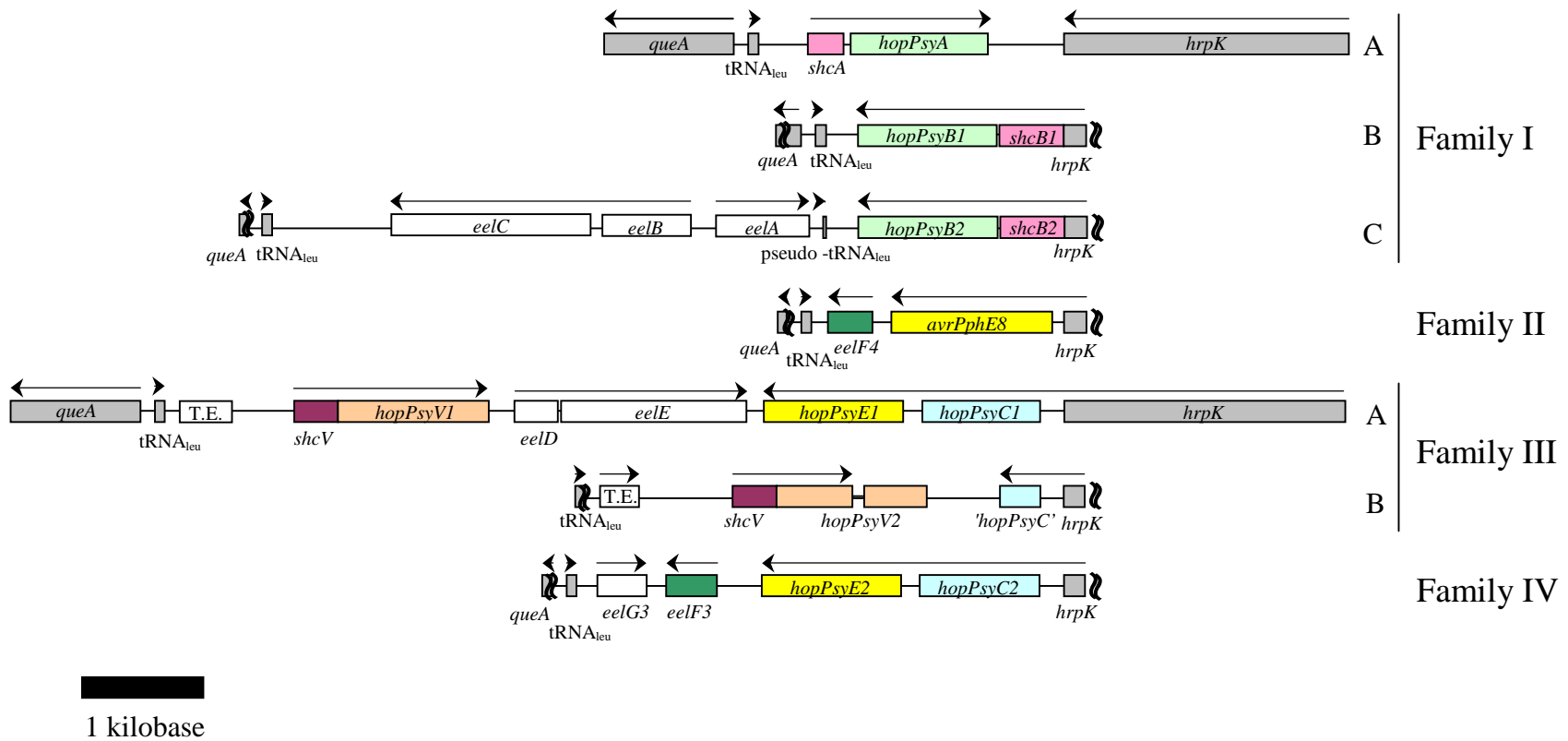


Figure 21. The EELs of Families I, II, III, and IV. The EELs are aligned by their *hrpK* sequences. *hrpK* and *queA* were not fully sequenced for the representatives of Families IB, IC, II, IIIB, or IV. Arrows indicate the direction of transcription. Conserved regions bordering the EELs are gray. Shared colors denote similar genes. Open boxes identify genes dissimilar from other genes in the EELs of Families I, II, III, and IV. Open reading frames that have similarity to transposable elements are labeled "T.E." Family representatives: IA, *Psy* 61; IB, *Psy* B5; IC, *Psy* 5D4198; II, *Pph* BK378; IIIA, *Psy* B728a; IIIB, *Psy* B452; IV, *Psy* W4N15.

Family V EEL

The EELs amplified from three strains—*Psy* NCPPB1053, *Pto* 2844, and *Ppe* 5846—contained an ORF immediately downstream of *hrpK* with high identity to *Pto* DC3000 *hopPtoB1*. The fragments amplified from *Psy* NCPPB1053 and *Pto* 2844 were about the same size as the fragment amplified from *Pto* DC3000, and the fragment amplified from *Ppe* 5846 was approximately 2.5 kb smaller than the fragment amplified from *Pto* DC3000. To test if the EELs of *Psy* NCPPB 1053 and *Pto* 2844 were likely equivalent to the *Pto* DC3000 EEL, the published nucleotide sequence of the *Pto* DC3000 EEL (locus AF232004) was used to identify primer pairs for amplifying internal regions of the *Pto* DC3000 EEL. A diagnostic PCR screen was used to differentiate these EELs (Figure 22). For *Psy* NCPPB1053 and *Pto* 2844, the diagnostic-PCR-amplified fragments were approximately the same size as the fragments amplified from *Pto* DC3000, indicating the EELs in these strains were likely equivalent to the *Pto* DC3000 EEL. The *Psy* NCPPB1053 and *Pto* 2844 EELs were classified with the *Pto* DC3000 EEL into Family VA (representative strain: *Pto* DC3000). No attempt was made to identify sequence polymorphisms in these EELs.

In the diagnostic-PCR screen, the fragment extending from *eelF1* to *queA* could not be amplified from the *Ppe* 5845 EEL. After completely sequencing the *Ppe* 5846 EEL, it was found that the deduced product of the first ORF downstream of *hrpK* retained 99% I to *Pto* DC3000 HopPtoB1 over the majority of its length, but was 138 residues shorter than HopPtoB1 at its carboxy terminus. This ORF was named HopPtoB5 (Figure 23A). Since the active domains of HopPtoB have not been

established, it is unclear whether HopPtoB5 likely functions in the same manner as HopPtoB1. Nonetheless, HopPtoB5 is probably secreted by the *hrp* TTSS given that it shared its 82 amino-terminal residues with HopPtoB1, and HopPtoB1 is a known effector. The next two ORFs downstream of *hopPtoB5* displayed high identity to transposable elements. The ORF immediately downstream of *hopPtoB5* was similar to a cointegrase from *Pseudomonas stutzeri* (Figure 24A). The other ORF was similar to a putative transposase from *P. stutzeri* (Figure 24B). These ORFs were not named. The remaining ORF in the *Ppe* 5846 EEL was similar to *eelG* found the EELs of Families IV and V, and was designated *eelG2* (Figure 23B). *EelG2* lacked the coding sequence for the 53 carboxy-terminal residues found in *Pto* DC3000 *EelG1*, but was 99% identical to the first 82 amino-terminal residues of *EelG1*. The remaining polypeptide retained 99% identity to *EelG1*. The intergenic region between *hrpK* and *hopPtoB5* in *Ppe* 5846 was nearly identical to the intergenic region between *hrpK* and *hopPtoB1* in *Pto* DC3000 (Figure 25). Additionally, the intergenic region downstream of the *tRNA_{Leu}* gene was nearly identical in the two EELs (Figure 26). This *Ppe* 5846 EEL (classified into Family IVB [representative strain: *Ppe* 5846]) was likely a deletion derivative of a Family IVA EEL (Figure 27).

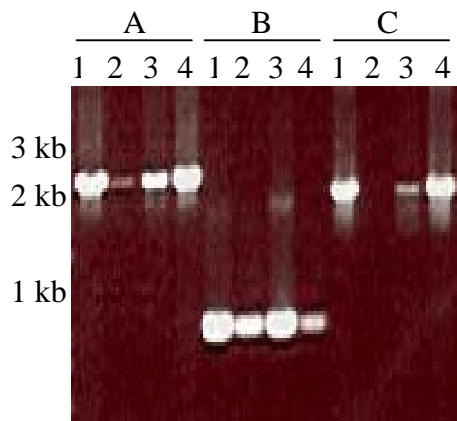


Figure 22. Amplification of diagnostic fragments in Family V EELs. The sizes of marker bands are labeled on the left. Primer pairs and strains: A, D1F and D2R, which amplify a 2059 bp fragment from *Pto* DC3000 extending from *hopPtoB1* to *eelH*; B, D2F and D3R, which amplify a 911 bp fragment from *Pto* DC3000 extending from *eelH* to *eelF1*; C, D3F and Q920, which amplify a 2065 bp fragment from DC3000 extending from *eelF1* to *queA*; 1, *Psy* NCPPB1053; 2, *Ppe* 5846; 3, *Pto* 2844; 4, *Pto* DC3000. The fragments amplified from *Psy* NCPPB1053 and *Pto* 2844 were about the same size as the fragments amplified from *Pto* DC3000. The fragment extending from *eelF1* to *queA* could not be amplified from *Ppe* 5846.

A.

```

HopPtoB1 1 MRPVGGPAPGYYPPTYEAERPTAQAGNDRARSSQASSSPAASVAPETPMLGDLKRFPAQ
HopPtoB5 1 MRPVGGPAPGYYPPTYEAERPTAQAGNDRARSSQASSSPAASVAPETPMLGDLKRFPAQ

HopPtoB1 61 RYPDMKVENIRLKI EGQEPGGKIGVKHTRRRKPDAGSSSHVHGGQSVASTSASAQSKALQ
HopPtoB5 61 RYPDMKVENIRLKI EGQEPGGKIGVKHTRRRKPDAGSSSHVHGGQSVASTSASAQSKALQ

HopPtoB1 121 DTNFKASDLAELARWCESPHPYALAPSKAAGKSSQLSANVVSILLQEGKHALEQRLEAQQ
HopPtoB5 121 DTNFKASDLAELARWCESPHPYALAPSKAAGKSSQLSANVVSILLQEGKHALEQRLEAQQ

HopPtoB1 181 LKLADV VVSEGRDHLHINLNYLEMDSCLGTSKGLWAPDSNDKLI AKAARYFDDFNAQKL
HopPtoB5 181 LKLADV VVSEGRDHLHINLNYLEMDSCLGTSKGLWAPDSNDKLI ARAARYFDDFNAQKL

HopPtoB1 241 PELAPLTKMKSKDSL GVMRELLRDAPGLVIGEGHNSTSSKRELINNMKSLKASGVTTLFM
HopPtoB5 241 PELAPLTKVKSKDSL GVMRELLRDAPGLVIGEGHNSTSSKRELINNMKSLKASGVTTLFM

HopPtoB1 301 EHLCAESHDKA LNNYLSAPKGSMPARLKNYLDLQSQGHQAPEELHTKYNF TTLVEAAKH
HopPtoB5 301 EHLCAESHDKV LNNYLSAPKGSPIVCGAD-----

HopPtoB1 361 AGLRVVSLDTTSTYMAPEKAEIKRAQAMNYAAEKIRLSKPEGKWVAVFGATHATSCDGV
HopPtoB5 -----

HopPtoB1 421 PGLAELHGVRSLVIDDLGLKSRATVDINVKNYGGKLNPDVRLSYKV
HopPtoB5 -----

```

B.

```

EelG1 1 MNKIVYVKAYFKPIGEEVSVKVPTGEIKKGFFGDKEIMKKETQWQQTGWSDCQIDGERLS
EelG2 1 MNKIVYVKAYFKPIGEEVSVKVPTGEIKKGFFGDKEIMKKETQWQQTGWSDSQIDGERLS

EelG1 61 KDVEDAVAQLNADGYEIQTVLPISGAYDYALKYRYEIRHNRTLSPGDQSYVFGYGYSF
EelG2 61 KDVEDAVAQLNADGYEIQTVLPYICCR-----

EelG1 121 TEGVTLVAKKFQSSAS
EelG2 -----

```

Figure 23. Sequence alignments with HopPtoB5 and EelG2. (A) Sequence alignment of HopPtoB1 and the deduced HopPtoB5 gene product. HopPtoB1 from the *Pto* DC3000 EEL and HopPtoB5 from the *Ppe* 5846 EEL were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background. (B) Sequence alignment of EelG1 from the *Pto* DC3000 EEL and the deduced EelG2 gene product from the *Ppe* 5846 EEL. EelG1 and EelG2 were aligned as in (A).

A.

```

cointegrase 1 MSWRAWPRPNHHPASKPTCRSRKPPSLTRRATTACVARPRRSVMRDLMAELKELRLHGMA
5846 ORF 1 -----
cointegrase 61 TAWAELTAQGESNTASSKWLLEHLLLEQEHTDRAMRSVSHQNMMAKLPMHRDLASFDFNAS
5846 ORF 1 -----
cointegrase 121 SADARLISELASLAFTDTAQNVLIGCPGTGKTHLATALAVSGITRHGKRVRFYSTVDLV
5846 ORF 1 -----MSCSSAAWHRKTHLASALAVSGITAYNKRVRFFSTVDLV
cointegrase 181 NLLEREKHDGKAGRIAQALLRMDLVILDELGYLPFSQAGGALLFHLLSKLYEHTSVVITTT
5846 ORF 40 NLLEREKYDGKAGRIAQALLRMDLVILDELGYLPFSQSGGALLFHLLSKLYEHTSVVITTT
cointegrase 241 NLSFANWSSVFGDAKMTTALLDRLTHHCHIVETGNEYSYRLQHSLLAAQAKIKSRERKRKG
5846 ORF 100 NLSFSNWSVFGDAKMTTALLDRLTHHCHIVETGNEYSYRLQHSLLAAQTKIKTRERKRKD
cointegrase 301 GQEPEDDEPF
5846 ORF 160 GNDIEDDEPF

```

B.

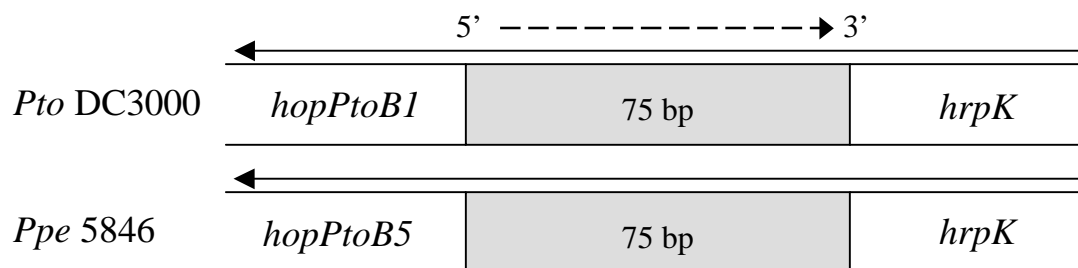
```

transposase 1 MISMEMMGKIRRMVFRDKLSLHEIAKRTGLARNTIRKWRRAPEAKQPVYQRRALFNKLSF
5846 ORF 1 MISMEMIGKIRRMVFRDKLSLHQIAKRTGLSRNTIRKWRRAPEANQPAYQRCAEFNKLNIP
transposase 61 FHVITPEQALKADSLRPKQORRSKAKALLAQIKADGYDGGYSQLTAFTRAWRGGQKASQAF
5846 ORF 61 FHEITLEQALKADSLRPHKHNRRSAKALFEQIKAFGYDGGYSQLTAFVRSWRCEQKSVRAF
transposase 121 VPLTFALGEAFQFDWSEGLLVGGIYRRMQVAHLKLCASRAFVLVAYPSQGHEMLFDAHT
5846 ORF 121 VPLTFALGEAFQFDWSEGLLVGGLFRRIOVSHMKLCASRAFVLVAYPSQGHEMLFDAHT
transposase 181 RSFGALGGVPRRGIYDNMKTAVDKVNKGKGRATVNRFAVMCAHYLFDPDFCNVAAGWEKG
5846 ORF 181 RSFGALGGVPRRGIYDNMKTAVDKVNKGKGRATVNRFSVMCAHYLFDPDFCNVASGWEKG
transposase 241 IVEKNVQDSRRRIWLDAQDCQFHSFEELNWLGCPLRALERADAPSIQRAECGRSA----
5846 ORF 241 IVEKNVQDSRRRIWLDAQNCMFTIFEELNVWLGQRCLTWAELVHPQYNCLTVAEVLELE
transposase -----
5846 ORF 301 QAEMMPMPTAFDGYVERTVRSSTCLISVARNRYSVPCERVQWVSSRLYPSRIVVIAD
transposase -----
5846 ORF 361 TVIASHERLFRDQVGFWDQHYIPLIERKPGALRNGAPFADLPKPLQLLKRGLRRHTNGD
transposase -----
5846 ORF 421 RIMMQVLAAVPIAGLEPVLVAVELVLESGSLADHILNVVARLTSTAPPCVETSLQLKV
transposase -----
5846 ORF 481 APVANTARYDRLRTTDEENRNA

```

Figure 24. Transposable elements in the *Ppe* 5846 EEL. (A) Sequence alignment of a cointegrase from *Pseudomonas stutzeri* and a *Ppe* 5846 gene product. The cointegrase (locus BAC55321, accession BAC55321) and the deduced *Psy* 5846 gene product were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background. This *Ppe* 5846 ORF was not named. (B) Sequence alignment of a putative transposase and a *Ppe* 5846 gene product. The putative transposase (accession BAC5522) from *P. stutzeri* and the *Ppe* 5846 gene product were aligned as above. This *Ppe* 5846 ORF was not named.

A.



B.

```

DC3000   1  GATTGAATCTCCGCGTACGAAAAATAGTGCCGAGCCCGGGCGTGACGCTGCCCGGGGCCCC
5846     1  GATTGAATCTCCGCGTACGAAAAATAGTGCCGAGCCCGGGCGTGACGCTGCCCGGGGACCC

DC3000   61  GACATTTTCAGTCAA
5846     61  GACGTTTCAGTCAA

```

Figure 25. Sequence alignment of the intergenic region between *hopPtoB1* and *hrpK* in *Pto* DC3000 and the intergenic region between *hopPtoB5* and *hrpK* in *Ppe* 5846. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.



B.

```

DC3000   1  ATTGAGAAAAGACCTTCAAATTCAAGGTCTTTTTTTTCGTCTGGTGGAAAGTGGTCTGAC
5846     1  ATTGAAAAGACCTTCAAATTCAAGGTCTTTTTTTTCGTCTGGGGAAAGTGGTCTGAC

DC3000   61  TGAGGCTGCGATCTACCCACCTGCCCGGAATGGCCGCGGAGCGCCAGGACTGCC TTC
5846     61  TGAGACTGCGATCTGCCACCTGCCCGGCACCTGGCCGCGGAGCGCCAAAGACTGCT TTC

DC3000   121 CAGCGCAGAGCGTCGGTACCCGGATCACACGACCAAGGATAACGCT
5846     121 CAGCGCAGAGCGTCGGTACCCGGATCACACACCAAGGATAACGCT

```

Figure 26. Sequence alignment of the intergenic region between the tRNA_{leu} gene and *eelG1* in *Pto* DC3000 and the intergenic region between the tRNA_{leu} gene and *eelG2* in *Ppe* 5846. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.

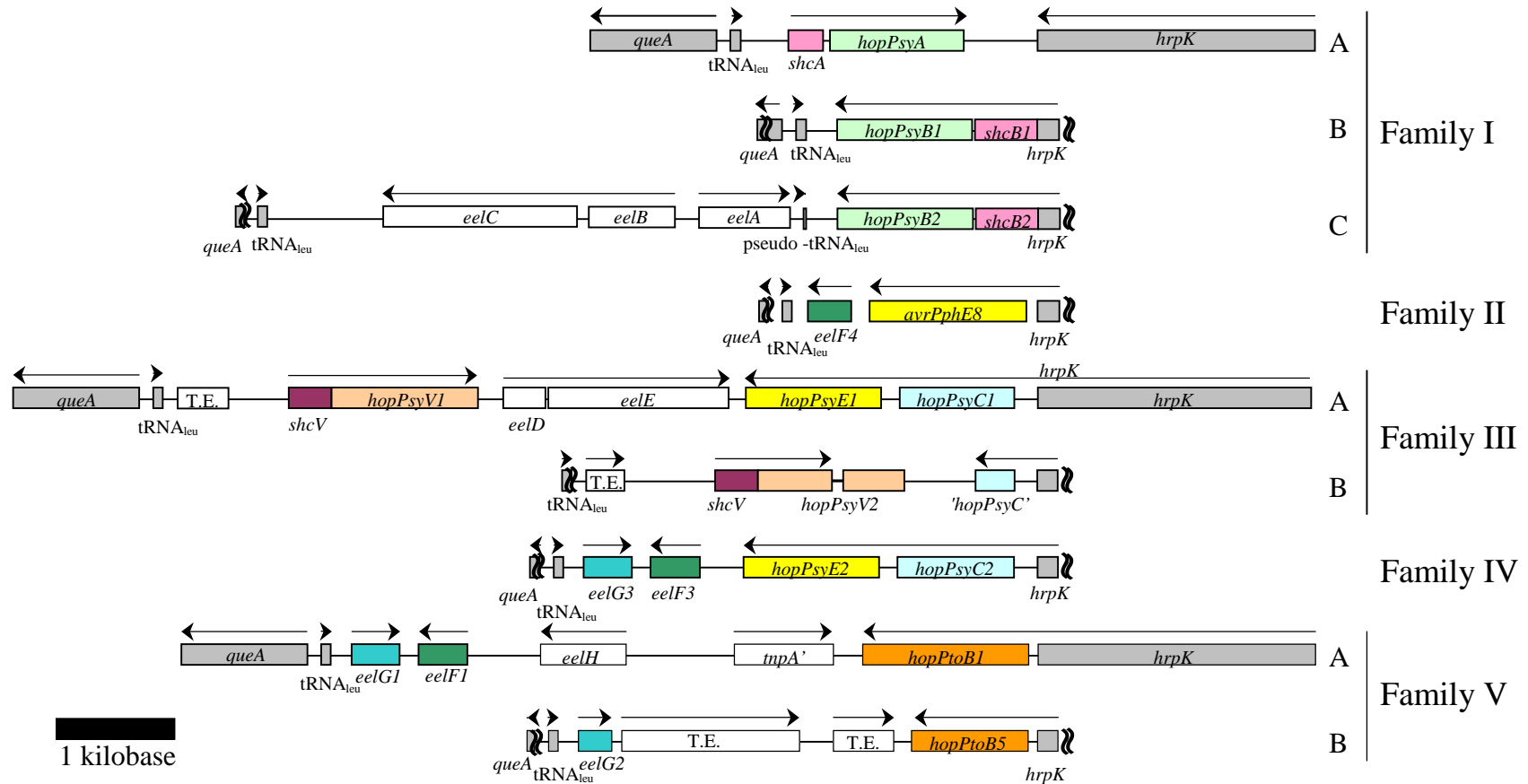


Figure 27. The EELs of Families I, II, III, IV, and V. The EELs are aligned by their *hrpK* sequences. *hrpK* and *queA* were not fully sequenced for the representatives of Families IB, IC, II, IIIB, IV, or V. Arrows indicate the direction of transcription. Conserved regions bordering the EELs are gray. Shared colors denote similar genes. Open boxes identify genes dissimilar from other genes in the EELs of Families I, II, III, IV, or V. Open reading frames that have similarity to transposable elements are labeled "T.E." Family representatives: IA, *Psy* 61; IB, *Psy* B5; IC, *Psy* 5D4198; II, *Pph* BK378; IIIA, *Psy* B728a; IIIB, *Psy* B452; IV, *Psy* W4N15; VA, *Pto* DC3000; VB, *Pph* 5846.

Family VI EEL

A distinct EEL was identified in *Psy* DH015. This EEL was less than 50% identical to the *Pto* DC3000 EEL across the 400 bp region downstream of *hrpK*, and was 2.5 kb smaller than the Family VA EEL. After completely sequencing the *Psy* DH015 EEL, it was found that the product of the first ORF downstream of *hrpK*, designated HopPsyG, was of similar size but only 50% identical to *Pto* DC3000 HopPtoB1 (Figure 28). Downstream of *hopPsyG* were alleles of *eelF* and *eelG*, which were designated *eelF5* and *eelF5*, respectively (Figure 29). Additionally, there was a short, likely not expressed, ORF between *hopPsyG* and *eelF* that has sequence similarity to a transposase from *Pto* DC3000 (Figure 30).

The intergenic region between *hrpK* and *hopPsyG* in *Psy* DH015 was dissimilar from the intergenic region between *hrpK* and *hopPtoB1* in *Pto* DC3000 and the intergenic region between *hrpK* and *hopPtoB5* in *Ppe* 5846 (Figure 31). Additionally, the intergenic region separating *hopPsyG* and *eelF5* was dissimilar from the intergenic region separating *hopPtoB1* and *mpA'* in *Pto* DC3000 (Figure 32). Thus, the *Psy* DH015 EEL was likely not a deletion derivative of a Family V EEL, and accordingly was classified into Family VI (representative strain: *Psy* DH015) (Figure 33). The Family VI EEL can be distinguished from the Family V EEL by the divergent *hopPsyG* and its flanking intergenic regions.

```

HopPtoB1      1  MRPVGGPAPGYYPPTYEAERPTAQAGNDRARSSQASSSPAASVAPETPMLGDLKRFPAG
HopPtoB5      1  MRPVGGPAPGYYPPTYEAERPTAQAGNDRARSSQASSSPAASVAPETPMLGDLKRFPAG
HopPsyG       1  MRPVGGPAPSYYPPEIERSVARSAPKELSSQASSSGTSSSESADTPKLGKIRHYAAG

HopPtoB1     61  RYPDMKVENIRLKI EGQEPG-GKDGVKHTRRRKPDAAAGSSHVHGGQSVASTSASAQSKAL
HopPtoB5     61  RYPDMKVENIRLKI EGQEPG-GKDGVKHTRRRKPDAAAGSSHVHGGQSVASTSASSQSKAL
HopPsyG      61  RYAHLAEAIVNINVSCH EQAEDK DAGRERKSERKSRAGNSTAVRCKQNA GSSSRSS----

HopPtoB1    120  QDTNFKASDLAELARWCESPHYALAPSKAAGKSSQLSANVVSILLQEGKHALEQRLEAQ
HopPtoB5    120  QDTNFKASDLAELARWCESPHYALAPSKAAGKSSQLSANVVSILLQEGKHALEQRLEAQ
HopPsyG     117  -DSRFKASSLAELASWSENTHSYSMAP---ASMFKGPVSAIDRFVSDAKTLEGLSEV

HopPtoB1    180  GLKLADV VVSEGRDHLHINLNYLEMDSCLGTSKGLWAPDSND---KKLIAKAARYFDDFN
HopPtoB5    180  GLKLADV VVSEGRDHLHINLNYLEMDSCLGTSKGLWAPDSND---KKLIARAARYFDDFN
HopPsyG     173  GVKLDDVIVS SRQTIINVNLNYTEMNEYLGRDQNLWVQPDRIINAGEKLRAKATLYFSDFS

HopPtoB1    237  AQKLPELAPLTKMKSKDSLGV MRELLRDAPGLVIGE GHNSTSSKRELINNMKSLKASGVT
HopPtoB5    237  AQKLPELAPLTKMKSKDSLGV MRELLRDAPGLVIGE GHNSTSSKRELINNMKSLKASGVT
HopPsyG     233  AQDKRPLCQLSKLKSKESELGV MRELLSGSPGLVIGE AHSSVASKRELIKNMKSLKADGVT

HopPtoB1    297  TLFMEHLCAESHDKALNNYLSAPKGSMPARLKNYLDLQSQGHQAPEELHITKYNFTTLVE
HopPtoB5    297  TLFMEHLCAESHDKV LNNYLSAPKGSPIVQAD-----
HopPsyG     293  TLFMEHLCADSHGKALDDYLKAPKGSMPARLKA YLDMQTKGNLCIGKVAEYNFTTLIR

HopPtoB1    357  AAKHAGLRVVS LDTTSTYMAPEKAEIKRAQAMNY YAAEKIRLSKPECKWVAFVGATHATS
HopPtoB5    -----
HopPsyG     353  AAKDAGLHVVP LDTAKTYETSLEDAETRYKVMNY YAAEKIRLDQPA CKWVAFVGS GHAAT

HopPtoB1    417  CDGVPGLAELHGVRSLVIDDLGLKSRATVDINVK NYGKLNPDVRLSYKV
HopPtoB5    -----
HopPsyG     413  CDGVPGLAELHGVRSLIIDDFGTKSRPDININAK KYADKINPDVRLSYKV

```

Figure 28. Sequence alignment of HopPtoB1, HopPtoB5, and HopPsyG. HopPtoB1 from the *Pto* DC3000 EEL, HopPtoB5 from the *Ppe* 5846 EEL, and the deduced HopPsyG gene product from the *Psy* DH015 EEL were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background.

A.

```

EelF1      1  MRAYKNLTAKIGGFLLALTIIGTSLEPAFAVNDCDLDNDNSTGATCGGNDKDLNDNDVTDA
EelF3      1  MRAYKTLTAKIGGLLLALTIIVGTSLEPYAVNDCMDNDNSTDATCGGNDKDLNDNDVTDA
EelF4      1  MRVYNALTTAKIGGLLLVLTVMVGTSLEPAFAVNDCMDNDNSTDARCGGNDKDLNDNDVTDA
EelF5      1  MRAYKTLTAKIGGLLLALSIVGTSLEPYAVNDCMDNDNSTDATCGGNDKDLNDNDVTDA

EelF1      61  AFGGNDKMDNDHHTDAAFGGNDKDLNDNHHTDAAFGGNDKDLNDNKTDAAFGGNDRDL
EelF3      61  AFGGNDKMDNDNHHTDAAFGGTDKDLNDNHHTDASFGGNDKMDNDHHTDAAFGGNDRDL
EelF4      61  AFGGNDKMDNDHHTDAAFGGTDKDLNDNHHTDAAFGC-----TDRDL
EelF5      61  AFGGNDKMDNDHHTDAAFGGNDKDLNDNHHTDAAFGGNDKMDNDHHTDAAFGGNDRDL

EelF1      121 DNDNNTDNYNGTPSAAAKK-
EelF3      121 DNDNNTDKYDGGASSAAAKK
EelF4      104 DNDNNTDKYHGSVPSAAAKK
EelF5      121 DNDNNTDKYDGGASSAAAKK

```

B.

```

EelG1      1  MNKIVYVKAYFKPIGEEVSVKVPTGEIKKGFFGDKETMKKETQWQQTGWSDCQIDGERLS
EelG3      1  MNKIVYVKAYFKPVGEEVTVKVPTGEIKQGGFFGEKEVMRKETRWQQTGWSDCQIDGERLS
EelG5      1  MNKIVYVKAYFKPVGEEVTVKVPTGEIKQGGFFGEKEVMRKETRWQQTGWSDCQIDGERLS

EelG1      61  KDVEDAVAQLNADGYEIQTVLPILSGAYDYALKYRYEIRHNRTELSFGDQSYVFGYGYSF
EelG3      61  KDVEDAVAQLNADGYEIQTVLPILSGAYDYALKYRYEMRHDRTDLSRDLSYVFGYGYSF
EelG5      61  KDVEDAVARLNADGYEIQTVLPILSGAYDYALKYRYEMRHDRTDLSRDLSYVFGYGYSF

EelG1      121 TEGVTLVAKKFQSSAS
EelG3      121 TEGVTLVAKKFQSSAR
EelG5      121 TEGVTLVAKKFQSSAS

```

Figure 29. Sequence alignments with EelF and EelG orthologs from the *Psy* DH015 EEL. (A) Sequence alignment of EelF1, EelF3, EelF4, and EelF5. EelF1 from the *Pto* DC3000 EEL, EelF3 from the *Pph* W4N15 EEL, EelF4 from the *Pph* BK378 EEL, and the deduced EelF5 gene product from the *Psy* DH015 EEL were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background. (B) Sequence alignment of EelG1, EelG2, and EelG5. EelG1 from the *Pto* DC3000 EEL, EelG3 from the *Psy* W4N15 EEL, and the deduced EelG5 gene product from the *Psy* DH015 EEL were aligned as in (A).

```

ISPsy5      1 MKSMPDNLPPDDLQLLKQMLAKMQSRVGFLEENALLRQRLFGRKSEQTADPATPQLALFN
DH015 ORF   1 -----

ISPsy5      61 EAESVVEAIDENAE EEVVTPAKRRRGRKRPADLPRIEVIHELPEHELTCVCGCRKHAIG
DH015 ORF   1 -----

ISPsy5     121 EEVSEQLEIVPMQIRVIKHKVRYGRRDCETAPVTADKPAQLIEKSMASPSVLAMLLTTK
DH015 ORF   1 -----

ISPsy5     181 YVDGLPLHRFEKVLGRHGIDI PRQTLARWVIQCGKHFQPLLNLMRDRLLESRFIHCDETR
DH015 ORF   1 -----

ISPsy5     241 VQVLKEPDREPSQSWMWVQTGGPPDRPVILFDYSTSRAQEVPMRLLDGYRGYVMTDDYA
DH015 ORF   1 -----

ISPsy5     301 GYNALGAQTGVERLGCWAHARRKFVEAQKVQPKGKTGRADIALNLINKLYGIERDLKASS
DH015 ORF   1 -----RLGRRAHACRKFVEAQKVQPKRKTGRADIALNLINKLYSIERDLNEGG

ISPsy5     361 DADRKIGRHEHSLPILLAQLKSWIEKTPQOVTAQNALGKAISYLASNWSKLERVVEEGLP
DH015 ORF   49 DEPRYEIRQKNSLPMLAQRHAWMEKTPQOGTAQNARCGARQHLSRYRDYLN-----GWP

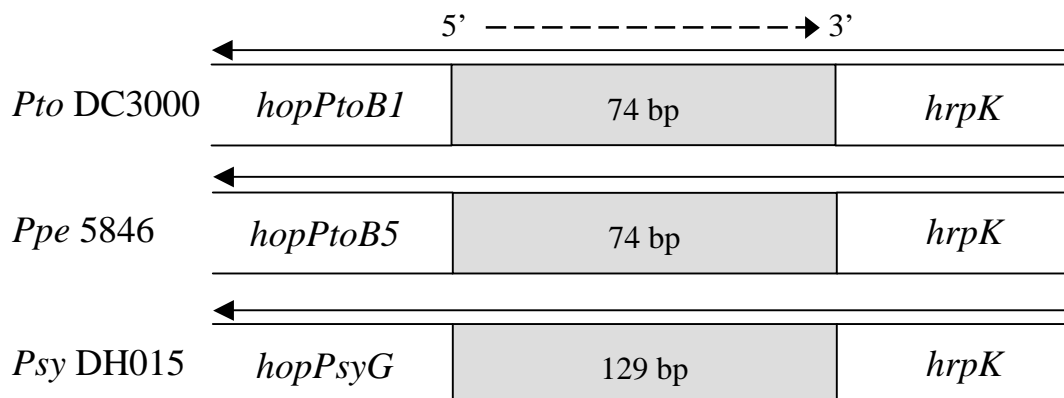
ISPsy5     421 LDNNTAERAIRPVIIGRKNWLFSDTPKGATASAQLYSLVETAKANGQEPYAWLRHALERL
DH015 ORF   103 GDSPVFGRTMIMRFAR-----

ISPsy5     481 PTATSVEDYEALLPWNCPEPLHS
DH015 ORF   -----

```

Figure 30. Sequence alignment of transposase ISPsy5 in *Pto* DC3000 with a short, likely not expressed, ORF in *Psy* DH015. ISPsy5 from *Pto* DC3000 (locus tag PSPTO3220, accession NP_793005) and the short ORF located between *hopPsyG* and *eelF5* were aligned using ClustalW version 1.8. Identical residues have a black background and similar residues have a gray background. This *Psy* DH015 ORF was not named.

A.



B.

```

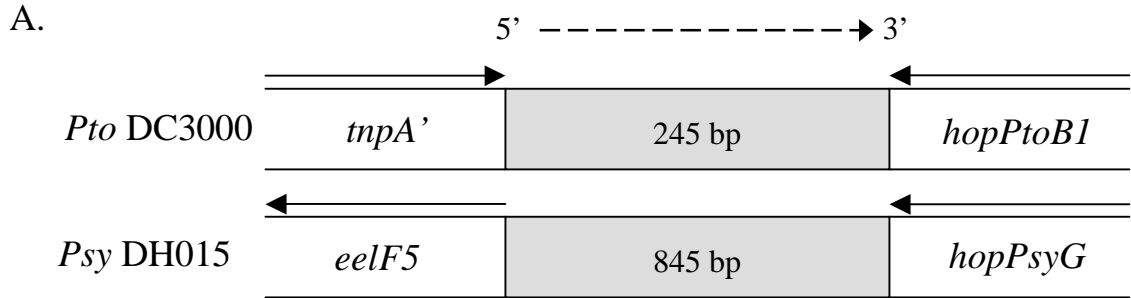
DC3000  1  GATTGAATCTCCGCGTACGAAAAATAGTGCCGAGCCCGGGCGTGACGC---TGCC--CGGG
5846    1  GATTGAATCTCCGCGTACGAAAAATAGTGCCGAGCCCAAGGCGTGACGC---TGCC--CGGG
DH015   1  GACTAAATCTCCGACT--GAACAAGACCAC-AGGACCGTCCCATCAAAGTGCATCATG

DC3000  57  CCCGACATTTTCAGTCAA-----
5846    57  CACCGACGTTTCAGTCAA-----
DH015   58  CAAAGAATATTTAGCCAAAGTCGTCCCTGCATTTTTGGCCAATAGATAGTGTCCAGACC

DC3000  -----
5846    -----
DH015   118 CGCGTGGGCCAA

```

Figure 31. Sequence alignment of the intergenic region downstream of *hrpK* in *Pto* DC3000, *Ppe* 5846, and *Psy* DH015. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.



B.

```

DC3000 1 -----TCCGGAAAA--ACGCCTTGGA---AACGTGCTGTTTCA--
DH015  1 CAATGTTTATTCCCTAAAGGTCGAGGAAAAGTACGAGGGGGATTGAAATCCGTGCTTCCTC

DC3000 34 AGGAAAATCGA-----TGACTTAAACAG-CGGAAAAACGTCTGACTATCT-GATC
DH015  61 AGTCAGACCGTCTGGCTCTACTTCATTTACCAGACGCTCCGAGCTCGCCCATCCCGAAC

DC3000 81 GG--GCC-----AGTTTTTTTCAACCTCAGCCATGAA---GGCATCAAAAATCGATG
DH015 121 GCTTGGCGAGTGCAGCTTTACAGTACCTGGCTCTATGACCTGGCTCTTGAAAAT-GAGT

DC3000 129 CTTACTTCAGACG-----TTCTTAACCTCAGTACCGAGGCCGGATAAACGA
DH015 180 CATACA-CAGACCGATGAATTTTTATTGCTTAACGTCGGAATCGATACATTATCATCGT

DC3000 176 --GTCC-----CTTTCTATGATGCTGT--TTCCAGTAAAC-----TGACAA
DH015 239 CCGTCCGAAGACCGGACTGTCCGCGGCCAGCCATTCAAAATAATCAGGATATCGTGACAA

DC3000 213 ATT---TCATGCACCTGCCGCCCGCTCTTCA--AGCGC-----
DH015 299 ATGCTGTCCGCGCCCGCAGCCGCATTTTGCACAGTGCCTTGAGGCTGTGTTTTCTCCAT

DC3000 -----
DH015 359 CCAGGCATGCCGTTGGGCCAGCATCGGCAAGCTGTTTTTCTGGCGAATTTTCATAACGCGC

DC3000 -----
DH015 419 CTCATCGCCGCCTTCGTTCAAATCCCCTCGATACTGTAGAGCTTGTGATCAAATTCAG

DC3000 -----
DH015 479 AGCAGTATCAGCCCACCGGTTTTGCGCTTGGGTTGCACTTTCTGCGCCTCAACAAATTT

DC3000 -----
DH015 539 GCGACACGCGTGGCCCTGCGGCTAAACGCTAAACACCCGCGTAGAGATAGACTTTTTT

DC3000 -----
DH015 599 CGACTTTTGGCTCGACTCACATCATGAGGTGGTGGCTCCGAAAAGAATCCGGATCACAG

DC3000 -----
DH015 659 CATCAGGCTTTAAATGTGGAATTCATGGAGCGGTTACAGAATAACGCTATGAAAAGCGCT

DC3000 -----
DH015 719 CGTTACAATCCACCTCGATCAACCCCTTTTTACGCGTGCTGCCTTTAGGATCAATTCTG

DC3000 -----
DH015 779 AGGTGAGCGATCAGGATGCGGCTTGATCACTACAACGCGCTTCTGATGAAGCGGCGC

DC3000 -----
DH015 839 TGTGCA

```

Figure 32. Sequence alignment of the intergenic region between *tnpA'* and *hopPtoB1* in *Pto* DC3000 and the intergenic region between *eelF5* and *hopPsyG* in *Psy* DH015. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.

Figure 33. The EELs of Families I, II, III, IV, V, and VI. The EELs are aligned by their *hrpK* sequences. *hrpK* and *queA* were not fully sequenced for the representatives of Families IB, IC, II, IIIB, IV, V, or VI. Arrows indicate the direction of transcription. Conserved regions bordering the EELs are gray. Shared colors denote similar genes. Open boxes identify genes dissimilar from other genes in the EELs of Families I, II, III, IV, V, or VI. Open reading frames that have similarity to transposable elements are labeled "T.E." Family representatives: IA, *Psy* 61; IB, *Psy* B5, IC, *Psy* 5D4198; II, *Pph* BK378; IIIA, *Psy* B728a; IIIB, *Psy* B452; IV, *Psy* W4N15; VA, *Pto* DC3000; VB, *Pph* 5846; VI, *Psy* DH015.

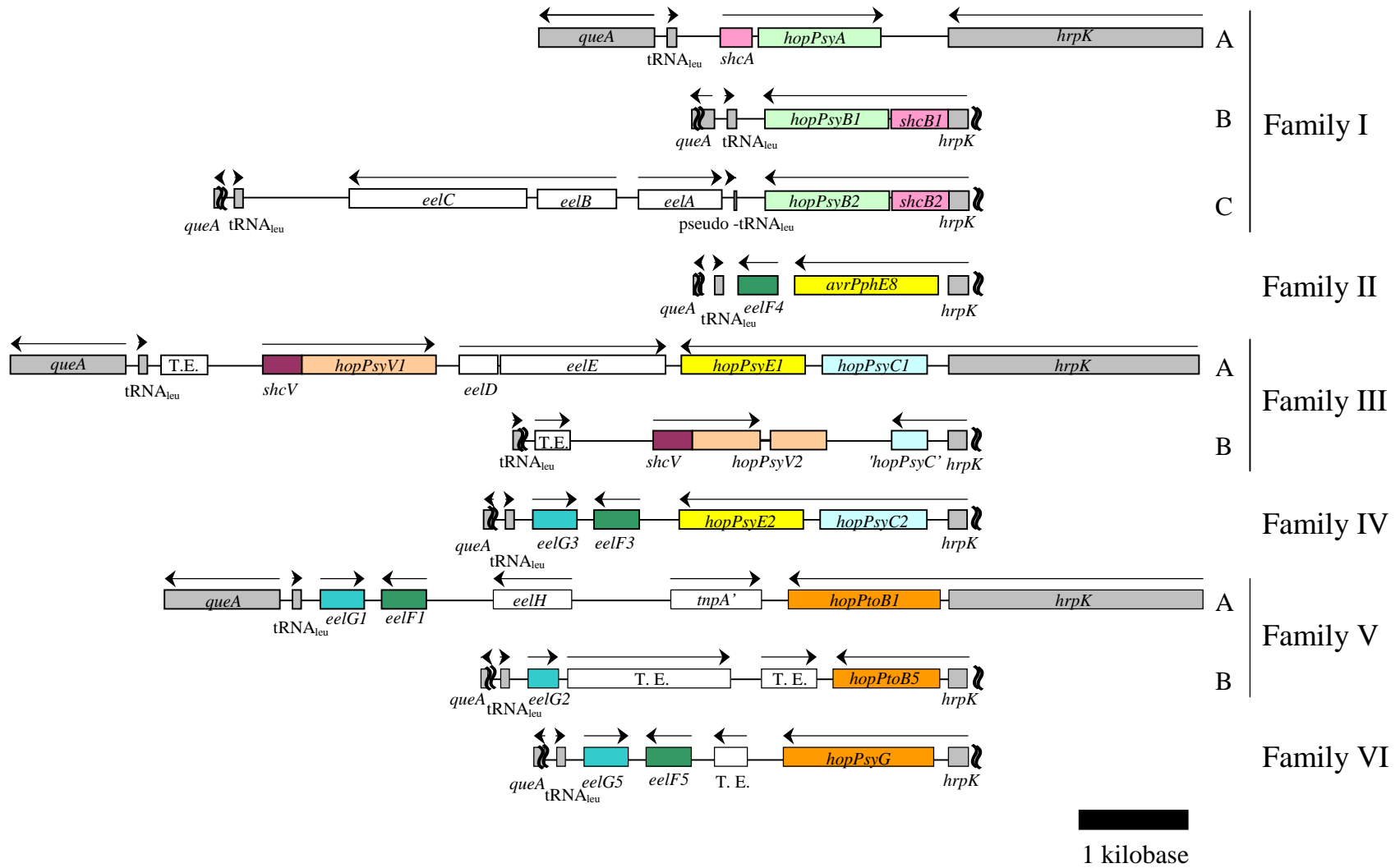


Figure 33 continued.

Identification of gene cassettes conserved among several EELs

Several sets of genes were present in multiple EEL families. A *hrpK*-*hopPsyC*-*hopPsyE* operon was in the EELs of Families IIIA and IV. Intergenic sequences surrounding *hopPsyC* and *hopPsyE* were aligned to determine if these two genes might be part of a conserved gene cassette. The intergenic region between *hrpK* and *hopPsyC* was 98% identical in *Psy* B728a and *Psy* W4N15 (Figure 34). For these two strains, the intergenic region separating *hopPsyC* and *hopPsyE* was 92% identical (Figure 35), and the 85 bp intergenic region downstream of *hopPsyE* was 98% identical (Figure 36). Thus, coding and surrounding intergenic regions of *hopPsyC* and *hopPsyE* were conserved in *Psy* B728a and *Psy* W4N15.

eelF and *eelG* were in the EELs of Families IV, VA, and VI (Figure 33). The intergenic region between *eelF* and *eelG* was greater than 90% identical in *Pto* DC3000, *Psy* W4N15 and *Psy* DH015 (Figure 37). Unexpectedly, the 300 bp region upstream of *eelF* in *Pto* DC3000 was only about 50% identical to the intergenic sequence upstream of *eelF* in *Psy* DH015 and *Psy* W4N15, which were >90% identical (Figure 38). The intergenic region between the tRNA_{leu} gene and *eelG* was more than 75% identical in *Pto* DC3000, *Ppe* 5846, *Psy* W4N15, and *Psy* DH015 (Figure 39, Figure 26). Thus, coding and surrounding intergenic regions of *eelF* and *eelG* were at least moderately conserved in *Pto* DC3000, *Psy* W4N15, and *Psy* DH015.

eelF4 in the Family II EEL appeared to have been derived from the putative *eelF*-*eelG* cassette. The 154 bp intergenic region upstream of *eelF4* was 85%

identical to the intergenic region upstream of *eelF* in *Psy* DH015 and *Psy* W4N15 (Figure 38). In the *Pph* BK378 EEL, the 129 bp region downstream of the tRNA_{leu} gene appeared to be a chimera of the intergenic region separating the tRNA_{leu} gene and *eelG* in Families IV and VI, and the intergenic region separating *eelF* and *eelG* in Families IV, VA, and VI (Figure 38 and 39). Notably, a region downstream of the tRNA_{leu} gene was conserved in most EEL families, including Family II (Figure 40), but was not similar to the intergenic region between *eelF* and *eelG*.

Additionally, although the sequenced portions of *hrpK* were $58.8 \pm 1.2\%$ guanine/cytosine (GC), and the sequenced portions of *queA* were $60.9 \pm 1.1\%$ GC, many of the genes identified in the characterized EEL were closer to 50% GC (Table 5). This suggested that the apparent cassettes (Figure 41) were acquired by horizontal transfer.

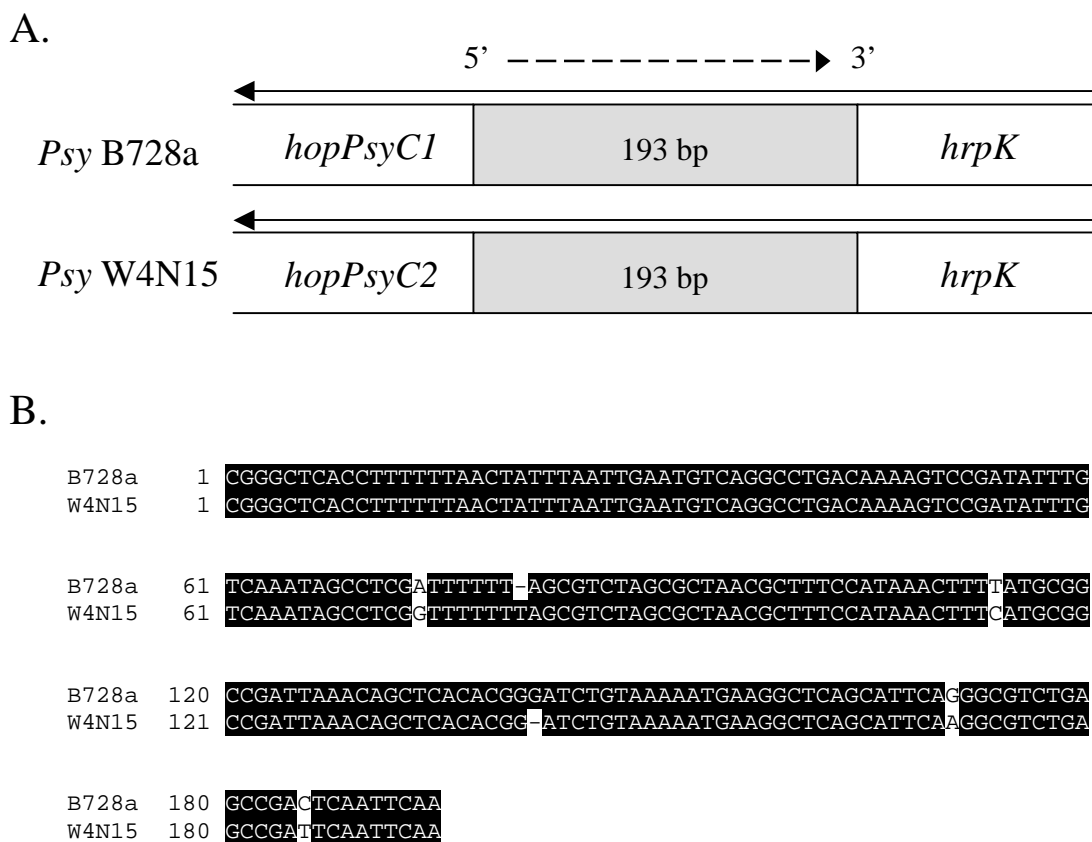


Figure 34. Sequence alignment of the intergenic region between *hopPsyC1* and *hrpK* in *Psy* B728a and the intergenic region between *hopPsyC2* and *hrpK* in *Psy* W4N15. *Psy* B728a represents the Family IIIA EEL and *Psy* W4N15 represents the Family IV EEL. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.

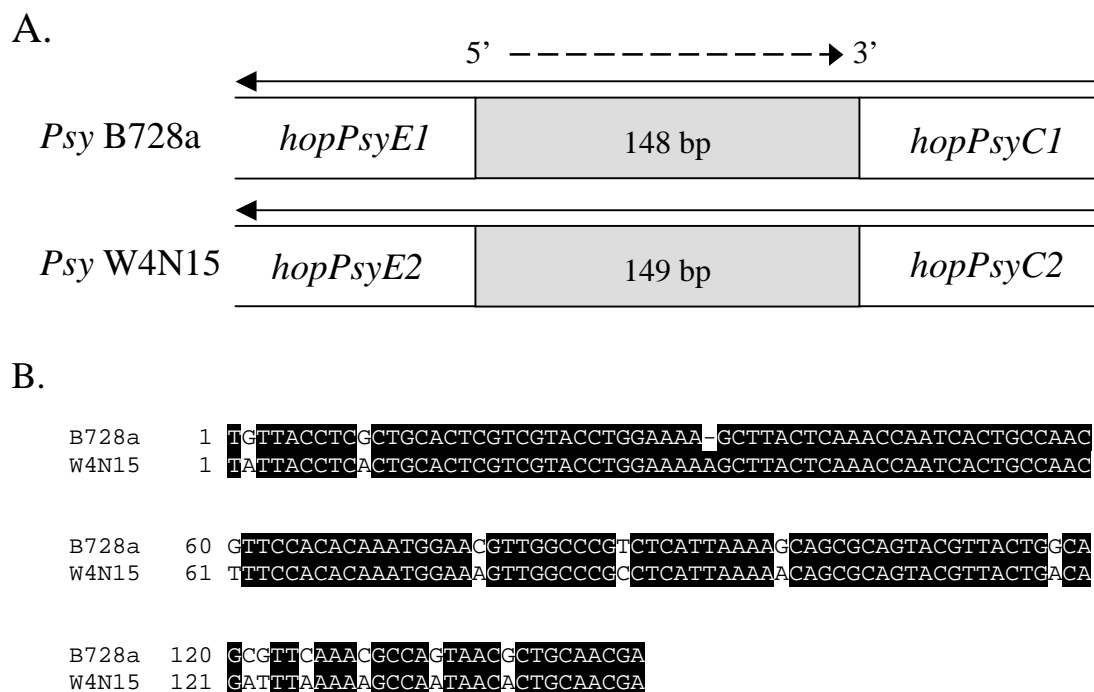
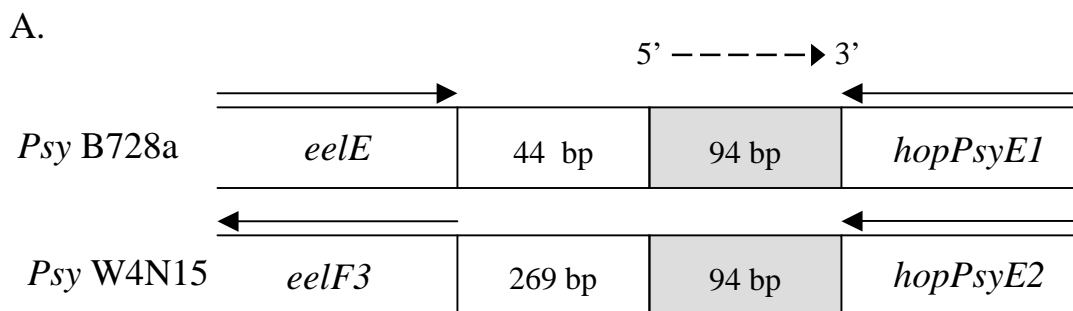


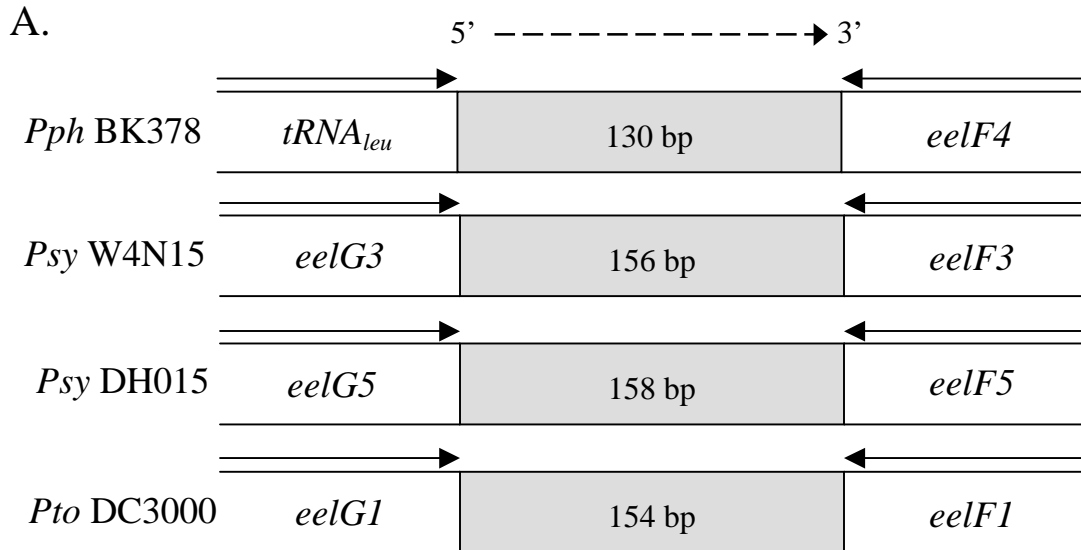
Figure 35. Sequence alignment of the intergenic region between *hopPsyE1* and *hopPsyC1* in *Psy B728a* and the intergenic region between *hopPsyE2* and *hopPsyC2* in *Psy W4N15*. *Psy B728a* represents the Family IIIA EEL and *Psy W4N15* represents the Family IV EEL. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.



B.

B728a	1	CCATTAAACAAT	TACGAAGTCATGAAAAATGCTGTCGCGTCCTGCCGCGGTACTTTCCG
W4N15	1	CCATTAAACAAT	CACGAAGTCATGAAAAATGCTGTCGCGTCCTGCCGCGGTACTTTCCG
B728a	61	TATGCCCTGAATACCACGCATACTCAACCGGTTT	
W4N15	61	TATGCCCTGAATACCACGCATACTCAACCGGTTT	

Figure 36. Sequence alignment of the intergenic region downstream of *hopPsyE* from *Psy* B728a and *Psy* W4N15. *Psy* B728a represents the Family IIIA EEL and *Psy* W4N15 represents the Family IV EEL. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The gray IGR shows the area used for the alignment. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.



B.

```

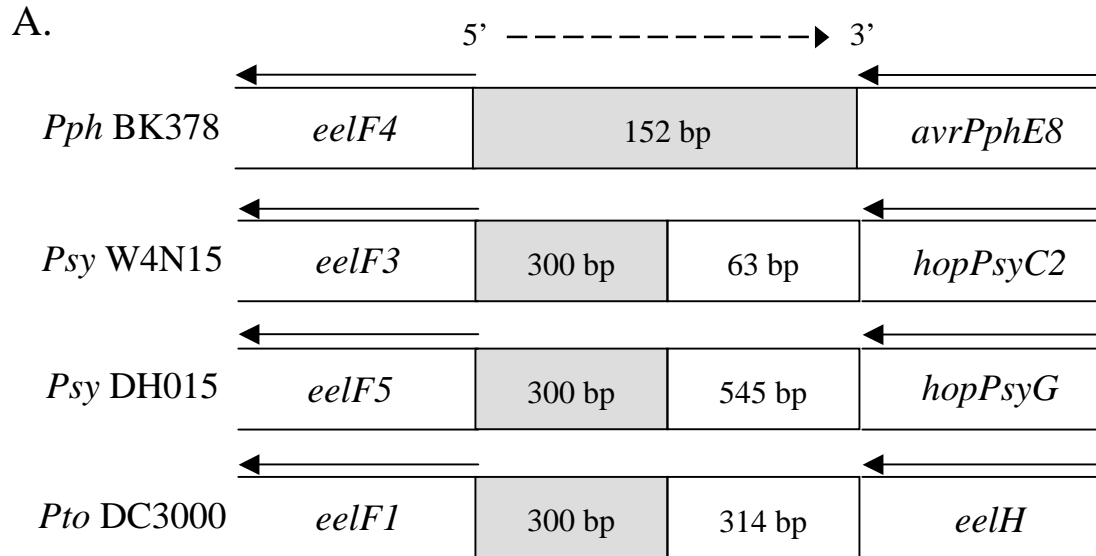
BK378      1  -----TTTAAAAAAG-----ACCTTGAAATTCAAGGTCITTTT
W4N15     1  ATCAATAATAGCGAACTCGTTTCACGGACGCCGCTCTGTCTCTGATACGAAAACGCTTG
DH015     1  ATCAATAATAGCGAACTCGTTTCACGGACGCCGCTCTGCCCTCAGATACGAAAACGCTG
DC3000    1  ----ATAATAGTGACCTCGTGGCACGGACGCCGCTCTGCCCTCTGATACGAAAACGCTT

BK378     34  TTTCGTCTGGTCAAAAGTGGTCTGACATAATCTGCATCTCCCCCTATTAACT-CAGAAT
W4N15     61  ACTCAACAAGAGGCAAGCG--CTCACGTG-CAAAGACCTGCCGTATCAGCAAGCGCAA
DH015     61  CCTCAACAAGAGGCAAGCGTCTCACGTG-CAAAGACCTGCCGTATCAGCAAGCGCAA
DC3000    57  CCTCAACAAGAGGCAAGCGTACTAACGTG-CACAAAGACCTGCCGTATCAGCAAGCGCAA

BK378     93  CAGGCTCGAAGC-ACGAAATAACAGTTTCGGCCGCGTTG
W4N15    118  GACGCTCACTTCCACGAAATAACAGATAGGTCGCGTTG
DH015    120  GACGCTCGCTTCCACGAAATAACAGATAGGTCGCGTTG
DC3000   116  GACGCTCGCTTCCACGAAATAACAGGATAGGTCGCGTTG

```

Figure 37. Sequence alignment of the intergenic region downstream of *eelF* from *Pph* BK378, *Psy* W4N15, *Psy* DH015, and *Pto* DC3000. *Pph* BK378 represents the Family II EEL, *Psy* W4N15 represents the Family IV EEL, *Psy* DH015 represents the Family VI EEL, and *Pto* DC3000 represents the Family VA EEL. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.



B.

```

BK378      1  CAAGGTTGATCCCTAAAGGTA GAGAAAAGAA -CGAGGGTGATTGAATAC -TGTGATCCG
W4N15     1  CAATGTTTATTCCTAAAGGTCGAGGAAAAGTA -CGAGGTG-ATTGAATAC -CGTGGTTCCG
DH015     1  CAATGTTTATTCCTAAAGGTCGAGGAAAAGTA -CGAGGGGGATTGAATAC -CGTGGTTCCG
DC3000    1  -CAGGTTTTCCCGGATAAGTGAAATGATGAACCAAGGGTTACTGAACACGTTTCGATCAG

BK378     59  TCAGTCAGACCGTCTGGCCATACTTCGTT -TACGAGACGCTCTGACACCGTCTAAATCCCG
W4N15     58  TCAGTCAGACCGTCTGGCTCTACTTCATT -TACCAGACGCTTCGAGCCCGTCCAATCCCG
DH015     59  TCAGTCAGACCGTCTGGCTCTACTTCATT -TACCAGACGCTCCGAGCTCCGCCAATCCCG
DC3000    60  TGACTAAAACAGTATGTAACTGCAGCCTTCTGCAAGACCGACAGAGGTCGACCAAACTGCG

BK378     118 AACGT-TTGCGCA---GTGCAGTTTACAGTACC GAACT-----
W4N15     117 AACGC-TTGCGCA---GTGCAGCTTTACAGTACCTGGCTCTATGACCTGCGTTCTTGAAA
DH015     118 AACGC-TTGCGCA---GTGCAGCTTTACAGTACCTGGCTCTATGACCTGCGCTCTTGAAA
DC3000    120 AGCCTGTTTCATACCCATCAATTTCTATAGCGACCGTTTCACACGACTCTCTTACC--GAT

BK378     -----
W4N15     173 ATTAGTCATACCCAGAGCGATGAGTTTTTTTATTGCTTAACGTGCGAATCGATACATTATC
DH015     174 ATGAGTCATACACAGAGCGATGAATTTTTTTATTGCTTAACGTGCGAATCGATACATTATC
DC3000    178 GCTGGGAGTACCAGAAAACCTTCCGCACCTGCATTTTTTGGCA---CTGTCCGATGGTTTGA

BK378     -----
W4N15     233 ATCGTCCGTCA GAAGAACGGACTTTCGCCGGGT CAGCCATTAAACAATCACGAAGTCAT
DH015     234 ATCGTCCGTCC GAAGACCGGACTTTCGCCGGGCCAGCCATTCAAATAATCACGATATCGT
DC3000    235 CCGCTTTTGGG GAGAAATTCCTCAAACGGAGAA CGATGAGTTTTTTGTTGCGTGCATGCT

BK378     -----
W4N15     293 GAAAAATG
DH015     294 GACAAAT-
DC3000    295 AATCGA--

```

Figure 38. Sequence alignment of the intergenic region upstream of *eelF* from *Pph* BK378, *Psy* W4N15, *Psy* DH015, and *Pto* DC3000. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The IGRs used in the alignment are gray. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.

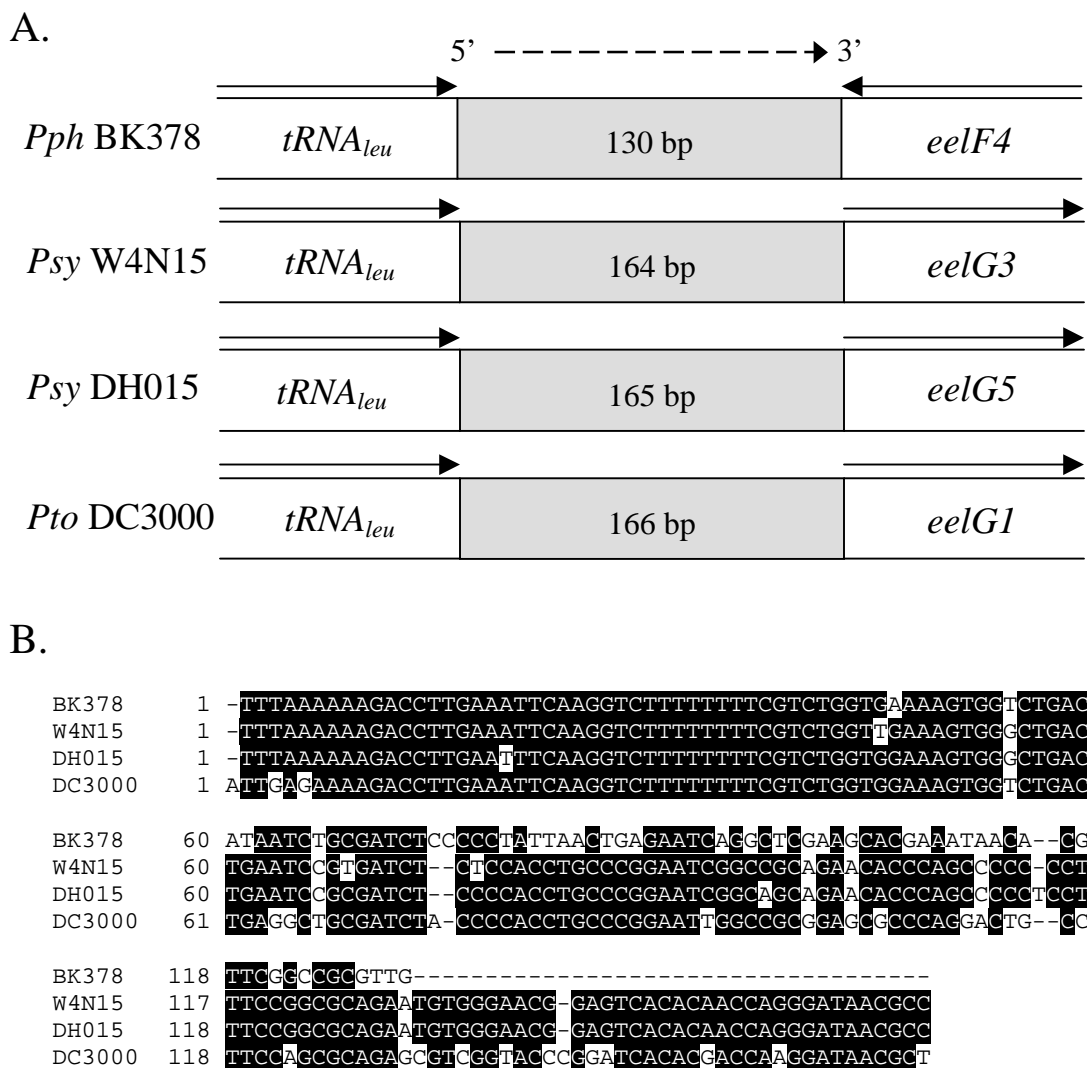


Figure 39. Sequence alignment of the intergenic region downstream of the *tRNA_{leu}* gene from *Pph* BK378, *Psy* W4N15, *Psy* DH015, and *Pto* DC3000. *Pph* BK378 represents the Family II EEL, *Psy* W4N15 represents the Family IV EEL, *Psy* DH015 represents the Family VI EEL, and *Pto* DC3000 represents the Family VA EEL. (A) Diagram showing the intergenic regions (IGRs) aligned in (B) and flanking genes. Solid arrows indicate the direction of transcription. The dashed 5' to 3' arrow over the IGRs denotes the direction of alignment. The number of base pairs (bp) in the IGRs are labeled. Note that the figure was not drawn to scale. (B) Alignment of the IGRs displayed in (A). Conserved nucleotides have a black background.

```

Psy 61      1  -TTTAAAAAAGACCTTGAAATTC AAGGTC TTTTTTTTCGTC TGGTGGAAAGTGCCTTGTT
Psy B5     1  -TTTAAAAAAGACCTTGAAATTC AAGGTC TTTTTTTTCGTC TGGTGGAAAGTGCCTTGTT
Pph BK378  1  -TTTAAAAAAGACCTTGAAATTC AAGGTC TTTTTTTTCGTC TGGTGGAAAGTGGTCTGAC
Psy W4N15  1  -TTTAAAAAAGACCTTGAAATTC AAGGTC TTTTTTTTCGTC TGGTGGAAAGTGGGCTGAC
Pto DC3000 1  ATTGAGAAAAGACCTTGAAATTC AAGGTC TTTTTTTTCGTC TGGTGGAAAGTGGTCTGAC
Ppe 5846   1  ATTGAGAAAAGACCTTGAAATTC AAGGTC TTTTTTTTCGTC TGGGGAAGTGGTCTGAC
Psy DH015  1  -TTTAAAAAAGACCTTGAAATTC AAGGTC TTTTTTTTCGTC TGGTGGAAAGTGGGCTGAC

Psy 61     60  GTCGTTTACCGTATGCATCTCCTGAAAA-AAATCTGCCTGGCTCAATCTCTAAAAACG-C
Psy B5     60  GTCGTTTACCGTATGCATCTCCTGAAAA-AAATCTGCCTGGCTCAATCTCTAAAAACG-C
Pph BK378  60  ATAATCTGCGATCTCCCCCTATTAAGTGAAGATCAGGC--TCGAGCACGAAATAACA--
Psy W4N15  60  TGAATCCGTGATCT-CTCCACCTGCCCG-GAATCGGCC--GCAGAACACCCAGCCCC-C
Pto DC3000 61  TGAGGCTGCGATCTACCCACCTGCCCG-GAATCGGCC--GGGAGCGCCAGGACTG--
Ppe 5846   61  TGAGACTGCGATCTGCCCCACCTGCCCG-GCACTGGCC--GGGAGCGCCCAAGACTG--
Psy DH015  60  TGAATCCGCGATCT-CCCCACCTGCCCG-GAATCGGCA--GCAGAACACCCAGCCCCCTC

Psy 61     118 CTTTC TCTGT CAGGGCAGGCTGCCTGAAACATGA--AACAGATT CACGGCAAGCCTGTGT
Psy B5     118 CTTTC TCTGT CAGGGCAGGCTGCCTGAAACATGA--AACAGATT CACGGCAAGCCTGTGT
Pph BK378  116 CGTTC-----GGCCG-----GTG-----
Psy W4N15  115 CTTTC-----CGGCGCAGAAATGTGGGAAC--GGAGTCACACAACCAGGGATAACGCC---
Pto DC3000 116 CTTTC-----CAGCGCAGAGCTCGGTACCCGGA-TCACACCACCAAGGATAACGCT---
Ppe 5846   116 CTTTC-----CAGCGCAGAGCTCGGTACCCGGA-TCACACAACCAAGGATAACGCT---
Psy DH015  116 CTTTC-----CGGCGCAGAAATGTGGGAAC--GGAGTCACACAACCAGGGATAACGCC---

Psy 61     176 TCGAGCAAGCGAAGGATACGCCGAGAAATGGCCCTCAGGTCCGTCTTCTGGGCCCGCG
Psy B5     176 TCGAGCAAGCGAAGGATACGCCGAGAAATGGCCCTCAGGTCCGTCTTCTGGGCCCGCG
Pph BK378  -----
Psy W4N15  -----
Pto DC3000 -----
Ppe 5846   -----
Psy DH015  -----

Psy 61     236 ACTTTGGGATATTCAGCAGTTCATAC
Psy B5     236 ACTTTGGGATATTCAGCAGTTCATAC
Pph BK378  -----
Psy W4N15  -----
Pto DC3000 -----
Ppe 5846   -----
Psy DH015  -----

```

Figure 40. Sequence alignment of the intergenic region downstream of the $tRNA_{leu}$ gene from the representative strains of Families IA, IB, II, IV, VA, VB, and VI. The intergenic regions downstream of the $tRNA_{leu}$ gene in Families IC and III had very low identity to the above sequences.

Table 5.
GC content of the genes in the EELs of the representative strains.

Family	Representative Strain	Gene	Length (bp)	% GC
IA	<i>Psy</i> 61	<i>shcA</i>	294	56.9
		<i>hopPsyA</i>	1127	52.6
IB	<i>Psy</i> B5	<i>shcB1</i>	528	51.7
		<i>hopPsyB1</i>	1149	53.2
IC	<i>Psy</i> 5D4198	<i>shcB2</i>	528	51.7
		<i>hopPsyB2</i>	1149	53.3
		<i>eelA</i>	771	49.3
		<i>eelB</i>	732	52.0
		<i>eelC</i>	1638	44.4
II	<i>Pph</i> BK378	<i>avrPphE8</i>	1323	55.2
		<i>eelF4</i>	369	50.4
IIIA	<i>Psy</i> B728a	<i>hopPsyC1</i>	972	50.5
		<i>hopPsyE1</i>	1149	57.9
		<i>eelD</i>	357	52.4
		<i>eelE</i>	1524	55.3
		<i>shcV</i>	363	51.5
		<i>hopPsyV1</i>	1236	49.0
IIIB	<i>Psy</i> B452	<i>shcV</i>	363	51.5
		<i>hopPsyV2</i>	623	49.2
IV	<i>Psy</i> W4N15	<i>hopPsyC2</i>	984	52.9
		<i>hopPsyE2</i>	1149	57.7
		<i>eelF3</i>	420	49.3
		<i>eelG3</i>	411	47.2
VA	<i>Pto</i> DC3000	<i>hopPtoB1</i>	1401	55.4
		<i>eelF1</i>	417	53.5
		<i>eelG1</i>	411	47.0
		<i>eelH</i>	726	50.5
VB	<i>Ppe</i> 5846	<i>hopPtoB5</i>	987	55.8
		<i>eelG2</i>	264	46.8
VI	<i>Psy</i> DH015	<i>hopPsyG</i>	1389	51.2
		<i>eelF5</i>	420	52.1
		<i>eelG5</i>	411	47.4

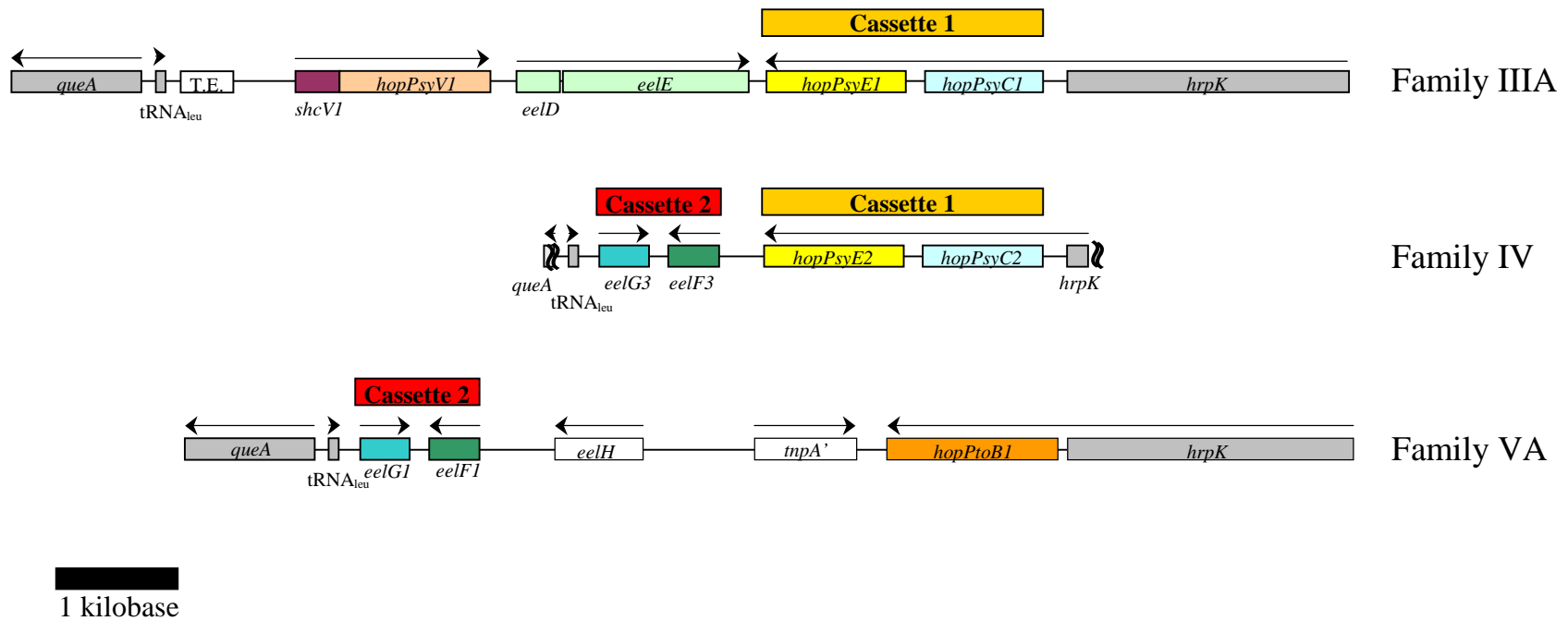


Figure 41. Two putative gene cassettes identified in the EELs of Families IIIA, IV, and VA. The EELs are aligned by their *hrpK* sequences. *hrpK* and *queA* were not fully sequenced for the representative of Family IV. Arrows indicate the direction of transcription. Conserved regions bordering the EELs are gray. Shared colors denote similar genes. Open boxes identify genes dissimilar from other genes in the EELs of Families I, II, III, IV, V, or VI. An open reading frame that has similarity to transposable elements is labeled "T.E." Family representatives: IIIA, *Psy* B728a; IV, *Psy* W4N15; VA, *Pto* DC3000.

Linkage between EEL, plant source, and geographical source

Strains that carried a Family IIIA EEL tended to be bean isolates from North America. Both strains that carried a Family IV EEL were isolated from Roseaceae plants. However, among the nine strains with a Family IA EEL were isolates from tomato, soybean, citrus, and *Prunus* spp. (Table 4). Moreover, strains isolated from tomato and *Prunus* spp. sometimes had a Family VA EEL, and strains isolated from citrus plants sometimes had a Family IIIB EEL. Thus, a strain's EEL does not appear to be linked to its plant source. This is consistent with current models predicting that multiple effectors control a strain's host range.

A possible correlation was observed between a strain's EEL and its geographical source. Strains that had a Family V EEL were either isolated or derived from strains isolated in the United Kingdom, France, or North Africa. Strains with other types of EELs were predominantly North American isolates. Since most strains surveyed in this study had a North American origin, additional studies would be necessary to confirm any relationship between a strain's EEL and its geographical source.

hopPsyB encodes a translocated effector

Each EEL in Families II, III, IV, and V, with the exception of the Family VB EEL, appeared to express at least one functional effector. HopPsyB in Family IB and IC, however, was just 31% identical to its only known ortholog, the effector HopPsyA. Its relatively high similarity to HopPsyA over the amino-terminal type III secretion domain, association with a ShcA-like chaperone, and probable expression from a

HrpL-dependent promoter upstream of *hrpK* suggested that *hopPsyB* might encode a translocated effector. A vector, named pSHB5EEL1-600, was obtained that expressed the *Psy B5 shcB1-hopPsyB1* operon from the P_{lacUV5} promoter in pDSK600 (Charity et al., 2003). When this construct was transformed into *Escherichia coli* MC4100 (pHIR11-2070)(pYXL2B), a strain that has been used previously to test for *hrp* TTSS-dependent effector activity, transformants exhibited an HR⁺ phenotype in tobacco leaves, consistent with effector activity (Figure 42). However, the genetics of plant-pathogen interactions between tobacco and *P. syringae* are not well established, which prevented further studies with tobacco to verify that HopPsyB1 is a TTSS-dependent effector.

Arabidopsis thaliana was chosen to confirm that HopPsyB1 was translocated into host cells by the *hrp* TTSS. A fusion between *hopPsyB1* and *avrRpt2* was constructed in pDSK519 enabling expression of the amino-terminal 78 residues of HopPsyB1 and the carboxy-terminal 175 residues of AvrRpt2. This construct, pJCB5EEL2-AR2, was transformed into *Pto* DC3000 to screen for HR activity in *RPS2* and *rps2* lines of *A. thaliana*. *Pto* DC3000 (pJCB5EEL2-AR2) elicited the HR in *A. thaliana* ecotype Columbia (*RPS2*) but not in the *rps2* derivative (Figure 43). These observations are consistent with translocation of HopPsyB into plant cells and indicate that HopPsyB, like HopPsyA, is a TTSS-dependent effector.

Additionally, *Pto* DC3000 transformants carrying pSHB5EEL1-600 exhibited moderated replication in *A. thaliana* (Figure 44) and slightly enhanced virulence in some *Brassica oleracea* varieties (Table 6). In contrast, growth of *Pto* DC3000 (pYXIL) expressing HopPsyA was suppressed in these plants (Table 6). *E. coli*

MC4100 expressing the *hrp* TTSS and HopPsyA or HopPsyB did not elicit the HR in *A. thaliana* or any of the *B. oleracea* lines (Table 6).

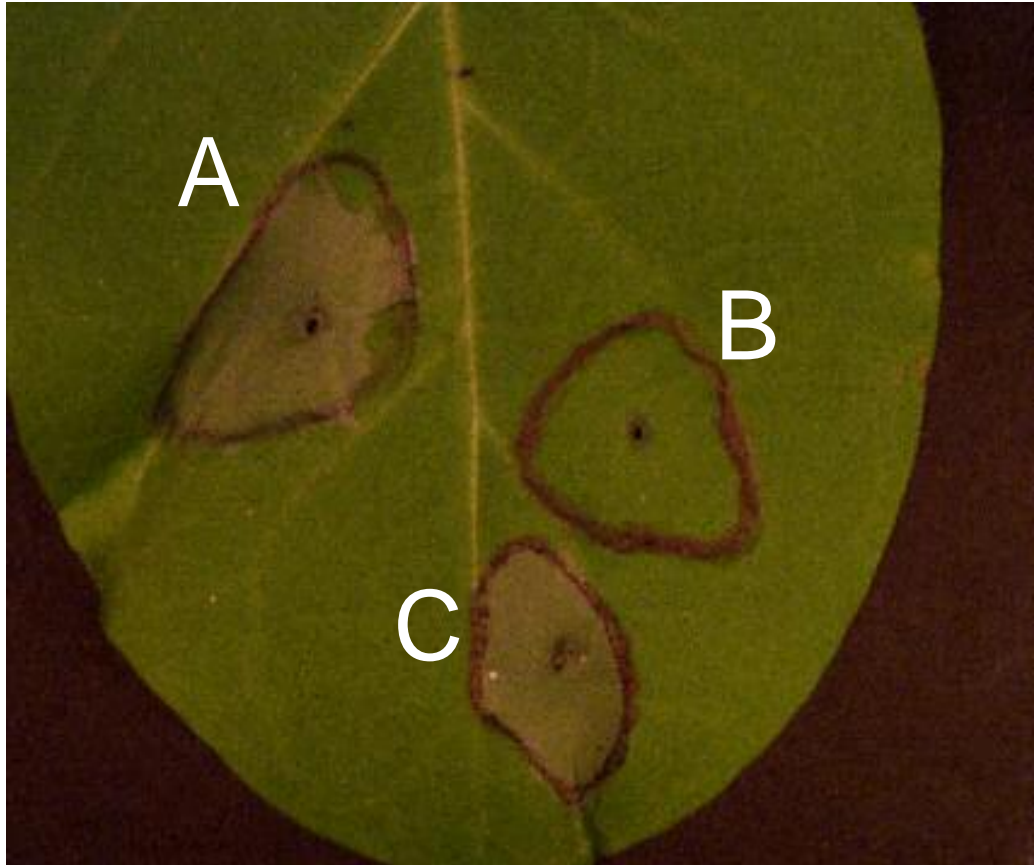


Figure 42. Activity of HopPsyB and HopPsyA in tobacco. Tobacco leaves were syringe-infiltrated with 10^9 cfu/mL of bacteria and scored for the hypersensitive response (HR) 24 hours later. A black pen was used to outline the inoculated sections of the leaf. (A) MC4100 (pHIR11-2070)(pYXL2B) expressing the *Psy* 61 *shcA-hopPsyA* operon produced the HR, (B) MC4100 (pHIR11-2070)(pYXL2B) containing the empty vector pDSK600 produced a null phenotype, (C) MC4100 (pHIR11-2070)(pYXL2B) expressing the *Psy* B5 *shcB1-hopPsyB1* operon produced the HR.

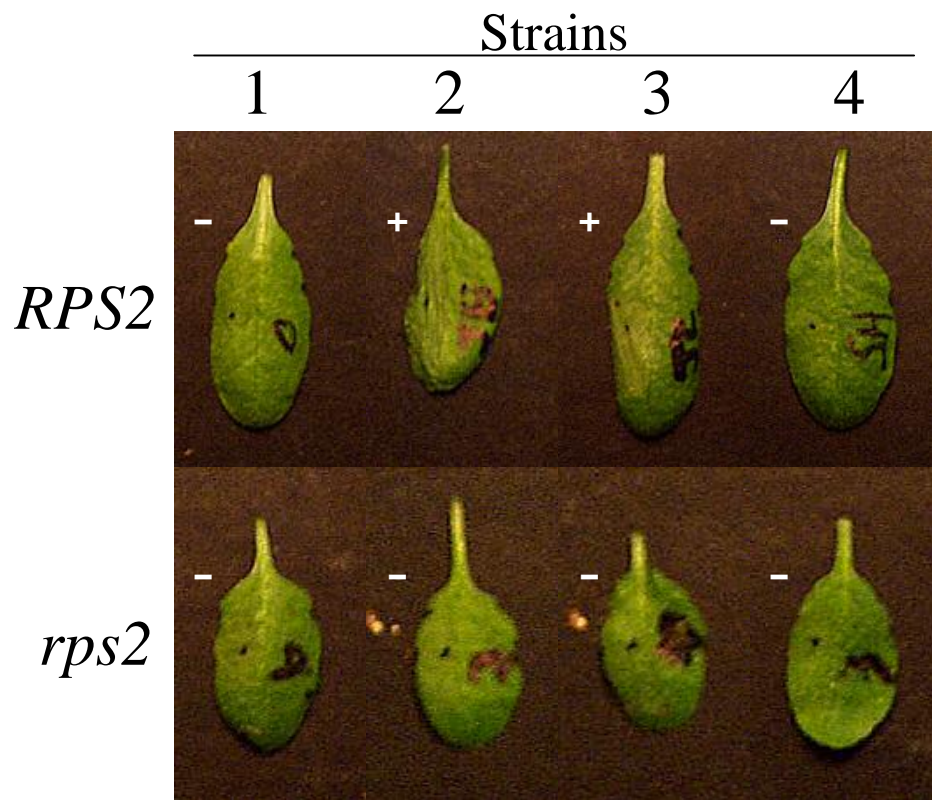


Figure 43. Evidence for type III secretion-dependent translocation of HopPsyB. *A. thaliana* leaves were syringe-infiltrated with 10^8 cfu/mL of bacteria and scored for the hypersensitive response (HR) 18 hours later. Panels: *RPS2*, *A. thaliana* ecotype Columbia (*RPS2*); *rps2*, *A. thaliana* ecotype Columbia (*rps2*); 1, *Pto* DC3000 carrying the empty vector pDSK519; 2, *Pto* DC3000 carrying pJBavrRpt2-600, which allows for expression of full-length *avrRpt2*; 3, *Pto* DC3000 carrying pJCB5EEL2-AR2, which allows for expression of a fusion of the 5'-terminus of *hopPsyB* with the 3'-terminus of *avrRpt2*; 4, *Pto* DC3000 carrying pSHB5EEL1-600, which allows for expression of full-length *hopPsyB* and *shcB*. The leaves in panels 2A and 3A display the HR, and are marked with a "+". All other leaves show a null response, and are marked with a "-".

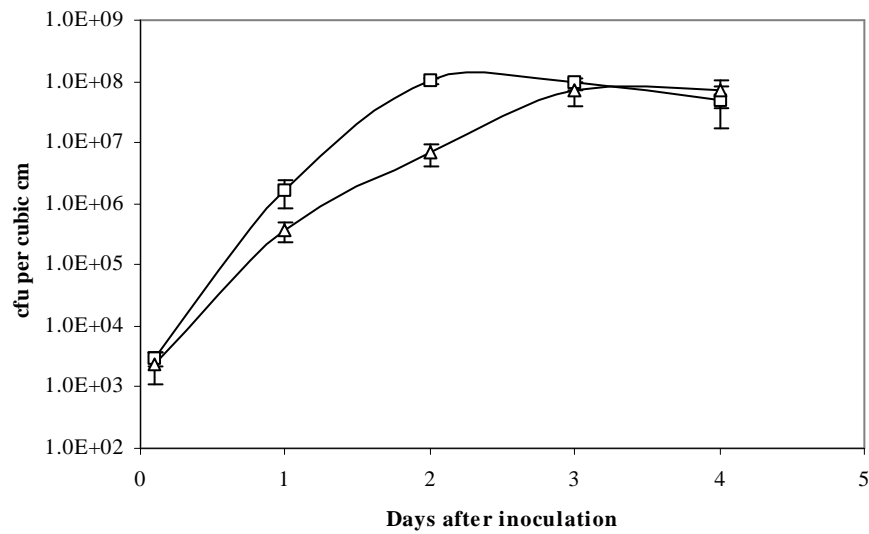
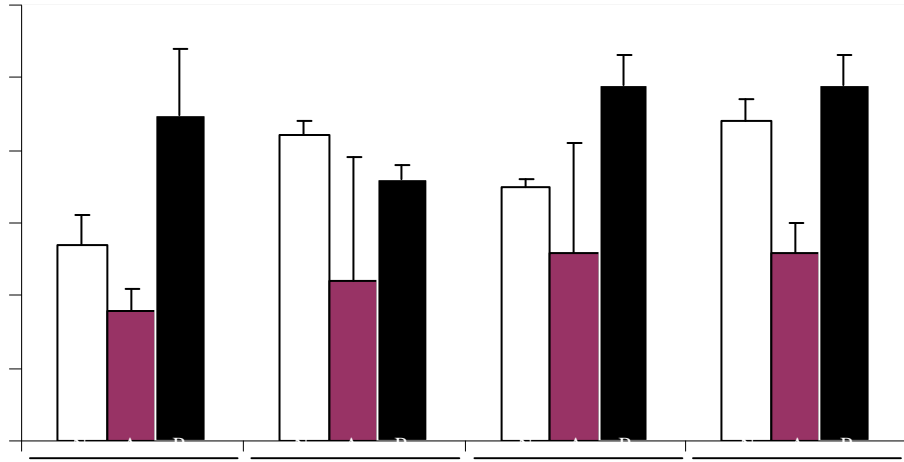


Figure 44. Multiplication of *Pto* DC3000 expressing and not expressing the *Psy* B5 *shcB-hopPsyB* operon in *A. thaliana* leaves. *Pto* DC3000 expressing the *hopPsyB-shcB* operon from the P_{lacUVA} promoter in pDSK600 (triangle) appeared to replicate slower than *Pto* DC3000 containing the empty vector pDSK600 (square). Data were generated by three trials and are shown as mean \pm standard deviation.



DISCUSSION

Before this study, the EEL could reasonably be characterized as a highly variable site for genes encoding effectors (Alfano et al., 2000; Hutcheson, 1999). However, the results presented herein indicate this locus might not be as heterogeneous as initially expected. Of the 33 strains whose EELs were amplified, 8 have an EEL equivalent to the *Psy* 61 EEL, 2 have an EEL equivalent to the *Pto* DC3000 EEL, and 2 have an EEL equivalent to the *Psy* B728a EEL. Those strains that do not have an EEL equivalent to the EEL of *Psy* 61, *Pto* DC3000, or *Psy* B728a have an EEL that shares components with these EELs. For instance, the Family IB EEL appears to be an inversion of the Family IA EEL; the Family II EEL contains an *avrPphE* allele similar to *hopPsyE1* in the *Psy* B728a EEL, and an *eelF* allele similar to *eelF1* in the *Pto* DC3000 EEL; the Family IIIB EEL appears to be a deletion derivative of the *Psy* B728a EEL; the Family IV EEL contains the *hopPsyC-hopPsyE* operon found in the *Psy* B728a EEL, and *eelF* and *eelG* alleles similar to *eelF1* and *eelG1* in the *Pto* DC3000 EEL; the Family VB EEL appears to be a deletion derivative of the *Pto* DC3000 EEL; the Family VI EEL contains a gene with low similarity to *hopPtoB1* in the *Pto* DC3000 EEL, and *eelF* and *eelG* alleles similar to *eelF1* and *eelG1* in the *Pto* DC3000 EEL. Thus, there appears to be a limit to the diversity of the EEL, which is further supported by the recent amplification of the EEL from other *P. syringae* strains (Deng et al., 2003).

The similarity of these newly characterized EELs to the EELs of *Psy* 61, *Pto* DC3000, and *Psy* B728a could be due to experimental bias from the PCR amplification strategy and the strains surveyed. Although published sequences of

queA and *hrpK* from various pathovars were used to design Q920 and K2688, the alleles from *Psy* 61, *Pto* DC3000, and *Psy* B728a were given higher priority, which might have predisposed the screen to amplify EELs similar to the EELs of these three strains. Additionally, the location of the *hrp/hrc* cluster has not been determined for the strains screened, and it is possible that more divergent EELs are associated with *hrp/hrc* clusters that are not closely linked to *queA*. The size of the EEL also influences the ability to amplify the region. Although a PCR protocol specifically designed for long-range applications was used, smaller loci, such as the Family IA EEL, are easier to amplify. Moreover, most of the strains screened were originally isolated from North America. Since a possible geographical linkage was detected among strains with similar EELs, it may be that populations indigenous to geographical regions not included in this study have novel types of EELs.

Although the examined EELs are not as varied as expected, genes were identified that could encode novel effectors. The following gene products are divergent from their respective orthologs, suggesting a possible altered activity: *Psy* B5 HopPsyB, *Psy* B452 HopPsyV2, *Psy* W4N15 HopPsyC2, *Ppe* 5846 HopPtoB5, and *Psy* DH015 HopPsyG. HopPsyV2 and HopPtoB5 are truncated relative to *Psy* B728a HopPsyV1 and *Pto* DC3000 HopPtoB1, respectively. While some effectors are tolerant of carboxy-terminal deletions (Gopalan et al., 1996; Mudgett and Staskawicz, 1999), others, such as *avrPphE* (Stevens et al. 1998), are not. HopPsyB, one of the more divergent newly discovered orthologs, was tested to determine if it had different properties than HopPsyA. HopPsyB, like HopPsyA, was able to initiate the HR in tobacco leaves, but allowed *Pto* DC3000 to grow to higher levels in some *Brassica*

oleracea varieties than HopPsyA. In addition, a fusion of the amino-terminus of HopPsyB to the carboxy-terminus of AvrRpt2 was able to initiate the HR in *A. thaliana* ecotype Columbia (*RPS2*), but not in *A. thaliana* ecotype Columbia (*rps2*). Thus, HopPsyB is an effector similar to, but distinct from HopPsyA. Since indicator plants for HopPsyV, HopPsyC, HopPtoB, and HopPsyG have not been established, it was not feasible to determine how their various orthologs affect plant phenotype.

Additionally, the EEL appears to have evolved by the accumulation of transposed gene cassettes. A *hopPsyC-hopPsyE* operon with conserved coding and intergenic sequences is in Families III and IV, and *eelF* and *eelG* are adjacent to one another with conserved coding and intergenic sequences in Families IV, V, and VI. The *shcA-hopPsyA* operon in Family IA appears to be an inverted and truncated derivative of the *shcB-hopPsyB* operon in Family IB/IC. It is also possible that the Family II *avrPphE8*, the Family V *hopPtoB*, and the Family VI *hopPsyG* could each be a gene cassette. The genes in the EELs appear to have a lower GC content than *queA* or *hrpK*, indicating the EEL is likely assembled by the acquisition of horizontally transferred genes before or after the integration of the *hrp* pathogenicity island.

The mosaic structure of the EEL is suggestive of integron-like assembly. Integrons were first identified during studies of the localized and sequential accumulation of promoterless antibiotic resistance genes in enteric bacteria, but have since been found in many taxonomic groups of bacteria (Hall and Collis, 1995). Integrons assemble through the activity of a trans-acting, site-specific integrase that mediates the insertion of gene cassettes through homologous recombination of *att* sites

(Grainge and Jayaram, 1999). An imperfect inverted repeat is a key feature of an *att* site (Grainge and Jayaram, 1999; Hall and Collis, 1995), and is retained after the transposition event. Many pathogenicity islands are associated with a tRNA locus that carries inverted repeats (Hacker and Kaper, 1999). Transposition of the LEE pathogenicity island in REPEC strains of *E. coli* is linked to the activities of a tRNA gene-linked integrase (Tauschek et al., 2002), and tRNA gene-linked excision of virulence factors has been demonstrated for other *E. coli* strains (Blum et al., 1994).

Adjacent to the *P. syringae* EEL is a tRNA_{leu} locus (Alfano et al., 2000) that could function as an *att* site, but with the exception of the EELs of Families IC and III, a 52 bp region 3' to the tRNA_{leu} gene is conserved. If recombination occurred within the tRNA gene, little sequence conservation would be expected in this region.

However, within the conserved region downstream of the tRNA_{leu} gene is 31 bp inverted repeat (5'-AAAAAAGACCTTGAAATTCAAGGTCTTTTTT).

Additionally, the *Pto* DC3000 genome contains at least two apparently functional homologs of a tRNA-linked integrase that could mediate excision and transposition of gene cassettes (Charity et al., 2003). tRNA-dependent excision of *avr* genes, an activity indicative of integrase activity (Hall and Collis, 1995), has been observed in some *P. syringae* strains (Jackson et al., 2000). The mosaic structure of the EEL, the association of the EEL with a candidate *att* site, and the presence of tRNA gene-associated integrases in the genome of at least one *P. syringae* strain indicate the EEL might assemble in an integron-like fashion.

In nine of the 10 unique EELs, at least one potential or known effector is expressed as part of the *hrpK* operon. This association might imply that *hrpK* is part

of the EEL. Four arguments, however, can be made against such a conclusion. Firstly, insertional inactivation or deletion of *hrpK* results in an Hrp⁻ phenotype (Huang et al., 1991); so *hrpK* is essential for the activity of the *hrp* TTSS. Secondly, the location of *hrpK* is conserved in all *hrp* clusters characterized to date, whereas the location of effectors, such as alleles of *avrPphE*, is variable (Jackson et al., 1999; Stevens et al., 1998). Thirdly, the GC content of *hrpK* is substantially higher than most genes in the EEL and is very similar to the GC content of genes in the central conserved region. Finally, sequence polymorphisms in *hrpK* only weakly correlate with the type of EEL present (Charity et al., 2003). Thus, the exclusion of *hrpK* from the EEL is justifiable.

The results presented herein suggest that the genetic diversity of the EEL is due to (1) integron-like assembly of horizontally-transferred gene cassettes, (2) accumulation of point mutations in effector genes, and (3) genetic rearrangements including duplication, deletion, inversion, and insertion of IS elements. However, characterization of the locus and its effectors is far from complete. Most of the EELs observed in this study were from *P. syringae* pv. *syringae* strains. Repeating the PCR-screen with other sets of primers might allow amplification of the EEL from more pathovars and from strains with different geographical origins, which would facilitate inferences about the evolution of the EEL. Additionally, a *P. syringae* strain could be transformed with the putative effectors *Psy* B452 HopPsyV2, *Psy* W4N15 HopPsyC2, *Ppe* 5846 HopPtoB5, or *Psy* DH015 HopPsyG and inoculated into a wide variety of plants to test how these proteins affect virulence. Furthermore, there are no data indicating that strains with equivalent EELs, even those isolated from different hosts,

are not the same or nearly identical. Strains could be tested to determine if variance within the EEL tends to accompany divergence of the *P. syringae* genome. If strains with equivalent EELs were very different across the rest of the genome, this would support an independent evolution of the EEL. Lastly, the potential tRNA-dependent integrases identified in the *Pto* DC3000 genome could be isolated and tested *in vitro* for their ability to transpose gene cassettes into the EEL. Given that the putative cassettes appear to be horizontally acquired, if this activity were established, it could reveal a mechanism for the rapid evolution of epiphytic bacteria to their local environment.

REFERENCES

- Agrios, G. N. 1997. Plant Pathology, 4th ed. Academic Press, pg 30.
- Alfano, J. R., Charkowski, A. O., Deng, W.-L., Badel, J. L., Petnicki-Ocwieja, T., van Dijk, K. and Collmer, A. 2000. The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bound by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. Proc. Natl. Acad. Sci. USA 97: 4856-4861.
- Alfano, J. R. and Collmer, A. 1997. The type III (*hrp*) secretion pathway of plant pathogenic bacteria: trafficking harpins, avr proteins and death. J. Bacteriol. 179: 5655-5662.
- Alfano, J. R., Kim, H.-S., Delaney, T. P. and Collmer, A. 1997. Evidence that the *Pseudomonas syringae* pv. *syringae* *hrp*-linked *hrmA* gene encodes an Avr-like protein that acts in a *hrp*-dependent manner within tobacco cells. Mol. Plant-Microbe Interact. 10: 580-588.
- Baker, C. J., Atkinson, M. M. and Collmer, A. 1987. Concurrent loss in Tn5 mutants of the ability to induce the HR and host plasma membrane K⁺/H⁺ exchange in tobacco. Phytopathol. 77: 1268-1272.
- Bertoni, G. and Mills, D. 1987. A simple method to monitor growth of bacterial populations in leaf tissue. Phytopathol. 77: 832-835.
- Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschape, H. and Hacker, J. 1994. Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. Infect. Immun. 62: 606-614.
- Boucher, C.A., Martinel, A., Barberis, P., Alloing, G., and Zischeck, C. 1986. Virulence genes are carried by a megaplasmid of the plant pathogen *Pseudomonas solanacearum*. Mol. Gen. Genet. 205:270-275.
- Bradbury, J.F. 1986. Guide to plant pathogenic bacteria. CAB International, Farnham Royal, Great Britain.
- Casadaban, M. J. 1976. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. J. Mol. Biol. 104: 541-555.
- Charity, J.C., Pak, K., Delwiche, C.F., and S.W. Hutcheson. 2003. Novel exchangeable effector loci associated with the *Pseudomonas syringae* *hrp* pathogenicity island: evidence for integron-like assembly from transposed gene cassettes. Mol. Plant-Microbe Interact. 16: 495-507.

- Chen, Z., Kloek, A. P., Boch, J., Katagiri, F. and Kunkel, B. N. 2000. The *Pseudomonas syringae* *avrRpt2* gene product promotes pathogen virulence from inside plant cells. *Mol. Plant-Microbe Interact.* 13: 1312-1321.
- Collmer, A., Lindegerg, M., Petnicki-Ocwieja, T., Schneider, D. and Alfano, J. 2002. Genomic mining type III secretion system effectors in *Pseudomonas syringae* yields new picks for all TTSS prospectors. *Trends Microbiol.* 10: 462-469.
- Cornelis, G.R. 1994. Yersinia pathogenicity factors. In *Bacterial Pathogenesis of Plants and Animals*, ed. JL Dangl, pp. 99-114. Berlin: Springer-Verlag.
- Cuppels, D. A. 1986. Generation and characterization of Tn5 insertion mutations in *Pseudomonas syringae* pv. *tomato*. *Appl. Environ. Microbiol.* 51: 323-327.
- Dangl, J. and Jones, J. 2001. Plant pathogens and integrated defence responses to infection. *Nature (London)* 411: 826-833.
- Deng, W.L., Rehm, A., Charkowski, A., Rojas, C.M., and Collmer, A. *Pseudomonas syringae* exchangeable effector loci: Sequences diversity in representative pathovars and virulence function in *P. syringae* pv. *syringae* B728a. *J. Bacteriol.* 185: 2592-2602.
- Denny, T., Gilmour, M. and Selander, R. 1988. Genetic diversity and relationships of two pathovars of *Pseudomonas syringae*. *J. Gen. Microbiol.* 134: 1949-1960.
- Doudoroff, M., and N.J. Palleroni. 1974. Genus I. *Pseudomonas* Migula, p. 217-243. In R. E. Buchanan and N. E. Gibbons (ed.), *Bergey's manual of determinative bacteriology*, 8th ed. Williams and Wilkins, Baltimore, Md.
- Fouts, D., Abramovitch, R., Alfano, J., Baldo, A., Buell, C., Cartinhour, S., Chatterjee, A., D'Ascenzo, M., Gwinn, M., Lazarowitz, S., Lin, N.-C., Martin, G., Rehm, A., Schneider, D., vanDijk, K., Tang, X. and Collmer, A. 2002. Genomewide identification of *Pseudomonas syringae* pv. *tomato* DC3000 promoters controlled by the HrpL alternative sigma factor. *Proc. Natl. Acad. Sci. USA* 99: 2275-2280.
- Gopalan, S., Bauer, D. W., Alfano, J. R., Loniello, A. O., He, S. Y. and Collmer, A. 1996. Expression of the *Pseudomonas syringae* avirulence protein AvrB in plant cells alleviates its dependence on the hypersensitive response and pathogenicity (Hrp) secretion system in eliciting genotype-specific hypersensitive cell death. *Pl. Cell* 8: 1095-1105.
- Grainje, I. and Jayaram, M. 1999. The integrase family of recombinases: organization and function of the active site. *Mol. Microbiol.* 33: 449-456.

Groisman E. A., Ochman, H. 1993. Cognate gene clusters govern invasion of host epithelial cells by *Salmonella typhimurium* and *Shigella flexneri*. EMBO J. 12:3779-3787

Guttman, D. and Greenberg, J. 2001. Functional analysis of the type III effectors AvrRpt2 and AvrRpm1 of *Pseudomonas syringae* with the use of a single copy genomic integration system. Mol. Plant-Microbe Interact. 14: 145-155.

Guttman, D., Vinatzer, B., Sarkar, S., Ranall, M., Kettler, G. and Greenberg, J. 2002. A functional screen for the type III secretome of the plant pathogen *Pseudomonas syringae*. Science 295: 1722-1726.

Hacker, J. and Kaper, J. 1999. The concept of pathogenicity islands. 1-11 in: Pathogenicity Islands and other mobile virulence elements. J. Kaper and J. Hacker. ASM Press. Washington, DC,

Hall, R. and Collis, C. 1995. Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. Mol. Microbiol. 15: 593-600.

He, S. Y. 1998. Type III protein secretion systems in plant and animal pathogenic bacteria. Annu. Rev. Phytopathol. 36:363-292.

Heu, S. and Hutcheson, S. W. 1993. Nucleotide sequence and properties of the *hrmA* locus associated with the *P. syringae* pv. *syringae* 61 *hrp* gene cluster. Mol. Plant-Microbe Interact. 6: 553-564.

Hirano, S., Charkowski, A., Collmer, A., Willis, D. and Upper, C. 1999. Role of the Hrp type III protein secretion system in growth of *Pseudomonas syringae* pv. *syringae* B728a on host plants in the field. Proc. Natl. Acad. Sci. USA 96: 9851-9856.

Hirano, S. S., and C. D. Upper. 1995. Ecology of ice nucleation-active bacteria, p. 41-61. In R. E. Lee, Jr., G. J. Warren, and L. V. Gusta (ed.), Biological ice nucleation and its applications. American Phytopathological Society, St. Paul, Minn.

Hirano, S. S., and C. D. Upper. 2000. Bacteria in the Leaf Ecosystem with Emphasis on *Pseudomonas syringae*--a Pathogen, Ice Nucleus, and Epiphyte. MMBR. 64:624-653.

Huang, H. C., Hutcheson, S. W. and Collmer, A. 1991. Characterization of the *hrp* cluster from *Pseudomonas syringae* pv. *syringae* 61 and Tn*phoA* tagging of exported or membrane-spanning Hrp proteins. Mol. Plant-Microbe Interact. 4: 469-476.

Huang, H. C., Schuurink, R., Denny, T. P., Atkinson, M. M., Baker, C. J., Yucel, I., Hutcheson, S. W. and Collmer, A. 1988. Molecular cloning of a *Pseudomonas*

syringae pv. *syringae* gene cluster that enables *Pseudomonas fluorescens* to elicit the hypersensitive response in tobacco. J. Bacteriol. 170: 4748-4756.

Hutcheson, S. 2001. The molecular biology of hypersensitivity to plant pathogenic bacteria. J. Plant Pathol. 83: 151-172.

Hutcheson, S., Bretz, J., Losada, L., Charity, J. and Sussan, T. 2003. Regulation and detection of effectors translocated by *Pseudomonas syringae*. IN PRESS in: Developments in Plant Pathology: *Pseudomonas syringae* pathovars and related pathogens. 2. N. Iacobelis. Kluwer Academic Publishers. Dordrecht, Hutcheson, S. W. 1998. Currents concepts of active defense in plants. Annu. Rev. Phytopathol. 36: 59-90.

Hutcheson, S. W. 1999. The *hrp* cluster of *Pseudomonas syringae*: a pathogenicity island encoding a type III protein translocation complex? 309-329 in: Pathogenicity islands and other mobile virulence elements. J. B. Kaper and J. Harker. Amer. Soc. Microbiol. Washington, D.C.,

Jackson, R., Mansfield, J., Arnold, D., Sesma, A., Paynter, C., Murillo, J., Taylor, J. and Vivian, A. 2000. Excision from tRNA genes of a large chromosomal region, carrying *avrPphB*, associated with race change in the bean pathogen, *Pseudomonas syringae* pv. *phaseolicola*. Mol. Microbiol. 38: 186-197.

Jackson, R. W., Athanassopoulos, E., Tsiamis, G., Mansfield, J. W., Sema, A., Arnold, D. L., Gibbon, M. J., Murillo, J., Taylor, J. D. and Vivian, A. 1999. Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in the bean pathogen *Pseudomonas syringae* pv. *phaseolicola*. Proc. Natl. Acad. Sci. USA 96: 10875-10880.

Jarvis K. G., Giron J. A., Jerse, A.E., McDaniel, T.K., Donnenberg, M.S., and Kaper J.B., 1995. Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. Proc. Natl. Acad. Sci. USA 92:7996-8000.

Hin, Q., Thilmony, R, Zwiesler-Vollick, J., and S. He. 1997. Microbes and Infection. 5:301-310.

Keen, N.T., 1992. The molecular biology of disease resistance. Plant Mol. Biol. 19:109-122.

Keen, N.T. Ersek, T., Long, M. Bruegger, B., Holliday, M. 1981. Inhibition of the hypersensitive reaction of soybean leaves to incompatible *Pseudomonas* spp. by blasticidin S, streptomycin or elevated temperature. Physiol. Plant Pathol. 18:325-357.

- Keen, N., Tamaki, J., Kobayashi, D. and Trollinger, D. 1988. Improved broad host range plasmids for DNA cloning in gram negative bacteria. *Gene* 70: 191-197.
- Kim, J. F., Charkowski, A. O., Alfano, J. R., Collmer, A. and Beer, S. V. 1998. Sequences related to transposable elements and bacteriophages flank avirulence genes of *Pseudomonas syringae*. *Mol. Plant-Microbe Interact.* 11: 1247-1252.
- King, E. O., Ward, M. K., and D. E. Raney. Two simple media for the demonstration of pyocyanin and fluorescein. *J. Lab. Clin. Med.* 44: 301-307.
- Lamb, C.J., Lawton, M.A., Dron, M., and R.A. Dixon. 1989. Signals and transduction mechanisms for activation of plant defenses against microbial attack. *Cell* 56:215
- Leach, J. E. and White, F. F. 1996. Bacterial avirulence genes. *Annu. Rev. Phytopathol.* 34: 153-179.
- Leister, R. T., Ausubel, F. M. and Katagiri, F. 1996. Molecular recognition of pathogen attack occurs inside of plant cells in plant disease resistance specified by the *Arabidopsis* genes *RPS2* and *RPM1*. *Proc. Natl. Acad. Sci. USA* 93: 3459-3464.
- Li, T.-H., Benson, S. A. and Hutcheson, S. W. 1992. Phenotypic expression of the *Pseudomonas syringae* pv. *syringae* 61 *hrp/hrm* gene cluster in *Escherichia coli* requires a functional porin. *J. Bacteriol.* 174: 1742-1749.
- Lindgren, P. B., 1997. The role of *hrp* genes during plant-bacterial interactions. *Annu. Rev. Phytopathol.* 35:129-152.
- Lindow, S. E. 1983. The role of bacterial ice nucleation in frost injury to plants. *Annu. Rev. Phytopathol.* 21:363-384.
- Mansfield, J., Jenner, C., Hockenull, R., Bennett, M. A. and Stewart, R. 1994. Characterization of *avrPphE*, a gene for cultivar-specific avirulence from *Pseudomonas syringae* pv. *phaseolicola* which is physically linked to *hrpY*, a new *hrp* gene identified in the halo-blight bacterium. *Mol. Plant-Microbe Interact.* 7: 726-739.
- Mudgett, M. B. and Staskawicz, B. J. 1999. Characterization of the *Pseudomonas syringae* pv. tomato AvrRpt2 protein: demonstration of secretion and processing during bacterial pathogenesis. *Mol. Microbiol.* 32: 927-941.
- Murillo, J., Shen, H., Gerhold, D., Sharma, A., Cooksey, D. A. and Keen, N. T. 1994. Characterization of pPT23B, the plasmid involved in syringolide production by *Pseudomonas syringae* pv. tomato PT23. *Plasmid* 31: 275-287.
- Perombelon, M.C.M., and G.P.C. Salmond. Bacterial Soft Rots, In Pathogenesis and Host Specificity in Plant Diseases. Vol I: Prokaryotes

Petnicki-Ocweija, T., Schneider, D., Tam, V., Chancey, S., Shan, L., Jamir, Y., Schechter, L., Janes, M., Buell, C., Tang, X., Collmer, A. and Alfano, J. 2002. Genomewide identification of proteins secreted by the Hrp type III protein secretion system of *Pseudomonas syringae* pv. tomato DC3000. *Proc. Natl. Acad. Sci. USA* 99: 7652-7657.

Pirhonen, M. U., Lidell, M. C., Rowley, D., Lee, S. W., Silverstone, S., Liang, Y., Keen, N. T. and Hutcheson, S. W. 1996. Phenotypic expression of *Pseudomonas syringae* *avr* genes in *E. coli* is linked to the activities of the *hrp*-encoded secretion system. *Mol. Plant-Microbe Interact.* 9: 252-260.

Reuter, K., Slany, R., Ullrich, F. and Kersten, H. 1991. Structure and organization of *Escherichia coli* genes involved in biosynthesis of the deazaguanine derivative of quinine, a nutrient factor for eukaryotes. *J. Bacteriol.* 173: 2256-2264.

Staskawicz, B. J., Dahlbeck, D., Keen, N. and Napoli, C. 1987. Molecular characterization of cloned avirulence genes from race 0 and 1 of *Pseudomonas syringae* pv. *glycinea*. *J. Bacteriol.* 169: 5789-5794.

Staskawicz, B. J., Dahlbeck, D. and Keen, N. T. 1984. Cloned avirulence gene of *Pseudomonas syringae* pv. *glycinea* determines race-specific incompatibility on *Glycine max* (L.) Merr. *Proc. Natl. Acad. Sci. USA* 81: 6024-6028.

Stevens, C., Bennett, M., Athanossopoulos, E., Tsiamis, G., Taylor, J. and Mansfield, J. 1998. Sequence variations in alleles of the avirulence gene *avrPphE.R2* from *Pseudomonas syringae* pv. *phaseolicola* lead to loss of recognition of the AvrPphE protein within bean cells and a gain in cultivar-specific virulence. *Mol. Microbiol.* 29: 165-177.

Tauschek, M., Strugnell, R. and Robins-Browne, R. 2002. Characterization and evidence of mobilization of the LEE pathogenicity island of rabbit-specific strains of enteropathogenic *Escherichia coli*. *Mol. Microbiol.* 44: 1533-1550.

van Dijk, K., Fouts, D. E., Rehm, A. H., Hill, A. R., Collmer, A. and Alfano, J. R. 1999. The Avr (Effector) proteins HrmA (HopPsyA) and AvrPto are secreted in culture from *Pseudomonas syringae* pathovars via the Hrp (Type III) protein secretion system in a temperature and pH-sensitive manner. *J. Bacteriol.* 181: 4790-4797.

van Dijk, K., Tam, V., Records, A., Petnicki-Ocweija, T. and Alfano, J. 2002. The ShcA protein is a molecular chaperone that assists in the secretion of the HopPsyA effector from the type III (Hrp) protein secretion system of *Pseudomonas syringae*. *Mol. Microbiol.* 44: 1469-1481.

Venkatesan M. M., Buysse, J. M., and Oaks, E.V. 1992. Surface presentation of *Shigella flexneri* invasion plasmid antigens requires the products of the *spa* locus. J. Bacteriol. 174:1990-2001.

Vivian, A. and Gibbon, M. J. 1997. Avirulence genes in plant-pathogenic bacteria: signals or weapons. Microbiol. 143: 693-704.

Whalen, M. C., Innes, R. W., Bent, A. F. and Staskawicz, B. J. 1991. Identification of *Pseudomonas syringae* pathogens of *Arabidopsis* and a bacterial locus determining avirulence on both *Arabidopsis* and soybean. Pl. Cell 3: 49-59.

Xiao, Y., Heu, S., Yi, J., Lu, Y. and Hutcheson, S. W. 1994. Identification of a putative alternate sigma factor and characterization of a multicomponent regulatory cascade controlling the expression of *Pseudomonas syringae* pv. *syringae* Pss61 *hrp* and *hrmA* genes. J. Bacteriol. 176: 1025-1036.

Xiao, Y. and Hutcheson, S. W. 1994. A single promoter sequence recognized by a newly identified alternate sigma factor directs expression of pathogenicity and host range determinants in *Pseudomonas syringae*. J. Bacteriol. 176: 3089-3091.

Young, J.M. 1991. Pathogenicity and identification of the lilac pathogen, *Pseudomonas syringae* pv. *syringae* van Hall 1902. Ann. Appl. Biol. 118:283-298.

Yuan, J. and He, S. Y. 1996. The *Pseudomonas syringae* Hrp regulation and secretion system controls the production and secretion of multiple extracellular proteins. J. Bacteriol. 178: 6399-6402.