Molecular Microbiology (2003) **47**(6), 1545–1562

Nucleotide sequence, functional characterization and evolution of pFKN, a virulence plasmid in *Pseudomonas syringae* pathovar *maculicola*

Laurence Rohmer,[1] Susanne Kjemtrup,[1†] Patrizia Marchesini[1†] and Jeffery L. Dangl[1–3*]
[1]*Department of Biology,* [2]*Curriculum in Genetics, and* [3]*Department of Microbiology and Immunology, Coker Hall 108, CB#3280, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280, USA.*

Summary

***Pseudomonas syringae* pv. *maculicola* strain M6 (*Psm* M6) carries the *avrRpm1* gene, encoding a type III effector, on a 40 kb plasmid, pFKN. We hypothesized that this plasmid might carry additional genes required for pathogenesis on plants. We report the sequence and features of pFKN. In addition to *avrRpm1*, pFKN carries an allele of another type III effector, termed *avrPphE*, and a gene of unknown function (ORF8), expression of which is induced *in planta*, suggesting a role in the plant–pathogen interaction. The region of pFKN carrying *avrRpm1*, *avrPphE* and ORF8 exhibits several features of pathogenicity islands (PAIs). Curing of pFKN (creating *Psm* M6C) caused a significant reduction in virulence on *Arabidopsis* leaves. However, complementation studies using *Psm* M6C demonstrated an obvious virulence function only for *avrRpm1*. pFKN can integrate and excise from the chromosome of *Psm* M6 at low frequency via homologous recombination between identical sequence segments located on the chromosome and on pFKN. These segments are part of two nearly identical transposons carrying *avrPphE*. The *avrPphE* transposon was also detected in other strains of *P. s.* pv. *maculicola* and in *P. s. tomato* strain DC3000. The *avrPphE* transposon was found inserted at different loci in different strains. The analysis of sequences surrounding the *avrPphE* transposon insertion site in the chromosome of *Psm* M6 indicates that pFKN integrates into a PAI that encodes type III effectors. The integration of pFKN into this chromosomal region may therefore be seen as an evolutionary process determining the formation of a new PAI in the chromosome of *Psm* M6.**

Accepted 28 November, 2002. *For correspondence. E-mail dangl@email.unc.edu; Tel. (+1) 919 962 5624; Fax (+1) 919 962 1625. †Present address: Paradigm Genetics, Inc., 104 Alexander Dr., PO Box 14528, RTP, NC 27709-4528, USA.

Introduction

Pathogens require an arsenal of pathogenicity factors to infect a host. Genes important for pathogenicity are often grouped in regions of the chromosome or on plasmids. These regions are referred to as pathogenicity islands (PAIs) (for a review, see Hacker and Kaper, 2000). PAIs typically carry virulence genes and sequences related to DNA mobility, such as genes encoding integrases or transposases, and insertion sequences (IS). It is very likely that PAIs are acquired through horizontal transfer (Hacker and Kaper, 1999). This is suggested by the fact that their G+C content often differs from that of the core genome. Moreover, in many cases, PAIs are flanked by tRNA genes. Some tRNA loci are known to be integration targets for DNA of bacteriophages or plasmids (Hou, 1999). The VPI (*Vibrio* pathogenicity island), for example, results from the integration of the bacteriophage VPIΦ (Karaolis *et al*., 1999) into the chromosome of *Vibrio cholerae*. In other cases, the PAI seems to have arisen from the integration of a plasmid. Indeed, Rajakumar *et al*. (1997) proposed that a 99 kb chromosomal element in *Shigella flexneri* 2a YSH6000 carrying a multiantibiotic resistance locus arose after the integration of an NR1-like plasmid.

The mosaic-like structure of numerous PAIs suggests that they are generated by a multistep process. This process may involve various mobile elements and multiple genomic rearrangements. For example, the SPI2 (*Salmonella* PAI 2) PAI contains separate 25 kb and 14.5 kb regions that differ from each other in their G+C content, their codon usage, the type of functions they encode and their respective distribution in other strains of *Salmonella* (Hensel *et al*., 1999). Similarly, the comparison of five locus of enterocyte effacement (LEE) PAIs from three rabbit-specific strains of enteropathogenic *Escherichia coli* (REPEC 83/39, 84/110-1 and RDEC-1) and human enteropathogenic or enterohaemorrhagic *E. coli* strains (E2348/69 and EDL933) shows that the five PAIs origi-

nated from the same source (Tauschek *et al*., 2002). However, each of them has evolved and diverged from the PAIs of the other strains through DNA insertions and potential deletions (Tauschek *et al*., 2002). On some occasions, a PAI may excise from the chromosome. The mechanism of excision is sometimes the same as the mechanism of integration. For example, an integrase encoded on the HPI (high pathogenicity island) of *Yersinia pseudotuberculosis* and *Yersinia pestis* is responsible for the integration of this PAI into its attachment site on the chromosome (Rakin *et al*., 2001). Rakin and coworkers showed that, when this integrase is expressed, the HPI could be excised.

In the plant pathogen *Pseudomonas syringae*, several PAIs have been identified (Alarcon-Chaidez *et al*., 1999; Jackson *et al*., 1999; 2000; Alfano *et al*., 2000; Lu *et al*., 2002). They carry genes encoding the type III secretion system apparatus (called the *hrp* cluster), type III effector proteins that serve as virulence factors or proteins catalysing the production of toxins. Thus, PAIs carry genes representing the major components of *P. syringae* pathogenicity. The *hrp* cluster is necessary for pathogenicity (Lindgren *et al*., 1986; 1988) and encodes a type III pilus (Jin and He, 2001).

The *P. syringae* type III secretion system delivers type III effector proteins to the plant cell (Cornelis and Van Gijsegem, 2000; Jin and He, 2001; Li *et al*., 2002). Although some type III effectors have demonstrable roles in virulence, very little is known about their host targets (Nimchuk *et al*., 2001). Target processes affected by *P. syringae* type III effectors were described recently (Mackey *et al*., 2002; Van der Hoorn *et al*., 2002). Some type III effectors are recognized inside the plant cell by the plant *R* gene system, and these trigger a disease resistance-associated programmed cell death known as the hypersensitive response (HR) (Dangl and Jones, 2001). Because the recognition of some type III effector proteins causes the plant to be resistant to a particular pathogen strain, most of the genes encoding *P syringae* type III effectors were originally named avirulence (*avr*) genes.

The *P. syringae hrp* PAI also encodes the HrpL alternative sigma factor. This factor regulates the expression of both the genes encoding the type III delivery pilus and those encoding type III effectors (Xiao *et al*., 1994). The promoters of these genes contain a consensus *hrp* box (Innes *et al*., 1993). Hence, many genes encoding putative type III effectors have been identified recently using sequence analysis methods on the whole genome of the *Pst* DC3000 strain and sequences of DNA fragments from other pathovars (Jackson *et al*., 1999; Alfano *et al*., 2000; Arnold *et al*., 2001; Boch *et al*., 2002; Fouts *et al*., 2002; Petnicki-Ocwieja *et al*., 2002).

The type III effector avrRpm1 (Ritter and Dangl, 1995) contributes to virulence on *Arabidopsis* plants that cannot recognize it (host genotype *rpm1*). It is encoded on a plasmid in several *P. syringae* pv. *maculicola* (*Psm*) strains, hereafter named pFKN. Several *P. syringae* plasmids have been partially or fully characterized. They often carry genes encoding virulence factors or functions enabling the bacteria to survive in given environments (for a review, see Vivian *et al*., 2001). We describe the analysis of the pFKN sequence and suggest that it carries at least one PAI. We investigated the mechanism by which the plasmid integrates and excises into and out of the chromosome of *Psm* M6, as well as the chromosomal region involved in the process. We also found that a piece of the pFKN mosaic structure is a transposon carrying an allele of a second type III effector, *avrPphE*. Finally, we describe the genomic context of this transposon in several *P. syringae* strains. Our work provides insights into the evolutionary process leading to the formation of new PAIs in *P. syringae*, and how they may be transferred from strain to strain.

Results

Sequencing of pFKN

The sequence of pFKN was determined to an average of ≈3.1× coverage using a shotgun approach. The complete nucleotide sequence of pFKN consists of 39 554 bp. Its overall G+C content is 53.53%. Twenty-nine putative open reading frames (ORFs) were identified, representing 75.7% of the sequence. To assign possible functions to these ORFs, we compared the sequence of each with available databases, using BLASTN and BLASTX. Twenty-seven ORFs had homology with genes present in the databases. We searched the translated products of each ORF for known functional domains and motifs and could assign a putative function to most of them (Table 1).

We organized the ORFs into various functional groups (Fig. 1). The plasmid encodes two type III effectors (*avrRpm1* and an allele of *avrPphE*). Five non-contiguous regions on the plasmid may be mobile elements or remnants thereof. Three other groups of genes were defined, according to the putative function that they encode.

Plasmid replication and maintenance. ORF1 shares high homology with genes that encode the replication proteins of plasmids belonging to the PT23A family in the P. syringae species (Sesma et al., 1998). Homologous replication proteins are found in other genera, such as the replication proteins from the ColE2 family in E. coli (Sesma et al., 1998). The product of ORF10 shares similarity with ParA, a protein involved in partitioning working in concert with a protein encoded by ParB. No ParB gene was found on pFKN. Therefore, it is not clear how the plasmid copies (around five per cell; data not shown) are distributed effi-

Table 1. ORFs and non-coding elements identified in *P.s.* pv. *maculicola* pFKN.

ORF element	Region on pFKN:	Number of codons	Predicted Protein size (kDa)	Start codon	Stop codon	Putative ribosome-binding site	Putative function of the product of the ORF according to homologies or closest homolog	Closest Homolog Accession no.	identity/ homology	length homology/ total length protein	Noticeable structures or domains	Putative cellular localization in the bacteria	fragment no. in the genome of *Pst* DC3000[a]	Gene expression[b]		
														KB	MM	*in planta*
ORF1	44 - 1354	437	48.3	atg	tga	TGGCGGT	replication protein (*P. syringae*)	CAA10908.1	94/96	437/437		inner membrane (but very weak)	4 (at least 2 copies)	+	-	NA
ORF2	1599 - 2750	384	41.4	atg	tga	TTGAGAT	homolog of AvrPphE in *P.s.*p.v *syringae*	AAF71495.1	72/80	380/382	Tyr kinase phosphorylation site 202-209 RSEYENIY	cytoplasm	1	-	+	+
ORF3	2888 - 4911	518		-	tga	AGGCGGC	Transposase of Tn21 (*E. coli*), truncated	BAA78805.1	88/94	655/988			1	-	+	NA
ORF4	4959 - 6944	662		atg	taa		ORFE+ORFD neighbouring *avrPpiA1* in *P.s.*p.v *pisi*, slight homology to phage proteins and transposases	CAB96969.1/ CAB96968.1	98/98	367/414	HELIX_LOOP_HELIX Myc-type, 'helix-loop-helix' dimerization domain signature (eukaryotic domain) (PDOC00038)	cytoplasm (weak)	at least 2 copies	+	+	NA
ORF5	7138 - 7797	295	24.5	atg	taa	AAAAGGT	AvrRpm1 (*P.s. maculicola*)	CAA48009	94/94	220/220	myristylation site	cytoplasm	none	-	+	+
ORF6	8116 - 9129	338		-	tga		UV light resistance prot RulB (*P. syringae*), truncated	AAC44639.1	85/90	345/430			at least 2 copies			NA
ORF7	10008-10436	142	16.3	gtg	tga		no homology (unknown function)					cytoplasm	none	+	+	NA
ORF8	10272 - 10904	211	23.7	gtg	tga		no homology (unknown function)				Internal Repeats (N-glycosylation motifs?)	inner membrane	none	-	+	+
ORF9	12262 - 13491	410	46.9	atg	tga	TGAGGC	transposase1 of IS801 (*P.syringae*)	CAA40540.1	94/95	410/410		cytoplasm	one copy	+	+	NA
	12532 - 13500	323		atg	taa			CAA40541.1	88/89	199/199						
ORF10	14114 - 14761	216	23.8	atg	taa	AGGAGGC	putative partition protein (homologous to ParA)	BAA84904.1	57/73	186/214			none	+	+	+
ORF11	15246 - 15884	213	24.0	atg	tga		ORF H0801 (*Halobacterium sp.* NRC-1)	AAC82856.1	34/47	205/238		cytoplasm	none	-	-	-
ORF12	16897 - 18930	683	73.0	atg	tga	CGAGAAC	methyl-accepting chemotaxis protein (*Pseudomonas aeruginosa*)	AAG05035	30/48	413/652	transmembrane regions	inner membrane	none	+	+	NA
ORF13	19220 - 20743	508	56.9	atg	tag	CGGCGGC	L0015 (*Escherichia coli*), carried on an IS element	AAC31494.1	37/54	502/512		cytoplasm	none	+	+	NA
ORF14	20830 - 21363	178	19.2	atg	taa		L0014 (*Escherichia coli*), carried on an IS element	AAC31493.1	39/56	98/117		cytoplasm	none	-	-	-
ORF15	23110 - 23490	126	13.3	atg	tga	AGGAGAA	CRCB protein (*E coli*), transmembrane protein	P37002	52/71	115/127			none	+	+	NA
ORF16	23860 - 24402	181	20.1	atg	tag	CGGAGCT	inorganic pyrophosphatase (*Pseudomonas aeruginosa*)	AAG07418.1	71/86	173/175		inner membrane	2	+	+	NA
ORF17	24396 - 25004	200	22.6	gtg	tag	AGGAGGC	conserved hypothetical protein (*Deinococcus radiodurans*)	AAF09614.1	32/49	162/199	Phosphatase domain 68-74 DGDPLDV (PDOC00325)	inner membrane	2	+	+	NA
ORF18	25019 - 26185	389	42.1	atg	tga	GGAGCGC	malic acid transport protein (*Methanococcus jannaschii*)	AAB98567.1	29/48	343/347	transmembrane domain	inner membrane	2	+	+	NA
ORF19	26303 - 27196	297	33.2	atg	tag	TRUNCATED?	putative transcriptional regulator LYSR-type (*Pseudomonas aeruginosa*)	AAG08767.1	53/68	225/297	bacterial regulatory helix-turn-helix protein, lysR family (PF00126)	cytoplasm	2	+	+	NA
ORF20	27294 - 28346	351	37.7	atg	tga		ORF492, surface antigen gene (*Methanosarcina mazei*)	CAA59198.1	37/56	280/491	repeats, signal peptide from aa 1 to aa 47	cytoplasm (weak)	2	+	+	NA
ORF21	28676 - 29053	126	13.4	atg	taa		putative lipoprotein (*Streptomyces coelicolor* A3(2))	AAG08744.1	37/46	125/317	ATP_GTP_A ATP/GTP-binding site motif A (P-loop) 49-56 GKDTSGKS (PDOC00017)	periplasmic space (strong)	2	+	+	NA
ORF22	29775 - 30734	337	35.5	atg	taa	TAGCGATG	integrase/recombinase ripX (*Bacillus subtilis*)	CAB14283.1	31/50	250/296		cytoplasm	2	+	+	NA
ORF23	30984 - 31682	233	25.8	atg	tag		putative gntR-family transcriptional regulator (*Streptomyces coelicolor* A3(2))	CAB50881.1	33/52	209/230		cytoplasm (weak)	3	+	+	NA
ORF24	31754 - 32944	397	43.3	atg	tga	AAAAGGA	putative transcriptional regulator (*Vibrio cholerae*)	AAF96060.1	50/68	401/403		inner membrane	3	-	-	-
ORF25	32988 - 34331	448	47.5	atg	tga	TGAAGGG	putative integral membrane sugar transporter (*Amycolatopsis orientalis*)	CAB45038.1	36/56	410/444		inner membrane	3	+	+	NA
ORF26	34626 - 34994	123	13.4	atg	tga	TCGGGGT	repressor protein (Bacteriophage Tuc2009)	AAA21825.1	34/54	94/286		membrane	3	-	+	NA
ORF27	35131 - 35694	188	21.6	atg	tag		transcriptional activator RFAH (*Salmonella typhimurium* LT2)	AAF33417	33/49	96(Nt)/162		inner membrane	3	+	+	NA
ORF28	36025 - 38628	868	94.8	atg	tga	AGGAGTT	Ct:Ct of TraE in *Escherichia coli* (strain:K-12) plasmid R721, Nt:VirB1 (*Brucella suis*)	Ct:AP002527 Nt:AF141604	Ct:37/54, Nt:37/51	Ct:327/785, Nt:227/238		periplasmic space	none	+	-	NA
ORF29	38759 - 39184	142	15.7	atg	tga	AGGAGAC	putative transcriptional regulator (*Streptomyces coelicolor* A3(2))	CAB45490.1	43/62	61/129		cytoplasm	4 (at least 2 copies)	+	+	NA

a. We identified 4 DNA fragments that were conserved between pFKN and the genome of *Pst* DC3000 (>85% nucleotide identity). The ORFs that have no homologues in the genome of *Pst* DC3000 are noted as 'none', the fragments are numbered 1–4.
b. The expression of each gene carried by the plasmid was assessed in KB and MM (*hrp*-inducing media) by RT-PCR, and by using GFP-translational fusions *in planta* for 7 of them (not shown). When the gene is expressed in one given condition, the table exhibits '+'. A lack of expression is marked by '–'. NA, not available.

ciently after cell division. ORF28 may encode a protein involved in the transfer of pFKN. The localization of the protein encoded by ORF28 is predicted to be in the bacterial periplasmic space. It has three domains. The N-terminus shows high homologies to VirB1 of Brucella suis (a mammalian pathogen) named after its homology to VirB1 of Agrobacterium tumefaciens. The central region shows homology to VirB4 of the same organism. The C-terminus shares homology with different DNA-binding proteins (DNA nickase, primase). However, no other gene on the plasmid seems to encode a function related to transfer. It is highly unlikely that this unique protein would allow transfer of pFKN. Therefore, ORF28 encodes an unknown function or could represent the remnant of a transfer system encoded on the plasmid before genomic rearrangements leading to its loss.

Type III effectors. The promoter regions of two genes on the plasmid contain an *hrp* box, suggesting that they are part of the type III regulon. ORF2 is 80% identical to the *avrPphE* allele carried by *P. s. syringae* (*Psy*) (Alfano *et al.*, 2000). The *avrPphE* gene had originally been identified in *P. s. phaseolicola* (*Pph*) and encodes a protein delivered to the host cell (Mansfield *et al.*, 1994).

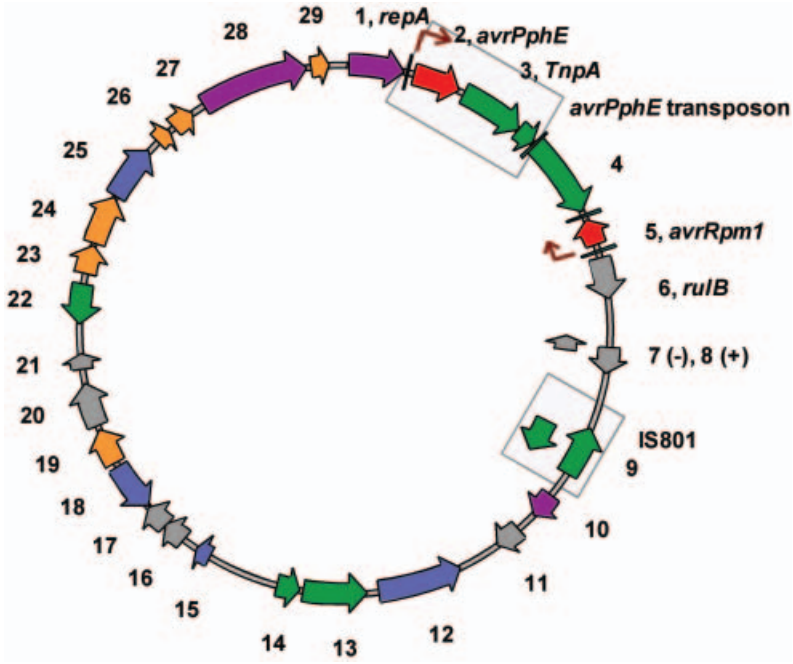

Fig. 1. Graphical map of pFKN and its 29 identified ORFs. The proposed direction of transcription is indicated by arrows. Red boxes indicate genes encoding for type III effectors. A black arrow line at the 5′ end of any genes indicates a possible *hrp* box. Genes related to those of mobile elements are represented by green boxes. The *avrPphE* transposon and the IS*801* homologue are framed. Transposon repeats surrounding the *avrPphE* transposon are marked by green lines. The *avrRpm1* gene is surrounded by repeats similar to the repeats of the transposon Tn*501*; these repeats are also marked by green lines. Purple boxes indicate genes with replication and partition functions. Genes encoding for putative transporters are indicated by blue boxes. The putative transcriptional regulators are indicated by orange boxes. The grey boxes represent genes that do not fall into any of the categories listed. The number of each ORF is noted in bold on the outside of the circle.

Since then, several alleles have been identified in other pathovars of *P. syringae* (Fouts *et al*., 2002; Guttman *et al*., 2002). They all contain an *hrp* box; in ORF2 of pFKN, this is found 51 bp upstream of the ATG at nt 1561 as 5′-GGAAC*TG*A-N13-CGAC*AT*A-3′ (Innes *et al*., 1993). Based on the homologies and the presence of the *hrp* box in the promoter, we hypothesized that the AvrPphE-like protein encoded by ORF2 is also a type III effector (see below). *AvrRpm1* (numbered ORF5) is identical to that defined in the strain *Psm* M2 (Dangl *et al*., 1992).

Mobile element-related genes. ORF3 (*TnpA*) encodes a putative truncated transposase. Importantly, *avrPphE* and *TnpA* are located between nearly identical inverted repeats of 66 bp (Fig. 1), thus defining a probable transposable element remnant. The sequences surrounding *avrRpm1* (ORF4 and ORF6) are homologous to a sequence identified in *P. s. pisi* by Arnold *et al*. (2001), with 96–97% nucleotide identity. The DNA fragment carrying these sequences in *P. s. pisi* may also be a transposon or a remnant. It is therefore likely that *avrRpm1* is also located on the remnant of a transposon.

ORF13 and ORF14 are related to ORFs found in the prophage 933 L on the LEE of the enterohaemorrhagic *E. coli* O157:H7, L0015 and L0014 respectively. These two ORFs are thought to belong to an IS element from the IS*66* family, primarily identified in *Sinorhizobium meliloti*, that includes IS elements from animal and plant pathogens. These two genes are usually associated with a third one, L0013 in the LEE, that is absent in pFKN (the sequence upstream of ORF14 does not contain any ORF and does not show homology to L0013). Moreover, the 5′ end of ORF14 is also missing. Therefore, the sequence encoding ORF13 and ORF14 may represent the remnant of an IS element belonging to the IS*66* family, not previously identified in *P. syringae*.

Transcription factors. Four pFKN genes share homology with transcriptional regulators. Some of these may function during interactions with eukaryotic hosts. For example, the closest homologue to ORF24 is a putative transcriptional factor from *V. cholerae*. ORF24 is also related to putative proteins of unknown function (along the whole length of the protein) found in organisms as diverse as *E. coli* (50% identity/67% homology), *Caenorhabditis elegans* (41%/57%) *Homo sapiens* (42%/58%), *Drosophila melanogaster* (40%/58%) and *Arabidopsis thaliana* (29%/42%). A domain characteristic of FAD-dependent oxidoreductase (in Pfam) is detected in ORF24. Therefore, the function of the translation product of this ORF may require electron transfer through FAD.

The protein encoded by ORF27 is a member of the small RfaH family. RfaH controls the transcription of *E. coli* and *Salmonella* operons that direct the synthesis, assembly and export of the lipopolysaccharide core, exopolysaccharide, F-conjugation pilus and haemolysin toxin (for a review, see Santangelo and Roberts, 2002). RfaH is a specific regulator of transcript elongation; its loss increases transcription polarity in these operons without affecting initiation from the operon promoters. In addition, in one case, the RfaH protein controls the transcription of

a single gene encoding an outer membrane haemin receptor, ChuA in pathogenic *E. coli*, as well as other genes involved in pathogenicity of the strain (Nagy *et al.*, 2001; 2002). The promoter of the genes or operons controlled by RfaH contains a conserved sequence called JUMPStart (Hobbs and Reeves, 1994). No JUMPStart sequence was detected on pFKN.

Transporters. ORF18 encodes a probable transmembrane protein that is highly homologous to sugar transporters. One is the product of *mae1*, a gene of the yeast *Schizosaccharomyces pombe*, encoding a protein involved in the uptake of L-malate, succinate and malonic acid. The product of ORF18 may provide the bacteria with a fitness advantage in certain conditions.

Other putative functions encoded on the plasmid. Other genes may encode proteins providing an advantage to the bacteria during the infection process. For example, the product of ORF12 displays all the characteristics of methyl-accepting chemotaxis proteins. The genes encoding the flagellar unit and the chemotaxis functions are often organized in clusters, which is not the case on pFKN. In this particular case, the flagellar unit may be provided *in trans*.

pFKN is composed of different fragments that indicate a multistep process leading to its formation

The G+C content of the pFKN ORFs is highly variable along the plasmid (Table 2). Yet several contiguous gene sets on the plasmid have similar G+C contents. This probably indicates that these genes were acquired together from an original source. The G+C content might allow us to track the multistep acquisition process resulting in pFKN. For example, ORFs 7, 8 and 10 share nearly identical G+C content (50.5%), but are separated by an IS element with a significantly different G+C content (60.3%). To determine the distribution of pFKN ORFs in other pathovars, and how this distribution is related to G+C content by ORF group, we performed sequence comparisons with the nucleotide sequence of the complete genomes of *P. syringae* pv. *tomato* DC3000 (http://tigrblast.tigr.org/ufmg/) and *P. syringae* pv. *syringae* B728a (http://www.jgi.doe.gov/JGI_microbial/html/index.html).
DNA fragments carrying from two to seven contiguous pFKN ORFs were detected in both genome sequences (85–100% identity). The four DNA fragments from *P. s. tomato* DC3000 carrying linked pFKN ORFs were numbered and added to Table 2. Of these, *Pst* DC3000 fragment 2 was also detected in *Psy* B728a (98% identical to pFKN). In addition, sequences homologous to *repA* and ORF13 were also detected in this *Psy* B728a. Table 2 shows that ORFs belonging to the same contiguous DNA fragment have a similar G+C content. For example, genes belonging to fragment 4 have a relatively high G+C content. Genes belonging to fragments 1 and 2 have similar G+C contents, and genes in fragment 3 exhibit a lower G+C content. Thus, the different fragments constituting pFKN were probably acquired by the three genomes in different steps, followed by subsequent sequence deletions and/or rearrangements with respect to *Psm* M6, as the entire collinear pFKN is not detected in either *Pst* DC3000 or *Psy* B728A.

Full virulence of PsmM6 requires pFKN and is essentially caused by avrRpm1

To determine the contribution to virulence of the plasmid pFKN, we cured *Psm* M6 of pFKN (see *Experimental procedures*) to generate strain *Psm* M6C. There was no obvious macroscopic difference in visible disease symptoms after inoculation of either wild-type or *Psm* M6C onto *rpm1-3 Arabidopsis* leaves (unable to recognize *avrRpm1*). At high levels of inoculation (5×10^6 cfu ml^{-1}), the inoculated leaves exhibited water soaking and tissue collapse starting at 24 h after inoculation. Inoculation at 5×10^5 cfu ml^{-1} led to chlorosis between 48 and 72 h after inoculation. Using this low-density inoculum, we measured pathogen growth over time in *rpm1-3* leaves. *Psm* M6C grew 10-fold less than the wild-type strain (Fig. 2A). Therefore, the cured strain is less virulent than the wild-type strain. We performed complementation experiments using various pFKN subclones in *Psm* M6C to determine which pFKN ORFs would partially or fully restore virulence. Only clones carrying *avrRpm1* could restore full virulence (Fig. 2A). This result is consistent with insertional mutation of *avrRpm1* in the *Psm* M2 strain (Ritter and Dangl, 1995). No other clone conferred enhanced virulence (data not shown).

The role of *avrPphE* in virulence could not be assessed using *Psm* M6C because a second copy of the gene is present on the chromosome of the cured strain [as assessed by polymerase chain reaction (PCR) amplification and sequencing of the PCR product, see below]. We replaced the *avrPphE* copy on the chromosome of *Psm* M6C with an omega fragment to create *Psm* M6CKO (see *Experimental procedures*). *Psm* M6CKO grew similarly to *Psm* M6C on *rpm1-3* plants (Fig. 2B), suggesting that *avrPphE* plays no major role in the virulence of *Psm* M6. However, our recent results indicate that expression of *avrPphE* from the *lac-nptII* promoter in PsmM6CKO does enhance virulence of PsmM6CKO, suggesting that the quantity of AvrPphE protein delivered to the host cell may be important in its function (J. Shock, J. Chang, Z. Nimchuk, L. Rohmer and J. L. Dangl, unpublished). As *avrPphE* was located in the vicinity of *avrRpm1* on pFKN, we tested the potential involvement of *avrPphE* in the viru-

Table 2. pFKN ORFs sorted by descending G+C content.

ORF element[a]	Region:	Putative function of the product of the ORF according to homologies or closest homolog	GC content	fragment # in the genome of *Pst* DC3000[b]
ORF1	44 - 1354	replication protein (*P. syringae*)	60.59	4 (at least 2 copies)
ORF9	12262 - 13491	transposase1 of IS801 (*P.syringae*)	60.32	one copy
ORF29	38759 - 39184	putative transcriptional regulator (*Streptomyces coelicolor* A3(2))	60.32	4 (at least 2 copies)
ORF28	36025 - 38628	Ct:Ct of TraE in *Escherichia coli* (strain:K-12) plasmid R721, Nt:VirB1 (*Brucella suis*)	60.02	none
ORF13	19220 - 20743	L0015 *(Escherichia coli)*, carried on an IS element	59.79	none
ORF14	20830 - 21363	L0014 *(Escherichia coli)*, carried on an IS element	59.79	none
ORF11	15246 - 15884	ORF H0801 (*Halobacterium sp.* NRC-1)	59.15	none
ORF22	29775 - 30734	integrase/recombinase ripX (*Bacillus subtilis*)	58.16	2
ORF6	8116 - 9129	UV light resistance prot RulB (*P. syringae)*, truncated	57.99	at least 2 copies
ORF18	25019 - 26185	malic acid transport protein (*Methanococcus jannaschii*)	56.64	2
ORF3	2888 - 4441	Transposase of Tn21 (*E. coli*), truncated	55.83	1
ORF4	4959 - 6944	ORFE+ORFD neighbouring *avrPpiA1* in *P.s.* p.v *pisi*, slight homology to phage proteins and transposases	54.58	at least 2 copies
ORF19	26303 - 27196	putative transcriptional regulator LYSR-type (*Pseudomonas aeruginosa*)	54.36	2
ORF20	27294 - 28346	ORF492, surface antigen gene (*Methanosarcina mazei*)	54.32	2
ORF21	28676 - 29053	putative lipoprotein (*Streptomyces coelicolor* A3(2))	53.7	2
ORF12	16897 - 18930	methyl-accepting chemotaxis protein (*Pseudomonas aeruginosa*)	53.05	none
ORF15	23110 - 23490	CRCB protein (*E coli*), transmembrane protein	53.02	none
ORF2	1599 - 2750	homolog of AvrPphE in *P.s.* p.v *syringae*	52.91	1
ORF27	35131 - 35694	transcriptional activator RFAH (*Salmonella typhimurium* LT2)	52.48	3
ORF16	23860 - 24402	inorganic pyrophosphatase (*Pseudomonas aeruginosa*)	52.46	2
ORF17	24396 - 25004	conserved hypothetical protein (*Deinococcus radiodurans*)	51.83	2
ORF23	30984 - 31682	putative gntR-family transcriptional regulator (*Streptomyces coelicolor* A3(2))	51.07	3
ORF7	10008-10436		50.5	none
ORF8	10272 - 10904		50.5	none
ORF10	14114 - 14761	putative partition protein	50.46	none
ORF24	31754 - 32944	putative transcriptional regulator (*Vibrio cholerae*)	48.53	3
ORF25	32988 - 34331	putative integral membrane sugar transporter (*Amycolatopsis orientalis*)	47.47	3
ORF26	34626 - 34994	repressor protein (Bacteriophage Tuc2009)	46.88	3
ORF5	7138 - 7797	*AvrRpm1*	44.29	none

a. Some pFKN ORFs are contiguous on the plasmid and are collinear in this table, indicating that they possess a similar G+C content. When this is the case, an empty row separates them from the next ORF in the table.
b. The DNA fragments from the *Pst* DC3000 genome that contain sequences homologous to ORFs from pFKN were colour-coded (#1: green, #2: red, #3: blue, #4 pink). The table indicates that fragment #4 has a relatively high G+C content, whereas fragment #3 has a relatively low content.

lence function of *avrRpm1*. We compared the growth of *Psm* M6C and *Psm* M6CKO when each carried *avrRpm1* on a plasmid. No difference in growth was observed (Fig. 2B), indicating that *avrPphE* does not affect *avrRpm1* virulence function.

Identification of additional genes putatively involved in pathogenicity based on their expression pattern

The expression of known *P. syringae* virulence factors is specifically induced *in planta*. Minimal media (MM) plus a sugar source seems to mimic the plant apoplast, leading to the expression of the type III pilus genes and effectors (Huynh *et al.*, 1989; Innes *et al.*, 1993). We compared the expression of the pFKN ORFs in MM and in rich medium (KB) using reverse transcription RT-PCR methods. When the expression of an ORF was observed in both media, we assumed that the gene was constitutively expressed. Expression of a gene in MM but not in KB may indicate a role for this gene during interaction with the host. Most of the pFKN genes are constitutively expressed. However, the expression of both *avrRpm1* (Ritter and Dangl, 1995) and *avrPphE* was detected in MM, but not in KB. Additionally, ORF26 is expressed in MM but not in KB. ORFs 14 and 24 were not expressed in either of the two conditions. The expression of ORF8

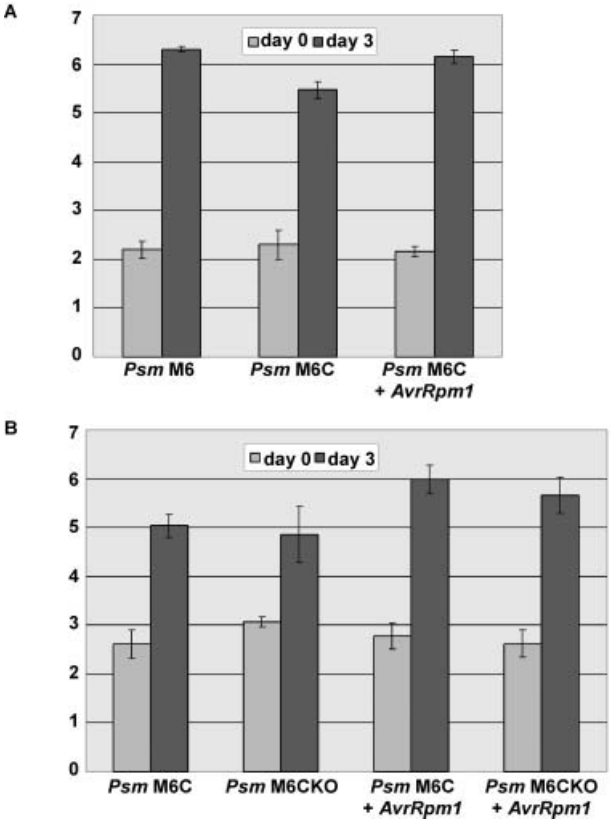

Fig. 2. *avrRpm1* encodes the determining virulence factor on pFKN.
A. The growth of *Psm* M6C on *rpm1-3* plants was compared with the growth of the wild-type *Psm* M6 strain and *Psm* M6C strains complemented with different fragments from pFKN. The pFKN subclone carrying *avrRpm1* restored full virulence to *Psm* M6C (the other subclones did not restore any virulence and are not shown). Leaves were hand inoculated with 10^5 cfu ml^{-1}, and a titration of bacteria in the leaves was performed 3 days after inoculation. Data points represent the mean ± standard error of four samples.
B. The *avrPphE* gene was deleted from the chromosome of *Psm* M6C to generate *Psm* M6CKO. To assess the virulence function of *avrPphE* or its influence on the virulence function of *avrRpm1*, the growth of *Psm* M6CKO on *rpm1-3* plants was compared with the growth of *Psm* M6C, *Psm* M6C + *avrRpm1* and *Psm* M6CKO + *avrRpm1*. Data points represent the mean ± standard error of four samples.

and ORF11 could not be assessed using this technique despite many attempts.

The six genes (*avrRpm1*, *avrPphE*, ORF8, ORF11, ORF14 and ORF24) for which no constitutive expression was observed by RT-PCR were selected for further study. We constructed transcriptional fusions between the 5′ end of the genes of interest and a gene encoding green fluorescent protein (GFP) and chose ORF10 as a constitutive control. The fluorescence of *Psm*M6C carrying the transcriptional fusions was assessed in KB, MM and *in planta* by exposing the bacteria grown in these conditions to blue light (488 nm) and observing them by fluorescence microscopy. A strain carrying promoterless GFP did not exhibit any fluorescence in the conditions tested, whereas all cells carrying the ORF10–GFP fusion did. Only the *avrPphE*, *avrRpm1* and ORF8 GFP fusions were active in plants (Table 1).

pFKN can integrate into the Psm *M6 chromosome*

One colony of *Psm* M6, derived during routine maintenance, exhibited unusual features. Although this strain triggered recognition (measured by HR) on *RPM1* plants and disease symptoms on *rpm1* plants, no plasmid could be isolated from it. *AvrRpm1* was still present in total DNA, as assayed by DNA blot (Fig. 3A). We therefore concluded that *avrRpm1* was now located on the *Psm* M6 chromosome. To determine whether additional portions of pFKN were also present on the variance chromosome, total DNA from the variant *Psm* M6 was digested with *Eco*RI, *Hin*dIII, *Eco*RV, *Sac*I, *Pvu*II, *Sma*I, *Bgl*II and *Nsi*I and probed with the DNA from different parts of the pFKN plasmid. The DNA of *Psm* M6 containing pFKN was used as a positive control. This analysis revealed a polymorphism between the *avrRpm1* gene in this strain (e.g. an 8.0 kb *Bgl*II fragment) compared with pFKN (e.g. an 8.8 kb *Bgl*II fragment; Fig. 3A). All restriction fragments from pFKN were detected in the DNA of *Psm* M6integron. Therefore, pFKN has integrated into the chromosome. For the purpose of clarity, the strain bearing *avrRpm1* in the chromosome will be referred to as *Psm* M6integron.

We have, on occasion, isolated a replicon from individual colonies of *Psm* M6integron (see next section) The plasmid purified from these bacteria always exhibited a restriction profile identical to that of pFKN. Therefore, pFKN can excise from the chromosome, and *Psm* M6integron can revert to *Psm* M6. Interestingly, the overall fitness of *Psm* M6 is affected by the integration status of pFKN into the chromosome. *Psm* M6integron grows in media (KB and MM) more slowly than either the original *Psm* M6 strain or a revertant from *Psm* M6integron that is cured of the plasmid (Fig. 3B). Because *Psm* M6 (with a pFKN replicon) and the plasmid-cured derivative of *Psm* M6integron both grew equivalently, we can conclude that it is the integration of pFKN *per se* that results in this growth impairment. This is not uncommon, as several plasmid replicons are known to slow growth of the organism when integrated (see *Discussion*).

The pFKN integration site in Psm *M6integron is the* avrPphE *transposon*

We noticed from the DNA blot that the genomic environment of *avrRpm1* is different in *Psm* M6 and *Psm* M6integron. This is also the case for both copies of *avrPphE* (data not shown), suggesting that the pFKN integration site is near *avrRpm1* and *avrPphE*, and that the

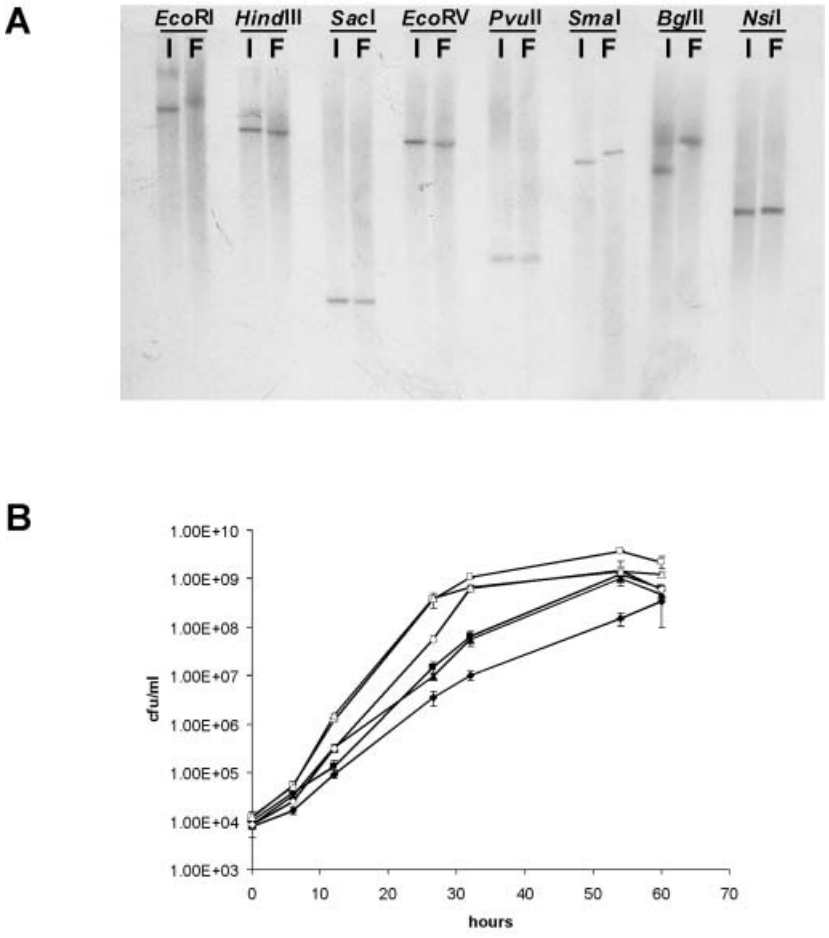

Fig. 3. pFKN can integrate into the *Psm* M6 chromosome.
A. Total DNA from *Psm* M6 (lanes F) and *Psm* M6integron (lanes I) was digested with *Eco*RI, *Hin*dIII, *Eco*RV, *Sac*I, *Pvu*II, *Sma*I, *Bgl*II and *Nsi*I and probed with various pFKN fragments. Here, the DNA blot was probed with the *avrRpm1* ORF. The two DNA samples exhibit a similar pattern except for a polymorphism (e.g. *Eco*RI, *Bgl*II and *Sma*I) near the site of pFKN integration into the chromosome.
B. The integration of the plasmid alters the overall fitness of *Psm* M6integron. The growth of *Psm* M6integron in KB (empty symbols) and MM (filled symbols) was compared with the growth of *Psm* M6 and *Psm* M6C. Bacteria were inoculated at 10^4 cfu ml^{-1}, and cultures were titrated at 0, 6, 12, 29, 32, 54 and 60 h after the inoculation. *Psm* M6pFKN is indicated by squares, *Psm* M6C is indicated by triangles, and rhombi indicate *Psm* M6integron. Data points represent the mean ± standard error of four samples (some error bars are not visible because too small). *Psm* M6pFKN and *Psm* M6C grew equally well in both minimal medium and KB. *Psm* M6integron grows slower than *Psm* M6pFKN and *Psm* M6C in both media.

chromosomal location into which the plasmid integrates is near the chromosomal copy of *avrPphE*. We first defined that the chromosomal *avrPphE* is carried on the same transposon found in the sequence of pFKN, using PCR amplification on total DNA of *Psm* M6C, with primers specific to sequence on the *avrPphE* transposon (Table 5). The PCR products were sequenced and compared with the corresponding sequenced regions of pFKN. As shown in Fig. 4A, the chromosomal copy of the transposon contains two sequences not present on pFKN. One 1244 bp sequence (region A) is located between *avrPphE* and the coding region of *TnpA*. It carries an ORF with a sequence that is nearly identical to various genes encoding resolvases of Tn*3* transposons. Additionally, a portion of 661 bp of this region encodes the 5′ portion of *TnpA*. Another region of 271 bp unique to the chromosomal copy is located within the coding region of the transposase (region C), and is 3′ from a region shared by both pFKN and chromosomal copies of *TnpA* (region B). The addition of regions A and C to the chromosomal copy results in a complete *TnpA* ORF correcting the frameshift at the 3′ end of *TnpA* carried on pFKN.

TnpA, *avrPphE* and the putative resolvase fragment are located between nearly identical inverted repeats of 66 bp on both the plasmid and the chromosome. Their organization is reminiscent of Tn*3*-like transposons (Sherrat, 1989). Moreover, this *TnpA* is 81% identical to those of the Tn*21* subfamily of Tn*3* elements (Diver *et al.*, 1983). As we suspected that the integration site is in the vicinity of *avrPphE* in the chromosome and the plasmid (i.e. in the two transposons), we analysed by PCR the two transposons in the genome of *Psm* M6integron. Figure 4B shows that the unique regions A and C are no longer located on the same transposon, but are instead separated by the entire plasmid sequence (Fig. 4C). This indicated that the integration site of the plasmid is located in the region between region A and region C, i.e. within the homology provided by region B (Fig. 4).

We used this sequence data to investigate the rate of excision occurring in *Psm* M6integron grown in different conditions by PCR. The rate of excision was estimated at around 10^{-4} (data not shown). No difference in rate of excision could be detected between bacteria

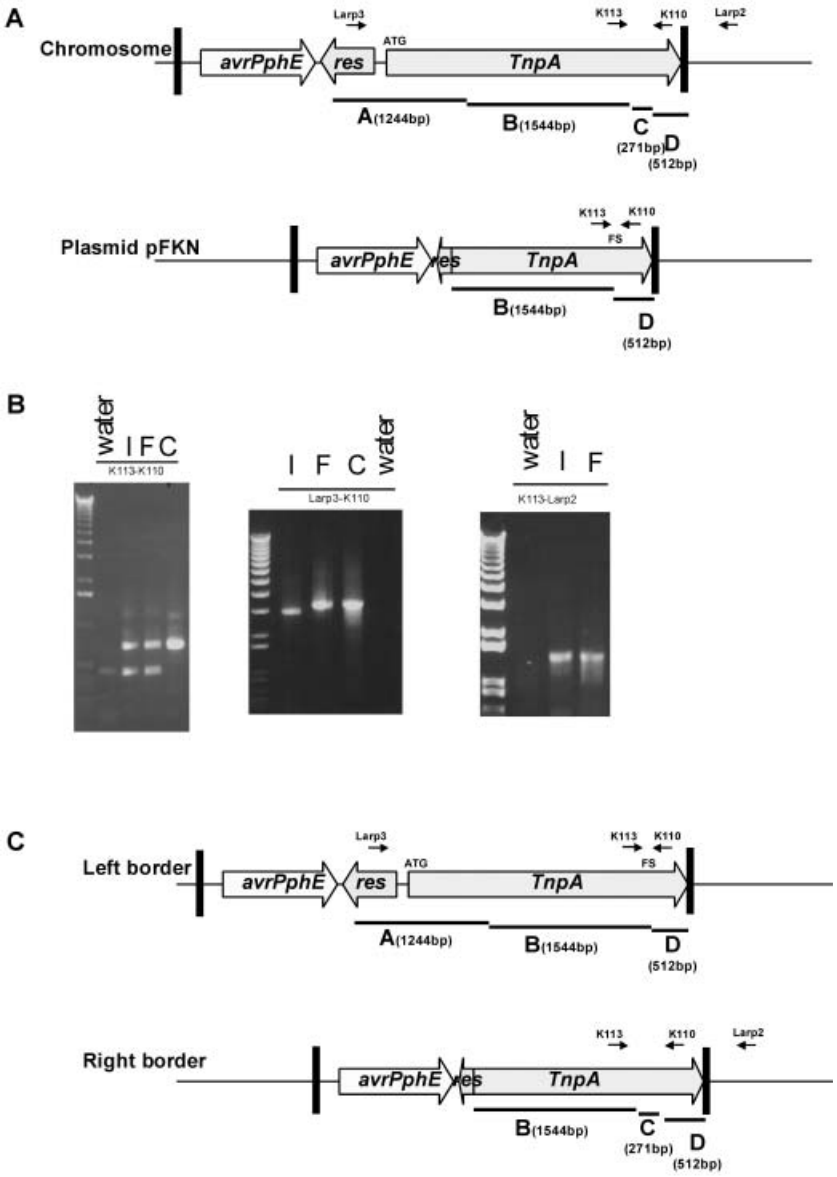

Fig. 4. The pFKN integration site is a second copy of the *avrPphE* transposon.
A. The region in the *Psm* M6 chromosome that hybridizes to the *avrPphE* and *TnpA* probes was amplified and sequenced. The sequence of the chromosomal region is 100% identical to the transposon in the plasmid, but contains two additional regions (regions A and C). *TnpA* encodes a transposase and *res* a resolvase. Based on homology, the chromosomal *avrPphE* transposon could be complete and functional. The *avrPphE* transposon carried on the plasmid lacks the 5′ end of *TnpA* and *res* (region A), and the absence of region C results in the truncation and a frameshift (marked FS) of the gene encoding the transposase. The primers used for PCR amplification (Fig. 4B) are indicated by arrows, and their names are in bold.
B. To investigate where regions A and C were distributed after integration of the plasmid into the chromosome, PCRs were performed on the DNA of *Psm* M6integron and *Psm* M6 using Larp3 and K110, and K110 and K113 (see Fig. 4A). The product amplified from *Psm* M6integron is run in lane I. The product amplified from *Psm* M6 is run in lane F. The product amplified from *Psm* M6C is run in lane C. The difference in size observed between the PCR products from *Psm* M6integron and *Psm* M6 using Larp3 and K110 indicates that regions A and C are not carried on the same putative transposon in the chromosome of *Psm* M6integron (in contrast to the chromosome of *Psm* M6). The PCR products from *Psm* M6integron and *Psm* M6 using K113 and K110 show that neither region A nor region C has been lost in the chromosome of *Psm* M6integron. The identical size of the PCR products from *Psm* M6integron and *Psm* M6 using K113 and Larp2 indicates that region C is on the right border of pFKN when integrated into the chromosome. The region A is on the left border, as it is carried by the other putative transposon.
C. After integration of pFKN into the chromosome, *TnpA* on the left border contains region A, but not region C. *TnpA* on the right border contains region C. The primers used for the PCR amplification are indicated by arrows, and their names are in bold (in Fig. 4A).

grown in KB, MM and on resistant or susceptible plants.

The genomic context of avrPphE *is different in various* P. syringae *strains*

We investigated the regions surrounding the *avrPphE* transposon on the chromosome of *Psm* M6 (Fig. 5) by TAIL-PCR. The left border shows 97% nt identity to the 5′ end of the insertion sequence ISPsy3 (AB063176) from *P. syringae* pv. *eriobotryae*. This element is interrupted by the *avrPphE* transposon. The translation product of the interrupted ORF found on this border is homologous to the IS*801* transposase (79% identity and 86% homology), and IS*801* has been associated with other type III effectors in various *P. syringae* pathovars (Kim *et al*., 1998). On the right border, the 200 bp sequence following the terminal repeat of the *avrPphE* transposon encodes a protein homologous (81% nucleotide identity) to the 3′ end of an unrelated putative transposase gene from *Pseudomonas aeruginosa*. The next 1062 bp sequence downstream shows high nucleotide homology to sequence identified in *P. syringae* pv. *pisi* (Arnold *et al*., 2001). This DNA fragment may be a transposon remnant and carries the *avrPpiA1* allele of *avrRpm1* from *P. syringae* pv. *pisi*, among other genes. In the chromosome of *Psm* M6, this sequence shows homology to the *P. s. pisi* genes ORFD (94% identity), encoding a putative trans-

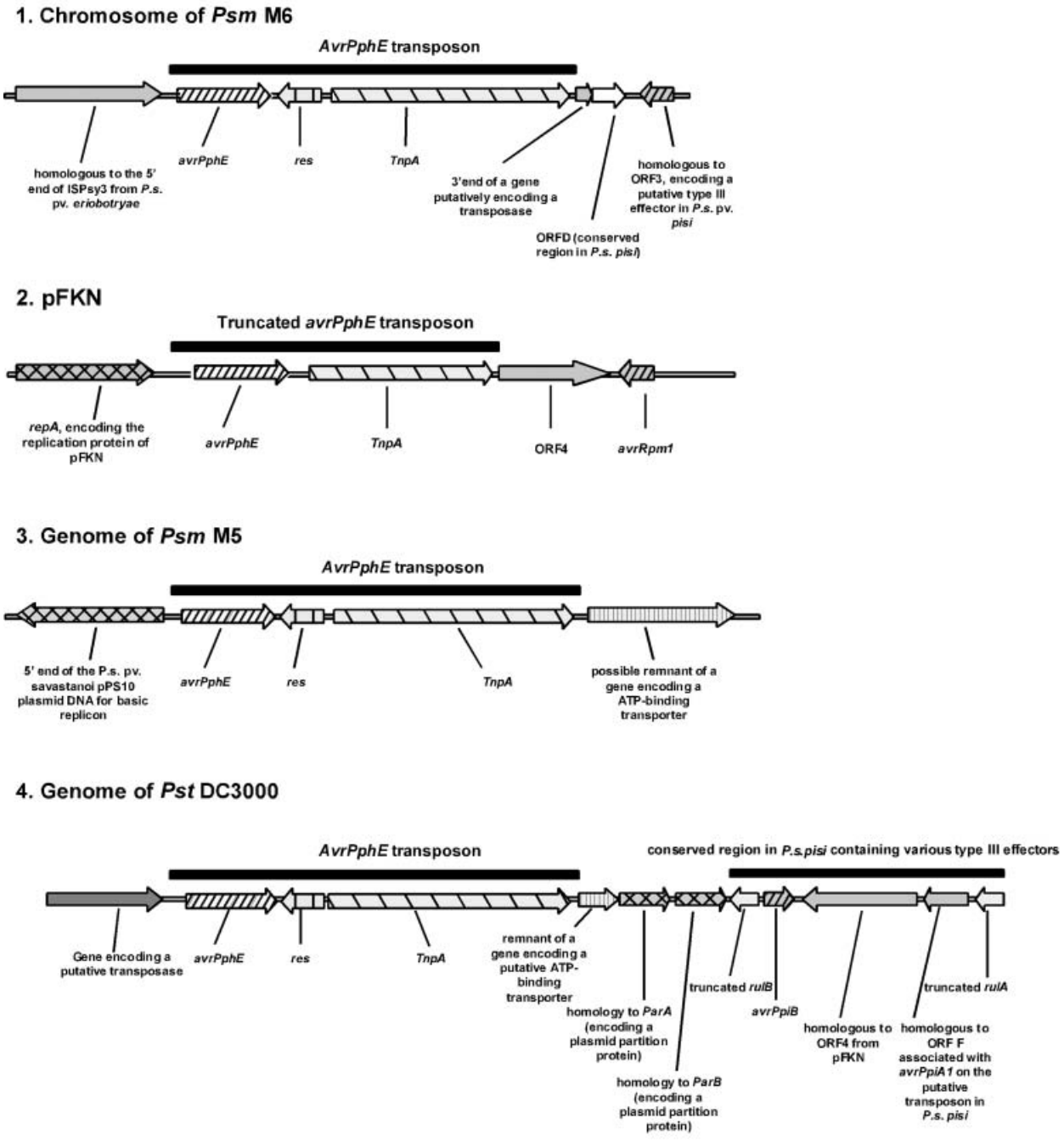

Fig. 5. Sequences surrounding the *avrPphE* transposon define four different genetic contexts. The figure is a graphical representation of the four genomic environments in which the *avrPphE* transposon was detected. The regions represented were characterized by sequence comparison using the BLASTX algorithm. It is not known whether the *avrPphE* transposon is located in the chromosome or on a plasmid in the genome of *Psm* M5 and *Pst* DC3000.

posase, and ORF3, encoding a putative type III effector (100% identity). Thus, this region in *Psm* M6 is a remnant of the putative transposon found in *P. s. pisi*.

We detected the *avrPphE* transposon in other strains of *P. syringae* pv. *maculicola* and in the sequence of the genome of *P. s. tomato* DC3000. In *Psm* M2, *Psm* M8, *Psm* M12, *Psm* m90 and *Psm* m10831, the borders are the same as in *Psm* M6 (data not shown). On the other hand, they differ in *Psm* M5 and *Psm* M9, and these all differ from *Pst* DC3000. In *Psm* M5 (Fig. 5) and *Psm* M9

(not shown), the left border consists of the 5′ end of a gene sharing 99% identity to the replicase of pPS10, a plasmid found in *P. syringae* pv. *savastanoi* (Nieto *et al*., 1992). This replicase is not related to that of pFKN (a PT23A-like plasmid; see above). The right border sequence is a remnant of a gene encoding a putative ABC transporter (23% identity and 43% homology to a putative sugar uptake ABC transporter ATP-binding protein from *Sinorhizobium meliloti*).

In the genome of *Pst* DC3000, we looked at 30 kb surrounding the transposon (Fig. 5). The proteins encoded on this 30 kb region share homology to either DNA mobility functions (transposases, plasmid partitioning proteins, replicases) or type III effectors (identical to *avrPpiB*; Cournoyer *et al*., 1995). The region also contains other sequences often found associated with plasmids and genes involved in pathogenicity, such as *rulA* and *rulB* (Sesma *et al*., 2000). Sequences directly neighbouring the transposon on the left side are homologous to a gene encoding the transposase of the IS*100* elements from *Yersinia pestis*, *Ralstonia solanacearum* and *Salmonella enterica*. On the right-hand side of the *avrPphE* transposon, the first homology is to a fragment of a putative ABC transporter, identical to the one identified in *Psm* M5. This suggests that this part of the two genomes is more recently diverged than either is from *Psm* M6.

No common sequence was identified in the four different environments surrounding the *avrPphE* transposon that would help to identify the integration target of the *avrPphE* transposon, although the proximal right borders in *Psm* M5 and *Pst* DC3000 are nearly identical. This suggests that the *avrPphE* transposon and this bordering sequence were acquired together by these two strains, and that subsequent events at both the distal right flank and on the left border are responsible for the differences in genomic context. In contrast, the insertion sites of the *avrPphE* transposon in *Psm* M5 and *Psm* M6 are completely different. In sum, the features of these integration sites are consistent with the model that PAIs evolve by repeated insertion and deletion into pre-existing PAIs (Hensel *et al*., 1999; Tauschek *et al*., 2002).

Discussion

We sequenced pFKN, an ≈40 kb plasmid that contributes to the virulence of *Psm* M6. The plasmid carries two type III effectors on remnants of transposons, contiguous to ORF8, itself neighbouring an element closely related to IS*801*. *avrRpm1* encodes a known virulence factor and *avrPphE*, a type III effector. The expression of ORF8 is induced *in planta* and, therefore, ORF8 may play a role in the interaction with the host plant. The portion of the plasmid carrying *avrRpm1*, *avrPphE* and ORF8 may thus be considered as a PAI. Interestingly, the G+C content of the ORFs carried on this region are very heterogeneous, and the three genes are flanked by sequences related to mobile elements and, therefore, are likely to have different origins. This is consistent with the idea that the formation of PAIs requires a multistep process (Hensel *et al*., 1999; Alfano *et al*., 2000; Tauschek *et al*., 2002). It is also consistent with the findings of Kim *et al*. (1998) showing that most of the known type III effectors are associated with this type of sequence. pFKN (and therefore the PAI that it carries) sometimes integrates into and can excise from the chromosome, as demonstrated by the discovery of the variant of the strain *Psm* M6. We identified the region of the chromosome into which the integration took place. This region is a transposon carrying a second copy of *avrPphE*.

Alleles of the *avrPphE* gene have been detected in many *P. syringae.* pathovars (Mansfield *et al*., 1994; Alfano *et al*., 2000; Fouts *et al*., 2002; Guttman *et al*., 2002), suggesting that it provides an important virulence function on a variety of host plants. Our findings are the first to associate *avrPphE* with a transposon. In *Psy* B728a and in *Pph* 1302A, for example, *avrPphE* resides directly next to the *hrp* locus, in a region called the 'exchangeable effector locus' or EEL (Alfano *et al*., 2000), and is not associated with a transposase. In fact, in *Psy* B728a, *avrPphE* is predicted to be the third gene in an operon controlled by the *hrpK* promoter (Alfano *et al*., 2000). This striking finding illustrates the dynamism of PAIs.

We determined by complementation assay that the partial loss of virulence in the strain cured of pFKN results from the absence of *avrRpm1*. The virulence function of *avrRpm1* has already been demonstrated in the strain *Psm* M2 (Ritter and Dangl, 1995). Recently, AvrRpm1 has been demonstrated to interact with RIN4, a protein involved in the regulation of some of the plant defence responses (Mackey *et al*., 2002). Mackey and coworkers proposed that the role of AvrRpm1 is to reduce the level of plant defence response triggered by the attack of the pathogen. This hypothesis might be consistent with the fact that only quantitative differences in growth were observed between *Psm* M6C and *Psm* M6. No difference in growth was observed between *Psm* M6C and *Psm* M6CKO (lacking *avrPphE*). Similar results were observed with the *avrPphE* allele of *P. s. phaseolicola* (Stevens *et al*., 1998). This observation could be explained by either a functional redundancy with another gene in the genome of the *Psm* M6 or the fact that the assay is inappropriate to assess the virulence function of this type III effector.

The sequence and the integration mechanism of pFKN into the *Psm* M6 chromosome provide insights on the evolutionary dynamics of the genomes of *P. syringae*

pathovars. First, the analysis of the G+C content of fragments of pFKN suggests that they may have different origins. This hypothesis is further supported by the presence of some defined fragments from pFKN in the genomes of *Pst* DC3000 and *Psy* B728a. The fragment 3 (see Table 2) encodes one putative transcriptional regulator, proteins predicted to be located in the bacterial inner or outer membrane by PSORT (inorganic pyrophosphatase, sugar transporter, putative protein) and a putative integrase/recombinase homologous to *ripX* from *Bacillus subtilis*. The presence of this latter gene suggests that fragment 3 may be the remnant of a bacteriophage.

The integration and excision mechanism is very probably a homologous recombination in the 1.5 kb sequence region B of *TnpA* carried on the *avrPphE* transposon (Fig. 6). It is not clear whether the integration/excision process relies on general recombination or on a site-specific recombination, dependent on an integrase/recombinase. As indicated by the sequence of the two transposons before and after integration of pFKN, no nucleotide sequences were altered at the site of exchange, and not a single nucleotide was lost or gained (this excludes a mechanism of transpositional site-specific recombination). General recombination requires a large DNA region, whereas the site-specific recombination requires as few as 240 bp for the lambda phage for example (Thompson and Landy, 1989). The rate of general recombination is unknown for *P. syringae*, but has been estimated at 10^{-3} in *Pseudomonas tolaasii* (and, when *recA* is inactivated, at 10^{-4}; Sinha *et al*., 2000). The rate of excision of pFKN has been estimated between 10^{-3} and 10^{-4} and would therefore be consistent with a mechanism of general recombination. Additionally, the region B is 1.5 kb long, enough to ensure general recombination. Older reports describe the integration of pMMC7501 into the chromosome of *P. s.* pv. *phaseolicola* (Curiale and Mills, 1982). An imprecise excision of pMMC7501 led to the formation of eight derivative plasmids, all sharing the same origin of replication (Szabo and Mills, 1984a). This excision occurred through recombination events between sets of repeats in the chromosome exhibiting sequence homology with IS elements (Szabo and Mills, 1984b; Poplawsky and Mills, 1987).

The integration of a virulence gene into the chromosome may be a first step towards its stabilization in the genome. Indeed, lack of selection can lead to the stochastic loss of the plasmid, and many steps in the pathogen life cycle probably do not require the presence of virulence functions. The loss of a plasmid carrying the type III effector *avrPpiB* was observed in *P. syringae* pv. *pisi* (Bavage *et al*., 1991). We propose that the integration of pFKN could lead to the stabilization of *avrRpm1* in the genome of *Psm* M6. Interestingly, a DNA segment of at least 42 kb carrying *avrPphB* in *P. syringae* pv. *phaseolicola* can excise from the chromosome, but it is not maintained in the strain as a plasmid. This region may represent an ancient plasmid integration event, followed by successive genomic rearrangements deleterious for plasmid maintenance (Jackson *et al*., 2000). The fact that we can still observe the excision of pFKN from the chromosome of *Psm* M6integron shows that the newly generated genetic unit is not rapidly stabilized in the genome, and that further events might be necessary to prevent excision. Furthermore, the integration of pFKN into the chromosome slightly impairs the fitness of *Psm* M6integron. The presence of the functional origin of replication of pFKN may interfere with the replication of the chromosome. Integra-

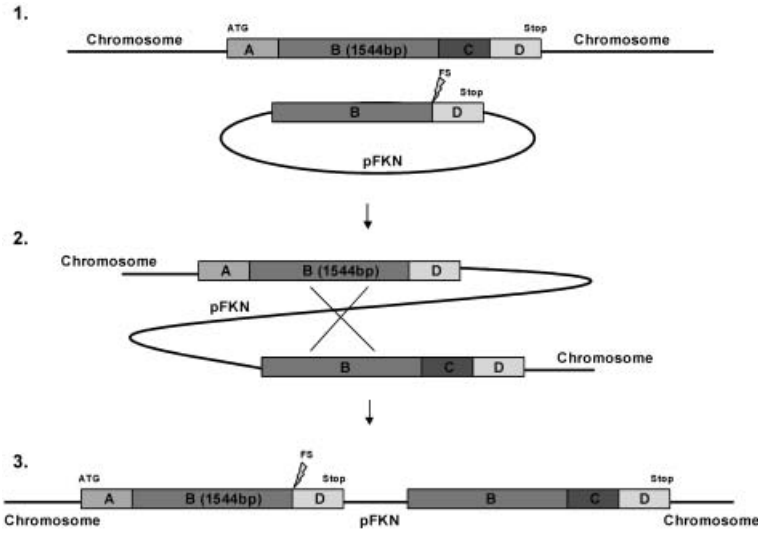

Fig. 6. The integration of pFKN may have occurred via homologous recombination between the regions B common to the two *TnpA* genes. The *TnpA* genes are constituted of four DNA sequence regions: (1) before the integration, the chromosomal copy of *TnpA* contains all four regions (A, B, C and D). The *TnpA* on the plasmid is missing the 5′ end of the gene (region A) and a fragment towards the 3′ end (region C), which results in a frameshift (FS). After the homologous recombination between two identical regions (region B) in the coding region of the *TnpA* genes (2), the region A is found in the *TnpA* of the left border and region C in the *TnpA* of the right border (3).

tion of multicopy plasmids is commonly lethal to *E. coli* unless the replication origin of the integrated element is suppressed (Yamaguchi and Tomizawa, 1980; Froehlich *et al*., 1983).

pFKN is not the only genetic unit potentially stabilized by integration. The chromosomal *avrPphE* transposon that acts as the insertion site for pFKN may be active before integration. The chromosomal *TnpA* has a full ORF, and we observed transcription in both KB and MM using RT-PCR (data not shown). The pFKN integration event disrupts the chromosomal *avrPphE* transposon, thus preventing autonomous transposition. As transposition can be detrimental, the stabilization of the *avrPphE* transposon could be advantageous for the bacterium.

Strikingly, pFKN integrates into a region of the *Psm* M6 chromosome already encoding two type III effectors (*avrPphE* and ORF3 identified in *Ppi*). The *avrPphE* transposon has been detected at four different loci (chromosomal and on pFKN in *Psm* M6, one in *Psm* M5 and one in *Pst* DC3000). In all cases, genes related to DNA mobility were identified in the vicinity. In at least three out of four cases (*Psm* M6 chromosome, pFKN and *Pst* DC3000), additional genes encoding type III effectors were also found in the vicinity of the *avrPphE* transposon (Fig. 5). The sequence available for *Psm*M5 is the shortest, and we detected no homology to virulence genes in the flanking sequence. However, part of the sequence on the right border is 100% identical to the immediate right border identified in *Pst* DC3000. It is likely that the larger genomic context of the *avrPphE* transposon in both strains has a common ancestor (or was transferred from one strain to the other). Finally, the transposon in *Pst* DC3000 and *Psm* M5 is 100% identical to the transposon in *Psm* M6. This suggests that the acquisition of the transposon was rather recent. As a comparison, *avrRpm1*, its allele *avrPpiA1* and their surrounding sequences are 97% identical in nucleotide sequence. The rather recent acquisition of *avrPphE* and its association with other genes related to pathogenicity and DNA mobility may suggest that the *avrPphE* transposon is part of a PAI in formation. The integration of pFKN into this site would be the next step in the process. The proposed mechanism of integration of pFKN suggests that this plasmid could integrate in the chromosome of any strain that carries the *avrPphE* transposon (provided that the strain acquires the plasmid). Transfer of genes related to pathogenicity or adaptation to various conditions may be facilitated or even explained by the integration of plasmids in the bacterial chromosome followed by the imprecise excision of plasmids. Excised plasmids would contain all or part of the integrated plasmid and regions from the chromosome. The transfer of these plasmids could explain how different strains share nearly identical regions in their respective genomes.

Experimental procedures

Bacterial strains and culture conditions

The bacterial strains used in this study are shown in Table 3 and the recombinant plasmids in Table 4. The *Pseudomonas* strains were grown in King's B media (KB) (King *et al*., 1954) shaken at 28°C or grown on KB agar plates at 28°C. The *E. coli* strains were grown in 2× YT broth or on Luria–Bertani (LB) broth agar plates at 37°C (Maniatis *et al*., 1989). For *Pseudomonas* strains, antibiotics were used at the following concentrations (mg l^{-1}): rifampicin, 25–50; kanamycin, 50; carbenicillin, 100; spectinomycin, 10. Plates contained cycloheximide (50 mg l^{-1}). For *E. coli* strains, antibiotics were used as follows (mg l^{-1}): ampicillin, 100; nalidixic acid, 10; kanamycin, 100; spectinomycin 100.

DNA manipulations

Plasmid DNA isolations were performed by alkaline lysis extraction as described by Maniatis *et al*. (1989). Genomic DNA isolations were performed as described by Syn and Swarup (2000). All molecular manipulations were done via standard procedures. Enzymes were purchased from New England BioLabs or Boehringer Mannheim and used according to the manufacturer's specifications. For Southern analysis, ≈2 μg of DNA was loaded per lane on agarose gel. DNA was transferred to Hybond N+ membranes (Amersham Pharmacia Biotech), and hybridization was performed in hybridization solution [1.5× SSC, 7% SDS, 10% polyethylene glycol (PEG) molecular weight 8000, 100 μg ml^{-1} sonicated and denatured herring sperm DNA, 250 μg ml^{-1} heparin] at 65°C. The DNA probes were labelled with α-ATP using reagents provided in the Prime-It II random primer labelling kit (Stratagene).

Growth curves and inoculations

To assess the growth of bacteria *in planta*, bacteria were inoculated at a density of 10^5 cfu ml^{-1} in 10 mM $MgCl_2$ on four leaves per plant, and four leaf discs from individual plants were collected at day 0 (to assess the original inoculum), day 1, day 3 and day 5. The discs were ground in 1 ml of 10 mM $MgCl_2$, and serial dilutions were plated on KB plates to titrate the bacteria. Four technical repetitions were performed per experiment (Ritter and Dangl, 1995).

Shotgun cloning and DNA sequencing

Purified pFKN DNA (2.25 μg) was partially digested with *Tsp*5091. Fragments ranging in size from 2 to 8 kb were isolated from an agarose gel with a QIAquick gel isolation kit (Qiagen). The extracted fragments were ligated into *Eco*RI-digested pUC18 (Pharmacia). The ligation was transformed into electrocompetent *E. coli* strain DH5α (Electro Cell Manipulator^R 600 from BTX, 50 μF, 129 Ω, 2.5 kV). This method has been described by Ansorge *et al*. (1997). The inserts of 150 randomly picked clones were sequenced using ABI dye terminator chemistry by the UNC-CH Automated DNA sequencing facility with M13F (TTCAGGGAGCCTGCG

Table 3. Bacterial strains used in this study.

Species or pathovar	Strain or race	Description	Source or accession no.
E. coli	Dh5α		Life Technology
P. syringae pv. *maculicola*	M6		LMG 5560
P. syringae pv. *maculicola*	M6integron	Variant of *Psm* M6 cured of pFKN in which the plasmid pFKN has integrated the chromosome	This study
P. syringae pv. *maculicola*	M6C	*Psm* M6 cured of pFKN	This study
P. syringae pv. *maculicola*	M2		LMG 5071
P. syringae pv. *maculicola*	M5		LMG 5559
P. syringae pv. *maculicola*	M8		IHR(W) 793
P. syringae pv. *maculicola*	M9		IHR(W) 1120B
P. syringae pv. *maculicola*	M12		IHR(W) 1809A
P. syringae pv. *maculicola*	M90		From Brian Staskawicz (University of California, Berkeley)
P. syringae pv. *maculicola*	M10832		From Brian Staskawicz (University of California, Berkeley)
P. syringae pv. *maculicola*	M6CAR1	*Psm* M6 cured of pFKN carrying pDSKAR1, encoding an HA-tagged version of AvrRpm1	This study
P. syringae pv. *maculicola*	M6CKO	*Psm* M6 cured of pFKN in which the chromosomal version of *avrPphE* has been replaced with an omega fragment	This study
P. syringae pv. *maculicola*	M6CKOAR1	*Psm* M6 cured of pFKN in which the chromosomal version of *avrPphE* has been deleted, carrying pDSKAR1	This study
P. syringae pv. *maculicola*	M6CpLaR9	*Psm* M6 cured of pFKN carrying pLaR9	This study
P. syringae pv. *maculicola*	M6CpLaR10	*Psm* M6 cured of pFKN carrying pLaR10	This study
P. syringae pv. *maculicola*	M6CpLaR11	*Psm* M6 cured of pFKN carrying pLaR11	This study
P. syringae pv. *maculicola*	M6CpLaR12	*Psm* M6 cured of pFKN carrying pLaR12	This study
P. syringae pv. *maculicola*	M6CpLaR13	*Psm* M6 cured of pFKN carrying pLaR13	This study

GTCC) and M13R (AACAGCTATGACCATG) primers. For gap closure, custom primers were designed from the ends of each contig and used to amplify PCR products with purified plasmid pFKN DNA as a template. The PCR products were purified using a High Pure PCR product kit (Hoffmann-La Roche) and sequenced directly by primer walking. Sequenced fragments were analysed and assembled using SEQUENCHER (Gene Codes). The coverage was on average 3.1× and minimum 2×. ORF determination was made manually and checked with GLIMMER (Delcher *et al*., 1999). The ORFs should meet the following criteria: (i) the start codon was ATG or GTG; (ii) the stop codon was TAA, TAG or TGA;

Table 4. Plasmids used in this study.

Plasmids	Description	Source or reference
pFKN	40 kb native plasmid in *Psm* M6	This study, AF359557
pUC18	cloning vector	Norrander *et al*. (1983)
pFK20	pFKN fragment (37174–1425) generated in the sequencing shotgun library, carrying the replication protein of pFKN	This study
pBBR1-MCS2	Broad-host-range vector	Kovack *et al*. (1995)
pRK2013	Helper plasmid	Figurski and Helinski (1979)
pDSKAR1	*avrRpm1* gene fused to an HA-tag carried on pDSK519	J. Chang, UNC
pLaR9	pFKN *Sca*I–*Sma*I fragment cloned into pBBR1-MCS2, carrying ORF7 to ORF11	This study
pLaR10	pFKN *Apa*I–*Bam*HI fragment cloned into pBBR1-MCS2, carrying ORF10 to ORF14	This study
pLaR11	pFKN *Eco*RI–*Bam*HI fragment cloned into pBBR1-MCS2, carrying ORF15 to ORF22	This study
pLaR12	pFKN *Eco*RI fragment cloned into pBBR1-MCS2, carrying ORF23 to ORF26	This study
pLaR13	pFKN *Sac*I–*Sac*II fragment cloned into pBBR1-MCS2, carrying ORF27 to ORF29	This study
pLaR24	Omega fragment carrying the spectinomycin resistance gene (*aadA+*) cloned into the sequences surrounding *avrPphE* in the chromosome or *Psm* M6, 1 kb on each side, carried on pBBR1-MCS2	This study
p186	vector for translation fusion (on the first reading frame) with the C terminus of AvrRpt2 and translational fusion with *GFP3* on a pBBR1-MCS2 backbone	J. Chang (UNC)

(iii) the size of the ORF was over 300 bp. We may not have predicted the correct initiation codon. Where possible, we chose an initiation codon that is preceded by an upstream ribosome binding site (RBS) sequence (optimally 5–13 bp before the initiation codon). The algorithms BLASTN and BLASTX were used to compare pFKN DNA sequences with those in the GenBank database (http://www.ncbi.nlm.nih.gov/blast/). The predicted proteins were also analysed for functionally important motifs using PROSITE and for functional and structural domains using PFAM (http://www.sanger.ac.uk/Software/Pfam/search.shtml), PRODOM (http://protein.toulouse.inra.fr/prodom/doc/prodom.html), SMART (http://smart.embl-heidelberg.de/) and BLOCKS (http://blocks.fhcrc.org/blocks/blocks_search.html). The sequence data have been submitted to the DDJ/EMBL/GenBank databases under the accession number AF359557. The sequence of pFKN was compared with the genome sequence of *P. syringae* pv. *tomato* DC3000, available at http://tigrblast.tigr.org/ufmg/, and *P. syringae* pv. *syringae* B728a, available at http://www.jgi.doe.gov/JGI_microbial/html/index.html.

Plasmid curing and complementation

The subclone pFK20 from the shotgun library carries the replicase gene (ORF1) of pFKN that is also an incompatibility determinant for plasmids belonging to the PT23A plasmid family (Gibbon *et al*., 1999). pFK20 was electroporated into *Psm* M6 to cure the strain of pFKN. Single transformants were grown in three consecutive liquid cultures without selection for pFK20 (carbenicillin). The bacteria retrieved after growth no longer contained pFK20, indicating that the fragment carried by pFK20 was not sufficient for plasmid maintenance. No pFKN sequences could be detected by DNA blot.

For complementation tests, fragments of pFKN were cloned into pBBR1-MCS2 (pLaR9, pLaR10, pLaR11, pLaR12 and pLaR13; see Table 4) and transferred to the cured strain by triparental mating, using the helper plasmid pRK2013.

We deleted the *avrPphE* gene from the cured strain by marker exchange as described by Hendrickson *et al*. (2000). In brief, *avrPphE* was replaced by an omega fragment carrying a gene conferring resistance to spectinomycin and streptomycin (Prentki and Krisch, 1984). Sequences flanking the gene on both sides (1 kb each) were cloned into pBBR1-MCS2, carrying *sacB* (a lethal gene when the bacteria is grown on sucrose; Zagorec and Steinmetz, 1990). Between the two flanking regions was inserted the omega fragment carrying a gene conferring resistance to spectinomycin and streptomycin (Prentki and Krisch, 1984). The resulting plasmid (pLaR24) was transformed by triparental mating into *Psm* M6C using the helper plasmid pRK2013. Deletion mutants were then selected on spectinomycin (for integration of the omega fragment) and sucrose (for loss of pLaR24).

RT-PCR analysis

RNA was extracted from exponentially growing *Psm* M6 as described by Innes *et al*. (1993). The bacteria were subjected to growth in the test medium for 3 h. The RNA was then extracted using Trizol^{R} LS reagent (Gibco BRL). RNA samples were treated with 10 U of RNase-free DNase I (Hoffmann-La Roche) for 1 h at 37°C according to the manufacturer's instructions. The enzyme was then inactivated for 5 min at 75°C. Reverse transcription reaction was performed using the RETROscript^{TM} kit (Ambion), according to the protocol provided by the manufacturer. PCRs included primers specific for each ORF to be tested and control primers allowing the amplification of part of the gene encoding for the glyceraldehyde-3P-dehydrogenase. The amplification with the control primers indicates that the non-amplification of the tested ORF results from the lack of expression of the gene and not a failed reaction. The primer sets used for this control were Larp17 (CAGCAACTTGCCGCTCAC) and Larp10 (ATGACTCTCCGTATCGC), giving rise to an amplification product of 1070 bp, or Larp17 and Larp11 (CTGCACCGT GAACTGGG), giving rise to a product of 406 bp. The following protocol was used: 3 min at 95°C, 30 cycles of 45 s at 94°C, 45 s at 54°C and 1 min at 72°C, and 5 min at 72°C. The primer sequences used for each ORF are available on request.

Transcriptional fusion and expression assay in planta

The 5′ ends of the genes for testing were cloned into p186 (J. Chang, personal communication) to fuse the N-termini of the genes for testing to the gene encoding GFP3. *Psm* M6C was transformed with the constructs by triparental mating between *Psm* M6C, DH5α carrying the construct or the helper plasmid pRK2013. Leaves of eight *rpm1-3* plants were inoculated with each transformant at a density of 5×10^7 cfu ml^{-1}. The bacteria were retrieved from the inoculated leaves by shaking the leaves in 10 mM $MgCl_2$ and 0.02% Silwet. They were resuspended in glycerol and immediately observed by fluorescence microscopy (Microscope Eclipse E800; Nikon). For the GFP to be activated, the bacteria were exposed to blue light (488 nm).

TAIL-PCR method

The TAIL-PCR was a modified version of the method described by Liu and Whittier (1995). Primers used are listed in Table 5. The degenerate primers (AD1 to AD8; Table 5) were provided by P. Tornero (UNC, Chapel Hill, NC, USA). On average, three to five degenerate primers yielded one specific amplification product. The biggest products were ≈1 kb. PCR products were isolated on a 0.8% agarose gel (QIAquick gel isolation kit; Qiagen) and sequenced directly. The surrounding sequences of the *avrPphE* transposon in the chromosome of *Psm* M6 are available in GenBank (AF544992), as well as the sequences from *Psm* M5 (AF544991).

Acknowledgements

We thank Drs Jeff Chang, Petra Epple, Sarah R. Grant, Gopal Subramaniam and Ben F. Holt, III, and Zack Nimchuk for their patience and advice during the course of this work. We are especially grateful to The Institute for Genome Research

Table 5. Primers used in this study.

Primer name	Sequence (5′–3′)	Purpose
fkp1405 <	GTGAAATCGCTTAGCGTACTTT	TAIL-PCR on the left border of the *avrPphE* transposon
fkp1444 <	AGCGGCAACAGCGCGGGGAAA	TAIL-PCR on the left border of the *avrPphE* transposon
fkp1509 <	GCCATAGCCGCTGAAGCAACATT	TAIL-PCR on the left border of the *avrPphE* transposon
fkp1559 <	TGTCGTCATTTGGCCGGGTC	TAIL-PCR on the left border of the *avrPphE* transposon
fkp4637 >	TGGGCGAAATCCGAGACCGTAGCT	TAIL-PCR on the right border of the *avrPphE* transposon
fkp4802 >	TGTCGCCGCTGGGCTGGG	TAIL-PCR on the right border of the *avrPphE* transposon
fkp6431 >	CGCCTCGGTCGTGACTGCATTG	TAIL-PCR on the right border of the *avrPphE* transposon
K110	TCGGCACTTTGTAGCCAGT	Sequencing of the *avrPphE* transposon
K113	GTCTTGCCCTGGCACCAC	Sequencing of the *avrPphE* transposon
Larp2	AATAGGTTTACACCTGTTTCACG	Sequencing of the *avrPphE* transposon
Larp3	TTTCACGCCCTTGGCCTTGG	Sequencing of the *avrPphE* transposon
AD1	NTCGASTWTSGWGTT	Degenerated primers for TAIL-PCR
AD2	NGTCGASWGANAWGAA	Degenerated primers for TAIL-PCR
AD3	WGTGNAGWANCANAGA	Degenerated primers for TAIL-PCR
AD4	NCTAGWASTWGSTTG	Degenerated primers for TAIL-PCR
AD5	NTGGCGWSATNTSATA	Degenerated primers for TAIL-PCR
AD6	TGWGNAGSANCASAGA	Degenerated primers for TAIL-PCR
AD7	AGWGNAGWANCAWAGG	Degenerated primers for TAIL-PCR
AD8	STTGNTASTNCTNTGC	Degenerated primers for TAIL-PCR

(TIGR) and the DOE-Joint Genome Initiative (JGI) for making the genome sequences of *P. syringae* DC3000 and B728a available before publication. Our work is supported by DOE Division of Energy Biosciences grant DE-FG05-95ER20187.

References

Alarcon-Chaidez, F.J., Penaloza-Vazquez, A., Ullrich, M., and Bender, C.L. (1999) Characterization of plasmids encoding the phytotoxin coronatine in *Pseudomonas syringae*. *Plasmid* **42:** 210–220.

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 bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. *Proc Natl Acad Sci USA* **97:** 4856–4861.

Ansorge, W., Voss, H., and Zimmermann, J. (1997) *DNA Sequencing Strategies: Automated and Advanced Approaches*. New York: Wiley.

Arnold, D.L., Jackson, R.W., Fillingham, A.J., Goss, S.C., Taylor, J.D., Mansfield, J.W., and Vivian, A. (2001) Highly conserved sequences flank avirulence genes: isolation of novel avirulence genes from *Pseudomonas syringae* pv. *pisi*. *Microbiology* **147:** 1171–1182.

Bavage, A.D., Vivian, A., Atherton, G.T., Taylor, J.D., and Malik, A.N. (1991) Molecular genetics of *Pseudomonas syringae* pathovar *pisi*: plasmid involvement in cultivar-specific incompatibility. *J Gen Microbiol* **137:** 2231–2239.

Boch, J., Joardar, V., Gao, L., Robertson, T.L., Lim, M., and Kunkel, B.N. (2002) Identification of *Pseudomonas syringae* pv. *tomato* genes induced during infection of *Arabidopsis thaliana*. *Mol Microbiol* **44:** 73–88.

Cornelis, G.R., and Van Gijsegem, F. (2000) Assembly and function of type III secretory systems. *Annu Rev Microbiol* **54:** 735–774.

Cournoyer, B., Sharp, J.D., Astuto, A., Gibbon, M.J., Taylor, J.D., and Vivian, A. (1995) Molecular characterization of the *Pseudomonas syringae* pv. *pisi* plasmid- borne avirulence gene avrPpiB which matches the R3 resistance locus in pea. *Mol Plant–Microbe Interact* **8:** 700–708.

Curiale, M.S., and Mills, D. (1982) Integration and partial excision of a cryptic plasmid in *Pseudomonas syringae* pv. *phaseolicola*. *J Bacteriol* **152:** 797–802.

Dangl, J.L., and Jones, J.D. (2001) Plant pathogens and integrated defence responses to infection. *Nature* **411:** 826–833.

Dangl, J.L., Ritter, C., Gibbon, M.J., Mur, L.A., Wood, J.R., Goss, S., *et al.* (1992) Functional homologs of the *Arabidopsis* RPM1 disease resistance gene in bean and pea. *Plant Cell* **4:** 1359–1369.

Delcher, A.L., Harmon, D., Kasif, S., White, O., and Salzberg, S.L. (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* **27:** 4636–4641.

Diver, W.P., Grinsted, J., Fritzinger, D.C., Brown, N.L., Altenbuchner, J., Rogowsky, P., and Schmitt, R. (1983) DNA sequences of and complementation by the tnpR genes of Tn21, Tn501 and Tn1721. *Mol Gen Genet* **191:** 189–193.

Figurski, D.H., and Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci USA* **76:** 1648–1652.

Fouts, D.E., Abramovitch, R.B., Alfano, J.R., Baldo, A.M., Buell, C.R., Cartinhour, S., *et al.* (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.

Froehlich, B.J., Tatti, K., and Scott, J.R. (1983) Evidence for positive regulation of plasmid prophage P1 replication: integrative suppression by copy mutants. *J Bacteriol* **156:** 205–211.

Gibbon, M.J., Sesma, A., Canal, A., Wood, J.R., Hidalgo, E., Brown, J., *et al.* (1999) Replication regions from plant-pathogenic *Pseudomonas syringae* plasmids are similar to ColE2-related replicons. *Microbiology* **145:** 325–334.

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

Hacker, J., and Kaper, J.B. (1999) The concept of pathogenicity islands. In *Pathogenicity Islands and Other Mobile Virulence Elements*. Hacker, J., and Kaper, J.B. (eds). Washington, DC: American Society for Microbiology Press, pp. 1–12.

Hacker, J., and Kaper, J.B. (2000) Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54:** 641–679.

Hendrickson, E.L., Guevera, P., Penaloza-Vazquez, A., Shao, J., Bender, C., and Ausubel, F.M. (2000) Virulence of the phytopathogen *Pseudomonas syringae* pv. *maculicola* is rpoN dependent. *J Bacteriol* **182:** 3498–3507.

Hensel, M., Nikolaus, T., and Egelseer, C. (1999) Molecular and functional analysis indicates a mosaic structure of *Salmonella* pathogenicity island 2. *Mol Microbiol* **31:** 489–498.

Hobbs, M., and Reeves, P.R. (1994) The JUMPstart sequence: a 39 bp element common to several polysaccharide gene clusters. *Mol Microbiol* **12:** 855–856.

Hou, Y.M. (1999) Transfer RNAs and pathogenicity islands. *Trends Biochem Sci* **24:** 295–298.

Huynh, T.V., Dahlbeck, D., and Staskawicz, B.J. (1989) Bacterial blight of soybean: regulation of a pathogen gene determining host cultivar specificity. *Science* **245:** 1374–1377.

Innes, R.W., Bent, A.F., Kunkel, B.N., Bisgrove, S.R., and Staskawicz, B.J. (1993) Molecular analysis of avirulence gene avrRpt2 and identification of a putative regulatory sequence common to all known *Pseudomonas syringae* avirulence genes. *J Bacteriol* **175:** 4859–4869.

Jackson, R.W., Athanassopoulos, E., Tsiamis, G., Mansfield, J.W., Sesma, A., Arnold, D.L., *et al.* (1999) Identification of a pathogenicity island, which contains genes for virulence and avirulence, on a large native plasmid in the bean pathogen *Pseudomonas syringae* pathovar *phaseolicola*. *Proc Natl Acad Sci USA* **96:** 10875–10880.

Jackson, R.W., Mansfield, J.W., Arnold, D.L., Sesma, A., Paynter, C.D., Murillo, J., *et al.* (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.

Jin, Q., and He, S.Y. (2001) Role of the Hrp pilus in type III protein secretion in *Pseudomonas syringae*. *Science* **294:** 2556–2558.

Karaolis, D.K., Somara, S., Maneval, D.R., Jr, Johnson, J.A., and Kaper, J.B. (1999) A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* **399:** 375–379.

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 Raney, D.E. (1954) Two simple media for the demonstration of phycocynin and fluorescine. *J Lab Clin Med* **44:** 301–307.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., II, and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166:** 175–176.

Li, C.M., Brown, I., Mansfield, J., Stevens, C., Boureau, T., Romantschuk, M., and Taira, S. (2002) The Hrp pilus of *Pseudomonas syringae* elongates from its tip and acts as a conduit for translocation of the effector protein HrpZ. *EMBO J* **21:** 1909–1915.

Lindgren, P.B., Peet, R.C., and Panopoulos, N.J. (1986) Gene cluster of *Pseudomonas syringae* pv. '*phaseolicola'* controls pathogenicity of bean plants and hypersensitivity of nonhost plants. *J Bacteriol* **168:** 512–522.

Lindgren, P.B., Panopoulos, N.J., Staskawicz, B.J., and Dahlbeck, D. (1988) Genes required for pathogenicity and hypersensitivity are conserved and interchangeable among pathovars of *Pseudomonas syringae*. *Mol Gen Genet* **211:** 499–506.

Liu, Y.G., and Whittier, R.F. (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* **25:** 674–681.

Lu, S.E., Scholz-Schroeder, B.K., and Gross, D.C. (2002) Characterization of the salA, syrF, and syrG regulatory genes located at the right border of the syringomycin gene cluster of *Pseudomonas syringae* pv. *syringae*. *Mol Plant–Microbe Interact* **15:** 43–53.

Mackey, D., Holt, B.F., Wiig, A., and Dangl, J.L. (2002) RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. *Cell* **108:** 743–754.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Mansfield, J., Jenner, C., Hockenhull, 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.

Nagy, G., Dobrindt, U., Kupfer, M., Emody, L., Karch, H., and Hacker, J. (2001) Expression of hemin receptor molecule ChuA is influenced by RfaH in uropathogenic *Escherichia coli* strain 536. *Infect Immun* **69:** 1924–1928.

Nagy, G., Dobrindt, U., Schneider, G., Khan, A.S., Hacker, J., and Emody, L. (2002) Loss of regulatory protein RfaH attenuates virulence of uropathogenic *Escherichia coli*. *Infect Immun* **70:** 4406–4413.

Nieto, C., Giraldo, R., Fernandez-Tresguerres, E., and Diaz, R. (1992) Genetic and functional analysis of the basic replicon of pPS10, a plasmid specific for *Pseudomonas* isolated from *Pseudomonas syringae* pathovar *savastanoi*. *J Mol Biol* **223:** 415–426.

Nimchuk, Z., Rohmer, L., Chang, J.H., and Dangl, J.L. (2001) Knowing the dancer from the dance: R-gene products and their interactions with other proteins from host and pathogen. *Curr Opin Plant Biol* **4:** 288–294.

Norrander, J., Kempe, T., and Messing, J. (1983) Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26:** 101–106.

Petnicki-Ocwieja, T., Schneider, D.J., Tam, V.C., Chancey, S.T., Shan, L., Jamir, Y., *et al.* (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.

Poplawsky, A.R., and Mills, D. (1987) Identification and mapping of regions that confer plasmid functions and of sites for excisive recombination of plasmid pMMC7105. *Mol Gen Genet* **206:** 265–272.

Prentki, P., and Krisch, H.M. (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene* **29:** 303–313.

Rajakumar, K., Bulach, D., Davies, J., Ambrose, L., Sasakawa, C., and Adler, B. (1997) Identification of a chromosomal *Shigella flexneri* multi-antibiotic resistance locus which shares sequence and organizational similarity with the resistance region of the plasmid NR1. *Plasmid* **37:** 159–168.

Rakin, A., Noelting, C., Schropp, P., and Heesemann, J. (2001) Integrative module of the high-pathogenicity island of *Yersinia*. *Mol Microbiol* **39:** 407–415.

Ritter, C., and Dangl, J.L. (1995) The avrRpm1 gene of *Pseudomonas syringae* pv. *maculicola* is required for virulence on *Arabidopsis*. *Mol Plant–Microbe Interact* **8:** 444–453.

Santangelo, T.J., and Roberts, J.W. (2002) RfaH, a bacterial transcription antiterminator. *Mol Cell* **9:** 698–700.

Sesma, A., Sundin, G.W., and Murillo, J. (1998) Closely related plasmid replicons coexisting in the phytopathogen *Pseudomonas syringae* show a mosaic organization of the replication region and altered incompatibility behavior. *Appl Environ Microbiol* **64:** 3948–3953.

Sesma, A., Sundin, G.W., and Murillo, J. (2000) Phylogeny of the replication regions of pPT23A-like plasmids from *Pseudomonas syringae*. *Microbiology* **146:** 2375–2384.

Sherrat, D. (1989) Tn3 and related TE: site-specific recombination and transposition. In *Mobile DNA*. Berg, D.E., and Howe, M.M. (eds). Washington, DC: American Society for Microbiology Press, pp. 163–184.

Sinha, H., Pain, A., and Johnstone, K. (2000) Analysis of the role of recA in phenotypic switching of *Pseudomonas tolaasii*. *J Bacteriol* **182:** 6532–6535.

Stevens, C., Bennett, M.A., Athanassopoulos, E., Tsiamis, G., Taylor, J.D., and Mansfield, J.W. (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.

Syn, C.K., and Swarup, S. (2000) A scalable protocol for the isolation of large-sized genomic DNA within an hour from several bacteria. *Anal Biochem* **278:** 86–90.

Szabo, L.J., and Mills, D. (1984a) Characterization of eight excision plasmids of *Pseudomonas syringae* pv. *phaseolicola*. *Mol Gen Genet* **195:** 90–95.

Szabo, L.J., and Mills, D. (1984b) Integration and excision of pMC7105 in *Pseudomonas syringae* pv. *phaseolicola*: involvement of repetitive sequences. *J Bacteriol* **157:** 821–827.

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

Thompson, J.F., and Landy, A. (1989) Regulation of the bacteriophage lambda site-specific recombination. In *Mobile DNA*. Berg, D.E., and Howe, M.M. (eds). Washington, DC: American Society for Microbiology Press, pp. 1–22.

Van der Hoorn, R.A.L., De Wit, P.J.G.M., and Joosten, M.H.A.J. (2002) Balancing selection favors guarding resistance proteins. *Trends Plant Sci* **7:** 67–71.

Vivian, A., Murillo, J., and Jackson, R.W. (2001) The roles of plasmids in phytopathogenic bacteria: mobile arsenals? *Microbiology* **147:** 763–780.

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.

Yamaguchi, K., and Tomizawa, J. (1980) Establishment of *Escherichia coli* cells with an integrated high copy number plasmid. *Mol Gen Genet* **178:** 525–533.

Zagorec, M., and Steinmetz, M. (1990) Expression of levansucrase-beta-galactosidase hybrids inhibits secretion and is lethal in *Bacillus subtilis*. *J Gen Microbiol* **136:** 1137–1143.