A dispensable region of the chromosome which is associated with an avirulence gene in *Pseudomonas syringae* pv. *pisi*

D. L. Arnold, J. Brown, R. W. Jackson and A. Vivian

*Pseudomonas syringae* pv. *pisi* comprises a number of races which fall into two phylogenetically distinct groups (designated I and II). Races are based on cultivar specificity in the host plant, pea (*Pisum sativum*), and are specified by the presence of avirulence genes. The avirulence gene *avrPpiA1* is present on the chromosome of all strains examined in race 2, which belongs to phylogenetic group II. A race 4B strain, from phylogenetic group I, lacks this avirulence gene and a comparative study was made of the chromosome in strains representing these two races. A race 2 cosmid clone (pAV270) carrying *avrPpiA1* was used as a basis for collinearity analysis of races 2 and 4B. A region of the chromosome amounting to 85 kb and including *avrPpiA1* was absent from race 4B compared with race 2. A fragment spanning the junction of the discontinuity in race 4B was isolated, cloned and used to delimit the extent of the additional DNA present in race 2. In both races the borders of the discontinuity contained DNA sequences which showed a high degree of conservation. A 7 bp slightly imperfect direct repeat (CCAGCT/,T) flanked the additional DNA in race 2, with a single copy in race 4B. The region flanking the additional DNA was present in all races of *P. syringae* pv. *phi*. These results confirm the phylogenetic groupings in *P. syringae* pv. *pisi*.

**Keywords:** genome mapping, fine structure comparison, phylogeny, pathogenicity islands

**INTRODUCTION**

*Pseudomonas syringae* pv. *pisi* (hereafter *Ppi*) causes bacterial blight of pea (*Pisum sativum*) and currently seven races of the pathogen are defined by their abilities to cause the disease in eight differential cultivated varieties (cvs) of pea. Gene-for-gene matching of specific resistance genes in the host and avirulence genes in the pathogen results in race/cultivar incompatibility and the induction of a hypersensitive reaction, which prevents spread and survival of the pathogen in the host plant. Currently, the system in pea and *Ppi* is postulated to involve the interaction of between four and six pairs of matching genes (Bevan *et al.*, 1995).

During attempts to introduce the broad-host-range plasmid RP4 (= RK2), the progenitor of the cosmid cloning vector pLAFR3, into certain races of *Ppi*, it was noted that two types of transconjugant were obtained, which differed in their ability to produce fluorescent pigment (Moulton *et al.*, 1993). Further characterization of the nonfluorescent strains showed them to have a race-4-like phenotype, even though they had apparently been derived from race 2 or race 6 cultures. Initially, it was not clear whether these race-4-like strains represented variation within a single strain or arose from an intimate mixture of two separate strains. Cournoyer *et al.* (1996) sequenced the region of the *hrpJKL* genes from the race 6 strain 1704B and from the 'race-4-like' strain PT10 and showed them to differ by ten substitutions scattered in the DNA sequences. This indicated that the two races could not be derived one from another by a simple genomic rearrangement. This conclusion was further confirmed by Arnold *et al.* (1996), who developed two pairs of specific oligonucleotide primers for the amplification of distinct and unrelated bands in each of these two races. The latter authors showed that strains of *Ppi* could be separated on the basis of these primer-
### Table 1. Bacterial strains, vectors and plasmids

<table>
<thead>
<tr>
<th>Strain, vector or plasmid</th>
<th>Relevant properties</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pseudomonas syringae pv. pisi</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>299A</td>
<td>Race 1, type strain</td>
<td>Taylor et al. (1989)</td>
</tr>
<tr>
<td>203</td>
<td>Race 2</td>
<td>Taylor et al. (1989)</td>
</tr>
<tr>
<td>870A</td>
<td>Race 3A, type strain</td>
<td>Taylor et al. (1989); D. L. Arnold &amp; A. Vivian, unpublished</td>
</tr>
<tr>
<td>895A</td>
<td>Race 4A, type strain</td>
<td>Taylor et al. (1989); Arnold et al. (1996)</td>
</tr>
<tr>
<td>PN8</td>
<td>Race 4B</td>
<td>Moulton et al. (1993); Arnold et al. (1996)</td>
</tr>
<tr>
<td>974B</td>
<td>Race 5, type strain</td>
<td>Taylor et al. (1989)</td>
</tr>
<tr>
<td>1704B</td>
<td>Race 6, type strain</td>
<td>Taylor et al. (1989)</td>
</tr>
<tr>
<td>1691</td>
<td>Race 7</td>
<td>HRI Wellesbourne, UK</td>
</tr>
<tr>
<td><strong>Escherichia coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DHSu</td>
<td>recA lacZΔM15</td>
<td>Hanahan (1983)</td>
</tr>
<tr>
<td>INVαF'</td>
<td>As DHSu except lacZΔM15Δ(lacZYA-argF) deoR' F'</td>
<td>Invitrogen</td>
</tr>
<tr>
<td><strong>Vectors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLAFR3</td>
<td>Tc'</td>
<td>Staskawicz et al. (1987)</td>
</tr>
<tr>
<td>pUC18</td>
<td>Ap', lacZ</td>
<td>Norrander et al. (1983)</td>
</tr>
<tr>
<td>pCRII</td>
<td>Ap', lacZ</td>
<td>Invitrogen</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pAV441</td>
<td>4.8 kb EcoRI fragment of pAV270 in pUC18</td>
<td>This study</td>
</tr>
<tr>
<td>pAV442</td>
<td>4.1 kb EcoRI fragment of pAV270 in pUC18</td>
<td>This study</td>
</tr>
<tr>
<td>pAV458</td>
<td>1.2 kb BamHI fragment of pAV270 in pUC18</td>
<td>This study</td>
</tr>
<tr>
<td>pAV459</td>
<td>3 kb PCR fragment (primers DA1 and DA2) from PN8 in pCRII</td>
<td>This study</td>
</tr>
</tbody>
</table>

Specific amplifications into two phylogenetic groups, designated I (containing races 1, 4B, 5, 7 and some nonfluorescent strains of race 3, now regarded as belonging to race 3B) and II (containing races 2, 4A, 6 and fluorescent strains of race 3, now regarded as race 3A). In this study the only race 3 strain included was the type race strain 870A, which should be regarded as representative of race 3A.

An avirulence gene, avrPpiA1, which matches the resistance gene R2 in pea, was characterized from race 2 strain 203 (Vivian et al., 1989; Dangl et al., 1992) and shown to be located on the chromosome. Races 5 and 7 also carry homologues of avrPpiA which are plasmid-borne and confer a similar specificity towards pea cultivars carrying the resistance gene R2 (Gibbon et al., 1997). This gene was not present in race 4B and we therefore sought to investigate the corresponding regions of the chromosomes of race 2 strain 203 and race 4B strain PN8 in relation to the divergence of the two phylogenetic groups.

### METHODS

**Bacterial strains and plasmids.** The bacterial strains and plasmids used in this study are shown in Table 1. *P. syringae* pathovars were grown at 25 °C on King's B agar (KB) (King et al., 1954) for 2 d or in Luria–Bertani (LB) broth (Miller, 1972) for 16 h. *Escherichia coli* strains were grown overnight on LB agar or LB broth at 37 °C. Antibiotics (Sigma) were used in the following concentrations (μg ml⁻¹): tetracycline, 10; ampicillin, 125.

**DNA manipulations.** DNA manipulations were performed as described by Sambrook et al. (1989). Total bacterial DNA was extracted from an overnight broth culture using a Puregene DNA isolation kit (Flowgen). Plasmid DNA was extracted using the alkaline lysis method or using a Qiagen midi-prep kit. Restriction enzymes, obtained from Gibco Life Tech-
nologies, were used according to the manufacturer's instructions. DNA fragments were ligated into pUC18 using T4 DNA ligase (Appligene) and transformed into E. coli strain DH5α. PCR products were cloned using the Original TA Cloning Kit (Invitrogen).

**Polymerase chain reaction.** The PCR reactions were carried out by combining the following reagents in a reaction mix: 10 × Taq buffer (HT Biotechnology), 0.2 mM of each dNTP (dATP, dCTP, dGTP and dTTP), the required concentration of each primer and 1:25 U Taq polymerase (HT Biotechnology) per reaction. The reaction mix was dispersed as 25 μl aliquots in 0.5 ml Eppendorf tubes and overlaid with mineral oil. The tubes were exposed to UV light (320 nm) from a transilluminator for 10 min prior to the addition of 10 ng genomic DNA or bacterial cells from a plate or broth culture. The samples were amplified through 1 cycle at 94 °C (5 min) followed by 30 cycles at 94 °C (30 s), 55 °C (1 min) and 72 °C (90 s), followed by one cycle of 72 °C (5 min) in a DNA Thermal Cycler 480 (Perkin Elmer). PCR products (20 μl) were separated on 1:5 % agarose gels (100 V); bands of interest were excised and the DNA purified using Qiaex (Qiagen).

**Hybridization.** DNA digests and PCR products for hybridization analysis were transferred from agarose gels to nylon membranes by vacuum blotting with a commercial apparatus (Appligene). Plasmid inserts and PCR bands, extracted from agarose gels, were labelled with [α-32P]dCTP using the random primer oligo-labelling method (Feinberg & Vogelstein, 1983) with a High Prime kit (Boehringer Mannheim). The blots were hybridized (65 °C, 16 h) using the labelled probes in the hybridization solution (Sambrook et al., 1989). The membrane was washed twice in 2× SSC (65 °C, 15 min), followed by 2× SSC/0.1% SDS (65 °C, 30 min) and finally 0.1× SSC/0.1% SDS (65 °C, 30 min). This constituted highly stringent conditions. The blot was placed in an X-ray cassette (GRI) between intensifying screens (GRI) and exposed to film at −70 °C as required.

**DNA sequencing.** Nucleotide sequence analysis was performed using the dideoxy chain-termination method (Sanger et al., 1977) with an Amplitise Sequencing Kit (Perkin Elmer) in a Macrophor sequencing system (Pharmacia). Inserts were sequenced in both directions using the pUC/M13 universal and reverse sequencing primers (Pharmacia) or oligonucleotides designed from the sequence. Oligonucleotides were synthesized by Gibco-BRL. Sequence analysis was performed at the Daresbury Laboratory via Seqnet using the GCG program.

**RESULTS**

**Comparison of the genomes of races 2 and 4B using the insert from pAV270 as a probe for collinearity**

The race 2 library cosmid, pAV270, contained approximately 25 kb of insert DNA, including seven EcoRI fragments designated E-a to E-g (Fig. 1). Bands E-d, E-e, E-f and E-g were extracted from the gel, labelled with 32P and used to probe EcoRI-digested total DNA from representative Ppi strains of each race except race 3B (Fig. 2). Fragments E-d, E-e and E-f were collinear in all races, including both the race 2 (203) and race 4B (PN8) genomes. To the right of this region in Fig. 1, the genomes of the seven races diverge in accordance with their phylogenetic groups. Thus, the fragment E-g (truncated due to the Sau3AI/BamHI junctions at the termini of the insert in the cosm id) produced a 7-4 kb band with races 1, 4B, 5 and 7 (group I) and a 10 kb band with races 2, 3A, 4A and 6 (group II). Use of a 1 kb EcoRI/BamHI (EB-a) fragment from the leftward end of fragment E-a to probe total EcoRI-digested DNA from races of Ppi showed hybridization which accorded with the two phylogenetic groups, namely to a 12 kb fragment in races 1, 4B, 5 and 7 (group I), and to a 3-5 kb fragment in races 2, 4A and 6 and to a slightly smaller fragment in race 3A (group II) (Fig. 3). However, in race 2, there was also strong hybridization to a 4-1 kb fragment, which appears to be unique to race 2, suggesting that the presence of avrPpiA1 in the chromosome is associated with additional DNA extending into fragment E-a (and E-c). Further digestion and subcloning from fragments E-a and E-c into pUC18 provided more focused probes to investigate the relative topologies of the genomes of races 2 and 4B, which indicated that the discontinuity between the two genomes resided in the 0-8 kb BamHI/PstI (BP-a) fragment within E-a and a 1-6 kb Strl/PstI (SP-c) fragment within E-c (Fig. 1).

**Isolation of the junction fragment from race 4B strain PN8**

In order to isolate the area of genome where the discontinuity occurred in the race 4B strain PN8, primers DA1 (5′ TTACCTCAAGATGTTACAGG 3′) and DA2 (5′ CGTCGCACTACGACCAAT 3′) were designed from race 2 (203) that could be used to amplify a fragment from race 4B. A subclone containing fragment BP-a (pAV458) was sequenced out of the vector, from the left-hand end, using the universal primer. A subclone containing fragment SP-c (pAV441) was sequenced out of the vector, from the right-hand end, using the universal primer. Primers DA1 and DA2 were designed from these sequences and used in a PCR reaction to amplify a 3 kb fragment from strain PN8. The resulting fragment was 32P-labelled and used to probe a blot of EcoRI-digested pAV270: hybridization was obtained to E-a and E-c but not to the intervening E-b fragment. Therefore the region spanning the discontinuity had been isolated from PN8 (data not shown). The junction fragment was cloned in pCRII (designated pAV459; Fig. 4a) and the insert used to probe EcoRI-digested total DNA of the eight races. The fragment hybridized to a 4-8 kb EcoRI fragment in races 2, 3A, 4A and 6 (group II) and a much larger 12 kb fragment in races 1, 4B, 5 and 7 (group I). Faint hybridization also occurred to a 4-1 kb fragment in race 2 indicating that in pAV459 both ends of the discontinuity had been cloned (Fig. 5).

**Sequencing the area of genomic discontinuity**

Limited sequencing of one DNA strand of plasmids pAV459 (race 4B; 1650 nt), pAV458 (race 2; 1203 nt) and pAV441 (race 2; 1366 nt) initially used the universal and reverse primers from the vector and then custom-made primers designed from the sequences generated. When the site of discontinuity in the DNA sequence between the two races was found, more comprehensive
Fig. 1. Restriction endonuclease map of pAV270 and subclones and fragments used as hybridization probes. E-a to E-g, EcoRI fragments; EB-a, EcoRI/BamHI fragment from E-a; BB-a, BamHI fragment from E-a; BP-a, BamHI/PstI fragment from E-a; SP-c, SstI/PstI fragment from E-c; A2, avrPpiA1.R2; *, vector sites in pLAFR3 (the boundaries of the insert in pAV270 are BamHI/Sau3AI hybrid junctions); E, EcoRI; H, HindIII; B, BamHI; P, PstI; S, SstI (not all restriction sites are shown).

Fig. 2. Hybridization analysis of Ppi DNA with fragments E-d to E-g of pAV270. Total genomic DNA of Ppi races was digested with EcoRI and separated on a 0.7% agarose gel; the resulting blot was probed using 32P-labelled EcoRI fragments E-d (1.3 kb), E-e (2.8 kb) and E-g (3.0 kb). Phylogenetic groups are shown above the number of the races and the strains shown are: race 1, 299A; race 2, 203; race 3A, 870A; race 4A, 895A; race 4B, PN8; race 5, 974B; race 6, 1704B; race 7, 1691. C, cosmid pAV270 digested with EcoRI.

Fig. 3. Hybridization analysis of Ppi DNA with fragment EB-a of pAV270 (see Fig. 1). Total genomic DNA of Ppi races was digested with EcoRI and separated on a 0.7% agarose gel; the resulting blot was probed using 32P-labelled fragment EB-a (1 kb). Phylogenetic groups are shown above the number of the races and the strains shown are: race 1, 299A; race 2, 203; race 3A, 870A; race 4A, 895A; race 4B, PN8; race 5, 974B; race 6, 1704B; race 7, 1691. C, cosmid pAV270 digested with EcoRI.

A J005681) were compared (together with the single-stranded sequences) with those in the EMBL/GenBank nucleotide sequence database using the GCG program FASTA: no significant matches were found. However, the sequences at the right-hand end of pAV459 (race 4B) were 84% identical to the first 121 bases of pAV441 (race 2) up to the border of the region of discontinuity.
Chromosome region associated with *avrPpiA1*

The left-hand end of pAV459 was 91% identical to 159 bases in pAV458 (race 2). A region of 8.5 kb of DNA was present in race 2 but not race 4B. At the border of discontinuity, a 7 bp slightly imperfect direct repeat (CCAGCT/TCAG) was found: one copy in pAV441 and one in pAV458 flanking the additional DNA in race 2 and one copy in pAV459 (from race 4B) at the junction (Fig. 4b).

**Fig. 5.** Hybridization analysis of *Ppi* DNA with the insert of pAV459. Total genomic DNA of *Ppi* races was digested with EcoRI and separated on a 0.7% agarose gel; the resulting blot was probed using 32P-labelled insert of pAV459. Phylogenetic groups are shown above the number of the races and the strains shown are: race 1, 299A; race 2, 203; race 3A, 870A; race 4A, 895A; race 4B, PN8; race 5, 974B; race 6, 1704B; race 7, 1691. C, cosmid pAV270 digested with EcoRI.

**Fig. 6.** PCR amplification of a 230 bp fragment from *Ppi* isolates using DA15 and DA16 primers. Phylogenetic groups are shown above the number of the races and the strains shown are: race 1, 299A; race 2, 203; race 3A, 870A; race 4A, 895A; race 4B, PN8; race 5, 974B; race 6, 1704B; race 7, 1691. P, pAV459; C, no cells.

**Examination of border regions**

Primers DA15 (5' TCGTCCAGAGACTCTCC 3') and DA16 (5' GAGTACGAGAAGGTTCG 3') were designed to allow the amplification of a 230 bp border region from pAV459 (race 4B). Similarly, primers DA15 and DA19 (5' GACGAGTTCGAGTTGTA 3') amplified a fragment of 260 bp which spanned the junction in pAV458 (race 2), and primers DA16 and DA20 (5' CACCGAGATGCCAGA 3') amplified a fragment of 200 bp at the junction in pAV441 (race 2). These
primers were used in pairs in PCR reactions with the eight races of Ppi. A 230 bp band was produced with primers DA15 and DA16 in races 1, 4B, 5 and 7 (Fig. 6). In race 2, a 260 bp band was produced with primers DA15 and DA19 and a 200 bp band with primers DA16 and DA 20 (data not shown). No bands were obtained with the remaining races.

**DISCUSSION**

The plant pathogen *Ppi* has been shown to comprise two phylogenetically distinct groups within the pathovar. All strains of races 1, 5 and 7 belong to group I, whilst all strains of races 2 and 6 belong to group II; races 3 and 4 comprised strains belonging to either group and those in group II are currently considered to belong to races 3A and 4A, with those in group I in races 3B and 4B (Arnold et al., 1996; D. L. Arnold & A. Vivian, unpublished). We examined corresponding regions of the chromosomes of races 2 and 4B, which are in groups II and I, respectively, of *Ppi*. The region chosen in race 2 contained a gene, *avrPpiA1*, which confers avirulence toward pea cultivars carrying the R2 resistance gene (Dangl et al., 1992; Gibbon et al., 1997). The race 4B strain lacks *avrPpiA* and a region of DNA associated with it. Two additional races of *Ppi*, races 5 and 7, also carry homologues of *avrPpiA1*, which are plasmid-borne (Gibbon et al., 1997). The distribution was suggestive of an ability for the *avrPpiA* alleles to move between the chromosome and plasmids; in consequence we sought evidence for this hypothesis by examining the sequences flanking the region of discontinuity between races 2 and 4B.

A region of 8.5 kb was present in race 2, but not in race 4B. Outside this region, the DNA sequence appeared to be conserved in races 2 and 4B and to follow one of two divergent patterns, corresponding to phylogenetic groups I and II (Arnold et al., 1996). Database searches failed to detect any evidence for the involvement of a known transposable element among the DNA sequences obtained. A 7 bp, slightly imperfect direct repeat was present at the boundaries of the 8.5 kb sequence in race 2, with a single copy being present in the corresponding region in race 4B. Pathogenicity islands (PAIs) have been found in Gram-negative as well as Gram-positive bacteria (Hacker et al., 1997) and they comprise large DNA regions (up to 200 kb of DNA) and often numerous virulence genes. The G+C content of these regions often differs from that of the surrounding bacterial genome and in most cases, they are flanked by specific DNA sequences, such as direct repeats or insertion sequence (IS) elements and can have a tendency to delete with high frequencies. The chromosomal region containing *avrPpiA1* shows some similarities to PAIs. For example, the presence of direct repeats and many virulence genes, including *avrPpiA1*, have a lower G+C ratio than the surrounding genome (Dangl et al., 1992; Vivian et al., 1997).

Recent evidence that Avr proteins are introduced into plant cells via a type III secretion system, similar to that responsible for the introduction of virulence proteins into animal cells, indicates that virulence and avirulence may be host-specific aspects of the same gene (Alfano & Collmer, 1997; Vivian & Gibbon, 1997). Whilst marker-exchange mutagenesis of *avrPpiA* in races 2 and 7 of *Ppi* did not affect pathogenicity towards the susceptible host as indicated by symptom development and bacterial growth in pea (Gibbon et al., 1997), a plasmid-borne homologue of *avrPpiA1*, *avrRpm1*, present in *P. syringae pv. maculicola*, has been shown to be required for full virulence at low inoculum levels on susceptible ascensions of *Arabidopsis thaliana* (Ritter & Dangl, 1995).

At present, just two of the five postulated *avr* genes that are the basis of the race structure have been fully characterized. In the case of *avrPpiB* (Cournoyer et al., 1995), this gene is present in races 1, 3A and 7 and in all cases is located on plasmid DNA. In contrast, *avrPpiA* is chromosomally located in group II race 2, whilst being plasmid-borne in group I races 5 and 7 (Gibbon et al., 1997). Thus the location of the latter gene appears to correlate precisely with the phylogenetic groups. It remains unclear whether the location of an *avr* gene affects it at a functional level in terms of the interaction with the host plant.

Specific PCR primers were used to amplify small fragments encompassing the two border regions in race 2 and the single one in race 4B. The regions in race 2 were exclusively amplified from that race suggesting that the area within the missing fragment is present only in race 2. The fragment covering the junction in race 4B was present in races 1, 4B, 5 and 7. However, sequences related to the insert of pAV459 were present in all races as shown by probing with this fragment. There could be several reasons why no fragment was amplified from the group II races. In race 2 it is likely that size (8.5 kb) prevents the recovery of a PCR product under the conditions used; in the remaining strains it is not clear why no fragment was obtained and may indicate absence or divergence of the primer sequences, or alternatively the insertion of an unrelated DNA sequence at this site.

In conclusion, the results presented here indicate that it is probable that *avrPpiA1* in race 2 represents an acquisition of this allele at some time in the past as evidenced by the presence of direct repeat sequences at the boundaries of the fragment involved. However, we did not find any evidence of insertion sequences which might indicate a continuing capability for the mobility of this gene.

**ACKNOWLEDGEMENTS**

Sequence analysis was performed using the Seqnet facility at BBSRC Daresbury Laboratory.

**REFERENCES**


Received 22 June 1998; revised 1 September 1998; accepted 21 September 1998.