

## Molecular Characterization of Avirulence Gene D from *Pseudomonas syringae* pv. *tomato*

D. Y. Kobayashi, S. J. Tamaki, and N. T. Keen

Department of Plant Pathology, University of California, Riverside 92521 U.S.A.  
Received 31 July 1989. Accepted 3 November 1989.

Avirulence gene D, cloned from *Pseudomonas syringae* pv. *tomato*, caused *P. s.* pv. *glycinea* to elicit a hypersensitive defense response on certain cultivars of soybean. Nucleotide sequence data for a 5.6-kb *Hind*III fragment containing *avrD* disclosed five long open-reading frames (ORFs) occurring in tandem. The phenotype conferred by *avrD* was expressed in *P. s.* pv. *glycinea* solely by the first of these ORFs (933 bases) that encoded a protein of 34,115 daltons. Neither a signal peptide sequence nor significant

regions of hydrophobicity were present that would indicate secretion of the protein or its membrane association. Hybridization analyses revealed that some but not all *P. syringae* pathovars contained DNA homologous to *avrD*. This included weak hybridization to all tested races of *P. s.* pv. *glycinea*, although none of them express the phenotype conferred by *avrD*. The *avrD* gene occurred on an indigenous 75-kb plasmid in several *P. s.* pv. *tomato* isolates.

*Additional keywords:* Southern blots, gene-for-gene interactions, gene overexpression, hypersensitive reaction, race specificity.

The hypersensitive response (HR) is an active plant defense mechanism occurring in response to infection by various pathogens. The HR involves rapid, localized cell necrosis followed by accumulation at the infection site of antimicrobial compounds called phytoalexins (Keen and Holliday 1982; Klement 1982). Induction of the HR is believed to result from specific plant recognition of the pathogen (Keen and Staskawicz 1988). For example, many phytopathogenic pseudomonads and xanthomonads (subdivided as pathovars according to the specialization of their host species) elicit the HR on nonhost plant species. However, a narrower level of specificity is illustrated by the recognition of some but not all biotypes (called races) of a single pathogen taxon by certain cultivars of a plant species.

Recognition in such cases has been shown to involve single genetic elements in both the host and pathogen, and has been termed gene-for-gene complementarity (Ellingboe 1976; Flor 1942). In this relationship, a single dominant gene for disease resistance in the plant is complemented by a single dominant gene for avirulence in the pathogen. It is only this combination of the respective alleles that leads to specific plant recognition and the HR. Avirulence genes, therefore, generally restrict the host range of the pathogen and provide a genetic basis for the classification of races.

Several avirulence genes have been cloned from bacterial

pathogens belonging to the *Pseudomonas syringae* van Hall and *Xanthomonas campestris* (Pammel) Dowson groups (Gabriel *et al.* 1986; Staskawicz *et al.* 1984, 1987; Swanson *et al.* 1988). Three avirulence genes from *P. s.* pv. *glycinea* (Coerper) Young *et al.*, the causal agent of bacterial blight of soybean, have been molecularly characterized (Napoli and Staskawicz 1987; Tamaki *et al.* 1988). Each of these genes encodes a single protein product that determines the race phenotype of the bacterium. One of these genes, *avrB*, corresponds to the *RpgI* resistance gene locus in soybean, and recent data also indicate that soybean contains single dominant resistance genes complementing *avrA*, *avrC*, and *avrD* (N. T. Keen and R. I. Buzzell, unpublished). Thus the *P. s.* pv. *glycinea*-soybean interaction is a gene-for-gene system that is useful for studying mechanisms determining race specificity and disease resistance.

Avirulence genes that function at higher levels of specificity by eliciting the HR on nonhost plants have recently been cloned from bacterial pathogens (Kobayashi *et al.* 1989; Whalen *et al.* 1988). In *P. s.* pv. *tomato* (Okabe) Young *et al.*, the causal agent of bacterial speck of tomato, three different avirulence genes were cloned that elicited the HR on certain soybean cultivars when expressed in *P. s.* pv. *glycinea* (Kobayashi *et al.* 1989).

Each gene differed according to the specific cultivars that reacted hypersensitively, analogous to previously cloned race-specific avirulence genes from *P. s.* pv. *glycinea*. Indeed, one *P. s.* pv. *tomato* avirulence gene was found to be identical to *avrA*, previously cloned from *P. s.* pv. *glycinea* race 6 (Napoli and Staskawicz 1987). The other two genes elicited patterns of cultivar specificity unlike any known *P. s.* pv. *glycinea* race. A DNA fragment carrying one of them, designated *avrD*, hybridized to all tested *P. s.* pv. *glycinea* races, although they did not express the phenotype conferred by *avrD*. To further investigate the molecular and biochemical events in *P. syringae*-soybean interactions, we chose to molecularly characterize the *P. s.* pv. *tomato* *avrD* gene.

Current address of D. Kobayashi: DNA Plant Technology, 6701 San Pablo Ave., Oakland, CA 94608 U.S.A.

Current address of S. Tamaki: Cleargene, Inc., University of California, Berkeley, Richmond Field Station, Richmond, CA 94804 U.S.A.

Address all correspondence to N. T. Keen.

Nucleotide and/or amino acid sequence data is to be submitted to GenBank as accession number J03681.

## MATERIALS AND METHODS

**Bacterial strains, plasmids, and media.** Bacterial strains and plasmids used or constructed in this study are listed

in Table 1. *Escherichia coli* strains were typically grown on Luria-Bertani (LB) agar medium (Maniatis *et al.* 1982) at 37° C. *P. syringae* strains were grown at 28° C on King's medium B (KMB) agar (King *et al.* 1954). Antibiotics were

**Table 1.** Bacterial strains, bacteriophage, and plasmids used in this study

Designation	Relevant characteristics <sup>a</sup>	Source or reference
<i>Escherichia coli</i> DH5 $\alpha$	F <sup>-</sup> <i>lacZ</i> M15 <i>endA1</i> <i>hsdR17</i> <i>supE44</i> <i>thi-1</i> <i>gyrA</i> <i>relA1</i> <sup>-</sup>	Bethesda Research Laboratories, Gaithersburg, MD
MV1193	( <i>lac-proAB</i> ) <i>thi</i> <i>supE44</i> ( <i>srl-recA</i> ) 306::Tn10 ( <i>tet</i> <sup>r</sup> ) (F' <i>traD36</i> ( <i>proAB lacZ</i> M15))	Vieira and Messing 1987
<i>Pseudomonas syringae</i> pathovars		
<i>P. s. pv. atropurpurea</i>		J. V. Leary, Univ. of California, Riverside
<i>P. s. pv. glycinea</i> race 0		B. Staskawicz, Univ. of California, Berkeley
<i>P. s. pv. glycinea</i> race 1		This laboratory
<i>P. s. pv. glycinea</i> race 4	<i>rif</i> <sup>r</sup> , <i>ap</i> <sup>r</sup>	This laboratory
<i>P. s. pv. glycinea</i> race 6		This laboratory
<i>P. s. pv. lachrymans</i>		This laboratory
<i>P. s. pv. mori</i> 0782-30		D. A. Cooksey, Univ. of California, Riverside
<i>P. s. pv. morsprunorum</i> 0782-28		D. A. Cooksey
<i>P. s. pv. phaseolicola</i> 0285-1		D. A. Cooksey
<i>P. s. pv. pisi</i>		This laboratory
<i>P. s. pv. savastanoi</i> 0185-8		D. A. Cooksey
<i>P. s. pv. syringae</i> from bean PS19		D. Gross, Washington State Univ., Pullman
<i>P. s. pv. syringae</i> from tomato 5D 4171		D. Gross
<i>P. s. pv. tabaci</i>		R. Durbin, Univ. of Wisconsin, Madison
<i>P. s. pv. tomato</i> PT21		Bender and Cooksey 1986
<i>P. s. pv. tomato</i> PT23		Bender and Cooksey 1986
<i>P. s. pv. tomato</i> PT24		Bender and Cooksey 1986
<i>P. s. pv. tomato</i> PT25		Bender and Cooksey 1986
Phage M13K07		Vieira and Messing 1987
Plasmids		
pUC118/pUC119	<i>E. coli</i> cloning plasmids	Vieira and Messing 1987
pRK415	Broad host range plasmid vector with pUC119 polylinker, Tc <sup>r</sup> Helper plasmid, Tra <sup>+</sup> , Km <sup>r</sup>	Keen <i>et al.</i> 1988
pRK2013		Ditta <i>et al.</i> 1980
pPT101	5.6-kb <i>Hind</i> III fragment positioning <i>avrD</i> open-reading frame downstream from the vector <i>lac</i> promoter of pRK415	Kobayashi <i>et al.</i> 1989
pPT102	5.6-kb <i>Hind</i> III fragment cloned in the opposite orientation to pPT101 in pRK415	This study
pPTD120	5.6-kb <i>Hind</i> III fragment cloned with <i>Pst</i> I site distal to the <i>lac</i> promoter in pUC119	This study
pPTD121	5.1- <i>Hind</i> III- <i>Pst</i> I fragment from the pPTD120 5.6-kb <i>Hind</i> III fragment cloned into pUC119	This study
pPTD122	0.5-kb <i>Pst</i> I- <i>Hind</i> III fragment from the pPTD120 5.6-kb <i>Hind</i> III fragment cloned into pUC118	This study
pPTD1211	Approximately 1.2-kb fragment from deletion A20 containing <i>avrD</i> cloned downstream from the vector <i>lac</i> promoter in pUC119	This study
pPTD1212	Deletion A20 containing the open- reading frame of <i>avrD</i> in antiorientation to the vector <i>lac</i> promoter in pUC118	This study
pPRD130	5.6-kb <i>Hind</i> III fragment cloned in pUC119 in the opposite orientation to pPTD120	This study
pPTD134	2.8-kb <i>Hind</i> III- <i>Bam</i> HI fragment from pPTD120 cloned in pUC119	This study
pPTD135	3.2-kb <i>Hind</i> III- <i>Bam</i> HI fragment from pPTD120 cloned into pUC119 in the orientation positioning the insert <i>Hind</i> III site distal to the vector <i>lac</i> promoter	This study

<sup>a</sup>Tc, tetracycline; Km, kanamycin; <sup>r</sup>, resistant; kb, kilobase.

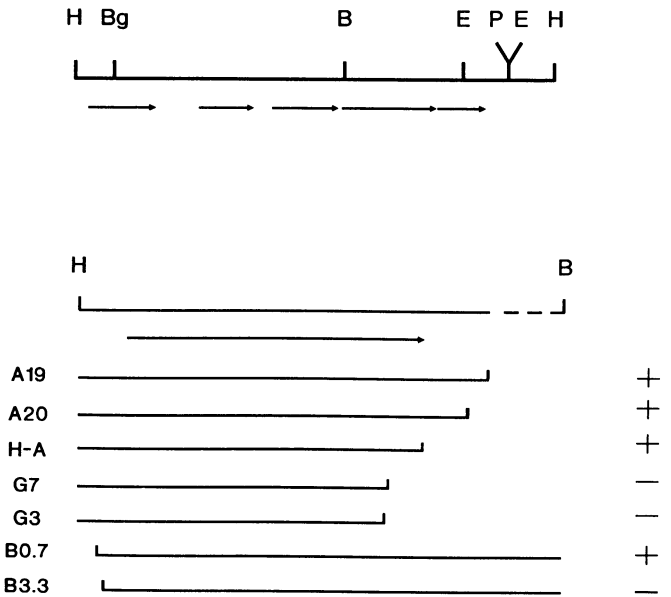
purchased from Sigma Chemical Co. (St. Louis, MO) and used at the following concentrations unless otherwise noted: tetracycline, 12.5 µg/ml; ampicillin, 50 µg/ml; rifampicin, 100 µg/ml; and kanamycin, 25 µg/ml.

**DNA manipulations.** Standard recombinant DNA methods were generally performed as described by Maniatis *et al.* (1982). Enzymes were purchased from New England Biolabs (Beverly, MA) or Bethesda Research Laboratories (Gaithersburg, MD).

For Southern blot analyses, DNA from several *P. syringae* pathovars was isolated as described for total DNA (Staskawicz *et al.* 1984) or for plasmid DNA (Bender and Cooksey 1986); 4.5 µg of total DNA was digested with the appropriate enzyme and electrophoresed in 0.7% agarose gels before transfer onto Zetabind nylon filters (AMF Cuno, Meriden, CT). Hybridizations with <sup>32</sup>P-labeled probes were performed in 50% formamide, 5× SSC, 1× Denhardt's solution, 0.02 M sodium phosphate, pH 6.7, and 0.1 mg/ml salmon sperm DNA with gentle shaking at 42° C. Blots were then washed in 2× SSC, 0.1% sodium dodecyl sulfate (SDS) at room temperature for 30 min, followed by 0.1× SSC, 0.1% SDS at 42° C for 2 hr before exposure to X-ray film.

Previously hybridized blots were stripped of probes in 0.4 M sodium hydroxide and then neutralized in 0.1× SSC, 0.5% SDS, and 0.2 M Tris-HCl, pH 7.5, each for 60 min at 42° C. Blots were then prehybridized for a minimum of 12 hr before hybridizing with a new probe.

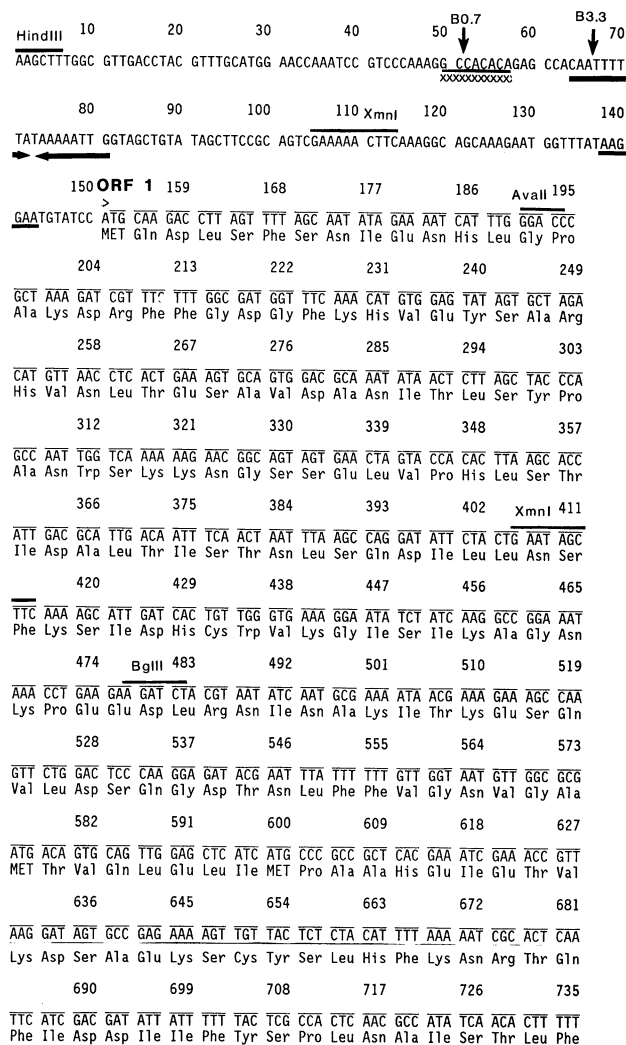
**DNA sequence analysis.** Exonuclease III (Henikoff 1984) was used to generate overlapping deletions on both ends



**Fig. 1.** Restriction map of the 5.6-kilobase *Hind*III fragment of *Pseudomonas syringae* pv. *tomato* (upper) indicating the location and reading direction of open-reading frames (ORFs) 1-5 as shown by the arrows. The expanded region (lower) indicates the left portion of the *Hind*III fragment containing ORF 1 (*avrD*) (arrow) with deletion clones used for its confirmation. The broken line denotes the extended fragment to the downstream *Bam*HI site. H, *Hind*III; Bg, *Bg*III; B, *Bam*HI; E, *Eco*RI; P, *Pst*I; Sp, *Spe*I; and H-A, *Hind*III-*Acc*I fragment: (+) = avirulence phenotype observed from the DNA fragment indicated when cloned in pRK415 and maintained in race 4 of *P. s.* pv. *glycinea*; (-) = no avirulence phenotype in *P. s.* pv. *glycinea* race 4.

of the 5.6-kilobase (kb) *Hind*III fragment containing the phenotype conferred by *avrD*. Appropriate restriction enzyme sites to conduct exonuclease III reactions on the entire *Hind*III fragment were not available in the pUC118/pUC119 polylinker. It was therefore necessary to divide the fragment into two subclones for both orientations.

For sequencing the first strand, pPTD120 (the *Hind*III fragment cloned in pUC119 orienting the *Pst*I site [Fig. 1] distal to the vector *lac* promoter) was digested with *Pst*I, releasing a 0.5-kb fragment. The vector with the remaining insert was religated back to itself, yielding pPTD121, and the 0.5-kb fragment was ligated into pUC118 in the proper orientation relative to the vector primer to yield pPTD122 (Table 1). To generate deletions throughout the first strand, pPTD121 was first digested with *Pst*I, creating a 3' overhang that was not susceptible to exonuclease III digestion. Before digestion with *Kpn*I for vector protection, the fragment was end-filled with T4 polymerase I to create



**Fig. 2.** Nucleotide sequence of the coding strand of a 5.6-kilobase *Hind*III fragment containing *avrD*. The predicted amino acid sequences are shown for open-reading frames (ORFs) 1-5. Selected restriction sites are noted by overscoring. Arrows indicate end points for selected exonuclease III deletions. Presumed ribosome binding sites are underscored. Underscored bases 50-58 and 64-81 indicate the nine base CA-rich sequence and the 18 base AT-rich palindromic sequence discussed in the text.

a susceptible end for exonuclease III digestion. Plasmid pPTD122, containing the remaining 0.5-kb *PstI-HindIII* fragment in pUC118, was sequenced in its entirety.

To subclone the opposite strand, pPTD130 (which contains the 5.6-kb *HindIII* fragment cloned in pUC119 in the opposite orientation to pPTD120) was digested with *BamHI*, releasing a fragment of 3.2 kb. The remaining insert fragment contained with the vector was self-ligated to create pPTD134, while the 3.2-kb *BamHI* fragment was ligated into pUC119 in the proper orientation and labeled pPTD135 (Table 1). Exonuclease III deletions were then generated in both pPTD134 and pPTD135.

Both strands of the 5.6-kb *HindIII* fragment were sequenced according to the dideoxy chain termination method of Sanger *et al.* (1977). Hydrophathy plots were obtained using the algorithm of Kyte and Doolittle (1982). Sequence analysis was performed using the Bionet system supplied by Intelligenetics Corporation, Mountain View, CA.

**Bacterial matings, plant growth conditions, and plant inoculations.** Conjugations were performed according to the method described by Ditta *et al.* (1980). Matings were incubated at 28° C for 5–8 hr followed by incubation at 4° C for 24 hr before streaking onto KMB agar supplemented with rifampicin, 25 µg/ml of ampicillin, and 25 µg/ml of tetracycline. *P. s. pv. glycinea* race 4 trans-conjugants were successively single-colony transferred on selective media to ensure purity before inoculation of soybean plants.

Soybean plants were grown from seed as previously described (Long *et al.* 1985). Bacteria were prepared for inoculation by resuspending cells in distilled water and adjusting the concentration to 10<sup>7</sup> cells per milliliter before infiltrating fully expanded primary leaves of soybean plants using the device of Hagborg (1970). Inoculations were scored daily for 1 to 5 days for the appearance of a visible HR, typically appearing within 24–30 hr, or water-soaked lesions, typically appearing after 48–72 hr.

**SDS-polyacrylamide gel electrophoresis.** *E. coli* cells containing the desired plasmids were grown in 5 ml of LB agar medium supplemented with ampicillin and 1 mM isopropyl-β-D-thiogalactopyranoside at 28° C for 14 hr. Whole cell proteins were extracted and run on 10% SDS-polyacrylamide gels before staining with Coomassie Brilliant Blue R 250 as described by Tamaki *et al.* (1988).

## RESULTS

**Sequence analysis.** Previous data indicated that a Tn5 insertion which mutated the phenotype conferred by *avrD* mapped close to a *Bg/II* site located at one end of the 5.6-kb *HindIII* fragment shown in Figure 1 (Kobayashi *et al.* 1989). However, since initial attempts to subclone the avirulence phenotype on a smaller DNA fragment failed, the entire *HindIII* fragment was sequenced. This data revealed five long open-reading frames (ORFs)

744		753		762		771		780		789	
GTC	GCT	TAT	GAT	AAA	GAG	CCC	CAT	TTT	TTA	CCT	GGC
Val	Ala	Tyr	Asp	Lys	Glu	Pro	His	Phe	Leu	Pro	Gly
798		807		816		825		834		843	
CCT	AAC	ATT	ATG	AAC	CCC	GTA	GAT	TCA	CTT	GTC	AGT
Pro	Asn	Ile	MET	Asn	Pro	Val	Asp	Ser	Leu	Val	Ser
852		861		870		879		888		897	
GCA	CTT	CTT	TAT	AAA	CTC	GAT	GGT	TTG	ACT	CGT	GAT
Ala	Leu	Leu	Tyr	Lys	Leu	Asp	Gly	Leu	Thr	Arg	Asp
906		915		924		933		942		951	
ATG	AGG	AGT	TTG	AAT	ATT	ATC	GCC	GAG	AAT	CCC	GCA
MET	Arg	Ser	Leu	Asn	Ile	Ile	Ala	Glu	Asn	Pro	Ala
960		969		978		987		996		1005	
CGA	TTA	CTA	GTA	ACC	GAA	CTA	AAG	CGT	GCT	AAT	ATT
Arg	Leu	Leu	Val	Thr	Glu	Leu	Lys	Arg	Ala	Asn	Ile
1014		1023		1032		1041		1050		1059	
AAC	TGG	CGA	ATA	GCG	GAA	GTG	GCT	GGA	CAT	ATG	AAT
Asn	Trp	Arg	Ile	Ala	Glu	Val	Ala	Gly	His	MET	Asn
1068		1077		1093		1103		1113		1123	
TCA	GTT	GCG	CAT	CTA	TTA	CCC	CTT	TAGTATACTT	TCGAAAAAAC	AGCTGCTGAT	TCCCAGAAAA
Ser	Val	Ala	His	Leu	Leu	Pro	Leu				
1133		1143		1153		1163		1173		1183	
TTAAACTTTA	TCAGTAGCTT	ATTCTATACA	TCATAGGGAG	GCGCAGATTT	ATTGATTTTT	TCGTCCTTGC					
1203		1213		1223		1233		1243		1253	
AACGCTCTGG	AGGCCTTGAT	TTATCAGGGG	GCAACAGCAG	GGTATTAGA	ATAAATCAGC	GTCTCCATAG					
1273		1283		1293		1303		1313		1323	
GTAGATTATT	TCGCGAATAG	TACACAGGGG	TGCAACATGA	ACGTTTCGAT	TGCCGCCTTG	GGAAACGTTT					
1343		1353		1363		1373		1383		1393	
TGTCGTCCTT	TGAGGTACACA	AATGAAGGCT	TTGATAACTG	CGCGTCATAT	AGAAGTCTG	TGCATCCAAT					
1413		1423		ORF 2		1433		1442		1451	
AATAATGCTA	AGTGACGACG	CTCG	ATG	ATC	ATT	GTC	ATT	ATC	CGA	CAT	GGT
			MET	Ile	Ile	Val	Ile	Ile	Arg	His	Gly
1469		1478		1487		1496		1505		1514	
CCA	CAA	AAT	TTG	CTT	GGC	GTG	TTT	CAG	GGG	CAG	TCT
Pro	Gln	Asn	Leu	Leu	Gly	Val	Phe	Gln	Gln	Ser	Asp
1523		1532		1541		1550		1559		1568	
GTA	GGG	ATA	GAT	CGG	TTC	AAG	GAA	ACT	GCC	AGA	ACG
Val	Gly	Ile	Asp	Arg	Phe	Lys	Glu	Thr	Ala	Arg	Thr
1577		1586		1595		1604		1613		1622	
GAT	GCT	ATT	TAT	AGC	TCT	AAC	TAT	AAA	CGC	TCA	CTG
Asp	Ala	Ile	Tyr	Ser	Ser	Asn	Tyr	Lys	Arg	Ser	Leu
1631		1640		1649		1658		1667		1676	
ACA	GTT	GAC	GTG	AAC	TTG	AGA	AGA	TTT	GTG	TCC	ACA
Thr	Val	Asp	Val	Asn	Leu	Arg	Arg	Phe	Val	Ser	Thr
1685		1694		1703		1712		1721		1730	
TTG	GGC	GCA	CTG	GAC	GGT	AAG	TCT	AAG	GAG	CTT	CTT
Leu	Gly	Ala	Leu	Asp	Gly	Lys	Ser	Lys	Glu	Leu	Ile
1739		1748		1757		1766		1775		1784	
TTA	TGG	AGA	AAG	TTA	ATA	ACA	CTT	GAG	TAT	ACC	CCA
Leu	Ser	Arg	Lys	Leu	Ile	Thr	Leu	Glu	Tyr	Thr	Pro
1793		1802		1811		1820		1829		1838	
CGT	TCA	GCG	TTG	GAG	CGT	TTT	GTG	CGT	GGA	ATA	CAT
Arg	Ser	Ala	Leu	Glu	Arg	Phe	Val	Arg	Gly	Ile	His
1847		1856		1865		1874		1883		1892	
CAA	GGC	GCG	GTG	ATT	GTG	TTT	TCT	CAT	GGA	GGT	ATT
Gln	Gly	Arg	Val	Ile	Val	Val	Ser	His	Gly	Gly	Ile
1901		1910		1919		1928		1937		1946	
CAC	ATG	TTA	GGG	GTG	CGA	CAG	ACT	TCC	TGC	CTT	CTA
His	MET	Leu	Gly	Val	Arg	Gln	Thr	Ser	Cys	Leu	Leu
1955		1964		1973		1982		1991		2000	
ATA	ATA	AAG	GTC	TCA	GGG	ACT	GAA	ATT	AGT	TTG	ATG
Ile	Ile	Lys	Val	Ser	Gly	Thr	Glu	Ile	Ser	Leu	MET
2009		2018		2027		2036		2045		2054	
AAT	TCA	ATA	GCG	GAA	GCG	ACG	TAC	TAT	GGA	AAA	TAT

Fig. 2 continued from previous page.

Fig. 2 continued on next page.

arranged in a cluster on one strand (Fig. 2). These ORFs, designated 1–5 according to their order on the *Hind*III fragment (Fig. 2), were preceded at the predicted ATG start codons by purine-rich sequences resembling ribosome binding sites (Shine and Dalgarno 1974). No ORFs of significant length were found on the opposite reading strand. Computer searches of DNA and protein sequence libraries through the Bionet resource did not yield known genes or proteins with significant homology.

ORF 1 was suspected of encoding the *avrD* gene product since it mapped to the region where the Tn5 insertion mutated the phenotype conferred by *avrD*. This ORF was comprised of 933 nucleotides and could encode a protein with a computer-predicted molecular mass of 34,368 daltons and an isoelectric point of 6.3. A consensus signal peptide sequence was not present at the amino terminus of the protein, and the hydropathy plot of the protein indicated no significant regions of hydrophobicity that would be expected of a membrane-spanning protein (data not shown).

The first ORF was separated from ORF 2 by 345 bases, the largest nontranslated region between any of the five ORFs (Figs. 1 and 2). The presumed start codon of ORF 2 was located at base 1428 and it terminated at base 2000. This ORF could encode a protein of 215 amino acids with a molecular mass of 23,848 daltons. ORF 3 began

at base 2287 and was separated from ORF 2 by 215 bases. This ORF consisted of 852 bases and could encode a protein of 284 amino acids with a molecular mass of 31,266 daltons. ORFs 3, 4, and 5 were more tightly clustered such that ORF 3 terminated at base 3138 and was separated by 18 bases from ORF 4 (Fig. 2). ORF 4 began at base 3159 and consisted of 1,032 bases. It was the largest of the five ORFs and could encode a protein of 344 amino acids with a molecular mass of 39,183 daltons. The terminus of ORF 4, which occurs at base 4190, overlapped the translational start of ORF 5 at base 4177. ORF 5 consisted of 594 nucleotides that could encode a protein of 198 amino acids with a molecular mass of 22,325 daltons. Neither signal peptide sequences nor significant regions of hydrophobicity were observed for any of the ORFs (data not shown).

**Deletion analysis of ORF 1.** Various 5' and 3' exonuclease III deletions generated for use as sequencing templates were screened for the phenotype conferred by *avrD* (Figs. 1 and 2). Deletions at the 3' end of the ORF were recovered as *Hind*III-*Eco*RI fragments and cloned into the broad host range plasmid, pRK415. These constructs were then conjugated into *P. s. pv. glycinea* race 4, followed by inoculation of the bacteria into appropriate soybean cultivars to screen for the phenotype conferred by *avrD*. Deletions A19 and A20, which have 3' end points mapping outside of ORF 1, expressed the phenotype

```

Asn Ser Ile Ala Glu Ala Thr Tyr Tyr Gly Lys Tyr Leu Asp Lys Gly Phe MET
2063      2072      2082      2092      2102      2112      2122
GGG CAG TGG GAG AGC ATC TAGAAAAATC AGATGCCCGC TCACCGGCTC AGCTGACACA TGACAGCGTG
Gly Gln Trp Glu Ser Ile
2132      2142      2152      2162      2172      2182      2192
ATGACCAAGC AGGGACTTTA TTTTCAGAGGG GTGTAATAG AGTCGGTACT CGCGTCTTGG GCCTTTGGAG
2202      2212      2222      2232      2242      2252      2262
CCTACGAAAG CGAACGAGCA CTCAATTAAT GGTTCCTGAT CATGAGGCCA TCCATGCAAG CGCTTCTAAA
2272      2282 ORF 3 2292      2301      2310      2319
CTAACTTTTT TATATGGGAT AGGT ATG CAA AGC CGA TTC AAT GGA TGG TCA ATG CAG
MET Gln Ser Arg Phe Asn Gly Trp Ser MET Gln
2328      2337      2346      2355      2364      2373
GTT CTT GAG GTG GAT GAT ACG GCA GCG GTT GGT CGA CAT ATT GAT CAG TTT GGT
Val Leu Glu Val Asp Asp Thr Ala Ala Val Gly Arg His Ile Asp Gln Phe Gly
2382      2391      2400      2409      2418      2427
TTC GCG ATC GTT TCG GGG GAA TGG AGA TTC GAT GCG TCT GAT TTT GAC CGC ATG
Phe Ala Ile Val Ser Gly Glu Trp Arg Phe Asp Ala Ser Asp Phe Asp Arg MET
2436      2445      2454      2463      2472      2481
GCC GCA CTT TAC GGC TTG GGC CCA ATG TAC CAG TCG GAT TTC AAC CCG CTT GAG
Ala Ala Leu Tyr Gly Leu Gly Pro MET Tyr Gln Ser Asp Phe Asn Arg Leu Glu
2490      2499      2508      2517      2526      2535
CAT GCA GAA GGT ATA GCA TCA TCG GGA ATT AAC CAG GTC GGA GGT CTG TCG AGC
His Ala Glu Gly Ile Ala Ser Ser Gly Ile Asn Gln Val Gly Gly Leu Ser Ser
2544      2553      2562      2571      2580      2589
GGC AGC CAT GTC GTG TTC AAC GGC GCT ACA GAC GTG CCG CTT CAT ACC GAT GGT
Gly Ser His Val Val Phe Asn Gly Ala Thr Asp Val Pro Leu His Thr Asp Gly
2598      2607      2616      2625      2634      2643
TCC TAT TTA CCT ATA GGC ACC ATC AAG ACG TCG ATC CTC TTT TGT AGA GAA TCT
Ser Tyr Leu Pro Ile Gly Thr Ile Lys Thr Ser Ile Leu Phe Cys Arg Glu Ser
2652      2661      2670      2679      2688      2697
GGC GCT CTC GGC GGG GAG TCC ATT CTG TTC GAT AGC GTG TCG GCA TTT CGA GCA
Ala Ala Leu Gly Gly Glu Ser Ile Leu Phe Asp Ser Val Ser Ala Phe Arg Ala
2706      2715      2724      2733      2742      2751
CTG AGC GAG GAT CAT CCT GAT CTT GCT CCG TCC TTG CTC GCC GAT AAT GCG TTC

```

Fig. 2 continued from previous page.

```

Leu Ser Glu Asp His Pro Asp Leu Ala Arg Ser Leu Leu Ala Asp Asn Ala Phe
2760      2769      2778      2787      2796      2805
AGG CCG CGA TCT ACT AGT ACG CGT TCG GGT AGG CAG TAT CAA CAC ATT GGG CCG
Arg Arg Arg Ser Thr Ser Thr Arg Ser Gly Arg Gln Tyr Gln His Ile Gly Pro
2814      2823      2832      2841      2850      2859
ATG TTT CTT CGT CCG GAA GAC GGA GAT ATT GTT GGC GGC TTC ACG CTC GAT ATC
MET Phe Leu Arg Arg Glu Asp Gly Asp Ile Val Gly Gly Phe Thr Leu Asp Ile
2868      2877      2886      2895      2904      2913
ACG GCT GAC TGG GAA TAC TCG CGT CGT ATG GAC GCA CCG GTG ATT GAC GCA GCG
Thr Ala Asp Trp Glu Tyr Ser Arg Arg MET Asp Ala Arg Val Ile Asp Ala Ala
2922      2931      2940      2949      2958      2967
GGC TAT CTC ATC CCG CTC GCC TCC GAA AAC AGC GAT TAC ACT CTG AAG TTT GGG
Ala Tyr Leu Ile Arg Leu Ala Ser Glu Asn Ser Asp Tyr Thr Leu Lys Phe Gly
2976      2985      2994      3003      3012      3021
TTG CAT AAA GGG CAG GTG CTG ATT ATG CGA AAC GAC CAG CTG TCG CAT GGT CGA
Leu His Lys Gly Gln Val Ile MET Arg Asn Asp Gln Leu Ser His Gly Arg
3030      3039      3048      3057      3066      3075
TGC TCA TAT GTC GAC GAC CCT GCC AGG CCT CGA ATC CTG TTT CGA GGA CTC TTT
Cys Ser Tyr Val Asp Asp Pro Ala Arg Pro Arg Ile Leu Phe Arg Gly Leu Phe
3084      3093      3102      3111      3120      3129
CTG TCC TCA CCA TGC GAT TCT GGT GCA CCA ACA GAC TTG GTC TGT ACC CGA GGT
Leu Ser Ser Pro Cys Asp Ser Gly Ala Pro Thr Asp Leu Val Cys Thr Arg Gly
3138      3148      3158ORF 4 3167      3176      3185
AGC CAA TCT TGACTGAGGG AATGTCACAT ATG CCG GAG CAA GCT GCG GGG GCG CAG
Ser Gln Ser MET Pro Glu Gln Ala Ala Gly Ala Gln
3194      3203      3212      3221      3230 BamHI 3239
TGG GCC TCA TGG ATA TCG TGT CCG GCC AGT AGT CAG CTT AAG TCG CCG ATC CAG ACC
Trp Ala Ser Trp Ile Ser Cys Arg Ala Ser Gln Leu Lys Ser Arg Ile Gln Thr
3248      3257      3266      3275      3284      3293
GAC CTT AAC GCC TAC TCA TTA AGC GAC GAC ATG AAA CAG GTC ATC GTT GAA TTT
Asp Leu Asn Ala Tyr Ser Leu Ser Asp Asp MET Lys Gln Val Ile Val Glu Phe
3302      3311      3320      3329      3338      3347
GTT CCG GAA TAT CCG AGG GGC GGT CCG GGC TGG AGC GAG AGG GGG ATG CCG GGC
Val Arg Glu Tyr Arg Arg Gly Gly Arg Gly Trp Ser Glu Arg Gly MET Arg Gly
3356      3365      3374      3383      3392      3401
GAC GAG GTC ATG GCG CCA ATG CTC GCC GTC AGG GAT ACC GCC CCT GAA CTG TTT

```

Fig. 2 continued on next page.

conferred by *avrD*. However, the avirulence phenotype was not observed for deletions G3 and G7, each having 3' end points that mapped within the ORF (Figs. 1 and 2). The 3' terminus of ORF 1 was further confirmed by subcloning a *HindIII*-*AccI* fragment into pRK415. This fragment, which extended four bases beyond the termination codon of ORF 1 (Fig. 2), expressed the phenotype conferred by *avrD* (Fig. 1).

ORF 1 initiates 151 bases from the left *HindIII* site (Fig. 1) and was positioned downstream from the vector *lac* promoter in plasmid pPT101. However, expression of the phenotype conferred by *avrD* was also observed in *P. s. pv. glycinea* race 4 cells harboring pPT102, which contains the 5.6-kb *HindIII* fragment cloned in the reverse orientation to pPT101 (data not shown). Thus, it was assumed that the transcriptional initiation point for *avrD* in *P. s. pv. glycinea* is contained within the 150 bases of 5' DNA. Two deletions, B0.7 and B3.3, generated from pPTD135 had 5' end points mapping between the left *HindIII* site and the predicted start codon of ORF 1 (Fig. 2). These deletions were isolated as *BamHI*-*PstI* fragments and subcloned into pRK415, positioning ORF 1 in the reverse orientation to the vector *lac* promoter. Avirulence gene function was observed with *P. s. pv. glycinea* race 4 cells containing B0.7, which had an end point beginning

at base 51. However, no activity was detected from B3.3, which had an end point located 16 bases 3' to B0.7 at base 67 (Fig. 2). Inspection of the region spanned by deletions B0.7 and B3.3 disclosed an 18-base pair AT-rich palindromic sequence that was preceded by CA repeats (Fig. 2).

**Expression of *avrD* in *E. coli*.** Deletion A20, which was constructed in pUC119 with ORF 1 positioned downstream from the vector *lac* promoter, was labeled pPTD1211. SDS-polyacrylamide gels of whole cell proteins from *E. coli* cells containing pPTD1211 disclosed a new protein band of 34 kDa, in close agreement to the calculated size of *avrD* from sequence data (Fig. 3). Deletion A20 subcloned into pUC118 yielded pPTD1212, which positioned ORF 1 in the reverse orientation to the *lac* promoter. This construct in *E. coli* did not produce a detectable band of similar size to the protein encoded by *avrD*, indicating that overexpression in *E. coli* is dependent on the vector *lac* promoter (Fig. 3). Although the reason is currently unclear, levels of the protein encoded by *avrD* were higher in *E. coli* cells grown at 28° C than at 37° C.

#### Southern blot hybridization of ORF 1 and ORF 5 to DNA of various *P. syringae* pathovars. Previous Southern

```

Asp Glu Val MET Ala Pro MET Leu Gly Val Arg Asp Thr Ala Pro Glu Leu Phe
3410      3419      3428      3437      3446      3455
GAA CTC GTT CGT ACG AAG ATC GCA CAG GCA GAG AAA GAT AAA TGG ATG GGC GGC
Glu Leu Val Arg Thr Lys Ile Ala Gln Ala Glu Lys Asp Lys Trp MET Gly Gly
3464      3473      3482      3491      3500      3509
AGC TTG GAC AGT AAT GTC ATG GCA GCG CAG TTG CAA TAT CGG CGA TTT GTC ATC
Ser Leu Asp Ser Asn Val MET Ala Ala Gln Leu Gln Tyr Arg Arg Phe Val Ile
3518      3527      3536      3545      3554      3563
GTA TCT ACG CCA CGT TCA GGA ACA CAT CTG CTG CGT ACG CTG CTT GGC TCG CAT
Val Ser Thr Pro Arg Ser Gly Thr His Leu Leu Arg Thr Leu Leu Gly Ser His
3572      3581      3590      3599      3608      3617
CCG TGT ATT GAG GTC CAC GGT GAG GCG TTT AAT CGA TTC GGT CAG CAC CTT TTG
Pro Cys Ile Glu Val His Gly Glu Ala Phe Asn Arg Phe Gly Gln His Leu Leu
3626      3635      3644      3653      3662      3671
CCT TAT TCG GTG CAG GAC ACG ACG GCA GCC GGG GTT CTT GAA AGG CAT CTA TTT
Pro Tyr Ser Val Gln Asp Thr Thr Ala Ala Gly Val Leu Glu Arg His Leu Phe
3680      3689      3698      3707      3716      3725
CGA CCG TAT TTC GAA TAT GTC GAA GCC GTC GGT TTC GTG CTC TTT CGC GAT CTC
Arg Pro Tyr Phe Glu Tyr Val Glu Ala Val Gly Val Leu Phe Arg Asp Leu
3734      3743      3752      3761      3770      3779
GAC ACC CAT TGG GCA GGC CAG AAC GTG TGG GGT GCC TTA GCA GAT GTC CGC GAC
Asp Thr His Trp Ala Gly Gln Asn Val Trp Gly Ala Leu Ala Asp Val Arg Asp
3788      3797      3806      3815      3824      3833
CTA AAA ATA ATC CTG CTC GAC CGA CGC AAC CGG CTG GAG CGA CTT GTG TCC GTA
Leu Lys Ile Ile Leu Leu Asp Arg Arg Asn Arg Leu Glu Arg Leu Val Ser Val
3842      3851      3860      3869      3878      3887
AAG AAA AGC CTG TGC GAT CAC GTT TGG TAC GTT GGC CGT GAA GAT AAG AGG TTG
Lys Lys Ser Leu Cys Asp His Val Trp Tyr Val Gly Arg Glu Asp Lys Arg Leu
3896      3905      3914      3923      3932      3941
AGG CCG CAT GTG GAA CTG TCG GTT CCG CTT CAT GAA CTC GTC GAC TTC ATT GAC
Arg Pro His Val Glu Leu Ser Val Pro Leu His Glu Leu Val Asp Phe Ile Asp
3950      3959      3968      3977      3986      3995
CGC GAT CTT GTA AAT CCG GCG CAA TTC TGC CAC CAG TTT CAC GGG CAC GAC ATA
Arg Asp Leu Val Asn Arg Ala Gln Phe Cys Asp Gln Phe His Gly His Asp Ile
4004      4013      4022      4031      4040      4049
TTG CCG ATC ACT TAC GAA GAG TTG CTC GCG ACT CCA GAA GTT GTG CAC GCC CGT

```

Fig. 2 continued from previous page.

```

Leu Pro Ile Thr Tyr Glu Glu Leu Leu Ala Thr Pro Glu Val Val His Ala Arg
4058      4067      4076      4085      4094      4103
ATG CTC AAG TTT CTT GGT GTG TCA GCT GCA ATG CTT CAG TCC GGA ACA GGT AAG
MET Leu Lys Phe Leu Gly Val Ser Ala Ala MET Leu Gln Ser Gly Thr Gly Lys
4112      4121      4130      4139      4148      4157
AAG GAG AAG GCG CCG GTT TCG GCG GTG GTC AAC AAC ATC GAT CAA TTG AAG TCA
Lys Glu Lys Ala Pro Val Ser Ala Val Val Asn Asn Ile Asp Gln Leu Lys Ser
4166      4175      4184
GAA CTT TCA GGT ACG AAA TAT GAG AGC TAT ATC
Glu Leu Ser Gly Thr Lys Tyr Glu Ser Tyr Ile
ORF 5 4185      4194      4203
ATG AGA GCT ATA TCT AAT ACG CCG CGA CTG TAT
MET Arg Ala Ile Ser Asn Thr Pro Arg Leu Tyr
4212      4221      4230      4239      4248      4257
TGG CCT CCC TAC ACA CCG ATG AAA ATG GAT GAT TGC AGG CCG AGC ATT GTT CGT
Trp Pro Pro Tyr Thr Pro MET Lys MET Asp Asp Cys Arg Pro Ser Ile Val Arg
4266      4275      4284      4293      4302      4311
GGY GAG GGG GTT TAT CTC TAC GAT GAT ACG GGT CGA CGC TAT ATT GAT GGT ATC
Gly Glu Gly Val Tyr Leu Tyr Asp Asp Thr Gly Arg Arg Tyr Ile Asp Gly Ile
4320      4329      4338      4347      4356      4365
TCT GGT AGC TAT AAT CAT TGC CTC GGA CAC TCG CAC TTC GGG CTT ATC GCG GCG
Ser Gly Ser Tyr Asn His Cys Leu Gly His Ser His Phe Gly Leu Ile Ala Ala
4374      4383      4392      4401      4410      4419
GTG AAG GAG CAA GTT GAT ACA CTC GTT CAT GCG TGT AAT ATY TCT TCA AAT ACG
Val Lys Glu Gln Val Asp Thr Leu Val His Ala Cys Asn Ile Ser Ser Asn Thr
4428      4437      4446      4455      4464      4473
GTA CTC CCC GAA GCG CTT GCC GAG AGA ATC AGT GGC AAA CTG GTG AAG GCC AGA
Val Leu Pro Glu Ala Leu Ala Glu Arg Ile Ser Gly Lys Leu Val Lys Ala Arg
4482      4491      4500      4509      4518      4527
CTY GTC CAT ACG TTT CTC GTC ATG AGT GGA AGT GAA GGT GTC GAG GCC GCC CTC
Leu Val His Thr Phe Leu Val MET Ser Gly Ser Glu Gly Val Glu Ala Ala Leu
4536      4545      4554      4563      4572      4581
AAG ATG GCG TGG CAG TAT CAG ATT AAC CCG GGG TGC CCT CAG CGA ACC AAA GTC
Lys MET Ala Trp Gln Tyr Gln Ile Asn Arg Gly Cys Pro Gln Arg Thr Lys Val
4590      4599      4608      4617      4626      4635
GTT GCG ATT GAT GGT GCG TAC CAC GGC TGC ACG CTG GGT GCG ATG ATA GCC ACG
Val Ala Ile Asp Gly Ala Tyr His Gly Cys Thr Leu Gly Ala MET Ile Ala Thr
4644      EcoRI 4653      4662      4671      4680      4689
CGY CGT GAA TTC ATC AAT GAA GGG GCA GCG CGT TAC TTG CCG CTC ACT CGA TTG

```

Fig. 2 continued on next page.



blot analyses (Kobayashi *et al.* 1989) indicated that several *P. syringae* pathovars, including all tested races of *P. s. pv. glycinea*, contained 5.6-kb *Hind*III fragments which were highly homologous to those from *P. s. pv. tomato*. The association of ORF 1 with the phenotype conferred by *avrD* permitted a more critical homology evaluation of the *Hind*III fragment in various pathovars.

A 300-base pair *Xmn*I fragment (internal to the transcript of *avrD*, extending from nucleotides 109 to 409, Fig. 2) was chosen as a gene-specific probe and hybridized to total DNA digested with *Hind*III from various *P. syringae* pathovars. No hybridization was observed to DNA from *P. s. pv. atropurpurea* (Reddy & Godkin) Young *et al.*, *P. s. pv. morsprunorum* (Wormald) Young *et al.*, *P. s. pv. pisi* (Sackett) Young *et al.*, *P. s. pv. tabaci* (Wolf and Foster) Young *et al.*, *P. s. pv. savastanoi* (Smith) Young *et al.*, and two isolates of *P. s. pv. syringae* van Hall. Strong hybridization was observed to 5.6-kb *Hind*III fragments from *P. s. pv. lachrymans* (Smith and Bryan) Young *et al.*, *P. s. pv. mori* (Boyer and Lambert) Young *et al.*, and *P. s. pv. tomato* (Fig. 4A), but all *P. s. pv. glycinea* races as well as the tested isolate of *P. s. pv. phaseolicola* (Burkholder) Young *et al.* yielded weakly hybridizing 5.6-kb *Hind*III bands. This suggested that sequence divergence within the *avrD* portion of the conserved 5.6-kb *Hind*III

```

Arg Arg Glu Phe Ile Asn Glu Gly Ala Ala Arg Tyr Leu Pro Leu Thr Arg Leu
4698      4707      4716      4725      4734      4743
CGA TGC CTA TTC CAT CAG GCC TTG AGG ATA TTT CAC ACT GGC GTG CGC TGC TGC
Arg Cys Leu Phe His Gln Ala Leu Arg Ile Phe His Thr Gly Val Arg Cys Trp
4752      4761      4777      4787      4797      4807      4817
CCG AAC AGG AAA CGA CCA TAGCTGCAAT CGTCGTGAG CCTGTATAGG CGATGGCCGG GACCCGAAT
Pro Asn Arg Lys Arg Pro
4827      4837      4847      4857      4867      4877      4887
TTCCGGATGG GTTCCTACGG GAATTGTCTG CGCTAGAACA GAAGTACGAT ATCCCGTTTA TCTGTGACGA
4897      4907      4917      4927      4937      4947      4957
GGTGATTTCG GGGGTAGGGC GCACAGGTGC ATTTTGTGAA TCCATTAACC AAGGTGCTAG CCCGGATATT
4967      4977      4987      4997      5007      5017      5027
GTCAATTTGA GCAAGTGCCT TGGTGGCGT TTTCCGATTA CCGCAGTGGT GACGACGGCA GACATGACCG
5037      5047      5057      5067      5077      5087      5097
ACAGTTTTCG TGCCCAATCT ATGCCACTTT TCAGGCATGG GCATACGCAG TCAGGCAATC TCCTTGCTG
5107      5117      5127      5137      5147      5157      5167
TCGTGACAGA TTGTTTCATTC TCGATTATCT AGACAGTCAT CGTAGTTATG AGGTAGTGGC AGCCGGTCCG

```

Fig. 2 continued from previous page.

```

5177      PstI 5187 EcoRI 5197      5207      5217      5227      5237
TTCCGGACTG CTGCAGGGAA TTCGCGAGAA CTTGCAGACG ACGAACAGTA TCGTGAGCGT ACAGGGGAAA
5247      5257      5267      5277      5287      5297      5307
GGGCTTATGT TGTGATTAC ATTGGAACG TCGCAAGCGT GCACCCGGCC CCAACTTCC GTTCGACAGC
5317      5327      5337      5347      5357      5367      5377
AGGGGGTTAT TGTGGTGGC GCCGATGGC ATCTGAAGTT GCGCCATCA GTTCTTGATC AGCGAACCCG
5387      5397      5407      5417      5427      5437      5447
AAGCGGATGA GTTGACCGAT CGCCTGGTCA GTTCAATTCC CAGTGTTCGC CAACAGTAAG CTCCCAGGTT
5457      5467      5477      5487      5497      5507      5517
CTGATAGTAA TAAAGGAGGG TCCAGTGTG TGGATACTTG GCGCTAAAC GGATTGCACA CGGCACTCTA
5527      5537      5547      5557      5567      5577      HindIII 5586
CATGACGTGC CGTTCAGGTG GGGACAGCTG TGCAACATAT GGTGTGACGA CACTATTGCG AATAAGCTT

```

Fig. 2 continued.

fragment could account for the absence of the phenotype conferred by *avrD* in any known *P. s. pv. glycinea* race, but raised the possibility that other pathovars might contain functional *avrD* genes. In addition, *P. s. pv. lachrymans*, but no other tested pathovar, had two hybridizing bands of 5.6 and 4.3 kb, indicating that sequences homologous to *avrD* are reiterated in this pathovar.

To determine the extent of conservation of the downstream ORFs on the 5.6-kb *Hind*III fragment, the blot in Figure 4A was stripped of the *avrD* probe and reprobed with the *Bam*HI-*Hind*III fragment from pPTD134 that contains a portion of ORF 4 and all of ORF 5 (Figs. 1 and 2). This probe strongly hybridized to the 5.6-kb *Hind*III fragments from all *P. syringae* pathovars tested (Fig. 4B). Of considerable interest is that hybridization to the 5.6-kb fragment in all *P. s. pv. glycinea* races was as intense as it was to the corresponding fragment from *P. s. pv. tomato*. *P. s. pv. lachrymans* yielded an additional hybridizing band of 4.3 kb, which is similar to results

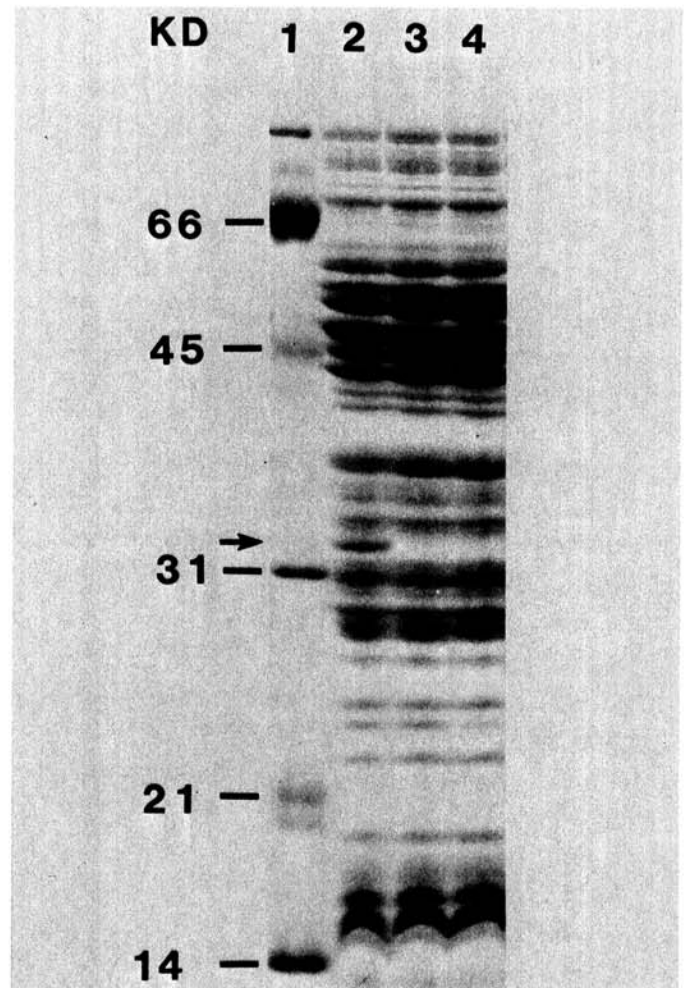


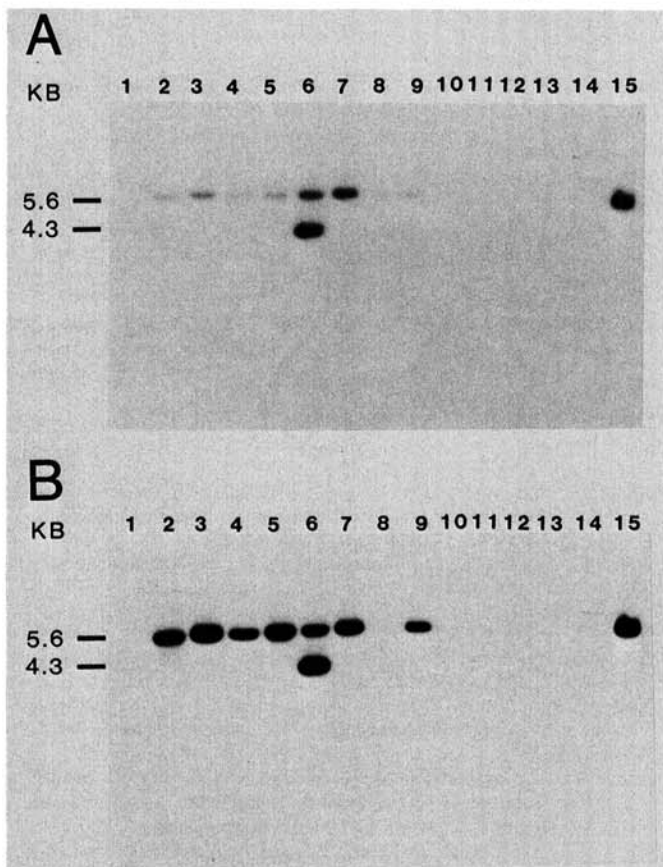
Fig. 3. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of *Escherichia coli* DH5 $\alpha$  cells containing pUC119, pPTD1211, or pPTD1212 grown with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside for 14 hr at 28 $^{\circ}$  C. Samples were prepared as described in the text, and 10  $\mu$ l was applied to each lane. Lane 1, size standards with sizes given in kilodaltons on the left; lane 2, pPTD1211; lane 3, pPTD1212; and lane 4, pUC119. The arrow denotes the presumed protein encoded by *avrD*.

obtained with the *avrD* probe (Fig. 4A). Hybridization was not observed to DNA from *P. syringae* pathovars that failed to show homology to the intragenic *avrD* probe.

**Location of *avrD* on an indigenous plasmid of *P. s. pv. tomato*.** Certain bacterial avirulence genes have been located on indigenous plasmids (Staskawicz *et al.* 1987; Swanson *et al.* 1988). Due to the distribution of *avrD* among several *P. syringae* pathovars, it was of interest to determine if this gene resided on a plasmid in *P. s. pv. tomato*. In Southern blot analyses, insert DNA from pPTD1211 (deletion A20 containing *avrD*) was probed to indigenous plasmid DNA from four different isolates of *P. s. pv. tomato*. Hybridization was observed only to a plasmid of approximately 75 kb (Fig. 5). This plasmid was previously designated as the B plasmid in *P. s. pv. tomato* (Bender and Cooksey 1986), but no phenotypes were assigned.

## DISCUSSION

Deletion analyses indicated that the phenotype conferred

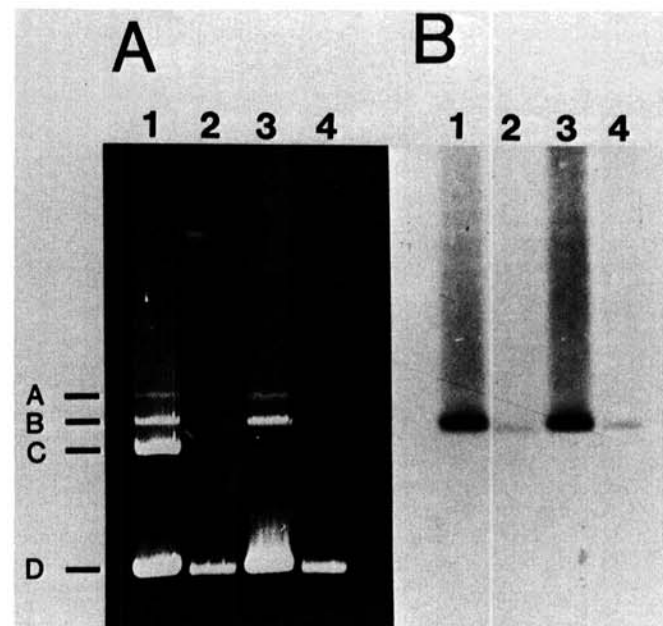


**Fig. 4.** Southern blot analyses of total DNA from several *Pseudomonas syringae* pathovars digested with *Hind*III and probed with a  $^{32}$ P-labeled *Xmn*I intragenic fragment from *avrD* (A) or with a  $^{32}$ P-labeled, 2.8-kilobase (kb) *Bam*HI-*Hind*III fragment containing a portion of open-reading frame (ORF) 4 and ORF 5, as discussed in the text (B). Lane 1, *P. s. pv. atropurpurea*; lane 2, *P. s. pv. glycinea* race 0; lane 3, *P. s. pv. glycinea* race 4; lane 4, *P. s. pv. glycinea* race 5; lane 5, *P. s. pv. glycinea* race 6; lane 6, *P. s. pv. lachrymans*; lane 7, *P. s. pv. mori*; lane 8, *P. s. pv. morsprunorum*; lane 9, *P. s. pv. phaseolicola*; lane 10, *P. s. pv. pisi*; lane 11, *P. s. pv. savastanoi*; lane 12, *P. s. pv. syringae* isolated from bean; lane 13, *P. s. pv. syringae* isolated from tomato; lane 14, *P. s. pv. tabaci*; and lane 15, *P. s. pv. tomato* PT23.

by *avrD* mapped to the first of five tandem ORFs present in the 5.6-kb *Hind*III fragment from *P. s. pv. tomato* (Figs. 1 and 2). The size of the protein predicted for ORF 1 by sequence data was confirmed when *E. coli* cells carrying *avrD* produced a new protein of 34 kDa (Fig. 3). However, it cannot be ruled out that ORFs 2, 3, 4, and 5 may function in conjunction with ORF 1 for expression of the phenotype conferred by *avrD* in *P. s. pv. glycinea*.

Unlike the similarity of *avrB* and *avrC* (Tamaki *et al.* 1988), *avrD* does not show significant homology at the DNA sequence level to any previously characterized *avr* gene. However, as noted by Kobayashi *et al.* (1990), *avrD* has considerable homology to a gene occurring in *P. s. pv. glycinea* that does not have the phenotype conferred by *avrD*. Some similarities also exist between *avrD* and previously characterized *avr* genes from *P. s. pv. glycinea*. For instance, the lack of a signal peptide sequence and of significant regions of hydrophobicity throughout the protein are similar to findings for *avrA* (Napoli and Staskawicz 1987) as well as for *avrB* and *avrC* (Tamaki *et al.* 1988). It therefore appears that the products of these *avr* genes are not membrane-associated or secreted from the bacteria, suggesting that they may not be directly involved in elicitation of the HR. This indeed appears to be the case for *avrD*, as shown in an accompanying research article (Keen *et al.* 1990).

Promoter or regulatory regions have not been identified for *avrD* by transcriptional studies, and the region 5' to *avrD* does not contain consensus sequences of known *Pseudomonas* promoters. Deletion analysis identified a CA-rich sequence that was followed by an AT-rich palindromic region, occurring in the 150 base pairs upstream from the translational start, which appears to



**Fig. 5.** Agarose gel electrophoresis of indigenous plasmid DNA from four different isolates of *Pseudomonas syringae* *pv. tomato* (A) stained with ethidium bromide and photographed under UV light, and Southern blot of the gel in A hybridized to the insert DNA of pPTD1211 containing *avrD* (B).



be important for expression of the avirulence phenotype in *P. s. pv. glycinea* (Fig. 2). Although the role of this AT-rich sequence in regulation is currently unclear, similar AT-rich sequences have been reported in *Pseudomonas* gene regulation (Deretic *et al.* 1987).

It should also be noted that the nine base sequence GCCACACAG present in the CA-rich region of *avrD* occurs in the 5' region upstream from the translational start of *avrA* (Napoli and Staskawicz 1987) and *avrB* (Tamaki *et al.* 1988). However, it is not known if this sequence is important for gene expression in *P. s. pv. glycinea*. Of interest is our preliminary observations in RNA slot blot and *lux* reporter gene experiments that indicated significantly increased expression of *avrD* when *P. s. pv. glycinea* cells were inoculated into soybean leaves and their growth was compared to growth on common laboratory culture media (D. Y. Kobayashi, H. Shen, and N. T. Keen, unpublished).

Southern blot analyses using two different DNA probes indicated that the entire 5.6-kb *HindIII* fragment containing *avrD* is conserved in some but not all *P. syringae* pathovars (Fig. 4). The different intensities of hybridization observed with the gene-specific *avrD* probe also showed that highly homologous sequences are present in some but not all pathovars and that *P. s. pv. glycinea* as well as *P. s. pv. phaseolicola* possess a sequence with less homology. Although similar findings have been reported with *avr* genes from *X. c. pv. malvacearum* (Smith) Dye (Gabriel *et al.* 1986), homologous but nonfunctional *avr* gene sequences have not been thus far observed in *P. s. pv. glycinea* races lacking a certain *avr* gene (Staskawicz *et al.* 1984, 1987). The weak homology of *avrD* to all tested *P. s. pv. glycinea* races may, therefore, reflect evolutionary changes in the nucleotide sequence of the *P. s. pv. glycinea* gene sufficient to cause the loss of the avirulence phenotype, as will be discussed in a subsequent research article (Kobayashi *et al.* 1990).

#### ACKNOWLEDGMENTS

Soybean seed was generously supplied by R. I. Buzzell and E. E. Hartwig. We thank Don Cooksey for discussion of the B plasmid of *P. s. pv. tomato* as well as several of the *P. syringae* isolates and a blot containing plasmid DNA from strain PT23.

Research was supported by National Science Foundation grant DCB-8709867.

#### LITERATURE CITED

Bender, C. L., and Cooksey, D. A. 1986. Indigenous plasmids in *Pseudomonas syringae* pv. *tomato*: Conjugative transfer and role in copper resistance. *J. Bacteriol.* 165:534-541.

Deretic, V., Gill, J. F., and Chakrabarty, A. M. 1987. Alginate biosynthesis: A model system for gene regulation and function in *Pseudomonas*. *Bio/Technology* 5:469-477.

Ditta, G., Stanfield, S., Corbin, D., and Helinski, D. R. 1980. Broad host range DNA cloning system for gram-negative bacteria: Construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. USA* 77:7347-7351.

Ellingboe, A. H. 1976. Genetics of host-parasite interactions. Pages 761-778 in: *Encyclopedia of Plant Pathology*, New Series, Vol. 4: Physiological Plant Pathology. R. Heitefuss and P. H. Williams, eds. Springer-Verlag, Heidelberg.

Flor, H. M. 1942. Inheritance of pathogenicity in *Melampsora lini*.

*Phytopathology* 32:653-669.

Gabriel, D. W., Burges, A., and Lazo, G. R. 1986. Gene-for-gene interactions of five cloned avirulence genes from *Xanthomonas campestris* pv. *malvacearum* with specific resistance genes in cotton. *Proc. Natl. Acad. Sci. USA* 83:6415-6419.

Hagborg, W. A. F. 1970. A device for injecting solutions and suspensions into thin leaves of plants. *Can. J. Bot.* 48:1135-1136.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.

Keen, N. T., and Holliday, M. J. 1982. Recognition of bacterial pathogens by plants. Pages 179-217 in: *Phytopathogenic Prokaryotes*, Vol. 2. M. S. Mount and G. H. Lacy, eds. Academic Press, New York.

Keen, N. T., and Staskawicz, B. 1988. Host range determinants in plant pathogens and symbionts. *Annu. Rev. Microbiol.* 42:421-440.

Keen, N. T., Tamaki, S., Kobayashi, D., and Trollinger, D. 1988. Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene* 70:191-197.

Keen, N. T., Tamaki, S., Kobayashi, D., Gerhold, D., Stayton, M., Shen, H., Gold, S., Lorang, J., Thordal-Christensen, H., Dahlbeck, D., and Staskawicz, B. 1990. Bacteria expressing avirulence gene D produce a specific elicitor of the soybean hypersensitive reaction. *Mol. Plant-Microbe Interact.* 3:112-121.

King, E. O., Ward, M. K., and Raney, D. E. 1954. Two simple media for the demonstration of phycocyanin and fluorescein. *J. Lab. Clin. Med.* 44:301-307.

Klement, Z. 1982. Hypersensitivity. Pages 149-177 in: *Phytopathogenic Prokaryotes*, Vol. 2. M. S. Mount and G. H. Lacy, eds. Academic Press, New York.

Kobayashi, D. Y., Tamaki, S. J., and Keen, N. T. 1989. Cloned avirulence genes from the tomato pathogen *Pseudomonas syringae* pv. *tomato* confer cultivar specificity on soybean. *Proc. Natl. Acad. Sci. USA* 86:157-161.

Kobayashi, D. Y., Tamaki, S. J., Trollinger, D. J., Gold, S., and Keen, N. T. 1990. A gene from *Pseudomonas syringae* pv. *glycinea* with homology to avirulence gene D from *P. s. pv. tomato* but devoid of the avirulence phenotype. *Mol. Plant-Microbe Interact.* 3:103-111.

Kyte, J., and Doolittle, R. F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.

Long, M., Barton-Willis, P., Staskawicz, B. J., Dahlbeck, D., and Keen, N. T. 1985. Further studies on the relationship between glyceollin accumulation and the resistance of soybean leaves to *Pseudomonas syringae* pv. *glycinea*. *Phytopathology* 75:235-239.

Maniatis, T., Fritsch, E. F., and Sambrook, J. 1982. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Napoli, C., and Staskawicz, B. 1987. Molecular characterization and nucleic acid sequence of an avirulence gene from race 6 of *Pseudomonas syringae* pv. *glycinea*. *J. Bacteriol.* 169:572-578.

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.

Shine, J., and Dalgarno, L. 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. USA* 71:1342-1346.

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

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

Swanson, J., Kearney, B., Dahlbeck, D., and Staskawicz, B. 1988. Cloned avirulence gene of *Xanthomonas campestris* pv. *vesicatoria* complements spontaneous race-change mutants. *Mol. Plant-Microbe Interact.* 1:5-9.

Tamaki, S., Dahlbeck, D., Staskawicz, B., and Keen, N. T. 1988. Characterization and expression of two avirulence genes cloned from *Pseudomonas syringae* pv. *glycinea*. *J. Bacteriol.* 170:4846-4854.

Vieira, J., and Messing, J. 1987. Production of single-stranded plasmid DNA. *Methods Enzymol.* 153:3-11.

Whalen, M. C., Stall, R. E., and Staskawicz, B. J. 1988. Characterization of a gene from a tomato pathogen determining hypersensitive resistance in non-host species and genetic analysis of this resistance in bean. *Proc. Natl. Acad. Sci. USA* 85:6743-6747.