Characterization of the *Pseudomonas syringae* pv. *syringae* 61 *hrpJ* and *hrpI* Genes: Homology of HrpI to a Superfamily of Proteins Associated with Protein Translocation

Hsiou-Chen Huang[1], Yingxian Xiao[2], Rong-Hwa Lin[1], Yang Lu[2], Steven W. Hutcheson[2], and Alan Collmer[3]

[1]Agricultural Biotechnology Laboratories, National Chung-Hsing University, Taichung, Taiwan 40227 Republic of China; [2]Department of Botany, University of Maryland, College Park, MD 20742 U.S.A.; [3]Department of Plant Pathology, Cornell University, Ithaca, NY 14853 U.S.A.

The *Pseudomonas syringae* pv. *syringae* 61 *hrpJ* and *hrpI* genes were sequenced and found to encode predicted proteins of 37,710 Da and 76,490 Da, respectively. The products of these genes were confirmed by T7 polymerase-dependent expression and sodium dodecyl sulfate-polyacrylamide gel analysis. HrpI belongs to a superfamily of proteins represented by *Yersinia pestis* LcrD.

Additional keywords: hypersensitive response, virulence.

The ability of bacteria to elicit the hypersensitive response (HR) in plants is controlled by *hrp* genes, which are so named because they are required for both the HR (on nonhost plants) and pathogenicity (on host plants) (Lindgren *et al.* 1986; Willis *et al.* 1991). *hrp* gene clusters have been demonstrated in *Pseudomonas syringae, Erwinia amylovora, P. solanacearum,* and *Xanthomonas campestris* (Lindgren *et al.* 1986; Boucher *et al.* 1987; Beer *et al.* 1991; Arlat *et al.* 1991; Bonas *et al.* 1991). The *hrp* gene clusters cloned from *P. syringae* pv. *syringae* 61 and *E. amylovora* Ea321 are unique in that they enable nonpathogenic bacteria like *P. fluorescens* and *Escherichia coli* to elicit the HR (Huang *et al.* 1988; Beer *et al.* 1991). Wei *et al.* (1992) recently reported that the *E. amylovora* Ea321 *hrp* cluster is responsible for the production of harpin, a 44-kDa protein that elicits the HR. An extracellular, proteinaceous factor that elicits an HR-like necrosis in tobacco has been reported for *P. solanacearum* (Genin *et al.* 1993).

DNA sequence analyses of *hrp* regions that are conserved among these bacteria suggest that some of the genes encode protein translocators. For example, the *P. syringae* pv. *syringae* 61 HrpH protein, identified by Tn*phoA* muta-

genesis as one of at least two envelope proteins encoded by the *hrp* cluster, is a member of a superfamily of outer membrane proteins involved in protein and phage secretion and is most similar to the YscC protein of *Yersinia enterocolitica,* which is required for the secretion of "Yop" virulence proteins (Michiels *et al.* 1991; Huang *et al.* 1992). Sequencing of the *hrp* clusters of *P. solanacearum* and *X. campestris* pv. *vesicatoria* has revealed putative open reading frames (ORFs) predicting similar proteins in these bacteria as well as homologs of the *Yersinia* YscJ and LcrD proteins (Fenselau *et al.* 1992; Gough *et al.* 1992). The second *P. syringae* pv. *syringae* 61 *hrp* gene identified by Tn*phoA* mutagenesis to encode an envelope protein (Huang *et al.* 1991) has been designated *hrpI* and is described here.

A 3,603-bp region of DNA spanning the contiguous *hrpJ* and *hrpI* genes was sequenced on both strands (Figs. 1,
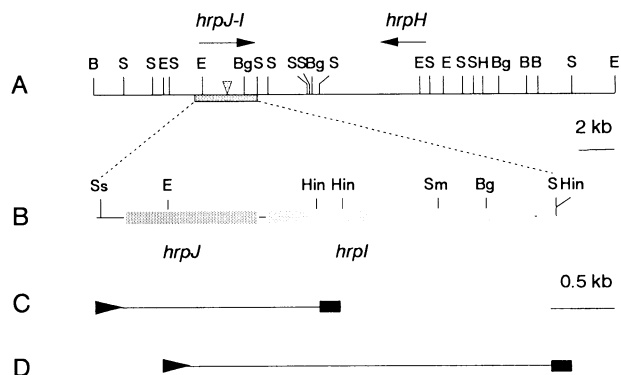

Fig. 1. Location in pHIR11 of *hrpJ-I* and *hrpH* and regions subcloned for T7 polymerase-dependent expression of *hrpJ* and *hrpI*. **A,** Restriction map of the pHIR11-derivative, pCPP2078, showing location of the PhoA⁺ Tn*phoA* insertion in *hrpI* (arrowhead). Arrows indicate the direction of transcription of *hrpJ-I* and *hrpH*. **B,** The region whose sequence is presented in Figure 2. **C,** The *hrpJ⁺* expression plasmid pYEXEX3. **D,** The *hrpI⁺* expression plasmid pYXEX3. The position and orientation of the vector T7 promoter in pYEXEX3 and pYXEX3 is represented by arrowheads, and the position of the vector Tφ terminator is represented by filled boxes. Restriction sites: Ba, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; Hin, *Hin*cII; S, *Sal*I; Sm, *Sma*I; Ss, *Ssp*I.

Corresponding author: Hsiou-Chen Huang.

```
   1  AGTGCAAGCCCTATAGCGATAACCATGCCAGCCAAAAAATATTTATAAGCATTTGTTTTATAAGACTATTTTAAAATTTGAAATTCTTTGAAATACATTT
 101  TTTTGCGCAACACGGTTTTTGGCTCTGCAGCTTTTGGCACAAACGCATAAGCCTTCGGCACCAGGGAACCGAACCGCGTCAATGACCCACTCAGCGGTCA
 201  TGACCAGGCCCAGAATCGTCCTATGAAAATCGTCGCCCCGCCTACATTACCCATTCGTCCGGTAGCGCCAATCCGCGCCATCACACCGGCTGCGCGAGCG
                       HrpJ  M  K  I  V  A  P  P  T  L  P  I  R  P  V  A  P  I  R  A  I  T  P  A  A  R  A
 301  ATTCCCGGCAGCGGGTTGCCTGATGAAAAAGGCACCTCGTCACTGCAGGTCTCGCGCTTCGCCGCAGCGCTGGTGCAGCACAGCCGGATACTGCGTGAGC
        I  P  G  S  G  L  P  D  E  K  G  T  S  S  L  Q  V  S  R  F  A  A  A  L  V  Q  H  S  R  I  L  R  E  R
 401  GAGAGCTGATCGCGAGCAGGAACGCGCTGCAGTCGCGGGCCGTCAAACTGGGCGAGCTTTATCAGTTGCTGATGAGCACCAGCGACACCGGGCTGGATAA
        E  L  I  A  S  R  N  A  L  Q  S  R  A  V  K  L  G  E  L  Y  Q  L  L  M  S  T  S  D  T  G  L  D  N
 501  CGCCGCCAGATTGCTGCGCAAAAAACTGCTGCAGGACAACGACGCCGATCTGGAACAGGTTCTGGAATTCGCTGACGGCGACGCTGCCAAGGCTCACGTG
        A  A  R  L  L  R  K  K  L  L  Q  D  N  D  A  D  L  E  Q  V  L  E  F  A  D  G  D  A  A  K  A  H  V
 601  GTGTTGCAGGCAGCGCGCAAACAGGCAGAAGACGATGGGGCTGAAGAAGAATATGCAGCGCTGACCCAGACGCTTAAACATCTGCGCCGCCGATTCGGTC
        V  L  Q  A  A  R  K  Q  A  E  D  D  G  A  E  E  E  Y  A  A  L  T  Q  T  L  K  H  L  R  R  R  F  G  P
 701  CGCGCGCCCGCGCCGGTATCAATACCGCAAGAGCCTTCGGGCGCCAGAACATCGATAACAAGCGCCGCACTGCCTTGCGCAATCTGTATGGCGTAGCGGT
        R  A  R  G  I  N  T  A  R  A  F  G  R  Q  N  I  D  N  K  R  R  T  A  L  R  N  L  Y  G  V  A  V
 801  GTCGGGACAACCCAACGTCACGGGGTTTGATCGAAGCCCTGATTGGCGAGCAACAGGAACCCGGCGAGTTCGACCTGAACCTGCGCGACATGCGTAGAGCC
        S  G  Q  P  N  V  T  G  L  I  E  A  L  I  G  E  Q  Q  E  P  G  E  F  D  L  N  L  R  D  M  R  R  A
 901  ATCGCCGACGATCTTTCGGCGATTACCCCGTCGGCTTCCCATGAACAACTGCGCACGCTGATGCATGGCCTGACCACCGCGCGCCATGTCACGACGTTGC
        I  A  D  D  L  S  A  I  T  P  S  A  S  H  E  Q  L  R  T  L  M  H  G  L  T  T  A  R  H  V  T  T  L  L
1001  TGCGCGGTTGCGAACACCTGCTGGGACGCATGCGCAAAAAGAATCCCAAACTGACCGTGGACCCTCCGGCCTTCCTCAAACACATGCTGACCCTGACGGC
        R  G  C  E  H  L  L  G  R  M  R  K  K  N  P  K  L  T  V  D  P  P  A  F  L  K  H  M  L  T  L  T  A
1101  CAATGGCATGAACGTGAATCAGACCTTGCAGCTGACCCAGCACATCGGGGGCAACAAACTCGAACACCAGTTGGCGTTTCTCAATGGCCTGCGGGCCGAT
        N  G  M  N  V  N  Q  T  L  Q  L  T  Q  H  I  G  G  N  K  L  E  H  Q  L  A  F  L  N  G  L  R  A  D
1201  GCTGATGCAACTGCCGATCCTGCTCTGGCGGGACCTGAAAAGCCGCCAGGCCGCGCTGAATAATCTACTCACGCTGATGGCCGAGCTTACGCATAAGGAA
        A  D  A  T  A  D  P  A  L  A  G  P  E  K  P  P  G  R  A  E  *
1301  CAGAAGCAACTATACGAGGGGCTTGCCTGATGAACCGGGTCATCAACTTTCTGAACATGGTCGCGCTGTCGGCCATGCGCCGCTCGGAACTGGTCGGTGC
                         HrpI  M  N  R  V  I  N  F  L  N  M  V  A  L  S  A  M  R  R  S  E  L  V  G  A
1401  GTTTTTTGTGATCGCCATCGTGTTCATGATGATCACGCCGCTGCCCACCGGGTTGATCGACGTGCTGATCGCCGTGAACATCTGTATTTCCTGCCTGCTG
        F  F  V  I  A  I  V  F  M  M  I  T  P  L  P  T  G  L  I  D  V  L  I  A  V  N  I  C  I  S  C  L  L
1501  ATCATGCTGGCGATGCACCTGCCAAGGCCACTGGCGTTCTCGACCTTCCCTGCCGTGCTGCTGCTGACCACGATGTTTCGTCTGGCCCTGTCGGTTTCCA
        I  M  L  A  M  H  L  P  R  P  L  A  F  S  T  F  P  A  V  L  L  L  T  T  M  F  R  L  A  L  S  V  S  T
1601  CCACGCGGTTGATTCTGCTCAACCAGGACGCCGGACACATCGTCGAGGCGTTCGGGCAGTTCGTGGTGGGCCCCAATCTGGCGGTGGGTCTGGTGATCTT
        T  R  L  I  L  L  N  Q  D  A  G  H  I  V  E  A  F  G  Q  F  V  V  G  P  N  L  A  V  G  L  V  I  F
1701  CCTGATCCTGACGGTGGTCAACTTTCTAGTGATTACCAAAGGCTCGGAACGGGTCGCCGAAGTCGGTGCGCGTTTTACCCTGGACGCGATGCCCGGCAAG
        L  I  L  T  V  V  N  F  L  V  I  T  K  G  S  E  R  V  A  E  V  G  A  R  F  T  L  D  A  M  P  G  K
1801  CAGATGTCCATCGACAGTGACCTGCGTGCCAACCTGATCACCGTCCACGAGGCCCGCAAACGACGCGCTGAACTGAACAAGGAAAGCCAGTTGTTCGGTG
        Q  M  S  I  D  S  D  L  R  A  N  L  I  T  V  H  E  A  R  K  R  R  A  E  L  N  K  E  S  Q  L  F  G  A
1901  CGATGGATGGCGCCATGAAGTTCGTCAACGGCGATGCGATCGCCAGCCTGATCATCGTTGCCATCAACATGATCGGCGGCATTTCCATCGGCGGTGCTGCA
        M  D  G  A  M  K  F  V  N  G  D  A  I  A  S  L  I  I  V  A  I  N  M  I  G  G  I  S  I  G  V  L  Q
2001  ACACAACATGGCGGCAGGCGACGCCTTGCAGCTGTATACCGTGCTGACCATCGGTGACGGCCTGATCGCGCAGATTCCTGCGGCTGCTGATTTCCGTGACC
        H  N  M  A  A  G  D  A  L  Q  L  Y  T  V  L  T  I  G  D  G  L  I  A  Q  I  P  A  L  L  I  S  V  T
2101  AGCGGCATGATCATCACCCGCGTGCCCAACACCGAGGCGGGTGTGGAAGCCAACATTGGCCGCGAGATCGCCGAGCAGATCACCAGCCAGCCCAAGGCCT
        S  G  M  I  I  T  R  V  P  N  T  E  A  G  V  E  A  N  I  G  R  E  I  A  E  Q  I  T  S  Q  P  K  A  W
2201  GGATCATCGCCTCGGTGGCCATGCTCGGCTTCGCGGCCCTGCCGGGAATGCCGACCGGTGTTTTCATTACCATCGCGATCATCTGCGGTGCCGGTGGACT
        I  I  A  S  V  A  M  L  G  F  A  A  L  P  G  M  P  T  G  V  F  I  T  I  A  I  I  C  G  A  G  G  L
2301  GCTGCAGTTGCAACGAGCAAAACCCAAGGCCGATGAGCAACGTACAGCCGCTGTGGCGCCAGAAATGAACGGCAAGGAAGACCTGCGTACTTTTTCGCCC
        L  Q  L  Q  R  A  K  P  K  A  D  E  Q  R  T  A  A  V  A  P  E  M  N  G  K  E  D  L  R  T  F  S  P
2401  AGCCGCCAGTTTGTGCTGCAATTCCACCCCGGTCAGGACAGCGCCACAGATCGAAGCGCTGGTCAGCGAAATACGCAAGCGCCGCAACCGTCTGGTGGTGC
        S  R  Q  F  V  L  Q  F  H  P  G  Q  D  S  A  Q  I  E  A  L  V  S  E  I  R  K  R  R  N  R  L  V  V  Q
2501  AGTACGGCCTGACCCTGCCTTCGTTCATCATCGAGCATGTCGATGACATTGCCCCGGACGAGTTTCGCTTCACCGTCTATGACGTACCGATGCTCAAGGC
        Y  G  L  T  L  P  S  F  I  I  E  H  V  D  D  I  A  P  D  E  F  R  F  T  V  Y  D  V  P  M  L  K  A
2601  CACCTTCACGCAGAGCCATGTCGCCGTGGAAGCGCGCCAGCTGGAAGGCGAAAACCTTCCAGCGGCCATCCCGGGAAATACCGACCGCCAGGAAGATCAA
        T  F  T  Q  S  H  V  A  V  E  A  R  Q  L  E  G  E  N  L  P  A  A  I  P  G  N  T  D  R  Q  E  D  Q
2701  TGGGTCTGGCTGCCTGCCGAGCAGTCTGGCGAATTGAACCCGGTCAGTTCCACGACGCTGCTCATCGAACGCATGGAGCGTGCATTGCAGTCCTGCGCGC
        W  V  W  L  P  A  E  Q  S  G  E  L  N  P  V  S  S  T  T  L  L  I  E  R  M  E  R  A  L  Q  S  C  A  P
2801  CGCAGTTCATCGGCTTGCAGGAAACCAAAGCGATCCTCAGCTGGCTGGAGTCCGAGCAACCGGAACTCGCTCAGGAAATGCAGCGGGTCCTCACCCTGAC
        Q  F  I  G  L  Q  E  T  K  A  I  L  S  W  L  E  S  E  Q  P  E  L  A  Q  E  M  Q  R  V  L  T  L  T
2901  GCGCTTCTCGGCCGTCCTGCAACGCCTGGCCTCGGAGTGTGTGCCGCTGAGGGCGATCCGGGTCATTGCCGAAACCCTGATCGAGCACTGCCAGCATGAA
        R  F  S  A  V  L  Q  R  L  A  S  E  C  V  P  L  R  A  I  R  V  I  A  E  T  L  I  E  H  C  Q  H  E
3001  CGCGATACCAATGTGCTGACCGACTATGTGCGCATTGCGCTCAAGTCGCAGATCTACCACCAGTACTGCGGTGCCGAAGGCCTGCAAGTGTGGCTCGTCA
        R  D  T  N  V  L  T  D  Y  V  R  I  A  L  K  S  Q  I  Y  H  Q  Y  C  G  A  E  G  L  Q  V  W  L  V  T
3101  CGCCGGAAAGCGAGGGGTTGTTGCGCGATGGCCTGCGCCAGACGCAAACCGAAACATTCTTCGCCGCTGAGCAACGAGACCAGCCAGATGCTCGTGCAACA
        P  E  S  E  G  L  L  R  D  G  L  R  Q  T  Q  T  E  T  F  F  A  L  S  N  E  T  S  Q  M  L  V  Q  Q
3201  GCTGCACATTGCCTTCCCGGTACGAGCTCCCGAACAGGCGGTTCTGCTGGTGGCACAGGATCTGCGCAGCCCTCTGCGTACCTTGTTGAGGGAGGAGTTC
        L  H  I  A  F  P  V  R  A  P  E  Q  A  V  L  L  V  A  Q  D  L  R  S  P  L  R  T  L  L  R  E  E  F
3301  TACCATGTACCCGTACTATGTTTTGCGGAAATCAGCAACGCGGCCAAGGTCAAGGTCATGGGTCGATTCGACCTTGAAGACGATCTCGAGCCGCTGGACA
        Y  H  V  P  V  L  C  F  A  E  I  S  N  A  A  K  V  K  V  M  G  R  F  D  L  E  D  D  L  E  P  L  D  N
3401  ACGAGCACGCTGCCTGAGCCTGCAAGGCCTGAATGCACGGAGACTTTCAGATGTTTGAATTACGCGTACTGAATGGCCAGCACCAAGGCGCGGCATTACC
        E  H  A  A  *
3501  GCTGATCGGTGAGCAATGGTCGATTGGCTCTGCCGGGCAACACGATCTGGCGCTGGACGATGCGGGCGTGGAAAGCCTGCATTGCCGCCTGCAACGTGTC
3601  GAC  3603
```

Fig. 2. Nucleotide sequence of a 3,603-bp region of *Pseudomonas syringae* pv. *syringae* 61 chromosomal DNA containing *hrpJ* and *hrpI*. The hatched region denotes Harp box sequences; potential ribosome binding sites preceding the *hrpJ* and *hrpI* ORFs are underlined.

2). Subclones were constructed (Sambrook et al. 1989) in pBluescript plasmids (Stratagene), partially deleted with either the Erase-a-Base kit (Promega, Madison WI) or the ExoIII/Mung Bean Deletion kit (Stratagene), and then sequenced from double-stranded templates by the chain termination method of Sanger et al. (1977) with the Sequenase version 2.0 sequencing kit (U.S. Biochemical, Cleveland OH). Both dGTP and dITP reaction mixtures were examined. Data were analyzed with the Genetics Computer Group version 7.1 (Devereaux et al. 1984) programs. The GenBank accession numbers are L11281 for hrpJ and L11582 for hrpI.

Predicted ORFs extended from positions 223 to 1260 for hrpJ and from 1330 to 3414 for hrpI. The direction of transcription was consistent with the orientation of TnphoA and Tn5-gusA1 insertions that possessed reporter enzyme activity (Huang et al. 1991; Xiao et al. 1992). A potential ribosome binding site (AGGCCCAGAA) was located 7 bp upstream of hrpJ; another (GAGGGG) was 9 bp upstream of hrpI. The first 13 nucleotides in the sequenced region agreed closely with the last 13 nucleotides of the "Harp box" consensus sequences found upstream of P. syringae pv. phaseolicola hrp genes (Fellay et al. 1991). No rho-independent terminator sequences were found downstream of hrpI. Instead, an ATG at position 3451 was preceded by a consensus ribosome binding site (GGAGA) and followed by an ORF that ran through position 3603. The molecular masses of the predicted proteins encoded by the two ORFs were 37,710 Da for hrpJ and 76,490 Da for hrpI.

To visualize the products of hrpJ and hrpI, plasmids pYXEX3 (hrpJ⁺) and pYXEX5 (hrpI⁺) were constructed (Fig. 1). First, a 1.7-kb SspI-HincII fragment containing the hrpJ region was subcloned into pBluescript II SK+. The new construct was digested by XbaI and KpnI to isolate the fragment with XbaI and KpnI ends. The T7 promoter vector pVEX11f⁺T8/24 (kindly provided by Adhya Sankar of NCI, NIH, Bethesda, MD) was used for expressing hrp genes. This derivative of pET-3a (Studier et al. 1990) is modified by insertion of a multicloning site. pVEX11f⁺T8/24 was digested with XbaI and KpnI to remove the Shine-Dalgarno sequence and ATG codon in the vector. The XbaI-KpnI fragment was then ligated into the vector to form pYXEX3, in which the native translation start sequence and the putative ORF were downstream of the vector T7 promoter. A 3.1-kb EcoRI-SalI fragment containing the hrpI region from a subclone of pHIR11 was also cloned into pVEX11f⁺T8/24 by the same procedure to form pYVEX5. pVEX11f⁺T8/24, pYXEX3, and pYXEX5 were then transformed into E. coli B121(DE3) (Studier et al. 1990).

The product of the hrp gene in each plasmid was specifically labeled by T7 promoter/polymerase-dependent expression in E. coli BL21 cells incubated in the presence of [³⁵S]-methionine (Studier et al. 1990). [³⁵S]-labeled proteins in the cell lysates were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and autoradiography (Fig. 3). Lysates of the hrpJ⁺ plasmid contained a unique band with an apparent molecular mass of 39 kDa, which agreed well with the mass of the predicted protein. The hrpI⁺ plasmid produced a protein with an apparent molecular mass of 67 kDa, which was somewhat less than that of the predicted protein but similar to the observed molecular mass of other members of the LcrD superfamily (Plano et al. 1991; Galán et al. 1992).

A search of the GenBank database with the FASTA program (Pearson and Lipman 1988) for comparisons of predicted amino acids revealed no proteins with similarity to hrpJ. However, hrpI showed strong similarity to several proteins in a superfamily of proteins that includes LcrD of Y. pestis (Plano et al. 1991), VirH of S. flexneri (Galán et al. 1992), InvA of S. typhimurium (Galán et al. 1992), FlbF of C. crescentus (Ramakrishnan et al. 1991), and two hrp homologs in P. solanacearum (Gough et al. 1992) and X. campestris pv. vesicatoria (Fenselau et al. 1992) (Fig. 4). Conservation of sequences in all of these proteins was highest in the hydrophobic, membrane-spanning regions in the amino-terminal portion of the protein (Fig. 4).

Gough et al. (1992) recently reported that Southern blot analysis of E. amylovora and P. syringae pv. phaseolicola DNA, probed with the P. solanacearum hrpO gene (the hrp homolog of the Yersinia lcrD), revealed the presence of hybridizing sequences. Our results confirm the presence of this gene and its protein product in P. syringae pv. syringae. Similarly, Beer et al. (1993) reported that the E. amylovora hrpI gene also encodes a protein in the LcrD superfamily.

Although the P. syringae pv. syringae 61 hrpI protein is required for the hrp⁺ phenotype, it is not known whether its primary function is in regulation or secretion. The function of members of the LcrD superfamily in other bacteria is similarly unclear, although they are all associated with processes involving protein translocation. As summarized
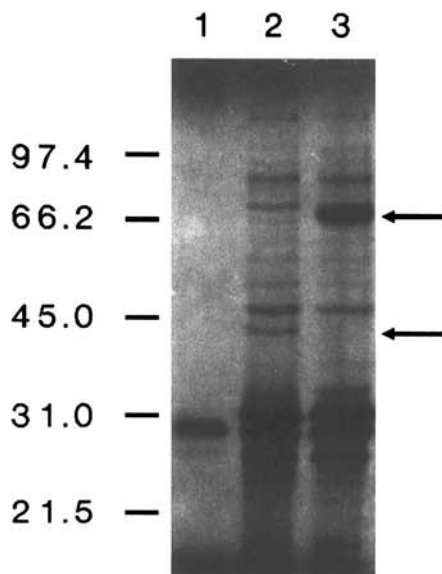
Fig. 3. Autoradiograph of the protein products of hrpJ and hrpI following T7 RNA polymerase-dependent expression in E. coli BL21(DE3) and SDS polyacrylamide gel electrophoresis. L-[³⁵S]methionine labeling and electrophoresis were performed as described in the text. Lanes: 1, BL21(DE3)(pVEX11f+T8/24); 2, BL21(DE3)(pYXEX3); 3, BL21(DE3)(pYXEX5). The positions of molecular mass standards (Bio-Rad) are indicated at the left in kilodaltons. The upper arrow indicates hrpI in lane 3; the lower arrow indicates hrpJ in lane 2.

```
HrpI    - - - - - - - M N R V I N F L N M V A L S A M R R S E L V G A F F V I A I V F M M I T P L P T G L I D    44
HrpO    - - - - - - - - - - - - M A K K N A I Q D F S G E I G I A A L V V A V V A L M V L P L P T M L I D    37
HrpC2   - - - - - - - - - - M L G D R V R A T R Y F A Y S G E V A I A A L V V A V I G L M I L P L P T P M I D    41
LcrD    - - - - - - - - M N P H D L E W L N R I G E - R K D I M L A V L L A V V F M M V L P L P P L V L D    41
VirH    - - - - - - - - - M I Q S F L K Q V S T K P E L I I L V L M V M I I A M L I I P L P T Y L V D    38
InvA    - - - - - - - - - M L L S L L N S A R L R P E L L I L V L M V M I I S M F V I P L P T Y L V D    38
FlbF    M A D A A A P N A S S M P S A K S L L D G L M R G E M G L A L G V V G I I V L L I P V P A P L L D    50

HrpI    V L I A V N I - C I S C L L I M L A M H L P R P L A F S - T F P A V L L L T T M F R L A L S V S T T    92
HrpO    A L L G L N I - T L S V V L L M V T M Y I P S A T S L S - A F P S L L L F T T L L R L S L N I A S T    85
HrpC2   T L I G I N I - T L S V V L L M V T M Y V P D S I S L S - S F P S L L L F T T L L R L S L N I A S T    89
LcrD    I L I A V N M - T I S V V L L M I A I Y I N S P L Q F S - A F P A V L L V T T L F R L A L S V S T T    89
VirH    F L I G L N I V V L A I L V F M G S F Y I E R I L S F S - T F P S V L L I T T L F R L A L S I S T S    87
InvA    F L I A L N I - V L A I L V F M G S F Y I D R I L S F S F T F P A V L L I T T L F R L A L S I S T S    87
FlbF    V L L A I S L - T G S V L I L M T A I L I K K P L E F T - S F P T V L L V T T L F R L G L N I A S T    98

HrpI    R L I L L N Q D - - - - - A G H I V E A F G Q F V V G P N L A V G L V I F L I L T V V N F L V I T K    137
HrpO    K S I L L H A D - - - - - A G H I I E S F G K L V V G G N L V V G L V V F L I I T V Q F I V I A K    130
HrpC2   K S I L L H A E - - - - - A G H I I E S F G K L V V G G N L V V G L V V F L I I T V Q F I V I A K    134
LcrD    R M I L L Q A D - - - - - A G Q I V Y T F G N F V V G G N L I V G I V I F L I I T I V Q F L V I T K    134
VirH    R L I L V D A D - - - - - A G K I I T T F G Q F V I G D S L A V G F V V F S I V T V V Q F I V I T K    132
InvA    R L I L I E A D - - - - - A G E I I A T F G Q F V I G D S L A V G F V V F S I V T V V Q F I V I T K    132
FlbF    R L I L S H G Q E G T G G A G A V I E A F G H L M M Q G N F V I G V I V F I I L I V V N F M V V T K    148

HrpI    G S E R V A E V G A R F T L D A M P G K Q M S I D S D L R A N L I T V H E A R K R R A E L N K E S Q    187
HrpO    G S E R V A E V G A R F T L D A M P G K Q M S I D A D L R A G H L S P E E A R K R R A L L A M E S Q    180
HrpC2   G S E R V A E V G A R F T L D A M P G K Q M S I D A D L R G G N L T A D E A R R K R A R L A M E S Q    184
LcrD    G S E R V A E V S A R F S L D A M P G K Q M S I D G D M R A G V I D V N E A R E R R A T I E K E S Q    184
VirH    G S E R V A E V A A R F S L D G M P G K Q M S I D A D L K A G I I D A A G A K E R R S I L E R E S Q    182
InvA    G S E R V A E V A A R F S L D G M P G K Q M S I D A D L K A G I I D A D A A R E R R S V L E R E S Q    182
FlbF    G S G R I A E V A A R F T L D S M P G K Q M A I D A D L S T G L I S Q D E A K I R R K E L E Q E S T    198

HrpI    L F G A M D G A M K F V N G D A I A S L I I V A I N M I G G I S I G V L Q H N M A A G D A L Q L Y T    237
HrpO    L H G G M D G A M K F V K G D A I A G L V I T L V N I L A G I V I G I T Y H N M T A G E A A N R F A    230
HrpC2   L H G G M D G A M K F V K G D A I A G L V I T M V N I L A G I V V G V T Y H G M T A F D A A N R F A    234
LcrD    M F G S M D G A M K F V K G D A I A G L I I I F V N I L G G V T I G V T Q K G L A A A E A L Q L Y S    234
VirH    L Y G S F D G A M K F I K G D A I A G I I I I F V N L I G G I S V G M S R H G M S L S G A L S T Y T    232
InvA    L Y G S F D G A M K F I K G D A I A G I I I I F V N F I G G I S V G M T R H G M D L S S A L S T Y T    232
FlbF    E F G A M D G A S K F V K G D A I A G L I I T A I N I I G G I I I G V V Q H K M P F G D A A S T Y T    248

HrpI    V L T I G D G L I A Q I P A L L I S V T S G M I I T R V P N - T E A G V E A N I G R E I A E Q I T S    286
HrpO    V L S I G D A M V S Q I P S L L I S V A A G V M I T R V S D E E Q A H K Q S S L G M E I V R Q L S T    280
HrpC2   I L S V G D A M V S Q I A S L L I S V A A G V M I T R V A N E N E T - R L S S L G L D I G R Q L T S    283
LcrD    I L T V G D G M V S Q V P A L L I A I T A G I I V T R V S S E D S - - - - S D L G S D I G K Q V V A    280
VirH    I L T I G D G L V S Q I P A L L I S I S A G F I V T R V N G D S D - - - - - N M G R N I M S Q I F G    277
InvA    M L T I G D G L V A Q I P A L L I A I S A G F I V T R V N G D T D - - - - - N M G R N I M T Q L L N    277
FlbF    I M T I G D G L V S Q I P A L I I S I A A G M V V S K A G V E G S A - - - - - - D K A L T T Q L A M    292

HrpI    Q P K A W I I A S V A M L G F A A L P G M P T G V F I T I A I I C G A G G L L Q L Q R A K P K A D E    336
HrpO    S A R A M F T A S A L L M G F A L V P G F P S F L F V A L A T L I F V F G Y T L R N R A K E G D G D    330
HrpC2   N A R A L M A A S V L L A C F A F V P G F P A V L F L L L A A A G A G G Y T I W R K Q R D I S G T    333
LcrD    Q P K A M L I G G V L L L L F G L I P G F P T V T F L I L A L L V G C G G Y M L S R K Q S R N D E A    330
VirH    N P F V L I V T S A L A L A I G M L P G F P F F V F F L I A V T L T A L F Y Y K - K V V E K E - - -    323
InvA    N P F V L V V T A I L T I S M G T L P G F P L P V F V I L S V V L S V L F Y F K F R E A K R S - - -    324
FlbF    N P V G L G M V S A S S G I I A L I P G M P I F P E A A M A L A A G A L A Y K R V Q D A K K P K A L    342

HrpI    Q R T A A V A P E M N G K E D L - - - - - - - - - - - - - - - - - - - - - - - R T F S P S R Q F V L Q F H P    367
HrpO    E G D A L P A L L R E G S K G K A P T I A - - - - - - - - - - - - - - E Q A P S F T V P V G V R L G A    367
HrpC2   D Q R K L P S A S R K G A K G E A P H I R - - - - - - - - - - - K N A P D F A S P L S M R L S P    370
LcrD    N Q D - L Q S I L T S G S G A P A A R T K A K T S G A N K G R L G E Q E A F A M T V P L L I D V D S    379
VirH    - - - - - K S L S D S S G Y T G T F D I D N S H D S S L A M I E N L D A I S S E T V P L I L L F A E    368
InvA    - - - - - A A K P K T S K G E Q P L S I E E K E G S S L G L I G D L D K V S T E T V P L I L L V P K    369
FlbF    D P A D L E A A A P S E P E E E P I S - - - - - - - - A S L A I D D V K I E L G Y G L L T L I - -    381
```
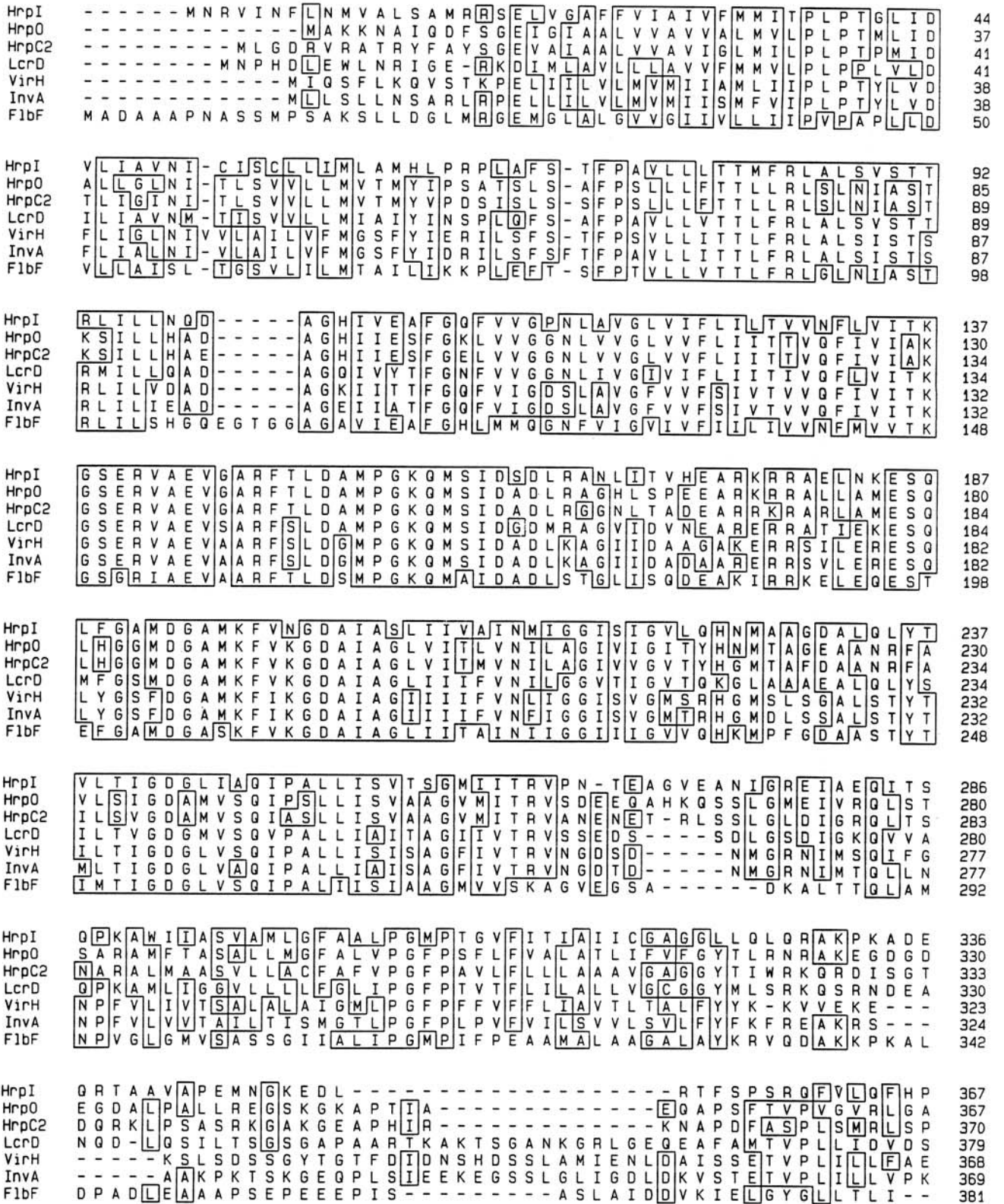
Fig. 4. Alignments (open boxes) in the predicted amino acid sequences of *P. syringae* pv. *syringae* HrpI, *P. solanacearum* HrpO, *X. campestris* pv. *campestris* HrpC2, *Y. pestis* LcrD, *S. flexneri* VirH, *S. typhimurium* InvA, and *C. crescentus* FlbF.

```
HrpI   G Q D S A - Q I E A L V S E I R K R R N R L V V Q Y G L T L P S F I I E H V D D I A P D E F R F T V   416
HrpO   E L A K G L D V P A L D T A F Q Q G R H A L A E A L G L P F P G I A I W K A D A L Q P D S Y E V R V   417
HrpC2  Q L A A L L D P A R L D Q A I E S E R R Q L V E L L G L P F P G I A I W Q T E S L Q G M Q Y E V L I   420
LcrD   S Q Q E A L E A N A L N D E t V R V R R A L Y L D L G V P E P G I H L R F N E G M G E G E Y I I S L   429
VirH   N K I N A N D M E G L - - - I E R I R S Q F F I D Y G V R L P T I L Y R T S N E L K V D D I V L L I   415
InvA   S R R E D L E K A Q L - - - A E R L R S Q F F I D Y G V R L P E V L L R D G E G L D D N S I V L L I   416
FlbF   - - - N D L D G R K L T D Q I R A L R K T L A S E Y G F V M P P V R I L D N M R L A N Q G Y A I R I   428


HrpI   Y D V P M L K A T F T Q S H V A V E A R Q L E G E N L P A A I P G N T D R Q E D Q W V W L P A E Q S   466
HrpO   H D I P G E P V A V P D G H L L I P D L P E A L R A Q A V - - - E A A G L P N H P A P H W I A P A H   464
HrpC2  H D V P E T R A E L E N T D - - - - D M Q A A L A R Q A I - - - S P - - - - - - - - - - - - - - -   447
LcrD   Q E V P V A R G E L K A G Y L L V R E S V S Q L E L L G I P Y E K G E H L L P D Q E A F W V S V E Y   479
VirH   N E V R A D S F N I Y F D K V C I T D E N G D I D A L G I P - - V V S T S Y N E R V I S W V D V S Y   463
InvA   N E I R V E Q F T V Y F D L M R V V N Y S D E V V S F G I N P T I H Q Q G S S Q - - Y F W V T H E E   464
FlbF   K E M E A G A G E V R L G C L M C M D P R G G Q V E L P G E H V R E P A F - - G L P A T W I A D D L   476


HrpI   G E L N P V S S T T L - - - - - - L I E R M E R A L Q S C A P Q F I G L Q E T K A I L S W L E S E Q   510
HrpO   V A Q D A A L S A T G Q R V E R V I A D H V V H V L R R S A H L F V G L Q E T Q W M L E R V T T D Y   514
HrpC2  - - - - - - - - - - - - - - - - - - L H A R A H L F V G I Q E T Q W M L E Q V A V D Y   472
LcrD   E E R L E K S Q L E F F S H S Q V L T W H L S H V L R E Y A E D F I G I Q E T R Y L L E Q M E G G Y   529
VirH   T E N L T N I D A K I K S A Q D E F Y H Q L S Q A L L N N I N E I F G I Q E T K N M L D Q F E N R Y   513
InvA   G E K L R E L G Y V L R N A L D E L Y H C L A V T V A R N V N E Y F G I Q E T K H M L D Q L E A K F   514
FlbF   R E E A T F R G Y T V V D P A T V L T T H L T E I L K E N M A D L L S Y A E V Q K L L K E L P E T Q   526


HrpI   P E L A Q E M Q R - V L T L T R F S A V L Q R L A S E C V P L R A I R V I A E T L I E H C Q H E R D   559
HrpO   P G L V A E A Q K - A V P A Q R I A D V L R R L L E E Q V P I R N M R A I L E S L V V W G P K E K D   563
HrpC2  P G L V A E V N K - A M P A Q R I A D V L R R L L E E R I P V R N I K S I L E S L V V W G P K E K D   521
LcrD   G E L I K E V Q R - I V P L Q R M T E I L Q R L V G E D I S I R N M R S I L E A M V E W G Q K E K D   578
VirH   P D L L K E V F R - H V T I Q R I S E V L Q R L L S E N I S V R N M K L I M E S L A L W A P R E K D   562
InvA   P D L L K E V L R - H A T V Q R I S E V L Q R L L S E R V S V R N M K L I M E A L A L W A P R E K D   563
FlbF   K K L V D D L I P G T V T A T T V Q R V L Q S L L R E R V S I R D L P Q I L E G V G E A A P H T A S   576


HrpI   T N V L T D Y V R I A L K S Q I Y H Q Y C G A E G - L Q V W L V T P E S E G L L R D G L R Q T Q T E   608
HrpO   T L M L V E Y V R G D L G R Q I A H Q A T G G T R Q M P A I L L D L S V E Q T V R Q A I K P T P A G   613
HrpC2  L L M L T E Y V R C D L G R Y L A H T A T A G T G Q L P A V M L D H A V E Q L I R Q S I R A T A A G   571
LcrD   V V Q L T E Y I R S S L K R Y I C Y K Y A N G N N I L P A Y L F D Q E V E E K I R S G V R Q T S A G   628
VirH   V I T L V E H V R A S L S R Y I C S K I A V S - G E I K V V M L S G Y I E D A I R K G I R Q T S G G   611
InvA   V I N L V E H I R G A M A R Y I C H K F A N G - G E L R A V M V S A E V D V I R K G I R Q T S G S   612
FlbF   V T Q L V E Q V R A R L A R Q L C W A N R G D D G A L P I I T L S A D W E Q A F A E A L I G P G D D   626


HrpI   T F F A L S N E T S Q M L V Q Q L H I A F - - - P V R A P E Q A V L L V A Q D L R S P L R T L L R E   655
HrpO   N F L T L D P Q Q V E A I I M R L R G I M Q G N P V E T P S A L A I V T S M D I R R Y V R R M I E P   663
HrpC2  N F L A L P P E Q A N Q L V E Q V E R I V G D H A Q H - - - P L A V V A S M D V R R Y V R R M I E A   618
LcrD   S Y L A L E P A V T E S L L E Q V R K T I G D L S Q I Q S K P V L I V - S M D I R R Y V R K L I E S   677
VirH   S F L N M - I E V S D E V M E T L A H A L R E L R N A K K N N F V L L V S V D I R R F V K R L I D N   660
InvA   T F L S L D P E A S A N L M D L I T L K L D D L L I A H K - D L V L L T S V D V R R F I K K M I E G   661
FlbF   K Q L A L P P S R L Q D F I R G V R D S F E R A A L A G E A P - V L L T S P G V R P Y V R S I I E R   675


HrpI   E F Y H V P V L C F A E I S N A A K V K V M G R F D L E D D L E P L D N E H A A   695
HrpO   H L Q A L N V Y S F Q E L G G Y V D L R P V G K L V L - - - - - - - - - - - - -   690
HrpC2  R L T W L G V Y S F Q E L G G S E V Q L Q P I G R V V V - - - - - - - - - - - -   645
LcrD   E Y Y G L P V L S Y Q E L T Q Q I N I Q P L G R I C L - - - - - - - - - - - - -   704
VirH   R F K S I L V I S Y A E I D E A Y T I N V L K T I I - - - - - - - - - - - - - -   685
InvA   R F P D L E V L S F G E I A D S K S V N V I K T I I - - - - - - - - - - - - - -   686
FlbF   F R G Q T V V M S Q N E I H P R A R L K T V G M V - - - - - - - - - - - - - - -   700
```

by Galán *et al.* (1992), current evidence suggests that the *S. flexneri* VirH and *S. typhimurium* InvA proteins are directly involved in protein translocation, whereas *Y. pestis* LcrD and *C. crescentus* FlbF may have a regulatory role. We have little clue of the function of *hrpJ*, but it is interesting to note that the protein has no homolog in the ORFs of the *hrp* cluster of *P. solanacearum* GMI1000 (C. A. Boucher, personal communication) or among the products of the *Y. enterocolitica ysc* operon, which is required for the secretion of Yop virulence proteins (Michiels *et al.* 1991).

LITERATURE CITED

Arlat, M., Gough, C. L., Barber, C. E., Boucher, C., and Daniels, M. J. 1991. *Xanthomonas campestris* contains a cluster of *hrp* genes related to the larger *hrp* cluster of *Pseudomonas solana-*

cearum. Mol. Plant-Microbe Intact. 4:593-601.

Beer, S. V., Bauer, D. W., Jiang, X. H., Laby, R. J., Sneath, B. J., Wei, Z.-M., Wilcox, D. A., and Zumoff, C. H. 1991. The *hrp* gene cluster of *Erwinia amylovora.* Pages 53–60 in: Advances in Molecular Genetics of Plant-Microbe Interactions. H. Hennecke and D. P. S. Verma, eds. Kluwer Academic Publishers, Dordrecht.

Beer, S. V., Wei, Z.-M., Laby, R. J., He, S.-Y., Bauer, D. W., Collmer, A., and Zumoff, C. 1993. Are harpins universal elicitors of the hypersensitive response of phytopathogenic bacteria? Pages 281-286 in: Advances in Molecular Genetics of Plant-Microbe Interactions, Vol. 2. E. W. Nester and D. P. S. Verma, eds. Kluwer Academic Publishers, Dordrecht.

Bonas, U., Schulte, R., Fenselau, S., Minsavage, G. V., Staskawicz, B. J., and Stall, R. E. 1991. Isolation of a gene cluster from *Xanthomonas campestris* pv. *vesicatoria* that determines pathogenicity and the hypersensitive response on pepper and tomato. Mol. Plant-Microbe Interact. 4:81-88.

Boucher, C. A., Van Gijsegem, F., Barberis, P. A., Arlat, M., and Zischek, C. 1987. *Pseudomonas solanacearum* genes controlling both pathogenicity on tomato and hypersensitivity on tobacco are clustered. J. Bacteriol. 169:5626-5632.

Devereaux, J., Haeberli, P., and Smithies, O. 1984. A comprehensive set of sequence analysis programs for the VAX. Gene 12:387-395.

Fellay, R., Rahme, L. G., Mindrinos, M. N., Frederick, R. D., Pisi, A., and Panopoulos, N. J. 1991. Genes and signals controlling the *Pseudomonas syringae* pv. *phaseolicola*-plant interaction. Pages 45-52 in: Advances in Molecular Genetics of Plant-Microbe Interactions. H. Hennecke and D. P. S. Verma, ed. Kluwer Academic Publishers, Dordrecht.

Fenselau, S., Balbo, I., and Bonas, U. 1992. Determinants of pathogenicity in *Xanthomonas campestris* pv. *vesicatoria* are related to proteins involved in secretion in bacterial pathogens of animals. Mol. Plant-Microbe Interact. 5:390-396.

Galán, J. E., Ginocchio, C., and Costeas, P. 1992. Molecular and functional characterization of the *Salmonella* invasion gene *invA*: Homology of InvA to members of a new protein family. J. Bacteriol. 174:4338-4349.

Genin, S., Gough, C. L., Arlat, M., Zischek, C., Gijsegem, F. V., Bargeris, P., and Boucher, C. A. 1993. Involvement of *Pseudomonas solanacearum hrp* genes on the secretion of a bacterial compound which induces a hypersensitive-like response on tobacco. Pages 259-266 in: Advances in Molecular Genetics of Plant-Microbe Interactions, Vol. 2. E. W. Nester and D. P. S. Verma, eds. Kluwer Academic Publishers, Dordrecht.

Gough, C. L., Genin, S., Zischek, C., and Boucher, C. A. 1992. *hrp* genes of *Pseudomonas solanacearum* are homologous to pathogenicity determinants of animal pathogenic bacteria and are conserved among plant pathogenic bacteria. Mol. Plant-Microbe Interact. 5:384-389.

Huang, H.-C., He, S. Y., Bauer, D. W., and Collmer, A. 1992. The *Pseudomonas syringae* pv. *syringae* 61 *hrpH* product: An envelope protein required for elicitation of the hypersensitive response in plants. J. Bacteriol. 174:6878-6885.

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

Huang, H. C., Schuurink, R., Denny, T. P., Atkinson, M. M., Baker, C. J., Yucel, I., Hutcheson, S. W., and Collmer, A. 1988. Molecular cloning of a *Pseudomonas syringae* pv. *syringae* gene cluster that enables *Pseudomonas fluorescens* to elicit the hypersensitve response in tobacco. J. Bacteriol. 170:4748-4756.

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 on nonhost plants. J. Bacteriol. 168:512-522.

Michiels, T., Vanooteghem, J.-C., de Rouvroit, C. L., China, B., Gustin, A., Boudry, P., and Cornelis, G. R. 1991. Analysis of *virC*, an operon involved in the secretion of Yop proteins by *Yersinia enterocolitica*. J. Bacteriol. 173:4994-5009.

Pearson, W. R., and Lipman, D. J. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85:2444-2448.

Plano, G. V., Barve, S. S., and Straley, S. C. 1991. LcrD, a membrane-bound regulator of the *Yersinia pestis* low-calcium response. J. Bacteriol. 173:7293-7303.

Ramakrishnan, G., Zhao, J.-L., and Newton, A. 1991. The cell cycle-regulated flagellar gene *flbF* of *Caulobacter crescentus* is homologous to a virulence locus (*lcrD*) of *Yersinia pestis*. J. Bacteriol. 173:7283-7292.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular Cloning, A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

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

Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. 185:60-89.

Wei, Z.-M., Laby, R. J., Zumoff, C. H., Bauer, D. W., He, S. Y., Collmer, A., and Beer, S. V. 1992. Harpin, elicitor of the hypersensitive response produced by the plant pathogen *Erwinia amylovora*. Science 257:85-88.

Willis, D. K., Rich, J. J., and Hrabak, E. M. 1991. *hrp* genes of phytopathogenic bacteria. Mol. Plant-Microbe Interact. 4:132-138.

Xiao, Y., Lu, Y., Heu, S., and Hutcheson, S. W. 1992. Organization and environmental regulation of the *Pseudomonas syringae* pv. *syringae* 61 *hrp* cluster. J. Bacteriol. 174:1734-1741.