

## Relatedness of Chromosomal and Plasmid DNAs of *Erwinia pyrifoliae* and *Erwinia amylovora*

Gayle C. McGhee,<sup>1\*</sup> Elise L. Schnabel,<sup>1</sup> Kimberly Maxson-Stein,<sup>1</sup> Beatrix Jones,<sup>2</sup>  
Verlyn K. Stromberg,<sup>3</sup> George H. Lacy,<sup>3</sup> and Alan L. Jones<sup>1</sup>

Department of Plant Pathology, Michigan State University, East Lansing, Michigan 48824-1312<sup>1</sup>; Department of Statistics, Pennsylvania State University, State College, Pennsylvania 16802<sup>2</sup>; and Department of Plant Pathology, Physiology, and Weed Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0330<sup>3</sup>

Received 18 June 2002/Accepted 20 September 2002

The plant pathogen *Erwinia pyrifoliae* has been classified as a separate species from *Erwinia amylovora* based in part on differences in molecular properties. In this study, these and other molecular properties were examined for *E. pyrifoliae* and for additional strains of *E. amylovora*, including strains from brambles (*Rubus* spp.). The nucleotide composition of the internal transcribed spacer (ITS) region was determined for six of the seven 16S-23S rRNA operons detected in these species with a 16S rRNA gene probe. Each species contained four operons with a tRNA<sup>Glu</sup> gene and two with tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup> genes, and analysis of the operons from five strains of *E. amylovora* indicated a high degree of ITS variability among them. One tRNA<sup>Glu</sup>-containing operon from *E. pyrifoliae* Ep1/96 was identical to one in *E. amylovora* Ea110, but three tRNA<sup>Glu</sup> operons and two tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup> operons from *E. pyrifoliae* contained unique nucleotide changes. When *groEL* sequences were used for species-specific identification, *E. pyrifoliae* and *E. amylovora* were the closest phylogenetic relatives among a set of 12 bacterial species. The placement of *E. pyrifoliae* distinct from *E. amylovora* corroborated molecular hybridization data indicating low DNA-DNA similarity between them. Determination of the nucleotide sequence of plasmid pEP36 from *E. pyrifoliae* Ep1/96 revealed a number of presumptive genes that matched genes previously found in pEA29 from *E. amylovora* and similar organization for the genes and origins of replication. Also, pEP36 and pEA29 were incompatible with clones containing the reciprocal origin regions. Finally, the ColE1-like plasmid pEP2.6 from strain Ep1/96 contained sequences found in small plasmids in *E. amylovora* strains IL-5 and IH3-1.

There are two subgroups of the fire blight pathogen *Erwinia amylovora* based on host range: one is able to cause disease on apple, pear, and related fruits, and the other is able to cause disease on raspberry, blackberry, and other bramble (*Rubus*) species. Although these organisms exhibit host specificity, they are considered identical species because DNA-DNA homology was high between the two subgroups (10). A bacterium isolated in Korea causes a shoot blight resembling symptoms of fire blight in shoots of Asian pear (23) and recently was designated a separate species, *Erwinia pyrifoliae* (21). It was differentiated from *E. amylovora* based on low relatedness (40 to 45%) in DNA-DNA similarity experiments; sequence divergence in the 16S-23S rRNA intergenic transcribed spacer (ITS) regions (21); and the number and size of bands in restriction studies of its plasmids (35).

The taxonomic placement of *E. pyrifoliae* strain Ep1/96 within the genus *Erwinia* was based on high 16S rRNA gene sequence identity to several species in the genus, including *E. amylovora*. Strain Ep1/96 was classified as a new species, *E. pyrifoliae*, in part because of its phylogenetic distance from *E. amylovora* strain Ea1/79 when sequences for single operons of the 16S-23S rRNA ITS region from the two bacteria were compared (21). PCR-ribotyping analysis had previously revealed length variability among the spacers amplified from a

collection of *E. amylovora* strains isolated from various hosts (16, 32); strains from tree fruit hosts were categorized into three ribotype groups, and strains from *Rubus* hosts were classified into a distinct fourth group. Therefore, one goal of this study was to evaluate the basis for the size polymorphism observed among the ITS regions of *E. amylovora* and *E. pyrifoliae* and its usefulness as a criterion for separating the two species.

Sequences of 16S rRNA genes have been shown to be of limited use in comparing relatedness in the genus *Erwinia* because of the low level of variation in this gene (25). Alternatively, *groEL* gene sequences likely possess a degree of nucleotide variation large enough for discrimination between closely related species (12). In *Escherichia coli*, *groEL* encodes a heat shock protein and is essential for cellular growth (8); the gene is highly conserved and exists as a single copy in most bacteria (39). Because *E. amylovora* and *E. pyrifoliae* were indistinguishable by 16S rRNA sequence analysis (21), we undertook to explore the use of *groEL* sequence analysis to establish relatedness among a number of species of *Erwinia* and other genera.

Naturally occurring strains of *E. amylovora* have a highly conserved 29-kb plasmid known as pEA29 (7). Sequence analysis of pEA29 from *E. amylovora* Ea88, a pear pathogen, predicted 21 open reading frames (ORFs); 13 ORFs were assigned presumptive functions based on similarity to known genes, and an origin typical of iteron-regulated plasmids was assigned (29). Some variation in pEA29 was detected in other strains, particularly strains isolated from bramble plants. PCR

\* Corresponding author. Mailing address: 103 Center for Integrated Plant Systems (CIPS), Department of Plant Pathology, Michigan State University, East Lansing, MI 48824. Phone: (517) 353-3272. Fax: (517) 353-5598. E-mail: mcgheeg@msu.edu.

TABLE 1. Strains of phytopathogenic bacteria used in this study and GenBank accession numbers for their *groEL* sequences<sup>a</sup>

Species and strain(s)	Source; origin	Reference	Accession no. for <i>groEL</i>	Presence of <i>repA</i> on Southern blots
<i>Brenneria alni</i>				
PVFi20	Italian alder; Italy	ATCC 700181	AF464136	No
<i>Brenneria quercina</i>				
670	Live oak acorn; California	CDFA	AF464126	No
1710	Live oak acorn; California	CDFA	AF464129	No
1727	Live oak acorn; California	CDFA	AF464128	No
EQ101	Live oak acorn; California	CDFA	AF464127	No
<i>Brenneria rubrifaciens</i>				
1015	Persian walnut; California	CDFA	AF464135	ND
EQ103	Persian walnut; California	CDFA; 43	AF464134	No
<i>Brenneria salicis</i>				
ES4	<i>Salix alba</i> var. <i>calva</i> (cricket-bat-willow); UK	CDFA	AF464133	No
ES102	<i>Salix alba</i> var. <i>calva</i> (cricket-bat-willow); UK	CDFA	AF464132	No
<i>Erwinia amylovora</i>				
Ea110	Apple; Michigan	36	AF258533	Yes
NZR5	Tree-fruit host unknown; New Zealand	32	AF258542	
CA3R	Apple; California	32	AF258532	ND
CA263	Tree-fruit host unknown; California	32	AF258531	ND
IH3-1	Indian hawthorn; Louisiana	15	AF258539	ND
Ea88	Pear; Washington	32	AF258535	ND
OR6	Pear; Oregon	32	AF258543	ND
CA1R	Apple; California	32	AF258530	ND
IL5	<i>Rubus</i> sp.; Illinois	32	AF258540	ND
MR1	<i>Rubus</i> sp.; Michigan	32	AF258541	ND
2-95	<i>Rubus</i> sp.; Nova Scotia	P. G. Braun	AF258534	ND
<i>Erwinia persicina</i>				
CDC 9108-82	Tomato; Japan	ATCC 35998; 11	AF274344	ND
<i>Erwinia pyrifoliae</i>				
Ep1/96	Asian pear; Korea	21	AF258537	Yes
Ep16/96 <sup>T</sup>	Asian pear; Korea	21	AF258536	Yes
Ep4/97	Asian pear; Korea	21	AF258538	ND
Ep8/95	Asian pear; Korea	21	ND	ND
<i>Erwinia rhapontici</i>				
ICPB ER 102	Rhubarb; England	ATCC 29283	AF464130	No
<i>Erwinia tracheiphila</i>				
NCPPB 2452	<i>Cucumis melo</i>	ATCC 33245	AF464131	No
<i>Pectobacterium cypripedii</i>				
ICPB EC 155	<i>Cypripedium</i> sp. (lady-slipper orchid); California	ATCC 29267	AF274343	No

<sup>a</sup> CDFA, from D. Opgenorth, California Department of Agriculture, Sacramento; ND, not done; UK, United Kingdom.

assays (3, 31) based on a region of pEA29 used in the identification of *E. amylovora* did not amplify DNA when they were applied to *E. pyrifoliae* (21, 35). Additionally, endonuclease restriction patterns for a 36-kb plasmid found in *E. pyrifoliae* differed significantly from those for pEA29, suggesting that the two plasmids were probably quite different (21, 35). We wondered if sequences found in pEA29 were conserved in the *E. pyrifoliae* plasmid.

Comparative studies of these two pear-pathogenic erwinias with diverse geographical origins were undertaken to confirm and expand our understanding of their genetic relationship, including their relationship to *Rubus*-pathogenic strains of *E. amylovora*. We examined the divergence of the 16S-23S rRNA operon between *E. amylovora* and *E. pyrifoliae*, their DNA-DNA relatedness as a measure of bacterial speciation (41), the utility of *groEL* sequences for differentiating these species from other erwinias, and the organization of their plasmids, especially of a 36-kb plasmid in *E. pyrifoliae* compared with plasmid pEA29 in *E. amylovora*.

#### MATERIALS AND METHODS

**General.** Strains of *Erwinia*, *Brenneria*, and *Pectobacterium* species used in this study are indicated in Table 1. *E. coli* strains JM109, TOP10F, and DH5 $\alpha$  were

used for cloning experiments. Genomic DNA was isolated from 1.5 ml of bacterial culture as previously described (44). All PCRs were performed in MJ Research thermal controllers (PTC100 or PTC-150; MJ Research, Watertown, Mass.) and, unless otherwise noted, consisted of 20 mM Tris-HCl at pH 8.4, 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM (each) deoxynucleotide triphosphates, 0.5  $\mu$ M concentrations of each primer, 0.04 U of *Taq* DNA polymerase (Invitrogen, Carlsbad, Calif./ $\mu$ l), and 10 to 100 ng of DNA or 10<sup>5</sup> to 10<sup>7</sup> CFU of bacteria as a template. The Genomics Technology Support Facility at Michigan State University performed all sequencing reactions.

**Southern hybridization.** To locate the ITS region, DNA was treated with RNase, extracted with phenol and chloroform, precipitated with ethanol, resuspended in distilled water, and digested overnight with various restriction enzymes. Digestion products were separated on 0.8% agarose gels in 0.5 $\times$  Tris-borate-EDTA (TBE) and transferred to a nylon membrane (37). Probes for the 16S rRNA and tRNA<sup>Glu</sup> genes were prepared from *E. coli* JM109 using PCR with the primers 16SF and 16SR (5'-AACCTGGGAACATGCATCTG and 5'-TGAATCACAAGTGGTAAGC) and GluF and GluR (5'-TCTAGAGCCCCA GGACACC and 5'-CACCCGAAGATGAGTTTTGAG), respectively. Reactions were as described above with the exception that a 35  $\mu$ M concentration of dTTP was replaced by a 35  $\mu$ M concentration of digoxigenin-labeled dUTP; cycling parameters were 95°C for 3 min followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min.

To produce *repA* and small plasmid probes, plasmid DNA was separated on a 0.8% agarose gel in 0.5 $\times$  TBE and transferred to a nylon membrane (37). Primers AJ758 (5'-CGCTCTACCGCTTTTTC) and AJ759 (5'-TTTCTCCCG CATCCAGTCAAG) were used to generate the 415-bp *repA* probe from *E. amylovora* Ea110 plasmid pEA29. A 119-bp probe for the origin of replication in the small plasmid pEP2.6 was prepared from *E. pyrifoliae* Ep1/96 by using the

PCR primers AJ800 (5'-CGAAGTGAGCGAATGAACGAGGAG) and AJ801 (5'-CACGGTGTGGGGATCAGGAAAGT). Both reactions contained 10 mM Tris-HCl at pH 8.3, 1.5 mM MgCl<sub>2</sub>, 0.2 mM (each) dCTP, dATP, and dGTP, 165 mM dTTP, 35 μM digoxigenin-labeled dUTP, 0.5 μM (each) primer, 0.04 U of *Taq* DNA polymerase (Invitrogen)/μl, and 10 to 100 ng of DNA as a template. Cycling parameters were 95°C for 4 min, followed by 40 cycles of 94°C for 60 s, 55°C for 30 s (*repA*), or 52°C for 60 s (small plasmid origin), and 72°C for 45 s. Probes were gel purified using the NucleoTrap Gel Extraction Kit (Clontech Laboratories, Palo Alto, Calif.) following electrophoresis on 1% agarose gels in 0.5× TBE, staining with ethidium bromide, and visualization with a UV lamp. Hybridization and detection of all probes were performed with CDP-Star according to the manufacturer's instructions (Roche Molecular Biochemicals, Indianapolis, Ind.).

**PCR-ribotyping.** The primers SpacerF (5'-TTGTACACACCGCCCGTCA) and SpacerR (5'-GGTACTTAGATGTTTCAGTTC) were used for PCR amplification of 16S-23S rRNA spacer regions as previously described (32). Reaction products were separated on 5% acrylamide gels run in 1× TBE buffer, stained with ethidium bromide, and visualized with UV illumination.

**Sequencing of 5S rRNA flanking regions.** A set of PCR libraries was constructed by digesting 2.5-μg aliquots of *E. amylovora* Ea110 DNA with *Pvu*II, *Sca*I, and *Ssp*I in 100-μl reactions for 18 h at 37°C. Digested DNA was extracted with phenol followed by chloroform, precipitated with ethanol, and resuspended in 20 μl of distilled water. Adaptors were ligated to digested DNA in 25-μl reactions consisting of 50 mM Tris-HCl (pH 7.6), 10 mM MgCl<sub>2</sub>, 1 mM ATP, 1 mM dithiothreitol, 5% (wt/vol) polyethylene glycol-8000, 5 μM adaptor, and 10 U of T4 DNA ligase for 18 h at 16°C. Heating to 70°C for 5 min terminated the reactions, and then 180 μl of 10 mM Tris-HCl at pH 8.0 and 1 mM EDTA were added (40).

PCRs (200 μl) consisted of 1× Expand Long Template PCR buffer 3 (Roche Molecular Biochemicals), 0.2 mM deoxynucleotide triphosphates, 0.5 μM (each) primer (23Sout, 5'-CATCACTGGTGTTCGGGTTGTCATGCCAATGG; and AP1, 5'-GGATCCTAATACGACTCACTATAGGGC), 5 U of *Taq* DNA polymerase (Invitrogen), and 1 μl of PCR library. Reactions were denatured at 94°C for 2 min and subjected to 30 cycles of 94°C for 30 s and 68°C for 4 min. Reaction products were cloned using the TOPO XL PCR Cloning Kit (Invitrogen) and sequenced. In some cases, reaction products were separated on a 1% low-melting-temperature agarose gel in 1× Tris-acetate-EDTA (TAE) buffer, and individual bands were purified using the Wizard PCR Preps DNA purification system (Promega, Madison, Wis.), directly sequenced or cloned into pGEM5zf(+), and then sequenced.

**Sequencing the 16S-23S spacers.** Fragments containing individual 16S-23S rRNA spacers were amplified from various strains by PCR using primer SpacerF paired with six different primers based on the six 5S flanking region sequences of *E. amylovora* Ea110 that had been identified. The 5Sflank1 to 5Sflank6 primers were the following: 5'-ATTTGTCGTTTATATATCTCTCATCTG, 5'-CCGCTCCACGATTCTGTA, 5'-CTGAAGGAAGGCTCACTAACTGAAG, 5'-TTC TGCCCTACTTTGTCTGGATT, 5'-CAGGCTGCAGATGCTAAACTCC, and 5'-CCCGGAGATGGTGAAAGAA, respectively. Reactions (volume, 50 μl) consisted of 1× Expand Long Template PCR buffer 3 (Roche Molecular Biochemicals), 0.2 mM deoxynucleotide triphosphates, 0.5 μM (each) primer, 0.5 μl of Expand Long Template PCR enzyme, and 100 ng of genomic DNA. Reaction products were diluted 1:1,000 and subjected to PCR as described above for PCR-ribotyping using the primers SpacerF or SpacerF2 (5'-GGGAGGGCGCTTACCCTTT) and SpacerR. The resulting PCR products were either purified directly from reactions or isolated on 1% low-melting-temperature agarose gels in 1× TAE using the Wizard PCR Preps purification kit and were sequenced.

**S<sub>1</sub> nuclease DNA similarity assay.** The S<sub>1</sub> nuclease procedure for free solution reassociation for DNA similarity assays was used for measuring heteroduplex formation (reviewed in references 17, 18, and 41). DNA isolation, French pressure cell fragmentation of DNA, hybridization, and the S<sub>1</sub> nuclease assays were conducted as previously reported (18). The random primers method (RadPrime DNA Labeling System; Invitrogen) was used to label probe DNA with [ $\alpha$ -<sup>32</sup>P] dCTP (Perkin-Elmer, Life Science Inc., Boston, Mass.). Probe and target DNAs were reassociated at 62.5°C for 24 h with 22.7% formamide and then incubated with S<sub>1</sub> nuclease. Beta emissions from the remaining heteroduplex incorporating labeled DNA were estimated with a scintillation counter from washed precipitates on glass fiber filters. Percent heterologous reassociation was determined by comparing the radioactivity detected to that obtained from homologous reassociations. Data for both homologous and heterologous reassociations were corrected for nonspecific heteroduplex formation based on controls with salmon sperm single-stranded DNA. Each reaction was repeated at least once. An average number was reported as percent DNA similarity.

**PCR and analysis of partial *groEL* fragments.** Fragments of the *groEL* gene were amplified from various strains with the primers groEL-A (5'-GAAGTKG CCTCTAAAGCGAAYGA) and groEL-B (5'-GCMACRCCACCACCAGCAACC) with cycling parameters of 94°C for 2 min followed by 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 60 s. The sequences for nucleotides 698 to 1207 were aligned with sequences from related species available in GenBank. Phylogenetic analysis was performed with the PHYLIP program package (version 3.573; J. Felsenstein, University of Washington, Seattle, Wash.) using the programs SEQBOOT, NEIGHBOR, DNADIST, and CONSENSE with 1,000 bootstrap replications. The *groES* sequence of *Erwinia carotovora* (accession no. AB008152), a *groEL* protein homologue, was used as the outgroup for the tree.

**Plasmid DNA isolation and analysis.** Plasmid DNA was isolated using the NucleoBond Plasmid Purification Midi kit (Clontech Laboratories) according to the manufacturer's instructions for low-copy-number plasmids. Samples containing multiple plasmids were separated using a 0.7% agarose gel run in 0.5× TBE buffer and visualized by ethidium bromide staining. Individual plasmid bands were excised and purified using the NucleoTrap gel extraction kit (Clontech Laboratories) according to the manufacturer's instructions. The plasmid pEP36 from *E. pyrifoliae* Ep1/96 was digested with *Kpn*I, *Bam*HI, and *Cla*I, and the fragments were shotgun cloned into pGEM3zf(+) or pGEM7zf (Promega, Madison, Wisc.). Subsequent to cloning, all mixtures were added to electroporation-competent *E. coli* DH5 $\alpha$  and pulsed at 25 μF, 200 Ω, and 2.5 kV in a Gene Pulser II apparatus (Bio-Rad, Hercules, Calif.) using a 0.2-cm-gapped electroporation cuvette. After incubation for 1 h at 37°C in SOC medium (37), the transformants were plated on Luria-Bertani agar amended with 50 μg of ampicillin per ml. Plasmids from selected transformants were analyzed by digestion with the appropriate restriction enzyme. Selected clones were initially sequenced with both T7 and SP6 primers (Promega) and subsequently sequenced by primer walking. Alternatively, the previously published PCR primer pairs AJ484 and AJ485 and AJ486 and AJ487, generated from the *E. amylovora* plasmid pEA29 (29), were used to amplify genes within the clones, and the resulting PCR products were sequenced. Orientations of sequences from each fragment were verified by sequencing across *Kpn*I restriction sites. Uncloned portions of the plasmid were sequenced directly from pEP36. Sequenced fragments were analyzed and assembled using the Lasergene software of DNASTAR Inc. (Madison, Wis.). The majority of the assembled sequence was derived from single sequencing runs, the sequences from which were proofread by eye. Sequence analysis was conducted using BLAST version 2.0 at the National Center for Biotechnology Information website (<http://www.ncbi.nlm.nih.gov/BLAST>) and the Pfam database of protein families search engine (<http://pfam.wustl.edu/>) (2).

**Plasmid curing.** The plasmid pC5, containing the origin region for pEP36 (Ep1/96) cloned into pGEM3zf(+), and the plasmid pC9, containing part of the origin region for pEA29 (Ea88) (29), were electroporated into competent cells of *E. pyrifoliae* and *E. amylovora*. Ampicillin-resistant transformants were selected and maintained on Luria-Bertani agar plus ampicillin. The PCR primers AJ275 (5'-GCTACCCACCAGGCCAGAAGA) and AJ267 (5'-AGACTACTACTAC TGACATCCTG), which amplify sequences from *betT*, were used to screen individual colonies to verify plasmid loss. Amplification parameters were 94°C for 2 min followed by 30 cycles of 94°C for 30 s, 53°C for 30 s, and 72°C for 60 s. The primer pairs CPS1L with CPS2Rc and AJ245 with AJ246 were used to verify the identities of cured bacteria as *E. pyrifoliae* or *E. amylovora*, respectively, as previously described (19).

**ColE1 plasmid isolation and analysis.** Small plasmids were isolated from *E. pyrifoliae* and *E. amylovora* as described for pEP36. The primers AJ796 (5'-GTGGCGAAACCCGACAG) and AJ797 (5'-GGAGCGCAGATACCAAACCT) (380 bp) were used in PCR assays for evaluating strains of *E. amylovora* for plasmids related to pEP2.6. The reaction mix contained 10 mM Tris-HCl at pH 8.3, 1.5 mM MgCl<sub>2</sub>, 0.2 mM (each) deoxynucleotide triphosphates, 0.5 μM (each) primer, 0.04 U of *Taq* DNA polymerase (Invitrogen)/μl, and 10 to 100 ng of DNA as a template. Cycling parameters were 95°C for 4 min, followed by 40 cycles of 94°C for 60 s, 56.9°C for 30 s, and 72°C for 30 s. The plasmid pEP2.6 was cut with *Kpn*I, cloned into pGEM3Zf(+), electroporated into *E. coli* DH5 $\alpha$ , reisolated, and sequenced by primer walking. *E. amylovora* small plasmids pEA2.8 and pEA1.7 were labeled with a kanamycin resistance gene according to the manufacturer's instructions using the GPS-1 genomic priming system (New England Biolabs, Inc., Beverly, Mass.) and electroporated into *E. coli* DH5 $\alpha$ , and Kan<sup>r</sup> colonies were selected. Following plasmid reisolation, the labeled plasmids were sequenced using GPS-1 genomic priming system kit primers S and N and subsequently sequenced by primer walking. The majority of the assembled sequence for small plasmids was derived from single sequencing runs, and the sequences were proofread by eye.

**Nucleotide sequence accession numbers.** The GenBank accession numbers for *groEL* sequences are listed in Table 1. Accession numbers for the six 16S-23S

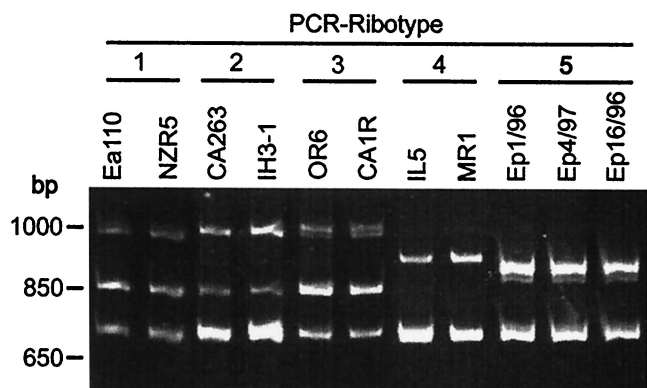


FIG. 1. PCR-ribotype fingerprints of *E. amylovora* strains isolated from tree fruits (Ea110, NZR5, CA263, OR6, and CA1R), hawthorn (IH3-1), and *Rubus* spp. (IL-5 and MR1) and three strains of *E. pyrifoliae* (Ep1/96, Ep4/97, and Ep16/96) isolated from Asian pear.

rRNA spacer regions are given parenthetically here for each of the following strains: *E. amylovora* strains Ea110 (AF449658 to AF449663), CA263 (AF449652 to AF449657), OR6 (AF449682 to AF449687), IL-5 (AF449670 to AF449675), and MR1 (AF449676 to AF449681); and *E. pyrifoliae* strain Ep1/96 (AF449664 to AF449669). Sequences of the following plasmids were entered into GenBank: pEP36 (AY123045), pEP2.6 (AY123048), pEA2.8 (AY123047), and pEA1.7 (AY123046).

**RESULTS**

**PCR ribotype.** *E. amylovora* strains Ea110, CA263, OR6, IL-5, and MR1 belong to ribotype groups 1, 2, 3, 4, and 4, respectively, as previously reported (32). A homogenous PCR pattern was observed for all *E. pyrifoliae* strains; this pattern was designated ribotype group 5 because it was distinct from those for *E. amylovora* (Fig. 1).

**Analysis of the 16S-23S rRNA intergenic spacer region of**

*E. amylovora*. Seven bands per strain were detected from the hybridization of a 16S rRNA gene probe against a Southern blot of *EcoRV*-digested chromosomal DNA of *E. amylovora* strains Ea110 and MR1 (data not shown). Hybridization of a tRNA<sup>Glu</sup> gene probe with four of the seven bands suggested that *E. amylovora* has four rRNA operons with a tRNA<sup>Glu</sup> gene in the spacer region. Primer pairs were developed for selectively amplifying six of the seven rRNA operons present in *E. amylovora* strain Ea110 and for direct sequencing of PCR fragments. Each primer pair was then used to amplify and sequence operons from other strains of *E. amylovora*.

Sequence analysis of the PCR products revealed polymorphism in the sizes of the spacers; they varied from 367 to 722 bp (Fig. 2). Alignment of the nucleotide sequences indicated that the spacers shared similar overall structure with, in order from the 16S to the 23S rRNA genes, a conserved 58- to 59-bp sequence, the tRNA gene region, a 98- to 99-bp sequence, a variable 25-bp region, an optional 139-bp sequence, and a 20-bp completely conserved sequence (Fig. 2). In each strain, four of the sequenced spacers contained a tRNA<sup>Glu</sup> gene, and two contained tandem tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup> genes. The tRNA<sup>Glu</sup> genes were of four types depending on the presence (X<sup>+</sup>) or absence (X<sup>-</sup>) of an *Xba*I restriction site and of a 139-bp region. Each tRNA<sup>Glu</sup> gene containing spacer found in strain Ea110 was different; the spacers were present in a 1:1:1:1 ratio according to type (X<sup>+</sup>, X<sup>-</sup>, X<sup>+</sup> + 139 bp, and X<sup>-</sup> + 139 bp, respectively; Fig. 2). The ratios for the corresponding ITS regions in strains CA263 and OR6 were 1:2:0:1 and 0:1:1:2, respectively (Fig. 2). In *Rubus* strains IL-5 and MR1 the ratio was 3:1:0:0. Four types of spacers containing tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup> genes were also observed (Y<sup>+</sup>, Y<sup>-</sup>, Y<sup>+</sup> + 139 bp, and Y<sup>-</sup> + 139 bp) in these strains. Y<sup>+</sup> + 139 bp- and Y<sup>-</sup> + 139 bp-type spacers were detected in tree fruit strains, Ea110,

ITS allele	ITS structure	Length (bp)	<i>Erwinia amylovora</i> strains				<i>Erwinia pyrifoliae</i>	
			Tree fruit			<i>Rubus</i> sp.		
			Ea110	CA263	OR6	IL5	MR1	Ep1/96
X <sup>+</sup>		367-372	1(2)	1(2)		3(1,2,6)	3(1,3,4)	4(1,2,4,5)
X <sup>-</sup>		376	1(1)	2(1,4)	1(1)	1(4)	1(2)	
X <sup>+</sup> 139		507-508	1(6)		1(2)			
X <sup>-</sup> 139		515	1(4)	1(6)	2(4,6)			
Y <sup>+</sup>		578				1(3)		
Y		583-584				1(5)	2(5,6)	2(3,6)
Y <sup>+</sup> 139		717	1(5)		2(3,5)			
Y 139		721-722	1(3)	2(3,5)				

FIG. 2. Organizational structure of 16S-23S rRNA ITS found in *E. amylovora* and *E. pyrifoliae*. Six rRNA operons were sequenced from each strain. The organizational structure of the spacers is indicated graphically. Regions C1, C2, and C3 are conserved regions of 58 to 60 bp, 98 to 99 bp, and 20 bp, respectively; region V is a variable 25-bp sequence; and region O is an optional 139-bp sequence. All X alleles contain a tRNA<sup>Glu</sup> gene, and all Y alleles contain tandem tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup> genes. X<sup>+</sup> indicates alleles that contain tRNA<sup>Glu</sup> genes with an *Xba*I restriction site. The number of copies of each allele is listed for each strain along with the primer (from 1 = 5SF1ank1 to 6 = 5SF1ank6) paired with the primer SpacerF, used to amplify the alleles given in parentheses.

CA263, and OR6, while Y<sup>+</sup> and Y<sup>-</sup>-type spacers were detected in the two *Rubus* strains of *E. amylovora* (Fig. 2).

**The 16S-23S spacer regions of *E. pyrifoliae*.** The six primer pairs developed for amplifying spacers from *E. amylovora* were used to amplify and sequence six ITS rRNA regions from *E. pyrifoliae* Ep1/96 (Fig. 2). Four spacers from Ep1/96 contained a tRNA<sup>Glu</sup> gene with an *Xba*I restriction site (X<sup>+</sup> alleles), and two spacers contained tandem tRNA<sup>Ile</sup> and tRNA<sup>Ala</sup> genes (Y<sup>-</sup> alleles). None of the spacers contained the 139-bp sequence found in *E. amylovora*. The nucleotide sequence for one of the X<sup>+</sup> spacers from *E. pyrifoliae* was identical to the X<sup>+</sup> spacer amplified with primer set 2 (5Sflank2 and SpacerF) from *E. amylovora* Ea110. The nucleotide sequences for the three remaining X<sup>+</sup> spacers from *E. pyrifoliae* Ep1/96 were identical, and these spacers were unique because they exhibited nucleotide changes in a region about 50 bp downstream from a putative box-A-like gene sequence (5'-TGCTCTTTAACAAT). The two Y<sup>-</sup> spacers from *E. pyrifoliae* Ep1/96 were similar in length to two spacers from *E. amylovora* MR1 and one spacer from IL-5, but they exhibited a few nucleotide changes not found in *E. amylovora*.

**DNA-DNA similarity assay.** In reciprocal pairings, *E. pyrifoliae* Ep1/96 formed a species-level DNA-DNA hybridization group ( $\geq 70\%$  similarity) (18, 42) separate from the *E. amylovora* strains Ea110, IL-5, and MR1; DNAs of the three strains of *E. amylovora* averaged 32.2% (standard deviation of  $\pm 8.8\%$ ;  $n = 23$  separate determinations), similar to the DNA of *E. pyrifoliae* Ep1/96. Among strains of *E. amylovora*, DNAs of *Rubus* strains IL-5 ( $88.0\% \pm 2.6\%$ ;  $n = 9$ ) and MR1 ( $70.6\% \pm 15.6\%$ ;  $n = 20$ ) were related at the species level to the DNA of tree fruit strain Ea110. DNAs of all erwinia strains (EP1/96, IL-5, MR1, and Ea110) were only  $2.7\% \pm 2.6\%$  ( $n = 38$ ) similar to the DNA of *Xanthomonas campestris* pv. *ionidii* strain ATCC17993, which was included as an outgroup.

**Variation in *groEL* sequences.** Sequences of a 489-bp region of the *groEL* gene were determined for PCR products amplified from several strains of *E. amylovora* and *E. pyrifoliae*, for some related bacteria in the genus, and for a few other bacteria in the family *Enterobacteriaceae* (Table 1, Fig. 3). A phylogenetic analysis using the partial *groEL* data set revealed that the strains of *E. amylovora* from *Rubus* and from tree fruit crops clustered together in all bootstrap replicates, and the three strains of *E. pyrifoliae* clustered together in all bootstrap replicates, indicating statistical support for grouping the two species in separate groups. The bootstrap replicate value for the branch dividing the clusters for *E. amylovora* and *E. pyrifoliae* from the other species in the analysis was 100, indicating that the two species were more closely related to each other than to other species in the analysis.

**Analysis of plasmid pEP36 in *E. pyrifoliae*.** DNA of *repA* from *E. amylovora* plasmid pEA29 was used as a probe in Southern hybridization experiments with plasmid DNA from a number of *Erwinia* species (Table 1). Only DNA from pEA29 and from a large plasmid in *E. pyrifoliae* strains Ep1/96 and Ep16/96<sup>T</sup>, designated as pEP36, hybridized with the *repA* probe, suggesting that the two plasmids had a similar origin of replication. Although *Bam*HI, *Kpn*I, and *Cla*I restriction digests of pEP36 were different from those reported for pEA29 (29), PCR assays with a number of individual pEA29 primer sets detected other pEA29 sequences in pEP36 DNA. Se-

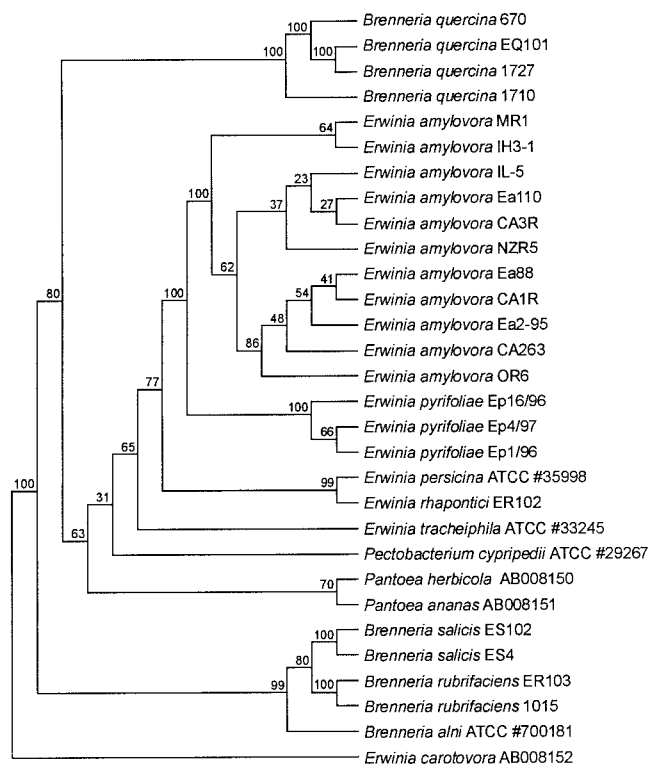


FIG. 3. Phylogenetic dendrogram based on a comparison of partial *groEL* gene sequences for *Erwinia* spp. and a set of bacteria closely related to *Erwinia*. The percentages at the nodes indicate the level of bootstrap support based on a neighbor-joining analysis of 1,000 resampled data sets. *E. carotovora* was used as the outgroup.

quence analysis was used to reveal whether pEP36 carried other pEA29 genes and how the genes were arranged relative to their order in pEA29 (29).

The nucleotide sequence of plasmid pEP36 from strain Ep1/96 contains 35,904 bp and has a G+C content of 49.6%. Like pEA29, the DNA of pEP36 has been methylated by the *dam* methylase encoded by chromosomal genes (27); 5 of the 11 *Cla*I restriction sites identified by sequencing were inactivated. One of the nine *Bam*HI restriction sites detected in pEP36 corresponded to the unique *Bam*HI site in pEA29 (Fig. 4); the pEP36 sequence was numbered from the first nucleotide in this restriction site. Analysis of the pEP36 nucleotide sequence by BLAST N algorithm (1) revealed the highest similarity to plasmid pEA29 from *E. amylovora* Ea88 (accession no. AF264948) (Fig. 4). Other significant nucleotide similarities were found to the *Yersinia pestis* CO92 genome (accession no. AJ414149) and a *Shigella flexneri* SHI-2 pathogenicity island (accession no. AF141323). Additional similarity was found to pEA29 from *E. amylovora* MR1 isolated from *Rubus* sp. (accession no. AF264949) that was distinct from pEA29 isolated from tree fruit and ornamental hosts (29) (Fig. 4). This region occurred at the same location in pEP36 as in pEA29 from *Rubus* strains of *E. amylovora*.

Fourteen of the thirty-two open reading frames (ORFs) found in pEP36 were detected previously in pEA29 from *E. amylovora* Ea88 (Table 2). Many of the genes common to both plasmids were present in the same order and orientation as in

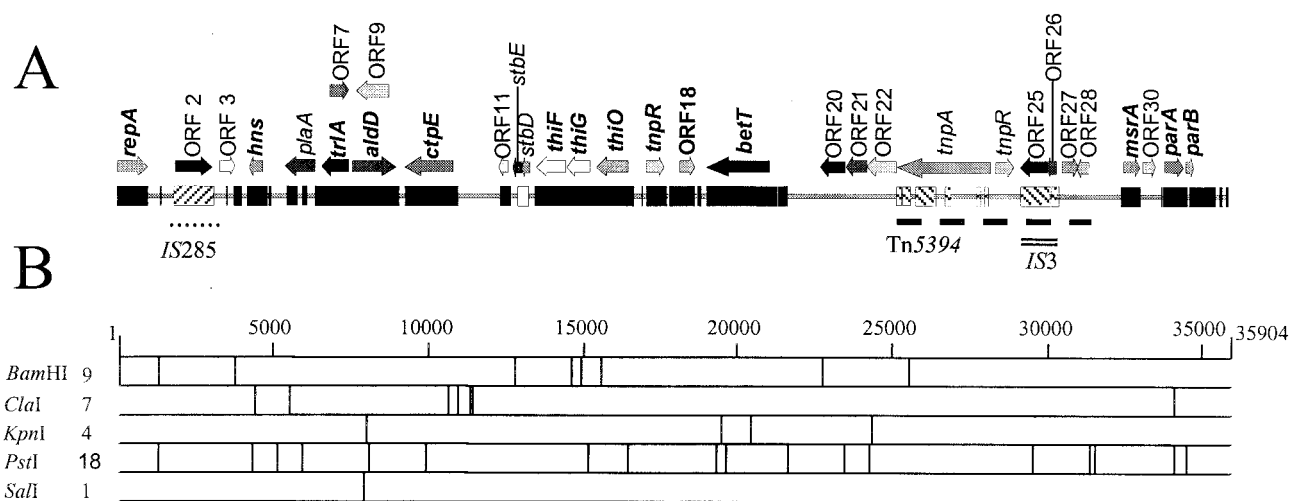


FIG. 4. Physical and restriction map of 35,904-bp plasmid pEP36 isolated from *E. pyrifoliae* strain Ep1/96. (A) Genetic map of pEP36. Solid black bar (■) indicates nucleotide homology to plasmid pEA29 isolated from *E. amylovora* strain Ea88. The forward-hatched box (▨), open box (□), and reverse-hatched box (▩) indicated regions of pEP36 with nucleotide similarities to the *Y. pestis* strain CO92 genome, *E. amylovora* plasmid pEA29 isolated from *Rubus* species, and an *S. flexneri* SHI-2 pathogenicity island, respectively; all nucleotide similarities reported are greater than 80%. The location of insertion elements IS285 and IS3 are shown by a dotted line (••••) and double line (==), respectively. The position of the transposon Tn5394 is designated by a dashed line (---). Potential ORFs greater than 375 nt and putative genes inferred from sequence data or with matches in GenBank are indicated by directional arrows. ORFs conserved between pEP36 and *E. amylovora* plasmid pEA29 are indicated in bold lettering. A key to the ORFs is found in Table 2. (B) Restriction map of pEP36. Numbers to the right of the restriction enzyme are indicative of the number of restriction sites found for each enzyme on the plasmid. Only nonmethylated *Cla*I sites are reported.

pEA29 (Fig. 4). Although the sequence between genes common to both plasmids had less than 80% nucleotide similarity, these intergenic sequences were still related to each other. Four ORFs with homology to transposases were found in pEP36. ORF 2 encoded a putative transposase previously identified in *Y. pestis* (34) and was located in a 1,317-bp insertion sequence that was 84% identical to IS285 previously described from *Y. pestis* (9). Interestingly, the right inverted repeat (nucleotides [nt] 3065 to 3114) had homology to pEA29 at nt 14764 to 14715 and 26885 to 26931. Three transposases were in the sequence extending from *betT* to *msrA*. This region contained a putative 6,339-bp transposon, designated Tn5394, with two 37-bp inverted repeats, one at each end, divergently transcribed transposase (ORF 23) and resolvase (ORF 24) genes, and a *res* site exhibiting dyad symmetry. Two other transposases, predicted products of ORF 25 and ORF 26, were located within an insertion sequence with similarity to IS3 from *E. coli* (5). This insertion element was complete with two 50-bp inverted repeats and was itself inserted into Tn5394. ORF 27 sits external to IS3 on Tn5394, and the amino acid sequence was 37% similar to a conserved putative membrane protein found in *Y. pestis* (34). ORF 28 is also located on Tn5394 and has no similarity to existing GenBank entries. In addition, the sequence extending from *betT* to *msrA* contained ORFs 20, 21, and 22; the three ORFs coded for putative ABC transporter proteins Yusa, Yusb, and Yusc identified previously for *Bacillus subtilis* (Table 2).

ORF 12 and ORF 13 encoded putative plasmid stability proteins StbE and StbD previously identified in *Morganella morganii* (14) and *Salmonella enterica* (28). Both sequences were absent in pEA29 from strain Ea88, but partial *stbE* and *stbD* sequences were found previously in pEA29 from *E. amylovora* strains MR1 and 2-95 (29). Pfam analysis predicted that

the putative product of pEP36 ORF 5 was a member of the omptin family of temperature-regulated outer membrane-associated serine proteases. Remnants of ORF 5 were also found on pEA29 in *E. amylovora* Ea88. Like pEP36, the pEA29 sequence similar to ORF 5 was located between the *hns* and *trlA* genes, but the sequence lacked a start codon and contained regions of sequence degradation that cause significant frame shifts in the gene. Finally, none of the sequences for ORF 3 and ORF 30 were related to pEA29 DNA, although ORF 30 had 87% similarity to a hypothetical protein found in *E. coli* (13).

**Origin of replication and repeat region.** The origin of replication in pEP36 had an organization similar to that of the origin found in pEA29 (Fig. 5). The *Bam*HI site at nt 1 was located 13 nt upstream from a replication initiation protein encoded by *repA* in both plasmids (93% identity). Putative partitioning proteins encoded by *parA-parB* and a series of 10 direct 8-bp repeats (5'-ATTCTGGG) at nt 33411 to 33490 were located immediately upstream from the iteron region in pEP36. The location of *parA* in pEP36 coincided with the position of a *parA*-related region in pEA29 (20, 29), and two of the 8-bp repeats in this region were detected in pEA29. Plasmid pEP36 was examined for a second group of 3 to 14 8-bp repeat sequences (5'-AATTACAG) found previously in pEA29 upstream from the *parA*-related region (20, 30); two of these repeats were found in this region of pEP36 although they were modified (5'-GATTACAG). Iterons found in pEP36 were homologous to iterons in pEA29, and they were located in the same relative position in the plasmids (Fig. 5).

The close similarity of the two origins of replication suggested that it should be possible to cure *E. pyrifoliae* of pEP36 by eviction as reported for *E. amylovora* and pEA29 (7, 26, 29). Using clone pC9 containing pEA29 DNA, pEP36 was evicted

TABLE 2. Predicted ORFs found in plasmid pEP36 from *E. pyrifoliae* strain Ep1/96

pEp36 ORF <sup>a</sup>	Strand	Predicted protein size (kDa)	Size in amino acids (putative gene function)	Putative protein/organism/ GenBank accession no.	Alignment region	Homologue alignment	% Identity
<b>Similar ORFs</b>							
<i>repA</i> /ORF 1	+	37.6	325 (replication protein)	RepA/ <i>Erwinia amylovora</i> /AAG31039.1	1–325	1–325 (325) <sup>b</sup>	93
<i>hns</i> /ORF 4	–	14.4	128 (histone-like nucleotid structuring protein, DNA binding)	H-NS/ <i>Erwinia amylovora</i> /AAG31058.1	1–128	1–128 (134)	63
<i>trlA</i> /ORF 6	–	31.1	285 (LysR transcriptional regulator)	MarR/ <i>Erwinia amylovora</i> /AAG31051.1	1–285	1–285 (285)	87
<i>aldD</i> /ORF 8	+	49.3	461 (aldehyde dehydrogenase)	AldD/ <i>Erwinia amylovora</i> /AAG31043.1	1–460	1–460 (461)	85
<i>ctpE</i> /ORF 10	–	54.3	507 (methyl-accepting chemotaxis protein)	CtpE/ <i>Erwinia amylovora</i> /AAG31044.1	1–507	1–507 (507)	84
<i>thiF</i> /ORF 14	–	34.3	321 (thiamine biosynthesis)	ThiF/ <i>Erwinia amylovora</i> /AAG31053.1	1–321	1–321 (321)	80
<i>thiG</i> /ORF 15	–	27.5	253 (thiamine biosynthesis)	ThiG/ <i>Erwinia amylovora</i> /AAG31045.1	2–253	1–252 (252)	90
<i>thiO</i> /ORF 16	–	40.1	359 (amino acid oxidase flavoprotein)	ThiO/ <i>Erwinia amylovora</i> /AAG31046.1	10–349	10–349 (349)	87
<i>tnpR</i> /ORF 17	+	21.3	194 (resolvase)	TnpR/ <i>Erwinia amylovora</i> /AAG31042.1	1–193	1–193 (194)	95
ORF 18	+	18.9	171 (conserved hypothetical protein)	ORF 15/ <i>Erwinia amylovora</i> /AAG31047.1	1–171	1–171 (171)	85
<i>betT</i> /ORF 19	–	76.2	676 (high-affinity choline transport gene)	BetT/ <i>Erwinia amylovora</i> /AAG31040.1	1–676	1–676 (676)	94
<i>msrA</i> /ORF 29	+	19.7	178 (peptide methionine sulfoxide reductase)	MsrA/ <i>Erwinia amylovora</i> /AAG31048.1	1–178	1–178 (178)	92
<i>parA</i> /ORF 31	+	22.2	208 (partitioning protein)	ParA/ <i>Erwinia amylovora</i> /AAG31050.1	4–208	2–206 (206)	50
<i>parB</i> /ORF 32	+	7.7	68 (partitioning protein)	ParB/ <i>Erwinia amylovora</i> /AAG31059.1	1–66	13–76 (76)	38
<b>ORFs not found</b>							
ORF 2	+	45.3	403 (transposase A)	Transposase mutator/ <i>Yersinia pestis</i> /NP_403677.1	1–403	1–402 (402)	90
ORF 3	+	17.9	159 (unknown)	Hypothetical protein			
ORF 5	–	34.7	312 (plasminogen activator)	Plasminogen activator precursor/ <i>Yersinia pestis</i> /A42928	1–312	1–312 (312)	71
ORF 7	+	24.0	213 (unknown)	Hypothetical protein			
ORF 9	–	39.4	344 (unknown)	Hypothetical protein			
ORF 11	–	10.9	94 (unknown)	Hypothetical protein/ <i>Yersinia pestis</i> /NP_404766.1	1–68	1–68 (81)	73
<i>stbE</i> /ORF 12	–	11.1	95 (plasmid stability)	Putative cytoplasmic protein/ <i>Salmonella enterica</i> serovar Typhimurium/NP_460510.1 or <i>stbE</i> / <i>Morganella morganii</i> /AAC84004.1	1–93	1–93 (94)	67
<i>stbD</i> /ORF 13	–	9.2	84 (plasmid stability)	Putative cytoplasmic protein/ <i>Salmonella enterica</i> serovar Typhimurium/NP_460511.1 or <i>StbD</i> / <i>Morganella morganii</i> /AAC84003.1	1–84	1–82 (82)	65
ORF 20	–	29.5	270 (ABC transporter)	Hypothetical protein/ <i>Staphylococcus aureus</i> /NP_371363.1 or <i>YusA</i> / <i>Bacillus subtilis</i> /B70020	25–267 25–269	33–271 (273) 33–273 (274)	51 48
ORF 21	–	23.7	222 (ABC transporter-binding protein-dependent transport system)	<i>YusB</i> / <i>Bacillus subtilis</i> /C70020	1–222	1–222 (222)	56
ORF 22	–	35.7	326 (ABC transporter-ATP binding protein)	<i>YusC</i> / <i>Bacillus subtilis</i> /D70020	13–326	28–341 (341)	57
ORF 23	–	113.5	998 (transposase)	Tnp/plasmid F/NP_061389.1	1–998	1–998 (1002)	76
ORF 24	+	21.7	193 (resolvase)	TnpR/ <i>Pseudomonas putida</i> /P30739	1–178	1–178 (181)	51
ORF 25	–	32.9	284 (transposase)	Integrase/ <i>Escherichia coli</i> /CAC43414.1	1–284	1–284 (284)	89
ORF 26	–	11.4	99 (transposase)	Tnp/ <i>Salmonella enterica</i> serovar Typhimurium LT2/NP_461695	1–99	1–99 (99)	83
ORF 27	+	19.3	178 (conserved hypothetical protein)	Putative membrane protein/ <i>Yersinia pestis</i> /NP_407480.1	1–166	12–179 (193)	37
ORF 28	–	14.0	126 (unknown)	Hypothetical protein			
ORF 30	+	15.3	138 (conserved hypothetical protein)	Hypothetical protein/ <i>Escherichia coli</i> /BAB33986.1	1–93	1–93 (135)	87

<sup>a</sup> Similar ORFs, ORFs similar to those found in *E. amylovora* plasmid pEA29; ORFs not found, ORFs not found in *E. amylovora* plasmid pEA29.

<sup>b</sup> Total length of ORF in GenBank.

from *E. pyrifoliae* Ep1/96 at low efficiencies. Conversely, clone pC5 containing the pEP36 origin readily evicted pEP36 from *E. pyrifoliae* Ep1/96 and pEA29 from *E. amylovora* Ea88 (data not shown). Eviction of both plasmids was confirmed by PCR using *betT* sequence from pEA29 as the target.

**Genetic organization of *E. pyrifoliae* plasmid pEP2.6.** In addition to pEP36, *E. pyrifoliae* Ep1/96 was found to contain three small plasmids as previously reported (35). Analysis of the sequence for one of these plasmids, designated pEP2.6, indicated that it was a 2,590-bp ColE1-like plasmid (Fig. 6).

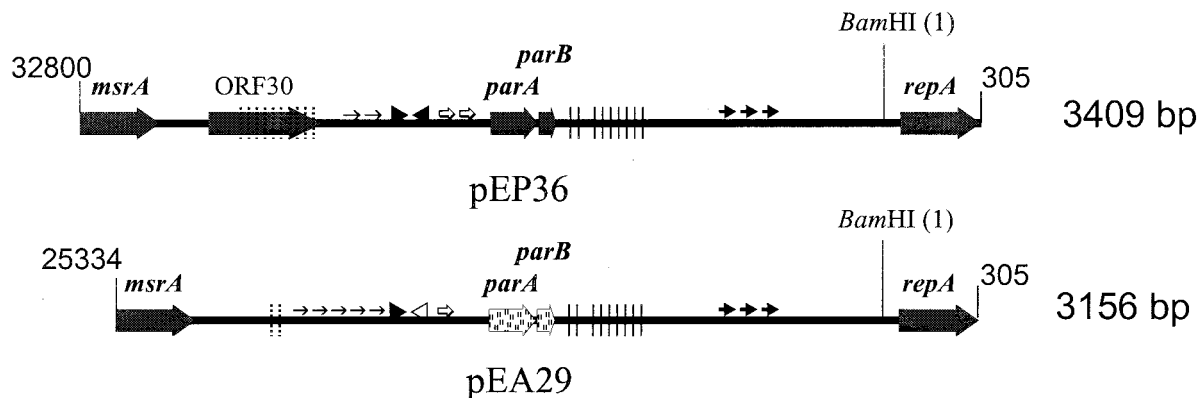


FIG. 5. Physical map comparing the organization of the origin of replication for plasmids pEP36 and pEA29 isolated from *E. pyrifoliae* strain Ep1/96 and *E. amylovora* strain Ea88, respectively. ORFs and their orientation are indicated by large directional arrows (►). Incomplete ORFs are indicated by flecked arrows (▨). The position and organization of discrete families of repeats are indicated by various symbols: dotted vertical lines (⋮) indicate an 8-bp repeat (ATTCTGGG) originally observed in pEP36, line arrows (→) designate a series of 3 to 14 8-bp repeats previously described in pEA29 (38). Inverted triangles (◄) show the position of an inverted repeat; the right repeat (◄) is incomplete in pEA29. Open arrows (⇨) indicate the position of two 23-bp repeats found on pEP36 and of a single copy found on pEA29. The nine iterons are designated by solid vertical lines (|). Solid arrows (➔) indicate three 20-bp repeats found on pEP36 and on pEA29. Numbers found at either end of the drawing are the position in nucleotides of the segment depicted. The length of each segment is indicated in base pairs to the right of the drawing. An identical *Bam*HI site is used to orient the plasmids. Maps are not to scale.

The origin of replication (*oriV*) was located at nt 1536 to 1540. The preprimer RNA II transcription start was located at nt 1017, and the antisense RNA I molecule regulating plasmid copy number was located on the complementary strand at nt 1123 to 1021. The positions of *oriV*, RNA I, and RNA II were located by analogy to ColE1-like replicons (accession no. AF014880 and NC\_001371). The plasmid also contained a 67-amino acid (aa) ORF (ORF 4) with 45% similarity to a RNA I inhibition modulator-like protein (*rom*) gene found in *E. coli* plasmid pEC157 (accession no. AF432497) and the *rom* product grouped with the Rop (repressor of primer) family of proteins by Pfam analysis. Three additional hypothetical ORFs were located on pEP2.6, but none had similarity to existing GenBank entries. A sequence with >73% amino acid similarity to a conserved hypothetical 13.8-kDa protein found in *E. coli* plasmid ColE1 (accession no. NP\_040360) was detected on pEP2.6 at nt 1560 to 1233; however, the gene was corrupted by an insertion sequence.

**Small plasmids in *E. amylovora*.** Southern blot analysis revealed three small plasmids in strain IL-5 with DNA that hybridized with the origin of replication from pEP2.6 used as a probe, and a PCR product was obtained from IL-5 using the primers AJ796 and AJ797; these results suggested that IL-5 contained ColE1-like plasmids. A 2,825-bp plasmid from IL-5 was sequenced. This plasmid, designated pEA2.8, contained RNA I-RNA II-*oriV* sequences with 84% nucleotide similarity to the origin of replication found in pEP2.6 (Fig. 6). ORF 1 in plasmid pEA2.8 had 95% amino acid similarity to the 13.8-kDa conserved hypothetical protein found in *E. coli* plasmid ColE1 (accession no. NP\_040360); residues of this protein were also found on pEP2.6 (Fig. 6). Plasmid pEA2.8 also contained a 281-aa ORF coding for the ampicillin resistance protein beta lactamase. ORFs 2 and 3 did not exhibit sequence homology to existing entries in GenBank.

Two small plasmids observed in agarose gels of plasmid DNA purified from Indian hawthorn strain IH3-1 did not hy-

bridize in Southern blots to the RNA I-RNA II-*oriV* probe generated from pEP2.6. However, sequence analysis of one of the two plasmids, a 1,711-bp plasmid designated pEA1.7, revealed an ORF with 74% nucleotide similarity to *rom* in pEP2.6 (Fig. 6). Curiously, no RNA I-RNA II origin of replication was detected in pEA1.7, and ORFs 1, 2, and 3 did not exhibit homology to GenBank entries by BLAST X algorithm.

## DISCUSSION

Methods for differentiating *E. pyrifoliae* and *E. amylovora* independent of host range are particularly important for proper identification and for understanding their biology. Our phylogenetic analysis of these and related species based on *groEL* sequence data confirmed a previous report that *groEL* analysis has sufficient resolving power for discriminating between related species within the same genus (12). The phylogenetic tree clearly showed that *E. pyrifoliae* was more closely related to *E. amylovora* than to the other species tested and that the two species were distinct. This conclusion was supported by DNA similarity data (21; also this study) and suggests that *groEL* sequence analysis may be an alternative to DNA reassociation assays when trying to discriminate between highly related species in the genus *Erwinia*.

Sequence analysis of 16S-23S rRNA operons was used to evaluate the utility of using PCR-ribotyping and ITS sequence data for grouping strains of *E. amylovora* and *E. pyrifoliae* and for differentiating the species. Many members of the family *Enterobacteriaceae*, including *E. coli* and *Salmonella* spp., have been shown to have seven rRNA operons, and seven copies were detected in *E. amylovora* (45; also this study). We were able to sequence six of the seven operons and show the presence of sequence and length variability in the operons, both within and between five strains of *E. amylovora*. This variation implies that recombination between copies of the spacers has occurred in the chromosomes of *E. amylovora*. This rearrange-



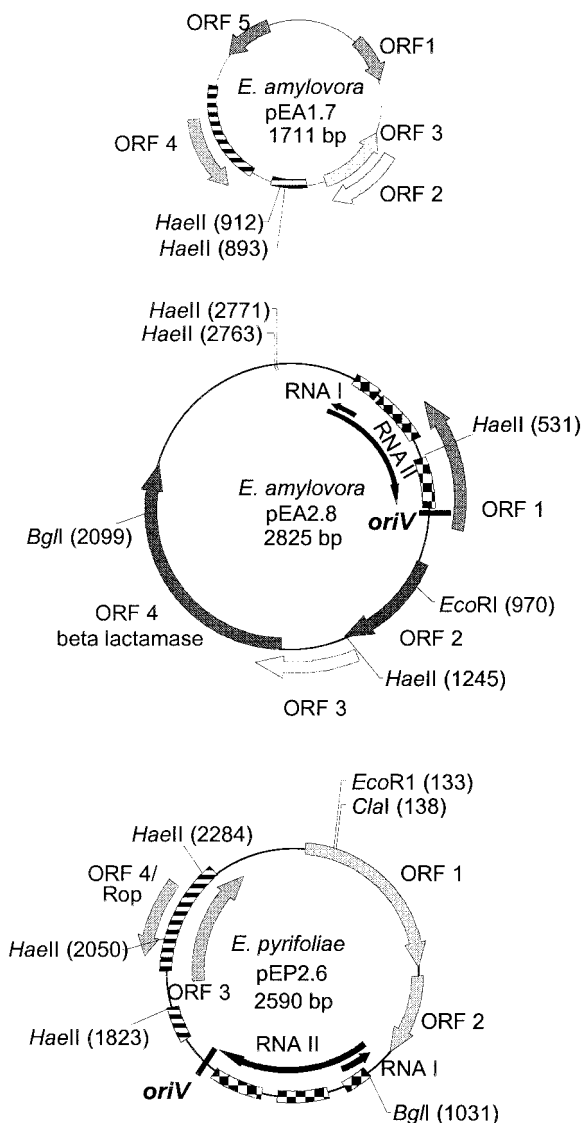


FIG. 6. Genetic maps of *E. pyrifoliae* plasmid pEP2.6 from strain Ep1/96 and of *E. amylovora* plasmid pEA2.8 from strain IL-5 and plasmid pEA1.7 from strain IH3-1. Regions of nucleotide similarity between pEP2.6 and pEA1.7 (78%) and between pEP2.6 and pEA2.8 (84 to 89%) are indicated by a striped bar (▨) and a checked bar (▩), respectively. The ORF coding for the RNA1 modulator protein was found on both pEP2.6 and pEA1.7 (ORF 4 on both plasmids). ORF 2 from plasmid pEA2.8 was similar to a 13.8-kDa hypothetical protein encoded on plasmid ColE1 in *E. coli* (accession number NP\_040360). ORF 4 from pEA2.8 had 98% similarity to a beta-lactamase protein that provides resistance to ampicillin. No other ORFs had identity to any existing GenBank entries.

ment may account for some of the genomic differences seen following pulsed-field gel electrophoresis analysis of genomic DNA after *Xba*I digests (45).

It was previously suggested, based on a phylogenetic analysis of 16S-23S ITS sequence data, that *E. pyrifoliae* was distantly related to *E. amylovora* (21). The use of ITS data for phylogenetic studies in the case of *E. amylovora* and *E. pyrifoliae* is questionable because of the high level of variability within the 16S-23S spacers. Some spacers contain optional blocks of nu-

cleotides rather than nucleotide differences resulting from single mutational events. Consequently, applying phylogenetic methods designed for nucleotide substitutions may produce misleading results when applied to differences due to recombination events (30). The 16S-23S ITS region of only a few strains of *E. pyrifoliae* has been examined; therefore, the level of heterogeneity between strains of *E. pyrifoliae* may be greater than currently recognized.

It was previously assumed that the 36-kb plasmid found in *E. pyrifoliae* and pEA29, a nontransferable highly conserved plasmid found in all natural strains of *E. amylovora* (7, 26), were not similar plasmids because DNA was not amplified using PCR-based assays designed to detect pEA29 and the plasmids exhibited different restriction digestion patterns (21, 35). But we found similarities between the two plasmids; pEP36 showed strong hybridization to a *repA* probe from pEA29, incompatibility with a pEA29 replicon, and high-level conservation of several pEA29 genes and gene order. Although pEP36 and pEA29 differ from one another in some ways, the two are unmistakably closely related and probably have similar functions in their respective species. Plasmid pEA29 is associated with increased virulence in *E. amylovora* (7, 26, 29), but the possibility that pEP36 contributes to the virulence of *E. pyrifoliae* remains to be determined.

A striking discovery from plasmid sequencing is the apparent importance of horizontal DNA exchange in their evolution. The presence of *ISEam1* in some derivatives of pEA29 (29) and of insertion sequences and the transposon Tn5394 in pEP36 suggests that horizontal transfer has played a role in the evolution of both plasmids. Genetic elements contribute to the larger size of pEP36 relative to pEA29 and account for some of the variation in their restriction patterns. The examination of other pEA29-like plasmids in strains of *E. pyrifoliae* and in other *Erwinia* species, such as those recently described from Japan (22), may provide further insight into the diversity and evolution of these plasmids and their hosts.

The detection of a plasmid in *E. pyrifoliae* that is closely related to the nontransmissible plasmid pEA29 raises the following question: how did pEP36 and pEA29 get established in their respective hosts? Although there is evidence of horizontal transfer of plasmids with genes for streptomycin resistance in *E. amylovora* (4), horizontal transfer of pEA29 has not been proven. Since no genes involved in conjugative transfer were found on either plasmid, and since putative functions could be assigned for a number of the genes detected on these plasmids, the presence of a transfer system unrelated to known systems is unlikely. A helper plasmid may have resulted in a rare transfer of one or both plasmids, but until natural horizontal transfer of pEP36 and pEA29 can be demonstrated, the original occurrence of pEP36 and pEA29 in their hosts by horizontal transfer is speculation.

The small plasmids in *E. pyrifoliae* and *E. amylovora* IL-5 were classified as ColE1-type plasmids based on the presence of an RNA I-RNA II-*oriV*-type origin of replication. None of these plasmids contained sequences homologous to sequences in a small plasmid (pEa8.7; accession no. AF017389) previously characterized from *E. amylovora* (33). ColE1 family plasmids have a narrow host range, a medium copy number, and lack a replication initiation protein. Proteins encoded by the host chromosome control ColE1 plasmid replication (6), and

antisense RNAs located on the plasmid control copy number (42). The Rop protein, encoded by *rom*, enhances the RNA I-RNA II duplex binding by increasing stability of the hybrid (24). The conservation of ColE1-type plasmids among some strains of *E. amylovora* and *E. pyrifoliae* indicates that small plasmid content is less useful for separating the species than previously thought (35).

Interestingly, the molecular characteristics of *E. pyrifoliae* resembled those of *Rubus* strains of *E. amylovora* more than tree fruit strains of *E. amylovora*. The length and types of 16S-23S rRNA spacers for *E. pyrifoliae* Ep1/16 and for *E. amylovora* IL-5 and MR1 (*Rubus* strains) were similar; and two rather than three bands were amplified from these strains by PCR-ribotyping. Previously derivatives of pEA29 from *Rubus* strains of *E. amylovora* were found to exhibit significant sequence variation in some regions when compared to the sequence of pEA29 from Ea88 (29). For example, partial sequences of the stability genes *stbE* and *stbD* found in pEP36 were previously detected on pEA29 found in two strains from *Rubus* but not in Ea88. Also, the divergent regions found in *E. amylovora* MR1 starts and stops in the same place in pEP36, but the region was shorter in pEP36, and the first 33 nt on MR1 were inverted compared to pEP36 (nt 3644 to 3676). Finally, ColE1-type plasmids are more commonly observed in *Rubus* strains of *E. amylovora* than in strains from apple and pear.

#### ACKNOWLEDGMENTS

This research was supported in part by the Michigan Apple Research Committee, the Michigan Agricultural Experiment Station, and by grants from USDA/CSREES.

We appreciate the assistance of Kathryn Bembas in sequencing *groEL*, Joseph M'Mwirichia in sequencing *E. amylovora* small plasmids, G. C. Foster for help in editing the manuscript, and Marlene Cameron for assistance with the graphics. We thank K. Geider for providing cultures of *E. pyrifoliae* and D. C. Opgenorth for contributing isolates of other necrogenic species in the amylovora group.

#### REFERENCES

- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**:3389–3402.
- Bateman, A., E. Birney, R. Durbin, S. R. Eddy, K. L. Howe, and E. L. L. Sonnhammer. 2000. The Pfam contribution to the annual NAR database issue. *Nucleic Acids Res.* **28**:263–266.
- Bereswill, S., A. Pahl, P. Bellemann, W. Zeller, and K. Geider. 1992. Sensitive and species-specific detection of *Erwinia amylovora* by polymerase chain reaction analysis. *Appl. Environ. Microbiol.* **58**:3522–3526.
- Chiou, C.-S., and A. L. Jones. 1991. The analysis of plasmid-mediated streptomycin resistance in *Erwinia amylovora*. *Phytopathology* **81**:710–714.
- Deonier, R. C., R. G. Hadley, and M. Hu. 1979. Enumeration and identification of IS3 elements in *Escherichia coli* strains. *J. Bacteriol.* **137**:1421–1424.
- Donoghue, D. J., and P. A. Sharp. 1978. Replication of colicin E1 plasmid DNA in vivo requires no plasmid-encoded proteins. *J. Bacteriol.* **133**:1287–1294.
- Falkenstein, H., W. Zeller, and K. Geider. 1989. The 29 kb plasmid common in strains of *Erwinia amylovora*, modulates development of fire blight symptoms. *J. Gen. Microbiol.* **135**:2643–2650.
- Fayet, O., T. Ziegelhoffer, and C. Georgopoulos. 1989. The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *J. Bacteriol.* **171**:1379–1385.
- Filippov, A. A., P. V. Oleinikov, V. L. Motin, O. A. Protsenko, and G. B. Smirnov. 1995. Sequencing of two *Yersinia pestis* IS elements, IS285 and IS100. *Contrib. Microbiol. Immunol.* **13**:306–309.
- Gardner, J. M., and C. I. Kado. 1972. Comparative base sequence homologies of the deoxyribonucleic acids of *Erwinia* species and other *Enterobacteriaceae*. *Int. J. Syst. Bacteriol.* **22**:201–209.
- Hao, M. V., D. J. Brenner, A. G. Steigerwalt, Y. Kosako, and K. Komagata. 1990. *Erwinia persicinus*, a new species isolated from plants. *Int. J. Syst. Bacteriol.* **40**:379–383.
- Harada, H., and H. Ishikawa. 1997. Phylogenetical relationship based on *groE* genes among phenotypically related *Enterobacter*, *Pantoea*, *Klebsiella*, *Serratia*, and *Erwinia* species. *J. Gen. Appl. Microbiol.* **43**:355–361.
- Hayashi, T., K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C.-G. Han, E. Ohtsubo, K. Nakayama, T. Murata, M. Tanaka, T. Tobe, T. Iida, H. Takami, T. Honda, C. Sasakawa, N. Ogasawara, T. Yasunaga, S. Kuhara, T. Shiba, M. Hattori, and H. Shinagawa. 2001. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* **8**:11–22.
- Hayes, F. 1998. A family of stability determinants in pathogenic bacteria. *J. Bacteriol.* **180**:6415–6418.
- Holcomb, G. E. 1998. First report of fire blight on Indian Hawthorn cultivar Olivia in Louisiana. *Plant Dis.* **82**:1402.
- Jeng, R. S., A. M. Svircev, A. L. Myers, L. Beliaeva, D. M. Hunter, and M. Hubbes. 2001. The use of 16S and 16S-23S rRNA to easily detect and differentiate common Gram-negative orchard epiphytes. *J. Microbiol. Med.* **44**:69–77.
- Johnson, J. L. 1984. Nucleic acids in bacterial classification, p. 8–11. *In* N. Krieg and J. G. Holt (ed.), *Bergey's manual of systemic bacteriology*, vol. 1. Williams & Wilkins Co., Baltimore, Md.
- Johnson, J. L. 1994. Similarity analysis of DNAs, p. 655–682. *In* P. Gerhardt, R. G. E. Murray, W. A. Woods, and N. R. Krieg (ed.), *Methods for general and molecular bacteriology*. ASM Press, Washington, D.C.
- Jones, A. L., and K. Geider. 2001. Gram-negative bacteria, *Erwinia amylovora* group, p. 40–45. *In* N. W. Schaad, J. B. Jones, and W. Chun (ed.), *Laboratory guide for identification of plant pathogenic bacteria*, 3rd ed. American Phytopathological Society, St. Paul, Minn.
- Kim, W.-S., and K. Geider. 1999. Analysis of variable short-sequence DNA repeats on the 29 kb plasmid of *Erwinia amylovora* strains. *Eur. J. Plant Pathol.* **105**:703–713.
- Kim, W.-S., L. Gardan, S.-L. Rhim, and K. Geider. 1999. *Erwinia pyrifoliae* sp. nov., a novel pathogen that affects Asian pear trees (*Pyrus pyrifolia* Nakai). *Int. J. Syst. Bacteriol.* **49**:899–906.
- Kim, W.-S., M. Hildebrand, S. Jock, and K. Geider. 2001. Molecular comparison of pathogenic bacteria from pear trees in Japan and the fire blight pathogen *Erwinia amylovora*. *Microbiology* **147**:2951–2959.
- Kim, W.-S., S. Jock, J.-P. Paulin, S.-L. Rhim, and K. Geider. 2001. Molecular detection and differentiation of *Erwinia pyrifoliae* and host range analysis of the Asian pear pathogen. *Plant Dis.* **85**:1183–1188.
- Kües, U., and U. Stahl. 1989. Replication of plasmids in gram-negative bacteria. *Microbiol. Rev.* **53**:491–516.
- Kwon, S.-W., S.-J. Go, H.-W. Kang, J.-C. Ryu, and J.-K. Jo. 1997. Phylogenetic analysis of *Erwinia* species based on 16S rRNA gene sequences. *Int. J. Syst. Bacteriol.* **47**:1061–1067.
- Laurent, J., M. A. Barny, A. Kotoujansky, P. Dufriche, and J. L. Vanneste. 1989. Characterization of an ubiquitous plasmid in *Erwinia amylovora*. *Mol. Plant Microbe Interact.* **2**:160–164.
- Marinus, M. 1987. DNA methylation in *Escherichia coli*. *Annu. Rev. Genet.* **21**:113–132.
- McClelland, M., K. E. Sanderson, J. Spieth, S. W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston, and R. K. Wilson. 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* **413**:852–856.
- McGhee, G. C., and A. L. Jones. 2000. Complete nucleotide sequence of ubiquitous plasmid pEA29 from *Erwinia amylovora* strain Ea88: gene organization and intraspecific variation. *Appl. Environ. Microbiol.* **66**:4897–4907.
- McGuire, G., F. Wright, and M. J. Prentice. 1997. A graphical method for detecting recombination in phylogenetic data sets. *Mol. Biol. Evol.* **14**:1125–1131.
- McManus, P. S., and A. L. Jones. 1995. Detection of *Erwinia amylovora* by nested PCR and PCR-dot-blot and reverse-blot hybridizations. *Phytopathology* **85**:618–623.
- McManus, P. S., and A. L. Jones. 1995. Genetic fingerprinting of *Erwinia amylovora* strains isolated from tree-fruit crops and *Rubus* spp. *Phytopathology* **85**:1547–1553.
- Palmer, E. L., B. L. Teviotdale, and A. L. Jones. 1997. A relative of the broad-host-range plasmid RSF1010 detected in *Erwinia amylovora*. *Appl. Environ. Microbiol.* **63**:4604–4607.
- Parkhill, J., B. W. Wren, N. R. Thomson, R. W. Titball, M. T. G. Holden, M. B. Prentice, M. Sebaihia, K. D. James, C. Churcher, K. L. Mungall, S. Baker, D. Basham, S. D. Bentley, K. Brooks, A. M. Cerdeno-Tarraga, T. Chillingworth, A. Cronin, R. M. Davies, P. Davis, G. Dougan, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Leather, A. V. Karlyshev, S. Moule, P. C. F. Oyston, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S. Whitehead, and B. G. Barrell. 2001. Genome sequence of *Yersinia pestis*, the causative agent of plague. *Nature* **413**:523–527.
- Rhim, S.-L., B. Völksch, L. Gardan, J.-P. Paulin, C. Langlotz, W.-S. Kim, and K. Geider. 1999. *Erwinia pyrifoliae*, an *Erwinia* species different from *Erwinia amylovora*, causes a necrotic disease of Asian pear trees. *Plant Pathol.* **48**:514–520.

36. **Ritchie, D. F., and E. J. Klos.** 1977. Isolation of *Erwinia amylovora* bacteriophage from the aerial parts of apple trees. *Phytopathology* **67**:101–104.
37. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. *Molecular cloning: a laboratory manual*, 2nd ed., vol. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
38. **Schnabel, E. L., and A. L. Jones.** 1998. Instability of a pEA29 marker in *Erwinia amylovora* previously used for strain classification. *Plant Dis.* **82**: 1334–1336.
39. **Segal, G., and E. Z. Ron.** 1996. Regulation and organization of the *groE* and *dnaK* operons in Eubacteria. *FEMS Microbiol. Lett.* **138**:1–10.
40. **Siebert, P. D., A. Chenchik, D. E. Kellogg, K. A. Lukyanov, and S. A. Lukyanov.** 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* **23**:1087–1088.
41. **Stackebrandt, E., and B. M. Goebel.** 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.* **44**:846–849.
42. **Wagner, E. G. H., and R. W. Simons.** 1994. Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* **48**:713–742.
43. **Wilson, E. E., F. M. Zeitoun, and D. L. Fredrickson.** 1967. Bacterial phloem canker, a new disease of Persian walnut trees. *Phytopathology* **57**: 618–621.
44. **Wilson, K.** 1997. Preparation of genomic DNA from bacteria, p. 2.4.1–2.4.2. *In* F. M. Ausubel et al. (ed.), *Current protocols in molecular biology*, vol. 1. John Wiley & Sons, Inc., New York, N.Y.
45. **Zhang, Y., and K. Geider.** 1997. Differentiation of *Erwinia amylovora* strains by pulsed-field gel electrophoresis. *Appl. Environ. Microbiol.* **63**:4421–4426.