Identification and Characterization of Transposable Elements of *Paracoccus pantotrophus*

Dariusz Bartosik,* Marta Sochacka, and Jadwiga Baj

Warsaw University, Institute of Microbiology, Department of Bacterial Genetics, 02-096 Warsaw, Poland

We studied diversity and distribution of transposable elements residing in different strains (DSM 11072, DSM 11073, DSM 65, and LMD 82.5) of a soil bacterium *Paracoccus pantotrophus* (α-*Proteobacteria*). With application of a shuttle entrapment vector pMEC1, several novel insertion sequences (ISs) and transposons (Tns) have been identified. They were sequenced and subjected to detailed comparative analysis, which allowed their characterization (i.e., identification of transposase genes, terminal inverted repeats, as well as target sequences) and classification into the appropriate IS or Tn families. The frequency of transposition of these elements varied and ranged from 10^{-6} to 10^{-3} depending on the strain. The copy number, localization (plasmid or chromosome), and distribution of these elements in the *Paracoccus* species *P. pantotrophus*, *P. denitrificans*, *P. methylutens*, *P. solventivorans*, and *P. versutus* were analyzed. This allowed us to distinguish elements that are common in paracocci (IS*Ppa2*, IS*Ppa3*—both of the IS5 family—and IS*Ppa5* of IS66 family) as well as strain-specific ones (IS*Ppa1* of the IS256 family, IS*Ppa4* of the IS5 family, and Tn*3434* and Tn*5393* of the Tn3 family), acquired by lateral transfer events. These elements will be of a great value in the design of new genetic tools for paracocci, since only one element (IS*1248* of *P. denitrificans*) has been described so far in this genus.

Insertion sequences (ISs), which are components of nearly all bacterial genomes, are the simplest forms of transposable elements (for reviews, see references 10 and 38). Although ISs do not encode genes that directly affect the phenotype of the bacterial host, they are not, however, eliminated in the course of evolution. It thus seems little likely that they are only "selfish" DNA that escapes natural selection due to more active replication related to transposition. The transposition of these elements usually involves structural changes in DNA that lead to the formation of mutations: insertions, deletions, inversions, and translocations (of even large DNA fragments), which may result in varied phenotypic effects (19). Insertion of ISs can lead to gene disruption or activation of adjacent cryptic genes by formation of upstream promoters (29). ISs are therefore the most recombinogenic factor of bacterial genomes. Their presence to a large extent determines that these genomes are not static, monolithic structures, but are subject to many structural rearrangements. These elements thus play the role of a factor that significantly enhances variability and consequently the adaptive and evolutionary capacities of their hosts.

In addition, ISs frequently occur on plasmids and bacteriophages, which can propagate them by lateral transfer between various bacterial populations. Some ISs are able to form composite transposons (Tns), which can carry various phenotypic traits (e.g., resistance to antibiotics or heavy metals and the ability to utilize different carbon sources) (1, 40). To date, more than 800 IS elements have been isolated from more than 200 prokaryotic species of both bacteria and archaea (http://www.is.biotoul.fr/is.html). The majority of transposable ele-

ments were identified from various sequencing projects. However, in most cases, their activity was not experimentally confirmed.

We initiated studies aimed at the identification and characterization of functional transposable elements of paracocci (α-*Proteobacteria*), which are physiologically among the most versatile bacteria performing a number of different growth modes. During the past decade, the genus *Paracoccus* has undergone serious taxonomic changes. Several new species were isolated, and the status of others was re-evaluated (2, 32). The genus currently embraces 17 species. Thus far, the only known transposable element of these bacteria is IS*1248* of *P. denitrificans* PdX13, classified as part of the IS5 family (3). There is a much greater wealth of information on ISs in the related genus *Rhizobium* (http://www.is.biotoul.fr/is.html).

In this paper, we present an analysis of the transposable elements of *P. pantotrophus*, which is a facultative chemolitoautotroph. It is able to grow with molecular hydrogen, sulfide, or thiosulfate as an electron donor under aerobic conditions. All three substrates can be used for mixotrophic growth. A large variety of organic compounds support aerobic and anaerobic (denitrifying) growth of these bacteria (32). *P. pantotrophus* strains show a great physiological heterogeneity, which is manifested, for instance, in the capacity of only some of them for methylotrophic growth (41). The high variability and plasticity of the species are indicated, for example, by the observation of the appearance within strain ATCC 35512 (unable to utilize methanol) of spontaneous mutants that were able to grow with methanol as the sole source of carbon and energy (17). Although there is no evidence linking the presence of transposable elements with important phenotypic characteristics of paracocci, however, that their physiological heterogeneity might result from various transposition events cannot be ruled out.

* Corresponding author. Mailing address: Warsaw University, Institute of Microbiology, Department of Bacterial Genetics, Miecznikowa 1, 02-096 Warsaw, Poland. Phone: 48 22 5541344. Fax: 48 22 5541404. E-mail: bartosik@biol.uw.edu.pl.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Characteristics	Source or reference
Strains		
P. pantotrophus		
DSM 11072	Contains pWKS1 (2.7 kb) and pWKS3 (>400 kb)	31
UWP1	Rif[r] derivative of DSM 11072 deprived of pWKS1	5
DSM 11073	Contains pKLW1 (~100 kb) and pKLW2 (>400 kb)	7, 31
DSM 11073R	Rif[r] derivative of DSM 11073	7
DSM 65	Contains pHG16-a (~70 kb) and pHG16-b (>400 kb)	7, 22
DSM 65R	Rif[r] derivative of DSM 65	5
LMD 82.5	Contains pPAN1 (~110 kb) and pPAN2 (>400 kb)	7
LMD 82.5R	Rif[r] derivative of LMD 82.5	5
P. denitrificans DSM 413	Contains pHG18 (>400 kb)	22, 41
P. methylutens DM12	Contains pMTH4 (~23 kb), pMTH1 (~40 kb), pMTH2 (100 kb) and pMTH3 (>400 kb)	2, 16
P. solventivorans DSM 6637	Contains pSOS1 (~70 kb)	2, 47
P. versutus UW1	Contains pTAV1 (107 kb)	6, 50
E. coli		
TG1	Host strain for plasmids	43
DH5α	Host strain for pRK2013	43
Plasmids		
pABW1	Km[r]; mobilizable cloning vector; ColE1 origin; *oriT* RK2	4
pGBG1	Cm[r]; contains a selective cartridge composed of *cI* gene of bacteriophage λ and tetracycline resistance gene under control of *pR* promoter	45
pWKS1	2.7 kb; natural plasmid of *P. pantotrophus* DSM 11072	5
pRK2013	Km[r]; helper plasmid carrying RK2 *tra* genes	15
pMEC1	Km[r]; shuttle vector composed of pABW1, replicator region of pWKS1, and selective cartridge of pGBG1	This study
pMEC1Δ	Km[r] Tc[r]; deletion derivative of pMEC1 deprived of functional *cI* gene	This study
pMEC100-199	Tc[r] derivatives of pMEC1 obtained in *P. pantotrophus* DSM 11072	This study
pMEC200-299	Tc[r] derivatives of pMEC1 obtained in *P. pantotrophus* DSM 11073	This study
pMEC300-399	Tc[r] derivatives of pMEC1 obtained in *P. pantotrophus* DSM 65	This study
pMEC400-499	Tc[r] derivatives of pMEC1 obtained in *P. pantotrophus* LMD 82.5	This study

With application of the entrapment vector pMEC1 (which carries a replicon of plasmid pWKS1 that is functional in paracocci) (5), we identified several novel transposable elements (some of which were most probably gained by lateral gene transfer). We showed that ISs and Tns are common in *P. pantotrophus*, which points to the high potential of these elements in determining internal genetic rearrangements of their hosts' genomes.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. All bacterial strains and plasmids used in this study are listed in Table 1. Bacteria were grown in Luria-Bertani (LB) medium (43) at 30°C (*Paracoccus* spp.) or 37°C (*Escherichia coli*). When necessary, the medium was supplemented with antibiotics as follows: kanamycin, 50 μg/ml; streptomycin, 50 μg/ml; rifampin, 50 μg/ml; and tetracycline, 0.5 to 3 μg/ml for different strains of *P. pantotrophus* (Table 2) and 20 μg/ml for *E. coli*.

DNA manipulations. Plasmid DNA was isolated according to the method of Birnboim and Doly (8) and, if required, purified by CsCl-ethidium bromide gradient centrifugation. Cloning experiments, digestion with restriction enzymes, ligations, and agarose gel electrophoresis were conducted in accordance with standard procedures as described by Sambrook and Russel (43). All enzymes were purchased from either Promega or Fermentas.

Entrapment vector construction. For the construction of the entrapment shuttle vector, the 1.8-kb *Eco*RI-*Pst*I restriction fragment carrying the replicator region of pWKS1 (5) was cloned into the corresponding multiple cloning site (MCS) of an *E. coli* specific mobilizable vector, pABW1 (4). The hybrid plasmid

was linearized with *Pst*I and ligated with a 2.9-kb entrapment cartridge (obtained from plasmid pGBG1) (45) composed of (i) a silent tetracycline resistance gene (*tetA*) under control of the *pR* promoter of bacteriophage λ and (ii) a gene coding for the λ CI repressor. Inactivation of a repressor gene or operator (e.g., through insertion of a mobile genetic element) results in constitutive expression of the tetracycline resistance. The resulting plasmid was designated pMEC1. A deletion derivative of pMEC1 (pMEC1Δ) was constructed by deletion of a 0.5-kb *Hin*dIII fragment coding for the terminal part of the *cI* gene. Disruption of this gene resulted in constitutive expression of tetracycline resistance.

TABLE 2. Analysis of Tc[r] mutants isolated from different *P. pantotrophus* strains by using pMEC1

P. pantotrophus strain	Tetracycline concn (μg/ml)	Frequency of Tc[r] mutants[a]	Frequency of mutation (%)[b]		
			Point mutations	Insert	
				<3 kb (IS)	>3 kb (Tn)
DSM 11072	3.0	5.0×10^{-6}	21	28	51[c]
DSM 65	0.5	8.9×10^{-5}	80	4	16
DSM 11073	1.0	1.1×10^{-3}	0	100[c]	0
LMD 82.5	1.0	1.2×10^{-3}	0	0	100[c]

[a] Calculated as the ratio between the number of Tc[r] clones obtained on selective medium and the total number of bacteria.
[b] One hundred Tc[r] clones were analyzed.
[c] All the inserts were identical in size.

Transformation. Competent cells of *E. coli* TG1 were prepared and transformed as described by Kushner (33). Transformants were selected on solidified LB medium supplemented with the appropriate antibiotic.

Triparental mating. Overnight cultures (spun down and washed to remove antibiotics) of the donor strain *E. coli* TG1 carrying a mobilizable vector, *P. pantotrophus* as the recipient, and *E. coli* DH5α carrying the helper plasmid pRK2013 were mixed 1:2:1. An aliquot of 100 μl of such mixture was spread on a plate with solidified LB medium. After overnight incubation at 30°C, the bacteria were washed off the plate, and suitable dilutions were plated on selective media containing rifampin (selective marker of the recipient strain) and kanamycin to select transconjugants. Spontaneous resistance of the recipient strains to kanamycin was undetectable under these experimental conditions.

Isolation of insertion mutants. The entrapment vector pMEC1 was introduced into recipient *P. pantotrophus* strains by triparental mating. The overnight culture of the Km^r transconjugant, carrying pMEC1, was spread on plates with solidified LB medium supplemented with tetracycline. Appropriate dilutions of the culture were also spread on tetracycline-free LB medium in order to determine the frequency of transposition. One hundred Tc^r colonies of each strain were further analyzed for plasmid content and restriction pattern. Spontaneous resistance of each strain to tetracycline was undetectable under these experimental conditions.

PCR amplification. For amplification of transposable elements, the following five pairs of forward and reverse primers (based on the sequence of a selective cartridge) were used: ALIS (5′-TTGTAATCAGCTATGCGCCG-3′) and ARIS (5′-TCTGGCTTGAGGTTGAAGGT-3′), BLIS (5′-TGGTGCGGTCATGGA ATTAC-3′) and BRIS (5′-GTATGCAGCCGTCACTTAGA-3′), CLIS (5′-TC CCTGCCTGAACATGAGAA-3′) and CRIS (5′-ACACAAGAGCAGCTTGA GGA-3′), DLIS (5′-TCTTGTCTGCGACAGATTCC-3′) and DRIS (5′-TTCA TACACGGTGCCTGACT-3′), and ELIS (5′-GGTTGCATGTACTAAGGAG G-3′) and ERIS (5′-GCAAGACTGGCATGATAAGG-3′). For amplification of the internal fragments of the transposable elements (used as probes in hybridization), the following primer pairs were used: LPPA1 (5′-GCGGCATATCAA GGCGGTGT-3′) and RPPA1 (5′-TTCTGACCGTCGAGCTTGCG-3′) (positions 35 to 1276 of IS*Ppa1* sequence), LPPA2 (5′-TGTACTGGCTGACCAAC GAG-3′) and RPPA2 (5′-TGGCTGAGAGGAACACCTTG-3′) (positions 75 to 787 of IS*Ppa2* sequence), LPPA3 (5′-TGGAAGTCCTACAACGATGC-3′) and RPPA3 (5′-ATGAGATGCGTTCACCGAAG-3′) (positions 107 to 899 of IS*Ppa3* sequence), LPPA5 (5′-TCGAGGATCTCATGATCGCA-3) and RPPA5 (5′-ACGTGC TCTTCCTTCCAGTC-3) (positions 1106 to 1863 of IS*Ppa5* sequence), L3434 (5′-ATCGGCAAGGCAGATTGACC-3′) and R3434 (5′-CAA GATGTTCACTGGCCGAG-3) (positions 964 to 2592 of Tn*3434* sequence), and LSTRAB (5′-ACGGCTTGCCTTCTATCTGC-3′) and RSTRAB (5′-AGA ATGCGTCCGCCATCTGT-3′) (positions 4483 to 5212 of Tn*5393* sequence). Amplification was performed in a Mastercycler (Eppendorf) with the synthetic oligonucleotides described above, *Thermus aquaticus* polymerase (Qiagen), and appropriate template DNAs. PCR products were separated by 0.8% agarose gel electrophoresis and if necessary purified with the Gel Out kit (DNA Gdansk II).

DNA sequencing and analysis. The nucleotide sequence was determined with a terminator sequencing kit and a Perkin-Elmer automatic sequencer (ABI 377). The transposable elements (present in pMEC1 derivatives) were sequenced starting with the appropriate sets of cartridge-specific starters (Fig. 1) and then with primers complementary to the previously determined sequence. Sequence analysis was done with programs included in the Genetics Computer Group (University of Wisconsin, Madison) GCG Package (14). Comparison searches were performed with IS Finder (http://www–is.biotoul.fr/is.html) and with the BLAST program provided by the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST).

Nucleotide sequence accession number. The nucleotide sequences of IS*Ppa1*, IS*Ppa2*, IS*Ppa3*, IS*Ppa4*, IS*Ppa5*, and Tn*3434* have been submitted to the GenBank database under accession no. AY177680, AY179508, AY177681, AY177682, AY225410, and AY232820, respectively.

RESULTS

Construction of entrapment vector pMEC1. In the first stage of our studies, we constructed a novel shuttle entrapment vector, pMEC1, which enables direct in vivo capture of transposable mobile elements in paracocci (Fig. 1). The vector is composed of (i) a replicator region of a small cryptic plasmid, pWKS1 (2.7 kb), of *P. pantotrophus* DSM 11072 (which ensures stable maintenance in all paracocci and is compatible with their natural plasmids) (5); (ii) *E. coli*-specific Km^r mo-
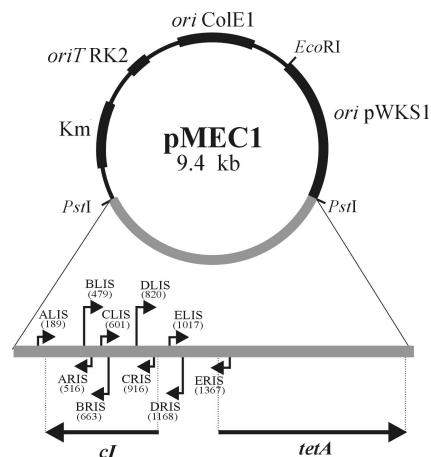

FIG. 1. Genetic organization of the shuttle entrapment vector pMEC1. A selective cartridge for positive selection of transposable elements, composed of the λ repressor gene (*cI*) and the silent tetracycline resistance gene (*tetA*), is shown below. The five pairs of nested primers (ALIS and ARIS, BLIS and BRIS, etc.) used for precise localization of the insertion site of the transposable elements by PCR are indicated. The positions of the primers within the selective cartridge are given in parentheses.

bilizable vector pABW1 (4); and (iii) an entrapment cartridge carrying a silent tetracycline resistance (*tetA*) gene (see Materials and Methods for details).

Both kanamycin and tetracycline markers (present in pMEC1) were suitable for selection in *Paracoccus* spp. However, the tetracycline concentration used for paracocci depends on the strain and copy number of the Tc^r plasmid (unpublished results). Therefore, we constructed a deletion derivative of pMEC1 constitutively expressing resistance to tetracycline (pMEC1Δ; see Materials and Methods for details), which allowed direct determination of appropriate tetracycline concentrations for the strains studied (Table 2).

Analysis of a pool of Tc^r mutants of *P. pantotrophus* strains. Plasmid pMEC1 (Km^r) was introduced into different strains of *P. pantotrophus* (DSM 11072, DSM 11073, DSM 65, and LMD 82.5), and then Tc^r mutants were selected as described in Materials and Methods. As shown in Table 2, Tc^r clones appeared with frequencies from 10^{-6} to 10^{-3}, depending on the strain. We analyzed the plasmid pattern of 100 Tc^r clones of each of the *Paracoccus* strains studied. For DSM 11072 and DSM 65, the size of the inserts (roughly estimated by restriction analysis) was differentiated (Table 2). In both cases, preliminary classes of plasmids carrying potential (i) ISs (below 3 kb), (ii) Tns (above 3 kb), and (iii) point mutations (plasmids of the size of pMEC1), were distinguished. In the case of two other strains tested, all of the Tc^r mutants carried one kind of insert of the sizes 2.8 kb (DSM 11073) and 5.5 kb (LMD 82.5).

Hybridization analysis. Initially we focused on the analysis of the insertion sequences "captured" in DSM 11072. For determination of their accurate size, we designed (on the basis of the sequence of the selective cartridge) five pairs of nested primers (Fig. 1) (nucleotide sequences are given in Materials and Methods). The primers together with pMEC1-derived plasmids (as template DNA) were used in PCRs. This additionally enabled precise localization of the transposable ele-

TABLE 3. Characteristic features of novel transposable elements of *P. pantotrophus*

Transposable element	Length (bp)	G+C content (mol%)	IR length (bp)	DR length in bp (sequence)	IS (or Tn) with highest level of homology	% Tnp identity/similarity	Special features	IS (or Tn) family/group
IS*Ppa1*	1,376	61	39	9 (CCAAAAAG)	IS*1490* (*B. cepacia*)	55/65	IHF binding site	IS*256*
IS*Ppa2*	832	60	14	2 (TA)	IS*1248* (*P. denitrificans*)	92/96 (ORF1), 90/90 (ORF2)	IHF binding site; frameshifting	IS5/IS427
IS*Ppa3*	1,055	61	24	9 (CTCATACTC)	IS*Aav2* (*A. avenae*)	55/70	Chi site	IS5/IS903
IS*Ppa4*	1,053	61	15	9 (AAGGTGATG)	IS*Aav2* (*A. avenae*)	56/63		IS5/IS903
IS*Ppa5*	2,829	63.5	32	8 (GAACATCC)	IS*Rm15* (*S. meliloti*)	56/65 (ORF1), 88/90 (ORF2), 71/76 (ORF3)	Translational coupling	IS66
Tn*3434*	3,659	59	38	6 (AGGGTA)	Tn*5393*	63/77	Cryptic	Tn*3*

ments in the cassette and the choice of the appropriate primers for sequencing of the entire element and its target site. We amplified all of the inserts, which confirmed that the insertions took place within the *cI* gene. On the basis of the sizes determined, four IS classes, of approximately 0.8, 1, 1.3, and 2.8 kb, were distinguished. Randomly selected representatives of each class (present in pMEC135, pMEC181, pMEC114, and pMEC156, respectively) were used for further analysis. To check whether the ISs classified in the individual classes are homologous, hybridization analysis was performed. For this purpose, the digoxigenin (DIG)-labeled internal fragments of the ISs (present in the four pMEC1 derivatives mentioned above) were probed, respectively, against the collection of previously PCR-amplified paracoccal transposable elements. The hybridization results confirmed that each of the distinguished IS classes contained highly homologous (probably identical) elements, which do not cross-hybridize with other elements tested (data not shown). The same probes were used for hybridization with PCR-amplified ISs of DSM 65 and DSM 11073. It appeared that the 1-kb-long IS of DSM 65 (present in pMEC382) did not hybridize with any of the probes. Each of the remaining three IS elements "captured" in this strain (present in pMEC306, pMEC355, and pMEC399) was 0.8 kb and gave weak hybridization signals with the IS-based probe of pMEC135, which indicates that they are slightly divergent (data not shown). On the other hand, all of the elements of DSM 11073 hybridized with the IS-based probe of pMEC156 (data not shown).

IS nucleotide sequences. The representative elements of each of the distinguished class (present in pMEC114, pMEC135, pMEC181, and pMEC156 of DSM 11072 as well as pMEC382 and pMEC399 of DSM 65) were sequenced. Comparison of the nucleotide sequences with those in databases revealed that only one of them (present in pMEC399) had been described earlier, as IS*1248a* of *P. denitrificans* PdX13 (52). The others are novel elements. Detailed analysis of the nucleotide sequences of the captured elements allowed to determine their structure, that is to identify (i) the transposase (Tnp) gene or genes, (ii) terminal inverted repeats (IRs) being the sites for Tnp binding and action, and (iii) target sequence (DR), which in most cases is duplicated upon insertion. Since all of the features mentioned (summarized in Table 3) were present, the captured elements were designated IS*Ppa1* (from pMEC114), IS*Ppa2* (from pMEC135), IS*Ppa3* (from pMEC181), IS*Ppa4* (from pMEC382), and IS*Ppa5* (from pMEC156). The G+C content of the sequences was in the range 61 to 63.5 mol%

(Table 3), which is slightly lower than previously determined for *P. pantotrophus* total DNA (64 to 68 mol%) (41).

(i) Characterization of IS*Ppa1*. The nucleotide sequence of IS*Ppa1* did not show any significant similarity to the sequences deposited in databases. Analysis of sequence data revealed the presence of one major open reading frame (ORF1), spanning 91% of the element. The presumptive ribosome binding site (RBS) is located within a purine-rich region (5′-GAGAAAG GGA-3′), 9 bp upstream of the initiation codon (ATG) of ORF1. No sequences homologous to the consensus of the *E. coli* promoters could be identified upstream of ORF1. ORF1 (nucleotides 96 to 1346) encodes a putative peptide of 416 amino acids (aa) with a predicted molecular mass of 47.3 kDa. The theoretical pI calculated from the deduced amino acid sequence of the putative protein is 9.76, which is typical for IS-encoded transposases (10). A comparison of this hypothetical protein with those in databases revealed the highest degree of homology with the product of *Magnetospirillum magneto-tacticum* gene *Magn5889* (accession no. ZP 00053133), putative transposases of IS*1490* and IS*1413* of *Burkholderia cepacia* (36), and the product of *Bradyrhizobium japonicum* chromosomal gene *id145* (accession no. AF322013). Significant similarities were also observed with transposases of IS*1632* of *B. japonicum* (30), IS*6120* of *Mycobacterium smegmatis* (25), IS*1166* of *Rhodococcus zopfii* (13), IS*T2* of *Thiobacillus fer-rooxidans* (54), and IS*666* of *Mycobacterium avium* (44). All of the elements listed above are classified as members of a distinct closely homologous subclass within the family IS*256*; however, for the majority of them, the transposition activity has not been experimentally determined (http://www-is.biotoul.fr/is .html). Of particular importance was the presence within the ORF1 product of three conserved domains (N2, N3, and C1) that encompass the DDE motif (in boldface; lowercase letters represent residues that are not highly conserved) (**DG** [65 aa] **DG**- -GF [106 aa] I-TSN- -**E**- - - - -VR), which has been shown to be conserved among members of the IS*256* family (**d**a/g [67 aa] **D**g/a- -gf/l [112 aa] i/l-t/st/nN- -**E**- - - - -v/lr/k) (38). This motif, composed of three acidic residues, was found in the catalytic domains of transposases of many bacterial mobile elements and integrases of retroviruses (10).

IS*Ppa1* has imperfect terminal IR sequences (IR1; Fig. 2A) of 39 bp (with 10 mismatches), which show similarities to corresponding regions of several members of IS*256* family (the best matches are shown in Fig. 2B). The sequences adjacent to IR1 of IS*Ppa1* encode additional two stretches of conserved inverted sequences IR2 (10 bp, with three mismatches) and
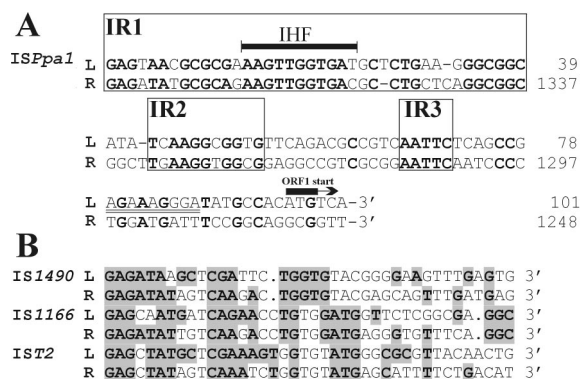
FIG. 2. (A) Comparison of the terminal nucleotide sequences of IS*Ppa1*. The identical residues are boldface. Three stretches of conserved sequences are boxed and indicated as IR1, IR2, and IR3. A putative site for binding IHF, present in IRL of IS*Ppa1*, is indicated by a thick bar. The start codon (ATG) of ORF1 is marked, and its transcriptional orientation is shown by an arrow. The numbers on the right refer to the deposited nucleotide sequence of IS*Ppa1* (accession no. AY177680). (B) Alignment of the nucleotide sequences of the left and right termini of related members of IS*256* family. Nucleotides identical to those of IS*Ppa1* are boldface and shaded. "L" denotes sequences at the 5′ (left) end, and "R" denotes complementary sequences at the 3′ (right) end of the elements.

FIG. 3. (A) Genetic organization of IS*Ppa2*. IRLs and IRRs are shown as solid boxes. The two consecutive overlapping ORFs (ORF1 and ORF2) and their transcriptional orientation are shown as arrows. The location of the DNA binding domain (HTH) of ORF1 and the DDE motif of a putative fusion protein are marked. The location of a putative frameshifting motif (5′-AAAAAG-3′) is indicated. (B) Alignment of the terminal nucleotide sequences of IS*Ppa2* and its closest relatives of the IS*427* group (IS*5* family). The identical residues of termini of IS*Ppa2* are boldface. Nucleotides of other sequences identical with those of IS*Ppa2* are boldface and shaded. A putative site for binding IHF, present in the IRL of IS*Ppa2*, is indicated by a thick bar. The putative IRs of IS*Ppa2* are depicted by a thin bar. "L" denotes sequences at the 5′ (left) end, and "R" denotes complementary sequences at the 3′ (right) end of the elements.

IR3 (5 bp) separated by a 13-bp spacer (Fig. 2A). These sequences might correspond to the internal IRs (being putative host factor binding sites), which have been previously reported for, e.g., IS*256* of *Staphylococcus aureus* and the related sequence, IS*1191* of *Streptococcus thermophilus*, but not for all members of IS*256* family (24). The IR1 located at the 5′ end of the element (IR1L) contains, within a conserved region, a sequence (5′-AAGTTGGTGAT-3′; Fig. 2A), which shows similarity (1 mismatch) to the consensus sequence of the integration host factor (IHF) binding site (5′-AANNNNTTGAT-3′) (21). Interestingly, within the cartridge, downstream of IS*Ppa1*, we localized additional sequence (5′-AAGCCAGTGAT-3′), partially located within the target site, matching the putative IHF binding site of the left-terminal IR (IRL) of IS*Ppa1*. A similar observation has been made for IS*1*, which has a preferred insertion site close to IHF binding site of pBR322 (21). IHF thus seems to be an important factor, which can participate in transposition of some transposable elements (19).

It has been shown that several members of the IS*256* family are able to create, after excision, circular forms, which are the intermediates in transposition (37). We were unable, however, to amplify by PCR (using outward-oriented primers based on the IS*Ppa1* sequence and total DNA of *P. pantotrophus* DSM 11072) a DNA fragment with the possible junction between IR1L and IR1R of IS*Ppa1* (data not shown).

(ii) Characterization of IS*Ppa2*. IS*Ppa2* is the shortest element identified in *P. pantotrophus* (Table 3). Similarly to the case in IS*Ppa1*, the IRL of IS*Ppa2* contains a putative IHF binding site (one mismatch) (Fig. 3B). However, no analogous sequences were identified in the proximity of the target site. IS*Ppa2* encodes two overlapping ORFs (159-bp overlap) (Fig. 3A). ORF1 (nucleotides 65 to 412) has an ATG initiation codon and encodes a potential protein of 115 aa with a pre-
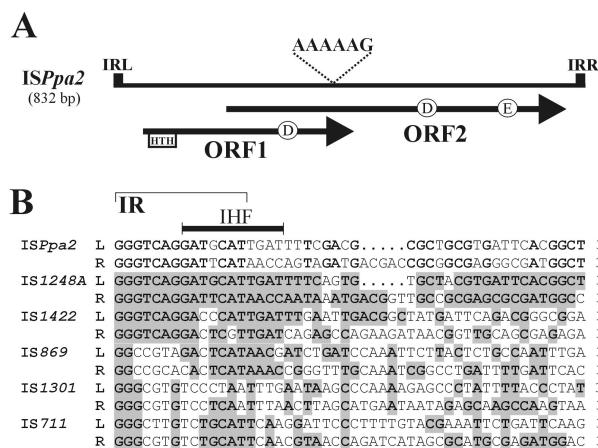
dicted molecular mass of 13.1 kDa (pI 10.18). A putative RBS (5′-GGAG-3′) is 7 bp upstream of ORF1. ORF2 (positions 253 to 822), with a GTG initiation codon and TGA stop codon placed within the right-terminal IR (IRR), has capacity to encode a 189-aa polypeptide of 21.8 kDa (pI 10.43). A potential frameshift motif (3′-AAAAAG-5′; position 389) located within a loop of a predicted mRNA stem-loop structure is present within the overlapping region (data not shown). It is probable that this motif (as shown for the IS*1* and IS*3* family members) (18, 26, 55) is able to promote a programmed translational frameshifting, which leads to the formation of a functional fusion transposase. Translational slippage in this case might result in the formation of a fusion protein (ORF1 plus ORF2) of 252 aa and a predicted molecular mass of 29.1 kDa (pI 10.62). The N-terminal part of ORF1 product (as well as of the putative fusion protein) carries a putative DNA binding helix-turn-helix (HTH) motif (data not shown). Additionally the fusion protein (but not the single ORFs) has a DDE motif (in boldface (V-I**D**AT [76 aa] LG**D**- -YD [44 aa] YK-R- -I**E**- -F-RLK) similar to transposases of IS*427* group of IS*5* family (v-I**d**st/s [76 aa] la**D**--YD [45 aa] Yk/r-R--i/v**E**--F-k/rLK) (38). The overall sequences of the ORF1 and ORF2 products showed the highest level of identity with putative transposases encoded by ORF1 and ORF2 of IS*1248b* (52, 53). Significant similarity was also observed with two corresponding ORFs of IS*1422* of *Ralstonia solanacearum* (accession no. BAA97980), as well as of IS*869* residing in plasmid pTiAB3 of *Agrobacterium tumefaciens* (39), IS*1301* of *Neisseria meningitidis* (27), and IS*711* of *Brucella ovis* (12)—all of which belong to the IS*427* group. The conservation of the IR sequences (Fig. 3B)

FIG. 4. Alignment of the terminal nucleotide sequences of IS*Ppa3*, IS*Ppa4*, and their closest relatives of the IS*903* group (IS*5* family). The identical residues of IS*Ppa3* sequences are boldface. Nucleotides of other sequences identical to those of IS*Ppa3* are boldface and shaded. The putative IRs of IS*Ppa3* are depicted by a thin bar. "L" denotes sequences at the 5′ (left) end, and "R" denotes complementary sequences at the 3′ (right) end of the elements.

FIG. 5. (A) Genetic organization of IS*Ppa5*. The ORFs and their transcriptional orientation are shown by arrows. IRLs and IRRs are shown as solid boxes. (B) Alignment of the terminal nucleotide sequences of IS*Ppa5* and its closest relatives of the IS*66* family. The putative IRs of IS*Ppa5* are depicted by thin bar. The identical residues of IS*Ppa5* sequences are boldface. Nucleotides of other sequences identical to those of IS*Ppa5* are boldface and shaded. "L" denotes sequences at the 5′ (left) end, and "R" denotes complementary sequences at the 3′ (right) end of the elements.

of IS*Ppa2* and the ISs listed above additionally confirmed the place of IS*Ppa2* within the subgroup of family IS*5* mentioned.

(iii) Characterization of IS*Ppa3* and IS*Ppa4*. Within the sequence of IS*Ppa3* of DSM 11072, we identified one long ORF (positions 68 to 1000), encoding a putative protein of 310 aa with a calculated molecular mass of 35.3 kDa (pI 10.81). No significant similarities were observed between the nucleotide sequence of IS*Ppa3* and sequences deposited in databases. However the predicted translation product of ORF1 was similar to several transposases of the IS*903* group of the IS*5* family and carried the DDE motif (boldface) (LLV**D**-TG [73 aa] T-**D**GAFDTR [67 aa] K- - -GYH-RSRI**E**ARMRCLK) highly conserved within this family (laI**D**-TG [71 aa] S-**D**GAYDTr [67 aa] K- - -gYh-RSls/i**E**TAMyRvK) (38). The most significant similarities were observed with putative transposases of plasmid pSD25 of *Ruegeria* sp. (accession no. AAN05200), IS*Aav2* of *Acidovorax avenae* subsp. *citrulli* (accession no. AF086815), IS*17* of the pAT484 plasmid of *Acinetobacter haemolyticus* (42), IS*Va2* of plasmid pJM1 of *Vibrio anguillarum* (51), and IS*1420* of *Ralstonia solanacearum* (accession no. AB028897). Interestingly, IS*Ppa3* encodes a sequence 5′-GCT GGTGG-3′ (position 421) identical to the Chi sites, which specify the place of action for RecBCD nuclease. This sequence of IS*Ppa3* might thus serve as a portable enhancer of recombination (19).

The nucleotide sequence of IS*Ppa4* of DSM 65 (1,053 bp) appeared to be highly homologous to IS*Ppa3* (73% of identity); however, no Chi-related sequence could be distinguished within this element. IS*Ppa4* contains one large ORF, encoding a putative peptide of 310 aa (pI 10.81) with 77% identity and 83% similarity to the putative transposase of IS*Ppa3*. The DDE motifs as well as the IRs of these two ISs are also highly conserved (data not shown and Fig. 4, respectively). Taking into account results of the comparative analysis, we can state that IS*Ppa3* and IS*Ppa4* can be classified as new closely related elements of the IS*903* group of the IS*5* family.

(iv) Characterization of IS*Ppa5*. IS*Ppa5* of DSM 11072, the longest IS identified in *P. pantotrophus* (Table 3), carries four similarly oriented ORFs (Fig. 5A) encoding putative proteins with predicted molecular masses of 14.1 kDa (ORF1, positions 100 to 483), 12.6 kDa (ORF2, positions 480 to 827), 60.3 kDa (ORF3, positions 890 to 2518), and 11 kDa (ORF4, positions 2518 to 2814), respectively. Database searches and comparison
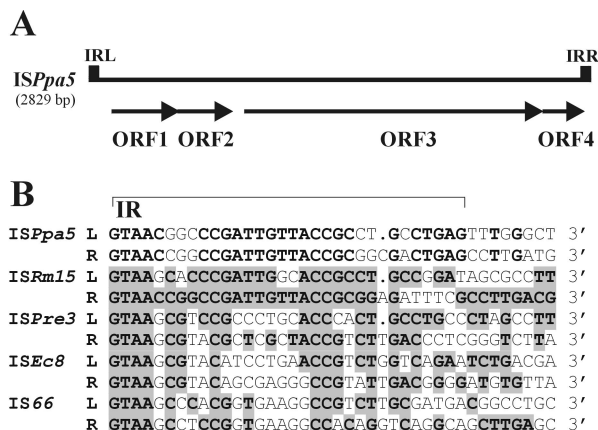
of their predicted amino acid sequences as well as comparison of the IR sequences (Fig. 5B) revealed that IS*Ppa5* belongs to the IS*66* family. Members of this family have been previously reported exclusively in bacteria belonging to the family *Rhizobiaceae* (genera *Agrobacterium* and *Rhizobium*) (38). More detailed analysis performed by Han et al. (28) demonstrated, however, that they are also widely distributed in other genera of gram-negative bacteria, including *Escherichia*, *Pseudomonas*, and *Vibrio*. Homology searches of the databases revealed that IS*Ppa5* is the most similar to several related ISs residing in the chromosome and symbiotic plasmid pSymA of *Sinorhizobium meliloti* (20). The highest similarities were observed with IS*Rm15* (72% identity at the nucleotide sequence level), which has been found to be inserted between the *nodI* and *nodQ* genes (nodulation) of *S. meliloti* (accession no. AF357834).

The majority of IS*66*-like elements encode three ORFs, corresponding to ORF1, ORF2, and ORF3 of IS*Ppa5*. Some of these ISs encode an additional ORF or ORFs, which are not, however, homologous to ORF4 of IS*Ppa5* (data not shown). In the case of IS*Ppa5* (and other IS*66* family members), the TGA termination codon of ORF1 overlaps the initiation codon of ORF2 (located in the −1 frame with respect to ORF1) within the 5′-ATGA-3′ sequence. The intergenic region between ORF2 and ORF3 of IS*Ppa5* is 61 bp long (usually 20 bp in the IS*66* family) (28). In contrast, the stop codon of ORF3 partially overlaps the start codon of ORF4 (located in the −1 frame with respect to ORF3) within the sequence 5′-TGATG-3′.

Database searches and comparison of the ORF3 product of IS*Ppa5* with over 40 homologous sequences revealed the presence of four conserved DDEE residues, which probably encompass a transposase-specific DDE catalytic triad (data not shown). These conserved residues are indicated in a PILEUP alignment of ORF3 and its closest relatives from IS*Rm15*, IS*Pre3* of *Pseudomonas resinovorans* (plasmid pCAR1; accession no. NC 004444), IS*Ec8* of *E. coli* EDL933 (associated with
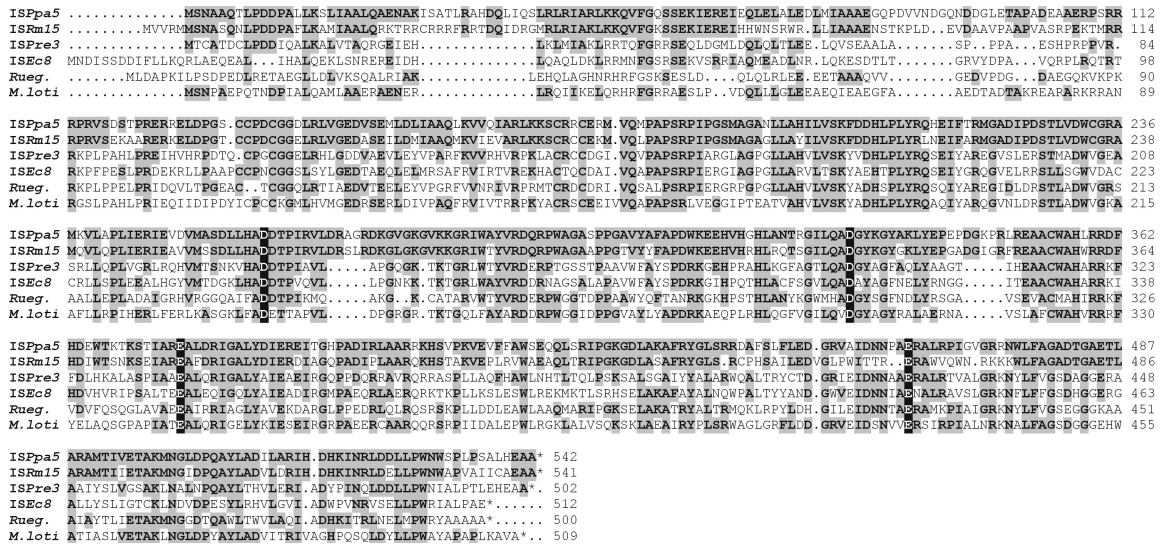
```
ISPpa5   .............MSNAAQTLPDDPALLKSLIAALQAENAKISATLRAHDQLIQSLRLRIARLKKQVFGQSSEKIEREIEQLELALEDLMIAAAEGQPDVVNDGQNDDGLETAPADEAAERPSRR 112
ISRm15   .........MVVRMMSNASQNLPDDPAFLKAMIAALQRKTRRCRRRFRRTDQIDRGMRLRIARLKKQVFGQSSEKIEREIHHWNSRWR.LLIAAAENSTKPLD..EVDAAVPAAPVASRPEKTMRR 114
ISPre3   .............MTCATDCLPDDIQALKALVTAQRGEIEH.............LKLMIAKLRRTQFGRRSEQLDGMLDQLQLTLEE.LQVSEAALA......SP..PPA..ESHPRPPVR. 84
ISEc8    MNDISSDDIFLLKQRLAEQEAL...IHALQEKLSNREPEIDH.............LQAQLDKLRRMNFGSRSEKVSRRIAQMEADLNR.LQKESDTLT.......GRVYDPA...VQRPLRQTRT 98
Rueg.    ....MLDAPKILPSDPEDLRETAEGLLDLVKSQALRIAK.............LEHQLAGHNRHRFGSKSESLD...QLQLRLEE.EETAAAQVV........GEDVPDG...DAEGQKVKPK 90
M.loti   .............MSNPAEPQTNDPIALQAMLAAERAENER.............LRQIIKELQRHRFGRRAESLP..VDQLLLGLEEAEQIEAEGFA......AEDTADTAKREARARKRRAN 89

ISPpa5   RPRVSDSTPRERELDPGS.CCPDCGGDLRLVGEDVSEMLDLIAAQLKVVQIARLKKSCRRCERM.VQMPAPSRPIPGSMAGANLLAHILVSKFDDHLPLYRQHEIFTRMGADIPDSTLVDWCGRA 236
ISRm15   RPRVSEKAARERKELDPGT.CCPDCGGELRLVGEDASEILDMIAAQMKVIEVARLKKSCRCCEKM.VQLPAPSRPIPGSMAGAGLLAYILVSKFDDHLPLYRLNEIFARMGADIPDSTLVDWCGRA 238
ISPre3   RKPLPAHLPREIHVHRPDTQ.CPGCGGELRHLGDDVAEVLEYVPARFKVVRHVRPKLACRCCDGI.VQVPAPSRPIARGIAGPGLLAHVLVSKYVDHLPLYRQSEIYAREGVSLERSTMADWVGEA 208
ISEc8    RKFPPESLPRDEKRLLPAAPCCPNCGGSLSYLGEDTAEQLELMRSAFRVIRTVREKHACTQCDAI.VQVPAPSRPIERGIAGPGLLARVLTSKYAEHTPLYRQSEIYGRQGVELRRSLLSGWVDAC 223
Rueg.    RKPLPPELPRIDQVLTPGEAC..TCGGQLRTIAEDVETEELEYVPGRFVVNRIVRPRMTCRDCDRI.VQVPAPSRLVEGGIPTEATVAHVLVSKYADHSPLYRQAQIYAREGIDLDRSTLADWVGRS 213
M.loti   RGSLPAHLPRIEQIIDIPDYICPCCKGMLHVMGEDRSERLDIVPAQFRVIVTRRPKYACRSCEEIVVQAPAPSRLVEGGIPTEATVAHVLVSKYADHLPLYRQAQIYARQGVNLDRSTLADWVGKA 215

ISPpa5   MKVLAPLIERIEVDVMASDLLHADDTPIRVLDRAGRDKGVGKGVKKGRIWAYVRDQRPWAGASPPGAVYAFAPDWKEEHVHGHLANTRGILQADGYKGYAKLYEPEPDGKPRLREAACWAHLRRDF 362
ISRm15   MQVLQPLIERIEAVVMSSDLLHADDTPIRVLDRSLRDKGLGKGVKKGRIWTYVRDQRPWAEAQLTSGILQADGYKGY.CKLYEPGADGIGRFREAACWAHWRRDF 364
ISPre3   SRLLQPLVGRLRQHVMTSNKVHADDTPIAVL.....APG.GK.TKTGRLWTYVRDERPTGSSTPAAVWFAYSPDRKGEHPRAHLKGFAGTLQADGYAGFAQLVAAGT.....IHEAACWAHARRKF 323
ISEc8    CRLLSPLEEEALHGYVMTDGKLHADDTPVCVL.....LPGNKK.TKTGRLWAYVRDDRNAGSALAPAVWFAYSPDRKGIHPQTHLACFSGVLQADAYAGFNELYRNGG.....ITEAACWAHARRKI 338
Rueg.    AALLEPLADAIGRHVRGGQAIFADDTPIKMQ.....AKG..K.CATARVWTYVRDERPWGGTDPPAAWVGTFANKGKHPSTHLANYKGWMHADGYSGFNDLYRSGA.....VSEVACMAHIRRKF 326
M.loti   AFLLRPIHERLFERLKASGKLFADETTAPVL.....DPGRGR.TKTGQLFAYARDRPWGGIDPPGVAYLYAPDRKAEQPLRHLQGFVGILQVDGYAGYRALAERNA.....VSLAFCWAHVRRRF 330

ISPpa5   HDEWTKTKSTIAREEALDRIGALYDIEREITGHPADIRLAARRKHSVPKVEVFFAWSEQQLSRIPGKGDLAKAFRYGLSRRDAFSLFLED.GRVAIDNNPAERALRPIGVGRRNWLFAGADTGAETL 487
ISRm15   HDIWTSNKSEIAREADLDRIGALYDIERDIAGQPADIPLAARCKHSTAKVEPLRVWAEAQLTRGYGLS.RCPHSAILEDVGLPNTTTR..ERAWVQWN.RKKKWLFAGADTGAETL 486
ISPre3   FDLHKALASPIAMEALQRIGALYAIEAEIRGQPPDQRRAVRQRRASPLLAQFHAWLNHTLTQLPSKSALSGAIYYALARWQALTRYCTD.GRIEIDNNAERALRTVALGRKNYLFVGSDAGGERA 448
ISEc8    HDVHVRIPSALTPEEALECIQGLYAIEADIRGWPAEQRLAERQRKTKPLLKSLESWLREKMKTLSRHSELAKAFAYALNQWPALTYYAND.GWVEIDNNIAENALRAVSLGRKNFLFFGSDHGGERG 463
Rueg.    VDVFQSQGLAVAEEAIRRIAGLYAVEKDARGLPPEDRLQLRQSKHSPLLDDLEAWLAAQMAIRPGEASKELAKATRYALTRMQKLRPYLDH.GILEIDNNTAERAMKPIAIGRKNYLFVGSEGGGKAA 451
M.loti   YELAQSGPAPIATEEALQRIGELYKIESEIRGRPAEERCAARQQRSRPIIDALEPWLRGKLALVSQKSKLAEAIRYPLSRWAGLGRFLDD.GRVEIDSNVVERSIRPIALNRKNALFAGSDGGGEHW 455

ISPpa5   ARAMTIVETAKMNGLDPQAYLADILARIH.DHKINRLDDLLPWNWSPLPSALHEAA* 542
ISRm15   ARAMTIIETAKMNGIDPQAYLADVLDRIH.DHKINRLDELLPWNWAPVAIICAEAA* 541
ISPre3   AAIYSLVGSAKLNALNPQAYLTHVLERI.ADYPINQLDDLLPWNIALPTLEHEAA*. 502
ISEc8    ALLYSLIGTCKLNDVDPESYLRHVLGVI.ADWPVNRVSELLPWNRIALPAE*..... 512
Rueg.    AIAYTLIETAKMNGGDTQAWLTWVLAQI.ADHKITRLNELMPWRYAAAAA*...... 500
M.loti   ATIASLVETAKLNGLDPYAYLADVITRIVAGHPQSQLDYLLPWAYAPAPLKAVA*.. 509
```

FIG. 6. Alignment of deduced amino acid sequences of the ORF3 translational product of IS*Ppa5*, IS*Rm15* of *S. meliloti*, IS*Pre3* of *P. resinovorans*, IS*Eco8* of *E. coli*, and two unnamed IS*66*-like elements of *Ruegeria* sp. (*Rueg.*) and *M. loti*. Dots indicate gaps introduced to maximize the alignment. Amino acids identical to the protein of IS*Ppa5* are boldface and shaded. The four conserved DDEE residues (encompassing the putative DDE motif), distinguished on the basis of comparative studies of 42 related sequences, are shown on a black background. Amino acid numbering is shown to the right of the sequence lines.

a pathogenicity island) (46), and two unnamed ISs of *Mesorhizobium loti* R7A (associated with a symbiosis island) (48) and *Ruegeria* sp. strain PR1b (plasmid pSD25; accession no. AF416331) (Fig. 6).

Identification of transposons of *P. pantotrophus*. As previously mentioned, we distinguished in *P. pantotrophus* a class of putative transposons (Table 2). None of these elements was a composite Tn flanked by any of the IS elements described above (as confirmed by hybridization analysis [data not shown]). We analyzed two transposons in detail. One of them was the only transposon caught by the entrapment vector in strain DSM 11072 (as was judged from the restriction analysis and then confirmed by DNA hybridization [data not shown]). The second one, isolated from LMD 82.5, seemed interesting because of its extremely high frequency of transposition (Table 2). The comparative analysis of the nucleotide sequence of the transposon of LMD 82.5 showed that it is identical to Tn*5393* (accession no. AJ431260) of the Tn*3* family. Transposition of this element generated a 6-bp duplication of the target sequence (5′-TGAACA-3′). Tn*5393*, besides genetic information required for transposition (transposase [*tnpA*] and resolvase [*tnpR*] genes, resolution sequence [*res*], and terminal IRs) carries two streptomycin resistance genes (*strA* and *strB*) placed downstream of the *tnpR* gene (Fig. 7A).

Analysis of the nucleotide sequence of the transposon of DSM 11072, designated Tn*3434* (Table 3), revealed a structure similar to that of Tn*5393*. Tn*3434* does not, however, contain the *str* module (Fig. 7A). It encodes two divergently oriented ORFs coding for a putative transposase of 960 aa and predicted mass of 107.3 kDa (ORF1, positions 48 to 2930) and a resolvase of 203 aa and predicted molecular mass of 22 kDa (ORF2, positions 3051 to 3662). The putative transposase and resolvase of Tn*3434* are similar to the corresponding proteins of Tn*5393* (63 and 77% identity and 77 and 88% similarity, respectively). The two ORFs are separated by a putative re-

combination site (*res*), which is involved in Tn*3*-like elements in cointegrate resolution and regulation of expression of the *tnpA* and *tnpR* genes (23).

The 38-bp IRs of Tn*3434* (with one mismatch) were of the length typically found in Tn*3*-type transposons (23). In contrast, IRs of Tn*5393* were much longer (81 bp with four mismatches), but their termini were homologous to those of Tn*3434* and other members of Tn*3* family (Fig. 7B).
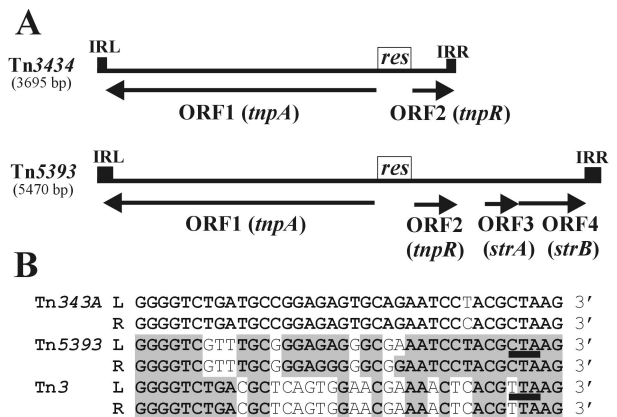

FIG. 7. (A) Genetic maps of Tn*3434* and Tn*5393*. The structural elements of these transposons (*tnpA*, transposase; *tnpR*, resolvase; *res*, recombination site; *strA* and *strB*, genes for aminoglycoside phosphotransferases) are indicated. The direction of transcription of the genes is indicated by arrows. (B) Comparison of the IR sequences of Tn*3434*, Tn*5393*, and Tn*3*. (The IRs of Tn*5393* are truncated, since they are 43 bp longer than those in other Tn*3* family members.) Nucleotides conserved in IRs of Tn*3434* are boldface. Nucleotides of related sequences identical to those of Tn*3434* are boldface and shaded. Stop codons of Tnps are underlined. "L" denotes sequences at the 5′ (left) end, and "R" denotes complementary sequences at the 3′ (right) end of the elements.
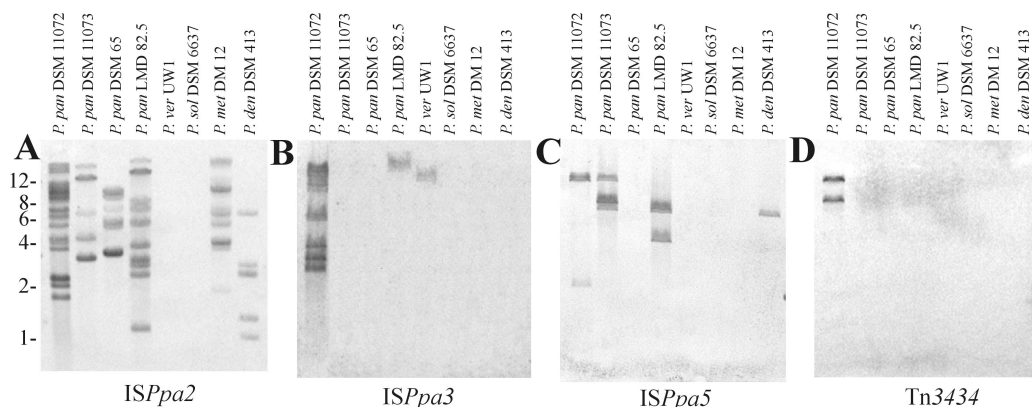
FIG. 8. Southern blot hybridization analysis of total DNAs from paracoccal strains. Samples of digested DNAs of *P. pantotrophus* (*P. pan*), *P. versutus* (*P. ver*), *P. solventivorans* (*P. sol*), *P. methylutens* (*P. met*), and *P. denitrificans* (*P. den*) were analyzed on agarose gel, transferred to nylon membrane, and hybridized with the appropriate probe, based on the transposable elements studied (marked below each panel). The total DNA was digested with *Pst*I (A), *Eco*RI-*Pst*I (B and C), and *Hin*dIII-*Bam*HI (D and E). Size markers (kilobases) are indicated on the left.

Distribution, copy number, and localization of the analyzed transposable elements. To investigate whether the identified transposable elements are specific only for the host strains or are widely spread among various *P. pantotrophus* isolates (or other paracoccal species), hybridization analysis was performed. For this purpose, PCR-amplified internal fragments of the elements (nucleotide sequences of the primers are given in Materials and Methods) were probed against total DNAs of *P. pantotrophus* strains and four other representatives of the genus *Paracoccus*: *P. versutus* UW1, *P. denitrificans* DSM 413, *P. solventivorans* DSM 6637, and *P. methylutens* DM12 (Fig. 8). The restriction enzymes for DNA digestion were chosen individually for each of the elements tested to avoid multiple hybridization signals derived from a single copy of a given IS. The number of hybridized DNA fragments was therefore equivalent to the minimum number of copies of a given element within the genome.

We detected a single copy of IS*Ppa1* only in one strain of *P. pantotrophus*: DSM 11072 (hybridization with one *Xba*I-*Bgl*II fragment of 9 kb) (data not shown). On the other hand, IS*Ppa2* is present in multiple copies in all strains of *P. pantotrophus* studied as well as in *P. methylutens* DM12 and *P. denitrificans* DSM 413 (Fig. 8A). The apparent copy number ranged from 5 to 14. The hybridization patterns of the strains tested were different. Some IS copies displayed different signal intensities, which reflects the presence of other closely related elements homologous to IS*Ppa2* (e.g., IS*1248*). At least seven copies of IS*Ppa3* could be detected in *P. pantotrophus* DSM 11072, and one copy could be detected in strain LMD 82.5 as well as in *P. versutus* UW1 (Fig. 8B). The related element IS*Ppa4* was shown to be present in one copy in parental DSM 65 (hybridization with one 11-kb *Pst*I fragment) (data not shown). In IS*Ppa5*, positive signals were detected in three strains of *P. pantotrophus* (two copies in the host strain DSM 11072, three copies in DSM 11073, and two copies in LMD 82.5) as well as one copy in *P. denitrificans* DSM 413 (Fig. 8C). The cryptic Tn Tn*3434* is present in two copies in parental strain DSM 11072 (Fig. 8D), while streptomycin-resistant Tn*5393* resides in a single copy only in LMD 82.5 (as judged from hybridization of the *strAB*-specific probe with a 6-kb *Bam*HI-*Hin*dIII restriction

fragment of the total DNA of this strain) (data not shown). This is in agreement with our observation that of all paracoccal strains tested, only LMD 82.5 was streptomycin-resistant (data not shown).

In order to localize the transposable elements within the hosts' genomes (plasmid or chromosome), additional hybridization analysis against DNAs of pHG16-a, pKLW1, pPAN1, pTAV1, pSOS1, and pMTH4 (harbored by the paracoccal strains tested) was performed. We found that pHG16-a of DSM 65 carries single copies of IS*Ppa2* (Fig. 9B, lane 2) and IS*Ppa4* (Fig. 9B, lane 3), while pKLW1 of DSM 11073 carries two copies of IS*Ppa2* (Fig. 9B, lane 1). The other elements were not present in plasmids occurring in these strains (data not shown).

DISCUSSION

We have identified and characterized several novel transposable elements residing in different strains of a soil bacterium, *P. pantotrophus* (Table 3). These are the first elements to
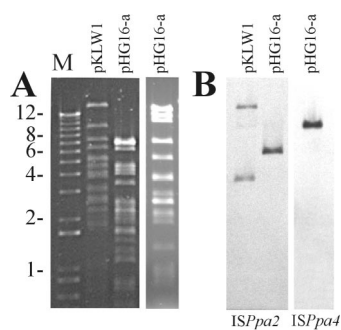

FIG. 9. DNA electrophoresis (A) and Southern blot hybridization analysis (B) of plasmids pKLW1 and pHG16-a of *P. pantotrophus* strains. Samples of plasmid DNAs were analyzed on an agarose gel, transferred to nylon membrane, and hybridized with the probe based on IS*Ppa2* or IS*Ppa4* (marked below the appropriate lanes). Plasmid DNAs were digested with *Eco*RI and *Pst*I (hybridization with IS*Ppa2*) or *Pst*I (hybridization with IS*Ppa4*). Size markers (kilobases) are indicated to the left of the gel (lane M).

be identified in this species and in the entire genus *Paracoccus* (with the exception of IS*1248* of *P. denitrificans*) (53). All of the elements have been identified by their integration into a selective cartridge of a novel vector, pMEC1, which proves that they are functional in transposition. These ISs and Tns may thus serve as a basis for the construction of novel tools for transposon mutagenesis, which may be very useful for the genetic analysis of this interesting group of bacteria.

In general, we can state that the transposable elements are abundant in this species and they represent four distinct families: IS*256* (IS*Ppa1*), IS*5* (IS*Ppa2*, IS*Ppa3*, and IS*Ppa4*), IS*66* (IS*Ppa5*), and Tn*3* (Tn*3434*, Tn*5393*). The genetic organization of IS*Ppa1*, IS*Ppa3*, and IS*Ppa4* is typical for the majority of known ISs. They encode a single ORF for a transposase protein with a family-specific DDE motif. The structure of IS*Ppa2* (IS*247* group of IS*5* family) differs, since it encodes two overlapping ORFs as well as the conserved frameshift motif, which is likely to promote the generation of a fusion protein (ORF1 plus ORF2) as a result of programmed translational frameshifting (9). The putative fusion protein of IS*Ppa2* encodes the transposase-specific DDE motif, which strongly suggests that ribosomal frameshifting may participate in the regulation of the transposition of this element. So far, the production of a transframe protein has been reported exclusively for the members of the IS*1* (18) and IS*3* families (26, 55).

A completely different mechanism of regulation of the expression of transposase genes seems to function in IS*Ppa5*. This element encodes two sets of slightly overlapping ORFs: ORF1 and ORF2 (analogous overlap was observed in all members of IS*66* family) and ORF3 and ORF4. IS*Ppa5* (and its relatives) does not encode the frameshifting motif, which precludes the possibility of generating transframe fusion transposases (9). It is thus probable that ORF1 and ORF2 of all these elements (as well as ORF3 and ORF4 of IS*Ppa5*) may be produced by a translational coupling mechanism (28). ORF3 of IS*Ppa5* (although it carries a putative DDE motif), unlike other bacterial transposases, has a relatively low isoelectric point (pI 6.75) and does not encode an HTH DNA binding motif. Interestingly, ORF1 (pI 8.77) encodes an HTH motif (data not shown), while ORF2 (the most conserved protein of IS*66*-family) encodes a basic protein (pI 10.41). We speculate that the proteins encoded by IS*66*-like elements may form multiprotein complexes with transposase activity. This is in agreement with the results of mutational analysis of IS*679* (identified in plasmid pB171 of enteropathogenic *E. coli* B171) (28), which proved that all three IS*679*-encoded proteins (homologous to ORF1 to ORF3 of IS*Ppa5*) are necessary for transposition. ORF4 of IS*Ppa5* is not conserved in other members of the IS*66* family: its role needs to be investigated.

Besides the ISs, we identified in *P. pantotrophus* two Tns, which are members of the Tn*3* family. These are a novel cryptic Tn designated Tn*3434* and Tn*5393* (homologous to the former one), which carry the streptomycin resistance module. Previous analysis demonstrated that Tn*5393* is widespread among many gram-negative bacteria isolated from agricultural habitats where streptomycin was used (34). Our observations of the high rate of transposition of Tn*5393* into the entrapment vector (Table 2) may explain the reason for its wide dissemination in many environmental isolates (11, 49). This points to the importance of soil bacteria as reservoirs of antibiotic resistance determinants in the environment.

In order to study the distribution and copy number of the identified elements, we performed hybridization analysis. It should be kept in mind that such analysis does not necessarily imply that all copies of a particular IS (or Tn) element detected are identical and functional. The only strain-specific transposable elements were IS*Ppa1*, Tn*3434* of *P. pantotrophus* DSM 11072, and Tn*5393* of LMD 82.5, as well as IS*Ppa4* of DSM 65 (residing on plasmid pHG16-a), a finding that suggests they might have been acquired by recent lateral transfer events.

Among the elements tested, IS*Ppa2* (closely related to IS*1248*) appeared to be the most widely distributed in paracocci. Since IS*Ppa2* and IS*1248* cross-hybridize, we were unable to distinguish these elements by hybridization (they are 87% identical at the nucleotide sequence level). It has been previously shown that IS*1248* is present in multiple copies in all strains of *P. denitrificans* tested (53). The presence of IS*Ppa2*-like sequences in plasmids of DSM 65 (pHG16-a) and DSM 11073 (pKLW1) suggests the possibility of their dissemination by lateral transfer. Members of the IS*5* family preferentially integrate within the YTAR sequence (Y = C or T; R = A or G) (often 5′-CTAG-3′). In the case of IS*Ppa2*, the sequence was slightly divergent: RTAR (5′-GTAA-3′). It is probable that preferential insertion of these ISs within the putative stop codons (TAG and TAA) may help to minimize the frequency of transpositional disruption of DNA coding regions, thus making possible the potentially broader dissemination of these elements (35).

IS*Ppa3* (IS*903* group of IS*5* family) and IS*Ppa5* homologues are unevenly distributed in the strains tested. Interestingly, hybridization and partial sequencing revealed that the only element "caught" by the entrapment vector in DSM 11073 was IS*Ppa5*. The reason for the unusually high frequency of transposition of this element in this strain (10^3 higher than in DSM 11072; Table 2) remains unclear. It cannot be excluded that its transposition might be stimulated by a host-encoded factor or factors or DNA sequence or sequences adjacent to the insertion site (e.g., transcriptional activation of IS*Ppa5*-encoded genes by a foreign promoter or promoters). We have initiated studies aimed at revealing the cause of this phenomenon. The high sequence similarities of IS*Ppa5* to the elements residing in chromosome and plasmid pSymA of *Sinorhizobium meliloti* suggest that the exchange of genetic material between members of the family *Rhizobiaceae* and *Paracoccus* spp. occurs frequently. This assumption can be supported by the results of our previous studies, in which we have shown that several paracoccal plasmids carry *repABC* replicons, homologous to that of pSymA and many other rhizobial megaplasmids (7).

The majority of the identified transposable elements seem to be of chromosomal origin. It is worth mentioning, however, that all of the paracoccal strains tested carry megaplasmids (2), which cannot be purified by the standard alkaline lysis procedure. The question of whether the transposable elements identified reside in one of these high-molecular-weight plasmids remains open.

The results presented in this paper are a preliminary part of our complex studies on the distribution and diversity of transposable elements of all species of the genus *Paracoccus*. We believe that the results obtained will serve to draw conclusions

of a more generalized nature regarding the frequency and directions of lateral transfer in this group of soil bacteria.

ACKNOWLEDGMENTS

REFERENCES

1. **Ajdič, D., W. M. McShan, R. E. McLaughlin, G. Savic, J. Chang, M. B. Carson, C. Primeaux, R. Tian, S. Kenton, H. Jia, S. Lin, Y. Qian, S. Li, H. Zhu, F. Najar, H. Lai, J. White, B. A. Roe, and J. J. Ferretti.** 2002. Genome sequence of *Streptococcus mutans* UA159, a cariogenic dental pathogen. Proc. Natl. Acad. Sci. USA **99:**14434–14439.
2. **Baj, J., E. Piechucka, D. Bartosik, and M. Wlodarczyk.** 2000. Plasmid occurrence and diversity in the genus *Paracoccus*. Acta Microbiol. Pol. **49:**265–270.
3. **Baker, S. C., S. J. Ferguson, B. Ludwig, M. D. Page, O.-M. Richter, and R. J. M. van Spanning.** 1998. Molecular genetics of the genus *Paracoccus*: metabolically versatile bacteria with bioenergetic flexibility. Microbiol. Mol. Biol. Rev. **62:**1046–1078.
4. **Bartosik, D., A. Bialkowska, J. Baj, and M. Wlodarczyk.** 1997. Construction of mobilizable cloning vectors derived from pBGS18 and their application for analysis of replicator region of a pTAV202 mini-derivative of *Paracoccus versutus* pTAV1 plasmid. Acta Microbiol. Polon. **46:**379–383.
5. **Bartosik, D., J. Baj, M. Sochacka, E. Piechucka, and M. Wlodarczyk.** 2002. Molecular characterization of functional modules of plasmid pWKS1 of *Paracoccus pantotrophus* DSM 11072. Microbiology **148:**2847–2856.
6. **Bartosik, D., J. Baj, and M. Wlodarczyk.** 1998. Molecular and functional analysis of pTAV320—*repABC* type replicon of a composite pTAV1 plasmid of *Paracoccus versutus*. Microbiology **144:**3149–3157.
7. **Bartosik, D., J. Baj, E. Piechucka, E. Waker, and M. Wlodarczyk.** 2002. Comparative characterization of paracoccal type *repABC* replicons. Plasmid **48:**130–141.
8. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acid Res. **7:**1513–1519.
9. **Chandler, M., and O. Fayet.** 1993. Translational frameshifting in the control of transposition in bacteria. Mol. Microbiol. **7:**497–503.
10. **Chandler, M., and J. Mahillon.** 2002. Insertion sequences revised, p. 305–366. *In* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. American Society for Microbiology, Washington D.C.
11. **Chiou, C.-S., and A. L. Jones.** 1993. Nucleotide sequence analysis of a transposon (Tn*5393*) carrying streptomycin resistance genes in *Erwinia amylovora* and other gram-negative bacteria. J. Bacteriol. **175:**732–740.
12. **Cloeckaert, A., M. Grayon, and O. Grepinet.** 2000. An IS*711* element downstream of the *bp26* gene is a specific marker of *Brucella* spp. isolated from marine mammals. Clin. Diagn. Lab. Immunol. **7:**835–839.
13. **Denome, S. A., and K. D. Young.** 1995. Identification and activity of two insertion sequence elements in *Rhodococcus* sp. strain IGTS8. Gene **161:**33–38.
14. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **11:**387–395.
15. **Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski.** 1980. Broad host-range DNA cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA **77:**7347–7351.
16. **Doronina, N. V., Y. A. Trotsenko, V. I. Krausowa, and N. E. Suzina.** 1998. *Paracoccus methylutens* sp. nov.—a new aerobic facultatively methylotrophic bacterium utilizing dichloromethane. Syst. Appl. Microbiol. **21:**230–236.
17. **Egert, M., A. Hamann, R. Komen, and C. D. Friedrich.** 1993. Methanol and methylamine utilization results from mutational events in *Thiosphaera pantotropha*. Arch. Microbiol. **159:**364–371.
18. **Escoubas, J. M., M. F. Prere, O. Fayet, I. Salvignol, D. Galas, D. Zerbib, and M. Chandler.** 1991. Translational control of transposition activity of the bacterial insertion sequence IS*1*. EMBO J. **10:**705–712.
19. **Galas, D. J., and M. Chandler.** 1989. Bacterial insertion sequences, p. 109–162. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington D.C.
20. **Galibert, F., T. M. Finan, S. R. Long, A. Puhler, P. Abola, F. Ampe, F. Barloy-Hubler, M. J. Barnett, A. Becker, P. Boistard, G. Bothe, M. Boutry, L. Bowser, J. Buhrmester, E. Cadieu, D. Capela, P. Chain, A. Cowie, R. W. Davis, S. Dreano, N. A. Federspiel, R. F. Fisher, S. Gloux, T. Godrie, A. Goffeau, B. Golding, J. Gouzy, M. Gurjal, I. Hernandez-Lucas, A. Hong, L. Huizar, R. W. Hyman, T. Jones, D. Kahn, M. L. Kahn, S. Kalman, D. H. Keating, E. Kiss, C. Komp, V. Lelaure, D. Masuy, C. Palm, M. C. Peck, T. M. Pohl, D. Portetelle, B. Purnelle, U. Ramsperger, R. Surzycki, P. Thebault, M. Vandenbol, F. J. Vorholter, S. Weidner, D. H. Wells, K. Wong, K. C. Yeh, and J. Batut.** 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science **293:**668–672.

21. **Gamas, P., M. G. Chandler, P. Prentki, and D. J. Galas.** 1987. *Escherichia coli* integration host factor binds specifically to the ends of the insertion sequence IS*1* and to its major insertion hot-spot in pBR322. J. Mol. Biol. **195:**261–272.
22. **Gerstenberg, C., B. Friedrich, and H. G. Schlegel.** 1982. Physical evidence for plasmids in autotrophic, especially hydrogen-oxidising bacteria. Arch. Microbiol. **133:**90–96.
23. **Grindley, N. D. F.** 2002. The movement of Tn*3*-like elements: transposition and cointegrate resolution, p. 272–302. *In* N. L. Craig, R. Craigie, M. Gellert, and A. M. Lambowitz (ed.), Mobile DNA II. American Society for Microbiology, Washington, D.C.
24. **Guedon, G., F. Bourgoin, M. Pebay, Y. Roussel, C. Colmin, J. M. Simonet, and B. Decaris.** 1995. Characterization and distribution of two insertion sequences, IS*1191* and iso-IS*981*, in *Streptococcus thermophilus*: does intergeneric transfer of insertion sequences occur in lactic acid bacteria cocultures? Mol. Microbiol. **16:**69–78.
25. **Guilhot, C., B. Gicquel, J. Davies, and C. Martin.** 1992. Isolation and analysis of IS*6120*, a new insertion sequence from *Mycobacterium smegmatis*. Mol. Microbiol. **6:**107–113.
26. **Haas, M., and B. Rak.** 2002. *Escherichia coli* insertion sequence IS*150*: transposition via circular and linear intermediates. J. Bacteriol. **184:**5833–5841.
27. **Hammerschmidt, S., R. Hilse, J. P. van Putten, R. Gerardy-Schahn, A. Unkmeir, and M. Frosch.** 1996. Modulation of cell surface sialic acid expression in *Neisseria meningitidis* via a transposable genetic element. EMBO J. **15:**192–198.
28. **Han, C.-G., Y. Shiga, T. Tobe, C. Sasakawa, and E. Ohtsubo.** 2001. Structural and functional characterization of IS*679* and IS66-family elements. J. Bacteriol. **183:**4296–4304.
29. **Hübner, A., and W. Hendrickson.** 1997. A fusion promoter created by a new insertion sequence, IS*1490*, activates transcription of 2,4,5-trichlorophenoxyacetic acid catabolic genes in *Burkholderia cepacia* AC1100. J. Bacteriol. **179:**2717–2723.
30. **Isawa, T., R. Sameshima, H. Mitsui, and K. Minamisawa.** 1999. IS*1631* occurrence in *Bradyrhizobium japonicum* highly reiterated sequence-possessing strains with high copy numbers of repeated sequences RSα and RSβ. Appl. Environ. Microbiol. **65:**3493–3501.
31. **Jordan, S. L., I. R. McDonald, A. J. Kraczkiewicz-Dowjat, D. P. Kelly, F. A. Rainey, J. C. Murrell, and A. P. Wood.** 1997. Autotrophic growth on carbon disulfide is a property of novel strains of *Paracoccus denitrificans*. Arch. Microbiol. **168:**225–236.
32. **Kelly, D. P., F. A. Rainey, and A. P. Wood.** 2000. The genus *Paracoccus*. *In* M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (ed.), The prokaryotes, 3rd ed. Springer Verlag, New York, N.Y. [Online.] http://www.springer-ny.com.
33. **Kushner, S. R.** 1978. An improved method for transformation of *E. coli* with ColE1-derived plasmids, p. 17–23. *In* H. B. Boyer and S. Nicosia (ed.), Genetic engineering. Elsevier/North-Holland, Amsterdam, The Netherlands.
34. **L'Abée-Lund, T. M., and H. Sørum.** 2000. Functional Tn*5393*-like transposon in the R plasmid pRAS2 from the fish pathogen *Aeromonas salmonicida* subspecies *salmonicida* isolated in Norway. Appl. Environ. Microbiol. **66:**5533–5535.
35. **Lee, Y.-A., S.-C. Fan, L.-Y. Chiu, and K.-C. Hsia.** 2001. Isolation of an insertion sequence from *Ralstonia solanacearum* race 1 and its potential use for strain characterization and detection. Appl. Environ. Microbiol. **67:**3943–3950.
36. **Lessie, T. G., W. Hendrickson, B. D. Manning, and R. Devereux.** 1996. Genomic complexity and plasticity of *Burkholderia cepacia*. FEMS Microbiol. Lett. **144:**117–128.
37. **Loessner, I., K. Dietrich, D. Dittrich, J. Hacker, and W. Ziebuhr.** 2002. Transposase-dependent formation of circular IS*256* derivatives in *Staphylococcus epidermidis* and *Staphylococcus aureus*. J. Bacteriol. **184:**4709–4714.
38. **Mahillon, J., and M. Chandler.** 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. **62:**725–774.
39. **Paulus, F., J. Canaday, F. Vincent, G. Bonnard, C. Kares, and L. Otten.** 1991. Sequence of the *iaa* and *ipt* region of different *Agrobacterium tumefaciens* biotype III octopine strains: reconstruction of octopine Ti plasmid evolution. Plant Mol. Biol. **16:**601–614.
40. **Providenti, M. A., and R. C. Wyndham.** 2001. Identification and functional characterization of CbaR, a MarR-like modulator of the *cbaABC*-encoded chlorobenzoate catabolism pathway. Appl. Environ. Microbiol. **67:**3530–3541.
41. **Rainey, F. A., D. P. Kelly, E. Stackebrandt, J. Burghardt, A. Hiraishi, Y. Katayama, and A. P. Wood.** 1999. A re-evaluation of the taxonomy of *Paracoccus denitrificans* and a proposal for the combination *Paracoccus pantotrophus* comb. nov. Int. J. Syst. Bacteriol. **49:**645–651.
42. **Rudant, E., P. Courvalin, and T. Lambert.** 1997. Loss of intrinsic aminoglycoside resistance in *Acinetobacter haemolyticus* as a result of three distinct types of alterations in the *aac*(6′)-*Ig* gene, including insertion of IS*17*. Antimicrob. Agents Chemother. **41:**2646–2651.
43. **Sambrook, J., and D. W. Russel.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

44. **Sangari, F. J., M. Bachli, L. E. Bermudez, and T. Bodmer.** 2000. Characterization of IS*666*, a newly described insertion element of *Mycobacterium avium*. Microb. Comp. Genomics **5:**181–188.

45. **Schneider, D., D. Faure, M. Noirclerc-Savoye, A. C. Barrière, E. Coursange, and M. Blot.** 2000. A broad-host-range plasmid for isolating mobile genetic elements in gram-negative bacteria. Plasmid **44:**201–207.

46. **Schneiker, S., B. Kosier, A. Puhler, and W. Selbitschka.** 1999. The *Sinorhizobium meliloti* insertion sequence (IS) element IS*Rm14* is related to a previously unrecognized IS element located adjacent to the *Escherichia coli* locus of enterocyte effacement (LEE) pathogenicity island. Curr. Microbiol. **39:**274–281.

47. **Siller, H., F. A. Rainey, E. Stackebrandt, and J. Winter.** 1996. Isolation and characterization of a new gram-negative, acetone-degrading, nitrate-reducing bacterium from soil, *Paracoccus solventivorans* sp. nov. Int. J. Syst. Bacteriol. **46:**1125–1130.

48. **Sullivan, J. T., J. R. Trzebiatowski, R. W. Cruickshank, J. Gouzy, S. D. Brown, R. M. Elliot, D. J. Fleetwood, N. G. McCallum, U. Rossbach, G. S. Stuart, J. E. Weaver, R. J. Webby, F. J. De Bruijn, and C. W. Ronson.** 2002. Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. J. Bacteriol. **184:**3086–3095.

49. **Sundin, G. W., and C. L. Bender.** 1995. Expression of the *strA-strB* streptomycin resistance genes in *Pseudomonas syringae* and *Xanthomonas campestris*

and characterization of IS*6100* in *X. campestris*. Appl. Environ. Microbiol. **61:**2891–2897.

50. **Taylor, B. F., and D. S. Hoare.** 1969. New facultative *Thiobacillus* and a reevaluation of the heterotrophic potential of *Thiobacillus novellus*. J. Bacteriol. **100:**487–497.

51. **Tolmasky, M. E., and J. H. Crosa.** 1995. Iron transport genes of the pJM1-mediated iron uptake system of *Vibrio anguillarum* are included in a transposonlike structure. Plasmid **33:**180–190.

52. **Van Spanning, R. J. M., A. P. De Boer, D. J. Slotboom, W. N. Reijnders, and A. H. Stouthamer.** 1995. Isolation and characterization of a novel insertion sequence element, IS*1248*, in *Paracoccus denitrificans*. Plasmid **34:**11–21.

53. **Van Spanning, R. J. M., W. N. M. Reijnders, and A. H. Stouthamer.** 1995. Integration of heterologous DNA into the genome of *Paracoccus denitrificans* is mediated by a family of IS*1248*-related elements and a second type of integrative recombination event. J. Bacteriol. **177:**4772–4778.

54. **Yates, J. R., and D. S. Holmes.** 1987. Two families of repeated DNA sequences in *Thiobacillus ferrooxidans*. J. Bacteriol. **169:**1861–1870.

55. **Zheng, J., and M. A. McIntosh.** 1995. Characterization of IS*1221* from *Mycoplasma hyorhinis*: expression of its putative transposase in *Escherichia coli* incorporates a ribosomal frameshift mechanism. Mol. Microbiol. **16:**669–685.