**Tn5047**: a chimeric mercury resistance transposon closely related to the toluene degradative transposon Tn4657


*Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 1231 82, Russia*

1 Present address: The Rockefeller University, New York, NY 10021, USA.
2 Present address: Microcide Pharmaceuticals, 850 Maude Ave, Mountain View, CA 94043, USA.

**Abstract**

This paper reports the discovery and characterization of Tn5047, a novel-type transposon vehicle for dissemination of mercury resistance in natural bacterial populations. Tn5047 (14876 bp), identified in a *Pseudomonas* strain from a mercury mine, is a Tn3 family mercury resistance transposon far outside the Tn21 subgroup. As in other Tn3 family transposons, Tn5047 duplicates 5 bp of the target sequence following insertion. Tn5047 apparently acquired its mer operon as a single-ended relic of a transposon belonging to the classical mercury resistance transposons of the Tn27 subgroup. The putative transposase and the 47 bp terminal inverted repeats of Tn5047 are closely related to those of the toluene degradative transposon Tn4651 and fall into a distinct subgroup on the fringe of the Tn3 family. The amino acid sequence of the putative resolvase of Tn5047 resembles site-specific recombinases of the integrase family. Besides the mer operon and putative transposition genes, Tn5047 contains a 4 kb region that accommodates a number of apparently defective genes and mobile elements.

**Keywords**: Tn5047, mer operon, Tn3 family, integrase, transposase

**INTRODUCTION**

Mercury resistance in bacteria is a widespread phenotype. As a rule, it is determined by several genes organized into a mercury resistance (mer) operon. These are often located on plasmids and transposons (Silver & Walderhaug, 1992; Misra, 1992). The first well-characterized mercury resistance transposons, Tn501 and Tn21, were from clinical isolates. Tn501 was isolated from a *Pseudomonas aeruginosa* strain from Australia (Stanisich et al., 1977). Tn21 was found in the plasmid NR1 in *Shigella flexneri* isolated in Japan (Womble & Rownd, 1988; Grinsted et al., 1990). Besides the mer operon and transposition genes Tn21 contains an integron carrying the antibiotic resistance genes sul and str (Brown et al., 1996). A plethora of transposons closely related to Tn21 and carrying different sets of antibiotic resistance genes has been detected in clinical isolates (for review see Grinsted et al., 1990). Tn3926, which is homologous to both Tn501 and Tn21, was identified in *Yersinia enterocolitica* isolated from raw milk in France (Lett et al., 1985). Tn21, Tn3926 and Tn501 define a compact group within the Tn3 family of transposons. They are bound by 38 bp inverted repeats (IRs) and contain two genes, *tnpR* and *tnpA*, which are involved in transposition (Grinsted et al., 1990; Osbourn et al., 1995).

Several mercury resistance transposons from environmental bacteria have been described less thoroughly. Tn1861 from *Pseudomonas putida* has a restriction pattern similar to Tn501 (Frielio & Chakrabarty, 1980). The apparently identical Tn3402 and Tn3403 (Radford et al., 1981) have a characteristic pattern of EcoRI and HindIII sites in the transposition genes, similar to that of Tn21 (A. J. Radford, cited by Stanisich et al., 1989). Thus, until recently, all mercury resistance transposons described from Gram-negative bacteria were apparently related to Tn21 and Tn501 with the possible exception...
of the poorly characterized Tn502 (Stanisich et al., 1989; Brown et al., 1996).

The first 'novel-type' mer transposon to be characterized was Tn5053 (Kholodii et al., 1993). Tn5053 apparently acquired its mer operon, merRTPFAD, as a single-ended relic of another transposon which is closely related to the classical mercury resistance transposons Tn21 and Tn501. The transposition module of Tn5053 is quite different from transposition modules of classical transposons of the Tn21 subgroup, but is closely related to a family of transposons carrying integrons (Kholodii et al., 1995; Brown et al., 1996). Tn5053 is bounded by 25 bp terminal inverted repeats and contains four genes involved in transposition, tniA, tniB, tniQ and tniR. Transposition of Tn5053 occurs via cointegrate formation mediated by the products of the tniABQ genes, followed by site-specific cointegrate resolution. This is catalysed by the product of the tniR gene at the res region, which is located upstream of tniR. Sequence analysis suggests that Tn5053 fell into a recently recognized superfamly of transposons including retroviruses, Tn7, and insertion sequences of the IS3 family (Kholodii et al., 1995).

In this communication we describe the isolation and complete nucleotide sequence of another 'novel-type' mercury resistance transposon, Tn5041, which was identified in a Pseudomonas strain from a mercury mine. As is the case with Tn5053, Tn5041 apparently acquired its mer operon from a transposon belonging to the Tn21 subgroup. The transposition module of Tn5041 belongs to the Tn3 family but is far outside the Tn21 subgroup.

METHODS

Bacterial strains and plasmids. The following bacterial strains and plasmids were used: Pseudomonas spp. KHP41 and KHP22 (Bogdanova et al., 1988); Escherichia coli K-12 strains AB1456 (a recipient in transposition experiments); JM83 (a host strain for pUC19-based recombinant plasmids); HB101 (a host for mer operon containing plasmids in mercury resistance testing); RP1 (Pansegrau et al., 1994); RP1:: Tn5041 (this paper): R831b (merB') (Schootel et al., 1974); pVS982 = pBR322:: Tn501 (merB) (Grimsted et al., 1982). pWWO containing Tn4651 (Tsuda et al., 1989) was obtained from Dr K. N. Timmis.

Mercury resistance testing. Susceptibility to HgCl₂ and phenylmercuric acetate (PMA) was determined by the disk diffusion method. LB agar (1% w/v) was overlaid with 5 ml 0.5% LB agar containing 0.1 ml late-exponential-phase bacterial culture of HB101 with or without a plasmid. Filter paper disks (6 mm diameter) loaded with 5-40 µg HgCl₂ or 2-20 µg PMA were placed on the plates. The growth inhibition zones were measured after incubation at 30 °C for 20 h. All the assays were repeated at least three times. With 10 µg PMA, inhibition zones were about 5 mm for PMA-sensitive strains (HB101 without plasmids or with pVS982, or RP1, or RP1:: Tn5041) and about 1.5 mm for the PMA-resistant strain (HB101 bearing R831b). Inhibition zones for HgCl₂ (at 20 µg) were about 6 mm for HB101 without plasmids or with RP1 and about 2 mm for HB101 with any other plasmid mentioned above.

Detection of mercury resistance transposons. This was done by introducing the broad-host-range plasmid RP1 into spontaneous rifampicin-resistant mutants of mercury-resistant Pseudomonas/Xanthomonas strains by conjugation from a rifampicin-sensitive E. coli donor strain. The matings were performed on solid LB medium. LB agar supplemented with kanamycin (50 µg ml⁻¹) and rifampicin (30 µg ml⁻¹) was used for selecting of RP1-containing transconjugants. After 2-3 serial transfers on selective medium containing kanamycin (50 µg ml⁻¹) and HgCl₂ (4 µg ml⁻¹) the transconjugants were mated with a streptomycin-resistant recipient E. coli strain AB1456 (str⁻) and RP1-Hg-r recombinant plasmid-containing clones were isolated on LB agar supplemented with HgCl₂ (4-5 µg ml⁻¹) and streptomycin (100 µg ml⁻¹).

RESULTS AND DISCUSSION

Detection of Tn5041 and related elements

For detection of mercury resistance transposons, we used a collection of mercury-resistant Pseudomonas/Xanthomonas strains isolated from a mercury mine in Kirgizia, Central Asia (Khesin & Karasyova, 1984; Bogdanova et al., 1988). In most of these strains, mer operons were located on 100 MDa narrow-host-range plasmids that were readily transferred by conjugation to their mercury-sensitive plasmid-cured derivatives but not to any Pseudomonas or E. coli strains tested (Khesin & Karasyova, 1984; Zh. Gorlenko & S. Z. Mindlin, unpublished). We surveyed this collection by introducing the broad-host-range plasmid RP1 and mating out to E. coli as described in Methods. As in the case of Gram-negative mercury-resistant strains described by Kelly & Reaney (1984), some strains from our collection (12 out of 23 tested) were refractory to introduction of RP1. For some strains, this may be due to incompatibility or surface exclusion by resident plasmids because plasmids of the same incompatibility group as RP1 can be found in some environmental bacteria (Amuthan & Mahadevan, 1994) and narrow-host-range plasmids showing incompatibility with RP1 are known (Smith & Thomas, 1987). For one strain (KHP22) tested, we demonstrated that the resistance to RP1 was not due
to the resident plasmid since a mercury-sensitive derivative of KHP22 which was cured of its single 100 MDa plasmid (Khesin & Karasova, 1984) was still refractory to introduction of RP1.

Four out of eleven RP1-accepting strains gave mercury-resistant transconjugants with frequencies about $10^{-5}$ per RP1-containing transconjugant. One of these strains contained a novel mercury resistance transposon, Tn5053, which was described previously by Kholodii et al. (1993, 1995). Restriction analysis of plasmid DNA of mercury-resistant transconjugants obtained from Pseudomonas sp. strain KHP41 revealed insertions of about 15 kb in a segment of RP1 located between oriV and Tn1. Sequencing of the junction between RP1 and the inserted element in three independent transconjugants revealed insertions of the same sequence after positions (numbering according to Pansegrau et al., 1994) 11019 of RP1 (the beginning of the klcB gene), 12025 and 12212 (a non-coding region at the border of oriV). Insertions of the element, Tn5041, resulted in duplications of 5 bp of the target sequence (GTATT, CTTCG, and GGAC, respectively).

Most mercury resistance determinants with a few possible exceptions (Witte et al., 1986; Mahler et al., 1986; Wang et al., 1987; Shiratori et al., 1989; Kholodii et al., 1993) were found on plasmids. A 4-2 kb plasmid, which was too small to accommodate Tn5041, and no large plasmids were found in strain KHP41 using both an alkaline extraction procedure (Birnboim & Doly, 1979) and agarose gel electrophoresis of crude extracts by the method described by Eckhardt (1978). Chromosomal DNA preparations from KHP41 showed hybridization with Tn5041-specific DNA probes while plasmid DNA preparations did not (data not shown).

Thus, Tn5041 was apparently localized chromosomally. As in other reported cases of apparent chromosomal location of mercury resistance determinants, the possibility was not excluded that Tn5041 was actually located on a megaplasmid which often go undetected by common methods (Barton et al., 1995). DNA elements showing hybridization with Tn5041 transposition genes were found in 100 MDa plasmids readily detected in three other Pseudomonas strains isolated from the same mine as KHP41 (see Methods). We failed to check transposability of these elements since their host strains were refractory to introduction of RP1. We demonstrated that about 10% of environmental Pseudomonas mercury-resistant strains collected in other parts of the world contained DNA elements showing hybridization with Tn5041 transposition genes. One of these strains contained a functional transposon similar to Tn5041 while other were either resistant to RP1 or transposition deficient. We found no elements related to Tn5041 in our collection of Acinetobacter and enterobacterial strains.

**Tn5041 belongs to the Tn3 family**

We determined the complete nucleotide sequence of Tn5041 on both strands as described in Methods. Results of sequence analysis are summarized in Fig. 1. Tn5041 is bound by 47 bp imperfect inverted repeats which are closely related to IRs of the toluene degradative transposon Tn4651 (Tsuda et al., 1989). The 38 outermost base pairs of these IRs are relics of mobile elements. Solid rectangles represent segments encoding amino acid sequences showing similarity to segments of a rat membrane protein located in the same order (Hallberg et al., 1993). The sequence of a putative res site is shown above the 141 bp spacer located between orfI and orfQ (bases forming a palindrome are underlined, those shared with the dif site are boxed). Restriction sites: B, BglII; E, EcoRI; EV, EcoRV; H, HindIII; N, Ncol; Nd, NdeI; P, PstI; Pv, PvuII; S, SacI; X, XbaI; Xh, XhoI.

---

**Fig. 1.** Genetic organization of Tn5041. The terminal 47 bp IRs are shown by solid arrowheads. The location and polarity of genes and ORFs is shown by open arrows. T, P, C and D indicate the corresponding mer genes, and 1 and 2 indicate homologues of urf-1 and urf-2, respectively. P denotes orfP encoding a porin-like protein. R' is a truncated copy of merR located in the same orientation as the functional copy of merR, α and γ are relics of mobile elements. Solid rectangles represent segments encoding amino acid sequences showing similarity to segments of a rat membrane protein located in the same order (Hallberg et al., 1993). The sequence of a putative res site is shown above the 141 bp spacer located between orfI and orfQ (bases forming a palindrome are underlined, those shared with the dif site are boxed). Restriction sites: B, BglII; E, EcoRI; EV, EcoRV; H, HindIII; N, Ncol; Nd, NdeI; P, PstI; Pv, PvuII; S, SacI; X, XbaI; Xh, Xhol.
G. YA. KHOLODII and OTHERS

Fig. 2. Comparison of terminal KY family transposons. Letters following names of transposons are showing locations of X03406; Tn1721, X61367. Other accession nos are listed in the same (It' indicates the et subgroup (including Tn22, Tn3926, Tn501 and Tn2722) (Jones, 1993) and Tn263 (Ulrich posases clustered around three previously recognized subgroups. The first cluster united a compact Tn22 transposases which we found in the databases. Analysis to find the exact position of Tn.5042 in the Tn3 family transposons from Gram-positive bacteria only: Tn917 (An & Clewell, 1991), Tn5422 (Lebrun et al., 1994), Tn1412 (accession no. L36547), Tn5401 (Baum, 1994) and Tn4556 (Siemiennick et al., 1990). Only three transposases were definitely outside the aggregate of the above clusters even when only the more conserved COOH-terminal part of the transposase protein were taken into account. These are a solitary transposase of IS1071 (Nakatsu et al., 1991) and a pair of closely related transposases of Tn5041 and Tn4651 (Fig. 3). Thus, sequence analysis of Tn5041 and Tn4651 transposases puts them on the fringe of the Tn3 family. Yet, Tn5041 and Tn4651 are genuine Tn3 family transposons. Their transposases contain most of the invariant amino acid residues found in other transposases of the Tn3 family including a triad of acidic residues, D-D-E, which are candidates for catalytic residues (Yurieva & Nikiforov, 1996).

Tn5041 encodes a protein showing homology to site-specific recombinases of the integrase family

Typically, transposons of the Tn3 family encode enzymes of the resolvase family which are used for resolution of cointegrates formed during transposition. The exceptions include IS1071 which lacks resolvase (Nakatsu et al., 1991) and Tn4430 and Tn5401, both

toluene degradative transposon Tn4651 (EMBL accession no. X83686; Dr R. Horak). Accordingly, we suggest that this ORF which we designated tnpA (Fig. 1) encodes the Tn5041 transposase.

To find the exact position of Tn5041 in the Tn3 family we aligned its putative transposase with 20 Tn3 family transposases which we found in the databases. Analysis of the alignment demonstrated that most of the transposases clustered around three previously recognized subgroups. The first cluster united a compact Tn21 subgroup (including Tn21, Tn3926, Tn501 and Tn1721) (Grinsted et al., 1990) with the more distantly related Tn2501 (Turner & Grinsted, 1987). Tn5393 (Chiu & Jones, 1993) and Tn163 (Urich & Puehler, 1994) can be also affiliated to this cluster on the basis of analysis of the more conserved COOH-terminal part of the transposase protein. The second cluster contained transposons from Gram-positive bacteria only: Tn917 (An & Clewell, 1991), Tn5422 (Lebrun et al., 1994), Tn1546 (Arthur et al., 1993), Tn4430 (Mahillon & Lereclus, 1988), and a putative transposon from Bacillus subtilis (Ivey et al., 1992). The third cluster included Tn3 and Tn1000 (Grinsted et al., 1990), Tn5396 (Elhai et al., 1994), Tn1412 (accession no. L36547), Tn5401 (Baum, 1994) and Tn4556 (Siemiennick et al., 1990). Only three transposases were definitely outside the aggregate of the above clusters even when only the more conserved COOH-terminal part of the transposase protein were taken into account. These are a solitary transposase of IS1071 (Nakatsu et al., 1991) and a pair of closely related transposases of Tn5041 and Tn4651 (Fig. 3). Thus, sequence analysis of Tn5041 and Tn4651 transposases puts them on the fringe of the Tn3 family. Yet, Tn5041 and Tn4651 are genuine Tn3 family transposons. Their transposases contain most of the invariant amino acid residues found in other transposases of the Tn3 family including a triad of acidic residues, D-D-E, which are candidates for catalytic residues (Yurieva & Nikiforov, 1996).

Tn5041 encodes a protein showing homology to site-specific recombinases of the integrase family

Typically, transposons of the Tn3 family encode enzymes of the resolvase family which are used for resolution of cointegrates formed during transposition. The exceptions include IS1071 which lacks resolvase (Nakatsu et al., 1991) and Tn4430 and Tn5401, both

![Fig. 2. Comparison of terminal IRs of (a) Tn5041 and (b) of an ky element located within Tn5041 with selected IRs of the Tn3 family transposons. Letters following names of transposons are showing locations of IRs in those cases when the IRs are not the same ('t' indicates the tnpA end and 'o' the other end). Dots represent bases identical with those in the upper sequence. Accession nos (or references): Tn5041, X83686; Tn4651 (Tsuda et al., 1988); avrB, M21965; Tn5396, U04360 and U04362; mksA, X16098; 'res501' (located between resII and tnpR in Tn501), X03406; Tn1721, X61367. Other accession nos are listed in the legend to Fig. 3.](image)

![Fig. 3. Evolutionary tree of Tn3 family transposons based on comparison of the COOH-terminal-third of their transposases. Accession nos: Tn21, X04891; Tn3926, X14236; Tn501, X03406; Tn2501, Y00502; Tn5393, M96392; Tn5401, X75779; Tn163, L14931; Tn917, M11180; Tn5422, L28104; Tn1546, M97297; Tn4430, X07651; Tn3, V00613; Tn1000, X60200; Tn1412, L36547; Tn5401, U03554; Tn4556, M29297; IS1071, M65135; Tn4651, X83686; Tn5041, X88999.](image)
from Bacillus thuringiensis, which use integrase-like proteins as resolvases (Mahillon & Lereclus, 1988; Baum, 1994, 1995). Database search revealed no Tn5041 genes encoding resolvase/invertase family proteins but that orfl, one of the two divergent ORFs located at the left arm of Tn5041, encoded a protein showing homology to phage P1 Cre and E. coli XerD proteins. These proteins are well known representatives of a large family of integrases involved in various site-specific recombination processes (Nash, 1996). Although the members of the integrase family exhibit a large diversity in their sequences, they share four invariant amino acid residues, the RHRY tetrad, in two characteristic regions, named domain 1 and domain 2 (Argos et al., 1986; Abremiski & Hoess, 1992). The tetrat is believed to be involved in catalysis (Chen et al., 1992). The predicted Orfl protein of Tn5041 shows 28–34% similarity to its nearest neighbours (including phage, chromosomal, integron and transposon integrases) in the two conserved domains (Fig. 4); similarity to more distant members of the family is about 20%. Tn5041 Orfl contains all of the RHRY tetrat residues in the proper positions (Fig. 4). On this basis we suggest that orfl encodes a functional protein involved in recombination, most probably in the resolution of the covalent linkage of the conserved Tyr residue with the 3' ends flanking the overlap region (for review see Nash, 1996). Inspection of the sequence upstream of orfl revealed a 34 bp non-perfect palindromic sequence as a candidate functional core of the Tn5041 resolution site (Fig. 1). This putative site has a striking similarity to the E. coli chromosomal dif site. This suggests that the Orfl-mediated strand cleavage/exchange occur 3' to the ATAA dyad portions (shared with dif and boxed in Fig. 1) separated by 6 bp central region. It should be noted that dif is operated by two proteins, XerD and XerC, showing 37% overall sequence identity (Blakely et al., 1993), but no such closely related partner of Orfl is encoded by Tn5041. Besides the functional core, the recombination locus may include extra binding sites for the recombinase and/or binding sites for accessory proteins. This is believed to provide additional asymmetry to recombinational loci to ensure directionality of the recombination process (Sherratt et al., 1995). For example, the Tn5401 locus involved in cointegrate resolution contains two extra binding sites for a transposon-encoded integrase-like Tnp1 protein. These extra sites are represented by direct repeats of single copies of the half-sites of the functional core (Baum, 1995). No extra core half-site copies or obvious binding sites for accessory proteins such as those found in ColEl cer (Stirling et al., 1988) were found around the putative resolution site of Tn5041. Further characterization of the putative resolution locus of Tn5041 requires a biochemical approach.

**Fig. 4.** Integrase-like domains (Argos et al., 1986; Abremiski & Hoess, 1992) of a putative resolvase of Tn5041 encoded by orfl. Orfl is aligned with its nearest neighbours from the integrase family in the order of closest similarity over the sequences shown. Dots represent bases identical with those in the upper sequence. Dashes represent bases absent in other sequences. Invariant amino acid residues which are believed to be involved in catalysis are marked by asterisks. Accession nos: Cre P1, X03453; XerD E. coli, M54884; XerC E. coli, P22885; Int of Integrons, J02967; TnpA Tn554, X03216; TnpI Tn4430, X07651.

<table>
<thead>
<tr>
<th>Domain 1</th>
<th>Domain 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orfl Tn5041</td>
<td>GFAFRESLIYELMDY</td>
</tr>
<tr>
<td>Cre P1</td>
<td>TLLAIAIAEYRVR</td>
</tr>
<tr>
<td>XerD E. coli</td>
<td>TLUV...G.T.BD</td>
</tr>
<tr>
<td>XerC E. coli</td>
<td>AGL.L...G.DIHR</td>
</tr>
<tr>
<td>Int integron</td>
<td>TPII...GLQ.RYED</td>
</tr>
<tr>
<td>TnpA Tn584</td>
<td>GLIG.I.VLS.RLED</td>
</tr>
<tr>
<td>TnpI Tn4430</td>
<td>TOV.I...ALSIK.ND</td>
</tr>
</tbody>
</table>

The sites of action of site-specific recombinases are often located upstream of their genes and comprise a functional core composed of a segment, called the overlap region, flanked by inverted repeats. For different members of the integrase family, the overlap region is 6–8 bp long while the repeats are 10–12 bp long. The invertases recognize the inverted half-sites and cleave and exchange DNA strands via the transient formation of a covalent linkage of the conserved Tyr residue with the 3' ends flanking the overlap region (for review see Nash, 1996). Operon of Tn5041 is 75% sequence identity showing 70% overall sequence identity (Blakely et al., 1993), but no such closely related partner of Orfl is encoded by Tn5041. Besides the functional core, the recombination locus may include extra binding sites for the recombinase and/or binding sites for accessory proteins. This is believed to provide additional asymmetry to recombinational loci to ensure directionality of the recombination process (Sherratt et al., 1995). For example, the Tn5401 locus involved in cointegrate resolution contains two extra binding sites for a transposon-encoded integrase-like Tnp1 protein. These extra sites are represented by direct repeats of single copies of the half-sites of the functional core (Baum, 1995). No extra core half-site copies or obvious binding sites for accessory proteins such as those found in ColEl cer (Stirling et al., 1988) were found around the putative resolution site of Tn5041. Further characterization of the putative resolution locus of Tn5041 requires a biochemical approach.

Tn5041 contains a functional mer operon showing 70% sequence similarity to the mer operons of Tn21 and Tn501

The central part of Tn5041 is occupied by a well-known mer operon (Fig. 1) showing 70% similarity to mer operons of classical mercury resistance transposons Tn21 and Tn501. The mer operon of Tn5041 is preceded by a 24 bp element showing 75% sequence similarity to the outermost part of Irs typical for Tn21-like transposons. As in Tn21-like transposons, the
genes and mobile elements. One such element Tn5041 contains a 4 kb region with relics of mobile elements and shortened genes into the family of IRs of the Tn3 family (Fig. 2b). We have identified previously unnoticed solitary 38 bp elements and shortened genes of Tn5041 does not confer organomercurial resistance (for details see Methods). Thus the function of an ORF, orfY, in this position, some mer operons contain a merB gene which encodes organomercurial lyase that confers resistance to organomercurials (Misra, et al., 1990). However orfY shows no homology to known merB (Reniero et al., 1995) or any other genes. Moreover, our experiments have demonstrated that the mer operon of Tn5041 does not confer organomercurial resistance (for details see Methods). Thus the function of orfY remains obscure. This is then followed by merD (which is thought to be involved in the fine regulation of mer operons) and two ORFs, urf-1 and urf-2, of unknown function. The latter contains a large deletion compared with Tn501.

Tn5041 contains a 4 kb region with relics of mobile elements and shortened genes

The 4 kb region between orfQ and the mer operon of Tn5041 contains traces of multiple rearrangements resulting in formation of apparently non-functional genes and mobile elements. One such element (κY) (Fig. 1) is bound by 38 bp non-perfect IRs which neatly fall into the family of IRs of the Tn3 family (Fig. 2b). We noticed that IRs of κY are closely related to the 38 bp sequence (res501 in Fig. 2b) which is located between the end of the resIII site and the start codon of trpR in Tn501 and Tn1721 but is absent in Tn21 and Tn3926. By searching nucleotide sequence databases we have identified previously unnoticed solitary 38 bp elements closely related to IRs of κY in the vicinity of the aubB gene of Pseudomonas syringae (Tamaki et al., 1988) and mkaD (mkfA) gene found in virulence plasmids of numerous Salmonella strains (Taira & Rhen, 1990; Matsui et al., 1990). Another defective element (ωλ in Fig. 1) is a truncated copy of an IS element showing 60-66% nucleotide sequence similarity to IS2 and 59-4% similarity to ISRm1 (Watson & Wheatcroft, 1991). A truncated copy of merR (R' in Fig. 1) showing 94-3% nucleotide sequence similarity to the functional copy merR of Tn5041 is located between ωλ and κY (Fig. 1). Traces of two membrane protein genes can be pinpointed in the 4 kb region. One of these is represented by an ORF, orfP, with a reasonable Shine-Dalgarno sequence (Fig. 1). The predicted translation product shows about 30% sequence similarity to a central segment of P. aeruginosa porins D2 and E1. No specific functions have been ascribed to this segment in the porin molecule (Yamano et al., 1993). Relics of another putative membrane protein gene are represented by four short reading frames. Segments of proteins translated from these frames show similarity to different segments of a rat integral membrane protein described by Hallberg et al. (1993) (Fig. 1). These speculations can be confirmed only by finding more closely related functional bacterial porins.

Tn5041 is a chimera composed of several modules that have different origins

In conclusion, our sequence analysis demonstrates that both the mer operon module and the transposition module of Tn5041 show relationships to known mobile elements. The mer operon module of Tn5041 displays 70% nucleotide sequence similarity to the mer operons of the well-known Tn21-like transposons and carries 3 to the stop codon of its merR a truncated copy of IR typical of these transposons. The transposase gene and IRs of Tn5041 are very similar to those of a toluene degradative transposon, Tn4651, but are very distantly related to those of Tn21-like transposons. Accordingly, we consider Tn5041 as a chimera which obtained its mer operon as a result of insertion of a transposon belonging to Tn21 subgroup into a transposon related to Tn4651 followed by deletion of primary (Tn21-type) trans-
position genes. Similar insertion–deletion events occurred during the genesis of quite different mercury resistance transposons, Tn5053 (Kholodii et al., 1993, 1995) and Tn501 (Grinsted & Brown, 1984). It seems that relays of transposition modules from the primary Tn21-like transposons to various other types is a common feature of the lifestyle of mer operons. Besides the mer operon and putative transposition genes, Tn5041 contains a 4 kb region of unknown origin and significance which accommodates a number of apparently defective mobile elements. This suggests that Tn5041-like transposition modules were used as vehicles by numerous passenger genes other than mercury resistance or tolune degradative operons.

ACKNOWLEDGEMENTS

This work was supported in part by the Russian State Program Frontiers in Genetics, and National Institutes of Health FIRCA Grant TW00396-02. We thank Dr R. Novick for support, Dr N. Brown and two anonymous referees for comments and helpful suggestions, and Drs K. N. Timmis and J. Grinsted for gifts of plasmids.

REFERENCES


Watson, R. & Wheatcroft, R. (1991). Nucleotide sequence of Rhizobium meliloti insertion sequence ISRm1: homology to IS2 from Escherichia coli and IS426 from Agrobacterium tumefaciens. DNA Seq 2, 163–172.


Received 20 November 1996; revised 10 March 1997; accepted 15 April 1997.