Complete nucleotide sequence of the self-transmissible TOL plasmid pD2RT provides new insight into arrangement of toluene catabolic plasmids

Jekaterina Jutkina a,*, Lars Hestbjerg Hansen b, Lili Li b, Eeva Heinaru a, Eve Vedler a, Merike Jõesaar a, Ain Heinaru a

a Department of Genetics, Institute of Molecular and Cell Biology, University of Tartu, Riia 23a, 51010 Tartu, Estonia
b Section of Microbiology, Institute of Biology, University of Copenhagen, Universitetsparken 15-1, 2100 Copenhagen OE, Denmark

ARTICLE INFO
Article history:
Received 19 March 2013
Accepted 20 September 2013
Available online 2 October 2013
Communicated by Georg Lipps

Keywords:
Pseudomonas migulae
Baltic Sea bacteria
TOL plasmid
Conjugal transfer
Toluene catabolic genes

ABSTRACT
In the present study we report the complete nucleotide sequence of the toluene catabolic plasmid pD2RT of Pseudomonas migulae strain D2RT isolated from Baltic Sea water. The pD2RT is 129,894 base pairs in size with an average G+C content of 53.75%. A total of 135 open reading frames (ORFs) were predicted to encode proteins, among them genes for catabolism of toluene, plasmid replication, maintenance and conjugative transfer. ORFs encoding proteins with putative functions in stress response, transposition and site-specific recombination were also predicted. Analysis of the organization and nucleotide sequence of pD2RT backbone region revealed high degree of similarity to the draft genome sequence data of the plant-pathogenic pseudomonad Pseudomonas syringae pv. glycinea strain B076, exhibiting relatedness to pPT23A plasmid family. The pD2RT backbone is also closely related to that of pGRT1 of Pseudomonas putida strain DOT-T1E and pBVIE04 of Burkholderia vietnamiensis strain G4, both plasmids are associated with resistance to toluene. The ability of pD2RT to self-transfer by conjugation to P. putida recipient strain PaW340 was experimentally determined. Genetic organization of toluene-degrading (xyl) genes and flanking DNA segments resembles the structure of Tn1721-related class II transposon Tn4656 of TOL plasmid pWW53 of P. putida strain MT53. The complete sequence of the plasmid pD2RT extends the known range of xyl genes carriers, being the first completely sequenced TOL plasmid, which is not related to well-studied IncP plasmid groups. We also verified the functionality of the catabolic route encoded by pD2RT by monitoring the expression of the xyle gene in pD2RT bearing hosts along with bacterial strains containing TOL plasmid of IncP-9 group. The growth kinetics of plasmid-bearing strains was found to be affected by particular TOL plasmid.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

A number of different toluene catabolic pathways have been elucidated, located either on chromosomes or on plasmids (Parales et al., 2008). The best characterized toluene-degrading plasmid is pWW0 (117 kb), which catabolic genes are organized in two clusters: the upper operon (xylUWCMABN), which enables to convert toluene to benzoate and the lower/meta operon (xylXYZLTEGFJQKIH), which codes for further conversion of benzoate and m-tolu-uate to the citric acid cycle intermediates, being under the control of the positive regulator genes xylR and xylS (Greated et al., 2002). Toluene catabolic plasmids that carry xyl genes highly similar to those on pWW0 are referred to as TOL plasmids. Numerous TOL plasmids that vary in size,
organization of catabolic area and the replicon type have been isolated from geographically distant locations (Chatfield and Williams, 1986; Sentchilo et al., 2000). Though, the complete sequences are available only for a few of them, namely the self-transmissible plasmid pWW0 (Greated et al., 2002) of the incompatibility group P-9, the non-self-transmissible plasmid pWW53 (Yano et al., 2007) and the self-transmissible plasmid pDK1 (Yano et al., 2010), the latter two belong to the IncP-7 plasmid family. Although several TOL plasmids of other yet unknown incompatibility groups have been isolated (Sentchilo et al., 2000), they have not been studied in such details as the TOL plasmids of the well-known IncP families.

A unique type of plasmid-encoded toluene-degradation pathway, with the key enzyme toluene ortho-monooxygenase (Tom), was found to be encoded by the plasmid DNA of Burkholderia cepacia strain G4 (Shields et al., 1995). This plasmid was originally called TOM, and later, during the release of the whole genome sequence data of the strain G4 (BioProject: PRJNA10696, CP000620), it was designated as pBVIE04. The TOM (pBVIE04) plasmid exhibits different genetic organization compared to the well-known TOL plasmids. Recently, a non-catabolic plasmid pGRT1 from P. putida strain DOT-T1E, related to pBVIE04, was sequenced and characterized by Molina and colleagues (Molina et al., 2011). Plasmid pGRT1 enables its host to tolerate toluene due to the efflux pump transport proteins, encoded by the tggHI operon (Molina et al., 2011). These two plasmids are distant from all known IncP families and have not been assigned to any incompatibility group.

In the present study we report the complete nucleotide sequence of the self-transmissible TOL plasmid pD2RT of Pseudomonas migulae strain D2RT isolated from Baltic Sea water. Presence of the conserved xyl gene clusters assigns pD2RT to the TOL plasmid family; however, its backbone reveals relatedness to the plasmids pBVIE04 and pGRT1. Thus, the plasmid pD2RT exhibits a novel combination of previously determined features, strongly associated with bacterial survival in the presence of toluene. At the same time, the catabolic region, which is closely related to that of pD2RT, was also found to be located on the IncP-9 TOL plasmid pD67 of P. migulae strain D67, revealing the presence of different xyl genes carriers in the bacterial community of Baltic Sea surface water.

2. Materials and methods

2.1. Bacterial strains and culture conditions

Isolation of m-toluate and toluene-degrading bacterial strains used in this study (Table 1) from Baltic Sea water and characterization of their biodegradative capacities have been previously described (Jutkina et al., 2011).

The bacterial strains were grown in liquid or solid minimal media containing M9 salts (Blair, 1959) and trace elements (Bauchop and Elsden, 1960) supplemented with m-toluate (2.5 mM), or in Luria–Bertani (LB) broth at 30 °C. Liquid cultures were grown on a rotary shaker (180 rpm). The ability of all studied strains to degrade toluene was verified by growing isolates on minimal medium in the presence of toluene vapors.

2.2. Plasmid DNA manipulation techniques and sequencing

The complete nucleotide sequence of the plasmid pD2RT was determined by using both the 454 pyrosequencing approach and the basic Sanger method.

Plasmid DNA, library preparation and subsequent pyrosequencing was applied as described previously by Burmolle and co-workers (Burmolle et al., 2012) with modifications for the XL+ sequencing platform according to manufacturer’s protocol (Roche diagnostics). In brief, plasmid DNA was purified using Plasmid Mini AX Kit (A&A Biotechnology), amplified to sufficient yield with Repli-g Kit (Qiagen) and sequenced on the GS FLX XL+ high throughput platform (Roche diagnostics).

In addition, plasmid DNA was purified with Plasmid Maxi Kit (Qiagen) according to manufacturer’s instructions. Then digested with restrictases (EcoRI, HindIII, PstI), cloned into appropriate pBluescript (Stratagene) sites and transformed into competent Escherichia coli DH5α cells (Inoue et al., 1990). Colonies were selected on LB agar media containing 150 mg ml⁻¹ ampicillin, 0.48 mg ml⁻¹ IPTG and 0.08 mg ml⁻¹ X-Gal. Selected clones were sequenced on a 3730xl DNA Analyzer (Applied Biosystems) using BigDye® Terminator 3.1. Cycle Sequencing Kit (Applied Biosystems) and universal M13 forward and reverse primers following the manufacturer’s instructions.

The nucleotide sequence of the catabolic region of the plasmid pD67 was determined based on the similarity with that of the plasmid pD2RT. Amplified products were sequenced by the Sanger approach as mentioned above.

2.3. Plasmid genome assembly, annotation and comparative analysis

The complete circular sequence of the plasmid pD2RT was derived from combining 3 contigs (74,456, 52,473 and 1364 bp) of 454 sequencing data, which were assembled with NEWbler 2.6 (Roche diagnostics), with sequences obtained from the Sanger approach and

<table>
<thead>
<tr>
<th>Table 1 The Baltic Sea water isolates used in this study.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacterial strain</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Pseudomonas migulae D2RT</td>
</tr>
<tr>
<td>Pseudomonas migulae D67</td>
</tr>
<tr>
<td>Pseudomonas stutzeri 2A54</td>
</tr>
<tr>
<td>Pseudomonas stutzeri 2A20</td>
</tr>
<tr>
<td>Pseudomonas sp. 2B49</td>
</tr>
<tr>
<td>Pseudomonas stutzeri B10v</td>
</tr>
<tr>
<td>Pseudomonas stutzeri 2D47</td>
</tr>
<tr>
<td>Pseudomonas stutzeri 2D49</td>
</tr>
<tr>
<td>Pseudomonas putida 2D61</td>
</tr>
</tbody>
</table>
assembled into contigs using BioEdit version 7.0.9.0 (Hall, 1999). Open reading frames were predicted using Glimmer 3.02 (Delcher et al., 1999) and Prodigal (Hyatt et al., 2010). To assign putative functions to all the predicted ORFs they were aligned to the nr, Swissprot, Pfam, KEGG, COG, Smart and Prk databases using BLAST and RPS-BLAST search tools (Altschul et al., 1997). ORFs that had a truncated sequence were identified as gene remnants.

The potential IHF-binding site was predicted by using E. coli binding sites matrix (Gama-Castro et al., 2011) and the matrix–scan program provided by Regulatory System Analysis Tools (http://embnet.ccg.unam.mx/rsa-tools/).

Comparative analysis of the plasmid pD2RT sequence was done with WebACT (Abbott et al., 2005) and visualized with ACT (Carver et al., 2005) software with default parameters.

2.4. Detection and multiple alignment analysis of xylA gene sequences

Since the xylA gene of pD2RT shares a low level of identity with the homologs from database (Table S1), the universal set of primers was created for the amplification of the xylA gene fragment. The primers xylA_F 5’-CCRYGATGCTGCTGCTAGC-3’ and xylA_R 5’-CATGCCMCGYTTYCCTTTCC-3’ were designed based on the alignment of the xylA gene sequences of the plasmids pD2RT and pWW0. No mismatches were allowed.

Amplification of the xylA fragment was performed in a volume of 25 μl containing 1× PCR buffer provided in the Fermentas TaqDNA polymerase Kit (with (NH₄)₂SO₄), 0.2 mM of each deoxyribonucleoside triphosphate, 2.5 mM MgCl₂, 0.4 μM of each primer, 0.5 U of TaqDNA polymerase (Fermentas). 5 μl of total DNA obtained by heating cell suspension at 95 °C for 15 min was added as a template. PCR amplification conditions were as following: initial denaturation at 95 °C for 5 min, then 35 cycles of 95 °C for 30 s, 58 °C for 30 s, and 72 °C for 30 s followed by final extension step at 72 °C for 5 min. Sequencing of the PCR products was done using the Sanger approach.

Multiple alignment of the obtained xylA gene sequences was done using CLUSTALX 2.0 software (Thompson et al., 2002). For multiple alignment additional xylA gene sequences were obtained from database: pSVS11 AF251324, pSVS15 AF251327, pWW0 AJ344068, pWW53 AB238971.

2.5. GenBank accession numbers

The complete nucleotide sequences of the plasmid pD2RT and the catabolic region of pD67 were deposited in GenBank under the accession numbers JX891462 and JX891461, respectively. The 203 bp long xylA sequences of the strains 2B49, 2D61, 2A20, B10v, 2D49, 2A54 and 2D47 were deposited under accession numbers KC255405–KC255412.

2.6. Assays of catechol 2,3-dioxygenase activity and growth kinetics

Bacterial strains were pregrown overnight in 25 ml liquid minimal medium supplemented with glucose. Fresh minimal medium supplemented with either glucose or glucose and m-toluate (2.5 mM) as an inducer was inoculated with 5 ml of pre-grown culture. After inoculation the optical density at 580 nm for all tested cultures was around 0.1. 4 h after inoculation 10 ml of the cultures were harvested by centrifugation and washed twice with phosphate buffer (50 mM, pH 7.5). Bacterial cells were re-suspended in the same buffer and sonically disrupted. Cell debris was removed by centrifugation 12,000 g for 30 min at 4 °C. Catechol 2,3-dioxygenase (C23O) enzyme activities were measured spectrophotometrically by monitoring the formation of reaction products at 375 nm (Hegeman, 1966). Protein concentrations were measured by the method of Bradford (Bradford, 1976) with bovine serum albumin as a standard. Similar conditions were applied for C23O activity measurements in strains grown on m-toluate as sole carbon and energy source. For transconjugant P. putida strain PaW340 (Franklin and Williams, 1980) tryptophan (40 g l⁻¹) was added to the growth medium.

The growth curves were assayed spectrophotometrically at 580 nm in a Tecan Infinite M200 microtiter plate reader using 96 well plate (final volume of 150 μl per well). In the latter case direct serial measurements were recorded every 15 min for 13 h in octicate. All experiments were run in triplicate.

2.7. Conjugal transfer test

For conjugation experiments donor and recipient (P. putida PaW340; m-Tol₁, Sm², Trp₁) cells were pre-grown over-night in 2 ml LB broth. Two volumes of donor culture and one volume of recipient culture in logarithmic phase of growth were mixed together and inoculated on LB agar. After incubation at 30 °C for 20 h the entire cell mass was suspended in 2 ml sterile saline (0.85% NaCl), then harvested by centrifugation, washed twice and re-suspended in 1 ml of sterile saline, from which serial dilutions were made.

The donor and transconjugant cells were enumerated by plating to solid minimal media supplemented with m-toluate, while in the latter case tryptophan (40 g l⁻¹) was also added to the growth medium. The frequency of conjugation was determined as the ratio of transconjugants (CFU ml⁻¹) to donors (CFU ml⁻¹).

2.8. Mutagenesis of the putative repA genes

The bacterial strains and plasmids used and constructed in the repA mutagenesis experiment are described in Table 2.

To disrupt the repA1 and repA2 genes in P. migulae strain D2RT, 1.9 kb DNA region containing repA1 and 2.3 kb DNA region containing repA2 were amplified by PCR from genomic DNA of the studied strain with corresponding pairs of primers: 70F1 (5’-GGATCTACAGATTGTGCTCTC-3’) – 70R2 (5’-GAATTGCTCTGCTCAGGCAC-3’) and 72F1 (5’-GGATCCGACGAACTCCCTCA-3’) – 72R2 (5’-GAATTCGATATA TCGCCGACACAC-3’). The amplified products were cloned into pTZ57R/T vector, resulting in repA1/pTZ57R and repA2/pTZ57R constructs. Then the repA1 region was excised from repA1/pTZ57R with Bst1107I and CpoI enzymes and repA2, in turn, was excised from repA2/pTZ57R using.
Table 2

Bacterial strains and plasmids used in mutagenesis experiment.

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Genotype or construction</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>supE44 ΔlacU169 (qφ80 lacZ ΔM15) recA1 endA1 hsdR17 thi-1 gyrA96 relA1</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>HB101</td>
<td>subE44 subF58 hsdR30 (r~ m~ ) recA13 ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 met-l</td>
<td>(Boyer and Roulland-Dussoix, 1969)</td>
</tr>
<tr>
<td>CC118::pir</td>
<td>Δ(ara-leu) araD ΔlacX74 galE galK pheA20 thi-1 rpsE rpoB argE (Am) recA1 pir phage lysogen</td>
<td>(Herrero et al., 1990)</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTZ57R/T</td>
<td>Cloning vector (Ap')</td>
<td>MBI Fermentas</td>
</tr>
<tr>
<td>pUTmini-Tn5 Km2</td>
<td>Delivery plasmid for mini-Tn5 Km2 (Ap' Km')</td>
<td>(de Lorenzo et al., 1990)</td>
</tr>
<tr>
<td>pGP704 L</td>
<td>Delivery plasmid for homologous recombination (Ap')</td>
<td>(Pavel et al., 1994)</td>
</tr>
<tr>
<td>pRK2013</td>
<td>Helper plasmid for conjugal transfer of pGP704 L (Km')</td>
<td>(Figurski and Helsinki, 1979)</td>
</tr>
</tbody>
</table>

Km' and Ap', resistance to kanamycin and ampicillin, respectively.

Ndel and Esp3l. Around 1 kb large fragments of both genes, repA1 and repA2, were deleted and replaced with the Km' gene, which was previously amplified by PCR from the plasmid pUTmini-Tn5 Km2 by using the primer KmSac (5'-CAGGAGCTCTTGTTATTATTTATCAACAAAGCC-3') (Hörak et al., 2004) and subsequently cleaved with Ecl136II. Before ligation of the Km' gene fragments into the opened constructs the Bst1107I-Cpnl and Ndel-Esp3l ends of the constructs were blunt ended. From obtained constructs, pTZ57RAREpA1::km and pTZ57RAREpA2::km, were excised with following enzymes, Sall-Sacl and Sall-KpnI, respectively, and inserted into plasmid vector pGP704, resulting in the formation of constructs pGP704AREpA1::km and pGP704AREpA2::km. These constructs were further transferred from E. coli CC118::pir (Herrero et al., 1990) into P. migulae strain D2RT by means of conjugation using the helper plasmid pRK2013 (Figurski and Helsinki, 1979).

Transconjugants were selected on plates supplemented either with kanamycin (50 μg mL⁻¹), carbenicillin (3000 μg mL⁻¹), or both 2.5 mM salicylate (the ability of bacterial strain P. migulae D2RT to use salicylate as sole carbon and energy source has been show previously by us (Jutkina et al., 2011)) and kanamycin (50 μg mL⁻¹). D2RTRepA1- transconjugants were verified by PCR analysis using the primers KmOc (5′-TGGGACGAGCTGATTCCC-3′) (Saumaa et al., 2006) and 70F1 and by growing on plates with 2.5 mM m-toluene. Conjugation experiments were conducted in duplicate.

3. Results and discussion

3.1. General features of the plasmid pD2RT

The plasmid pD2RT is 129,894 base pairs in size with an average G+C content of 53.75%. A total of 135 open reading frames (ORFs) were predicted to encode proteins. Among them, 21 were predicted to be genes for toluene catabolism, 39 for plasmid replication, maintenance and conjugative transfer, 14 for transposition and site-specific recombination. The remaining ORFs encode proteins with putative functions in stress response and proteins with other known and unknown functions. The relevant data regarding all predicted ORFs including their exact location, protein length, GC content, putative function and the closest homologs from databases is revealed in Table S1. The circular map of pD2RT is depicted in Fig. 1.

The overall genetic organization of the plasmid pD2RT indicates two clearly distinguishable regions: one encompasses the plasmid backbone while the other so called accessory part of the plasmid, is responsible for enhancing host cell catabolic properties (74,363–129,853 bp). The later region of pD2RT containing toluene-degrading (xyl) genes and their flanking DNA segments shows remarkable similarity to Tn1721-related class II transposon Tn4656 of plasmid pWW53 (AB062597) (Tsuda and Genka, 2001). The xyl genes are organized as four transcriptional units: the upper pathway operon (xylUWCMABN), the meta pathway operon (xylXYZLTEGFJQKIH) and the two transcriptional regulatory genes, xylS and xylR. Transposon-related genes trpR and trpA and the resolution site res were found downstream of the upper pathway gene cluster towards the right inverted repeat IR-R. Two identical inverted copies of IS1492b element are found to be inserted in the terminal repeats of this transposon, resulting in the formation of a large composite IS-transposon. Besides the xyl genes the putative transposon comprises genetic determinants for environmental fitness of host bacterial strain, plasmid stability within the cell and genes encoding transcriptional regulators. ORF97 encodes a putative oxidative stress-induced OsmC-like protein (PF02566), which suggested role is to detoxify organic peroxides (Dubbs and Mongkolsuk,
ORF103 is predicted to encode a FAD-dependent D-lactate dehydrogenase that catalyzes the oxidation of D-lactate to pyruvate. The ability to utilize D-lactate could serve as carbon and energy source or contribute to a pH reduction (Siezen et al., 2005). Two additional accessory genes, ruvA and ruvB, that have been incorporated into the plasmid backbone (44,883–46,584) in close proximity to the partition-associated region presumably confer resistance to ultraviolet light, thereby enhancing the fitness of P. migulae strain D2RT.

The main structure of the pD2RT backbone is represented by several gene clusters: trbC-trbN-DNA.topoisomerase III (5226–13,942), traK-mobE (31,178–38,546), parBA (47,152–49,321), repA1-repA2 (51,932–55,315) and traM-traY (56,293–69,228) that are separated from each other by a range of ORFs with yet unknown functions, and single ORFs with functional assignment like methylase and zeta toxin encoding genes (Fig. 1).

Analysis of the plasmid pD2RT architecture and gene content revealed the high degree of similarity to the Contig065 of plant-pathogenic multi-plasmid bacterial strain Pseudomonas syringae pv. glycinea B076 (Pseudomonas savastanoi pv. glycinea B076) draft genome, which is presumably a plasmid-derived sequence (Qi et al., 2011). The majority of ORFs exhibit more than 90% identity between corresponding gene products and high conservation of their location and transcriptional orientation in relation to each other (Fig. 2a). The second closest homolog was found to be the backbone region of the plasmid pGRT1 of P. putida strain DOT-T1E, although the extent of sequence similarity is considerably lower compared to that of Contig065. Backbones of these two plasmids share the main structural components and also display similar organization of particular gene clusters within the genome (Fig. 2b), however the identity level between single gene products is mostly in the range of 50–80%. The remarkable
difference in their backbone genetic content is the presence of kfrA gene in the plasmid pGRT1 and its absence in pD2RT, which can be associated with plasmid partitioning. The plasmid pD2RT is also distantly related to the plasmid pBVIE04 of B. vietnamiensis strain G4 exhibiting the presence of common genes that are required for plasmid stability and replication as well as gene clusters that are crucial for conjugal transfer, although sequence similarity between corresponding proteins is low and generally restricted to conserved protein domains. Comparison of pD2RT and pBVIE04 genomes brings out certain plasticity of the common backbone structure, since the conserved gene clusters have different order and orientation within particular genomes (Fig. 2c). Schematic representation summarizing the comparative analysis of Contig065, pD2RT, pGRT1 and pBVIE04 is represented on Fig. 3. To conclude, from the performed analysis it could be deduced that the new TOL plasmid pD2RT is likely formed by incorporation of the Tn4656-like transposon into the plasmid vehicle which shared the last common ancestor with a plasmid of a phytopathogenic pseudomonad, and the formation of plasmids pD2RT and pBVIE04 occurred through independent acquisition of both accessory and backbone-associated gene clusters by different carriers.

3.2. Replication, segregation and maintenance of pD2RT

Two genes, repA1 (ORF70) and repA2 (ORF72), relevant to plasmid replication function have been identified on pD2RT. The ORF70 encodes a putative replication initiation protein A (pfam010134), which belongs to the RPA superfamily (cl02339). Members of this family are single-stranded DNA binding proteins, participating in the processes of DNA replication, recombination and repair. The ORF70-encoded putative RepA is most closely related (99% identity) to the hypothetical protein (EFW77624) of P. syringae pv. glycinea strain B076. It also shows 75% and 71% identity with the putative replication protein (YP_004750582) of the plasmid pGRT1 from P. putida strain DOT-T1E and hypothetical protein (YP_001109991) of the plasmid pBVIE04 from B. vietnamiensis strain G4, respectively.

The putative RepA encoded by ORF72 displays the potential ability to initiate plasmid replication due to the presence of a Replicative family domain (pfam03090) and a Primase domain (pfam08708). The ORF72-encoded replicase shows the highest degree of identity (99%) with the replicase of P. syringae pv. glycinea strain B076 and around 62% with those of pPT23A family plasmids, originating from different Pseudomonas syringae pathovars (Zhao et al., 2005). Plasmids belonging to the pPT23A family are shown to be universally distributed among Pseudomonas syringae pathovars being important agents for virulence (Sundin, 2007). The ORF72-encoded RepA is also similar (62% identity) to the RepA of pGRT1 (YP_004750586), which participation in plasmid replication was experimentally verified. Four almost identical direct repeats (DR) of 17 bp were detected upstream of ORF72: GTGTCCGGCATATGCA (55,834–55,850 bp); GGTGCCGGATATGCC (55,856–55,872 bp); GGTGCCGGATATGCA (55,878–55,894 bp); GTGTCCGGCATATGCA (55,901–55,917 bp). These DRs could serve as iterons, i.e., multiple RepA binding sites, present at the putative origin of plasmid pD2RT replication (oriV). Based on the similar architecture of the replication-related areas of the plasmids pD2RT and pGRT1 we proposed that most likely the ORF72-encoded protein is responsible for the initiation of plasmid pD2RT replication.

To verify whether RepA1 or RepA2 serves as an active pD2RT replication initiator protein in P. migulae strain D2RT, we firstly created two constructs, pGP704Arepa1::km and pGP704Arepa2::km, where the both PCR-amplified repA genes were inactivated by Km’ gene insertions, subsequently ligated into plasmid vector pGP704L molecules and electroporated to the E. coli CC118pir cells. Then by means of conjugation both constructs were introduced into the strain P. migulae D2RT resulting only in the formation of D2RT derivative strain D2RTrepA1−, harboring the pD2RT plasmid, where the repA1 gene (ORF70) was as expected to be replaced by homologous recombination with its mutated form. We could not observe D2RT derivatives, carrying the plasmid with the repA2 mutated gene, which refers to the crucial role of ORF72-encoded RepA2 in the pD2RT replication. Thus, the conducted mutagenesis experiment confirms our preliminary assumption that the RepA2 protein is essential in the initiation of plasmid pD2RT replication process in the strain P. migulae D2RT. However, it cannot be entirely ruled out that the product of ORF70 might also have some accessory role in the replication process and/or it could become essential in certain environmental conditions/host strains.

Two genes, parA (ORF65) and parB (ORF64), crucial for plasmid stability and maintenance, have been identified. ORF64 and ORF65 encode a centromere-binding protein ParB and a ParA ATPase, respectively, which amino acid sequences are most closely related (98% and 99% identity) to the homologs of P. syringae pv. glycinea strain B076 and also show 73% and 49% identity with those of pPT23A family plasmid pDC3000A (Zhao et al., 2005) from Pseudomonas syringae pv. tomato strain DC3000. The pD2RT partitioning system is related to the type Ia partitioning loci, assuming an active movement of replicated plasmid copies to dividing bacterial daughter cells (Gerdes et al., 2000). Two almost identical 19 bp palindrome sequences TTGTTGCTAGTAGCAACAC (48,171–48,189 bp) and GTGTTGCTACTAGCAACAA (48,142–48,160 bp) were detected between the predicted parA and parB genes, revealing probable location of centromere-like region parS. Putative binding site for the central integration host factor (IHF) was also predicted near the palindromic sequences on the complement strand: ATCAACCACTA (48,104 bp – 48,115 bp). It has been shown that binding of IHF to the parS site of the plasmid P1 creates a bend in the DNA and enables formation of the partitioning complex (Funnell, 1991).
Fig. 2. Pairwise BLAST alignments of the plasmid pD2RT with (a) Contig065 from the draft genome sequence data of *Pseudomonas syringae* pv. glycinea strain B076 (NZ_AEGG010000065); (b) plasmid pGR1 of *Pseudomonas putida* strain DOT-T1E (NC_015855); (c) plasmid pVIE04 of *Burkholderia vietnamiensis* strain G4 (NC_009228). The red bars indicate collinear homologous regions, while the blue bars indicate inverted regions of homology. BLAST comparison analysis was done with WebACT and visualized with ACT software with default parameters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
(TIGR02384), while the ORF105-encoded product is a toxin of RelE/StbE family (TIGR02385). A putative DNA methylase encoded by ORF26 likely contributes to the segregational stability of pD2RT, acting as a component of a restriction-modification system (Kulakauskas et al., 1995). The plasmid pD2RT contains also a bacteriophage abortive infection (Abi) system, which is encoded by ORF90 and belongs to Abi_2 superfamily (cl01988), that is thought to limit phage infection within a bacterial population through causing death of infected bacterial cells (Labrie et al., 2010).

3.3. Transferability of pD2RT

The ability of the plasmid pD2RT to self-transfer was determined in mating experiments with recipient P. putida strain PaW340, exhibiting transfer frequency of around $2.6 \times 10^{-5}$ transconjugants per donor cell.

The conjugation system of pD2RT belongs to the family of type IVB secretion systems (T4SS-IVB), revealing the presence of three main gene clusters: traKJIH, trbC-DNA primase-trbBAN and traMNOPQRTUWXY (Sexton and Vogel, 2002). The genetic organization of the respective region of pD2RT and BLAST search analysis of the corresponding gene products show the highest degree of similarity of this region to that of the plasmids pGRT1 and pBVIE04, genomic DNA of P. syringae pv. glycinea strain B076. The conjugal transfer in general encompasses DNA processing, its further recruitment to the translocation channel and translocation reaction, mediated by mating pair formation apparatus (Alvarez-Martinez and Christie, 2009). The DNA processing of pD2RT is putatively conducted by the mobABCDE gene cluster with a key player enzyme relaxase, encoded by the mobA gene (ORF41). Relaxase binds and nicks the plasmid DNA at the origin of transfer (oriT). Mutation of mobA reduces self-transferability of the plasmid pGRT1, while mobD is dispensable for DNA transfer (Molina et al., 2011). Following the classification of conjugative transfer systems proposed by Garcillan-Barcia and colleagues (Garcillan-Barcia et al., 2009), which is based on phylogeny of relaxase proteins, the plasmid pD2RT belongs to the MOBp family. MobA showed the closest homology to the plasmid pGRT1-encoded MobA (61% identity) of the clade MOBp12 and to the plasmid pBVIE04-encoded homologous protein (48% identity) of the clade MOBp13. The putative oriT of pD2RT is located between the mobB (ORF42) and mobC (ORF43) genes encompassing inverted repeats followed by a nick site (Fig. 4). The oriT was predicted based on the high level of sequence similarity with the respective region of the plasmid pRA2 of Pseudomonas alcaligenes NCIB 9867.
This region is highly conserved among pD2RT, pGRT1, pBVIE04 and several other plasmids in database, although the oriT of pGRT1 and pBVIE04 has not been identified (Fig. 4).

3.4. Mobile genetic elements on pD2RT

A number of mobile genetic elements (MGEs) and their remnants have been identified on the plasmid pD2RT, all enclosed within the region 74,363–129,853 bp, that resembles the structure of a class II transposon, highly similar to the toluene catabolic transposon Tn4656 of Tn1721-related transposon group (Fig. 5). Similarly to Tn4656 it contains two basic modules: the toluene-degrading xyl genes and the transposition-related segments, IRs and res-tnpR-tnpA. Although the res-tnpR-tnpA region of the two transposons shows high degree of nucleotide sequence similarity (over 94% identity), the putative gene products vary significantly. The 6 bp deletion at the 3’ terminus of the resolvase-coding gene (tnpR) of pD2RT abolishes the stop codon resulting in an extended protein consisting of 210 aa compared to 186 aa resolvases of Tn4656 and other Tn1721-related transposons. The transposase of Tn4656 encoded by a single tnpA gene is a large protein of 988 aa. Analysis of the corresponding region of pD2RT revealed the presence of a 4 bp deletion, which causes frameshift mutation in the tnpA gene, resulting in the formation of two ORFs assigned as tnpAa and tnpAb. Thus, it can be concluded, that most likely the functional activity of the transposase–resolvase complex of pD2RT has been lost or minimized due to the remarkable changes in its coding sequences.

Tsuda and Genka (Tsuda and Genka, 2001) have shown that the insertion of Tn4656 causes 5 bp duplication of the target site. The analysis of the plasmid pD2RT revealed the presence of a pentanucleotide TTCTT flanking the both sides of the proposed class II transposable element, the termini of which contain 39 bp imperfect inverted repeats (IR) highly similar to those of Tn4656. Two identical inverted copies of IS1492b are inserted in the terminal IRs of the putative Tn4656-like transposon of pD2RT, resulting in the formation of a large IS-composite transposon structure. Both IS1492b elements have 12 bp sub-terminal inverted repeats with additional 6 and 3 bp sequences outside the left and right IRs, respectively (Fig. 5). Although a number of catabolic composite transposons have been identified, the mobility of only a few of them has been experimentally proven (Tsuda et al., 1999). To prove the ability of the entire region of pD2RT to transfer as an IS-composite transposon remains for further research.

In addition, four intact mobile genetic elements, namely an IS30-related element, an ISPpu4-like element, an insertion sequence of IS116/IS110/IS902 family and a retron-type reverse transcriptase are found to be incorporated into the area located upstream of the catabolic gene clusters. A number of remnant genes, mostly encoding MGEs, have been found in this plasmid region, referring

Fig. 4. Comparison of the putative oriT region of the plasmid pD2RT with homologous sequences of the plasmids pGRT1, pBVIE04 and the identified oriT region of the plasmid pRA2. Inverted repeat sequences of pRA2 are marked with arrows, and the nic site is indicated by /). The corresponding sequences of four plasmids are shown with shaded boxes.

Fig. 5. Genetic organization of the putative Tn4656-like class II transposon of pD2RT. The order of major determined ORFs and genetic elements are outlined. Single ORFs and ORFs clusters representing mobile genetic elements and catabolic genes are designated with rectangles. The res site is circled. Sequences of detected IRs of identified Tn and IS elements are triangled and boxed, respectively, and the location of these MGEs within the plasmid is also provided. The positions of insertions of both IS1492b elements into the left and right IRs (IR-L and IR-R) of the putative class II transposon are indicated with a short vertical line.
to instability and consistent rearrangements of pD2RT’s genome.

3.5. Toluene catabolic genes on pD2RT

BLAST analysis revealed that the meta operon of pD2RT is most similar to that of Pseudomonas sp. ST41, which was isolated from pristine soil of Antarctica and was shown to be effective oil degrader at low temperatures (Stallwood et al., 2005), exhibiting 95–100% amino acid sequence identities in case of different proteins. It is also closely related to the meta gene cluster of the plasmids pDTG1 (Dennis and Zylstra, 2004) and pNAH20 (Heinaru et al., 2009), which are involved in the degradation of naphthalene. Identities with the homologs of plasmid pWW53 vary from 89% (XylFI) to 100% (XyII).

The upper (xyl) operon of pD2RT apparently consists of different gene blocks evolved independently and most likely separately recruited to the ancestral form of this gene cluster. This can be concluded based on the fact that the homologies between the corresponding gene products encoded by the upper operons of pD2RT and other TOL plasmids vary significantly, showing 81–99% identity depending on the particular gene product (Table S1). Compared to other gene products of the upper operon, the XylA protein of pD2RT has remarkably low similarity to the homologs of Pseudomonas sp. TW3 (81% identity), pWW0 (80% identity), pWW53 (79% identity) and pDK1 (79% identity). However, BLAST search revealed the high degree of similarity (99% and 95% identity) between the XylA encoded by pD2RT and the corresponding sequences of pSVS11 from Pseudomonas sp. SV11 and pSVS15 from Pseudomonas sp. SV15, respectively. Both plasmids pSVS11 and pSVS15 have been isolated in Belarus from different oil-contaminated sites and bear genes for degradation of toluene (Sentchilo et al., 2000). Similarly to pD2RT, pSVS11 and pSVS15 carry single upper and meta pathway gene clusters.

The first two genes of the upper operon, xylU and xylW, are not required for degradation of toluene, nevertheless they have remained persistent in the composition of the xyl operon of TOL plasmids (Williams et al., 1997). XylU encoded by pD2RT is a 150 aa hypothetical protein, which is bigger in size compared to its 132 aa homologs of pWW53 and pDK1. Analysis of the nucleotide sequence of the xylW gene revealed two nucleotide substitutions that lead to a premature termination and truncation of the benzyl alcohol dehydrogenase at codon position 250 compared to 348 aa long xylW proteins of other TOL plasmids. The mutation destroys the C-terminal domain of the protein, apparently abolishing the expression of a functional xylW gene.

To conclude, the obtained results show remarkable differences in sequences of xyl genes and the corresponding proteins of the plasmid pD2RT compared to other TOL plasmids, which indicate continuous changes and rearrangements in catabolic xyl gene clusters.

3.6. Conservation of xyl genes on different plasmid types

Extensive study on the diversity of TOL plasmids of pseudomonads conducted by Sentchilo and colleagues (Sentchilo et al., 2000) highlighted the structural differences in organization of catabolic areas of different xyl genes carriers. Three general types of xyl operons were revealed based on the comparative analysis of either meta or upper pathway genes. Lower operon of known TOL plasmids resembles either meta I or meta II gene cluster of plasmid pWW0 or that of archetypical TOL plasmid pWW0. Distinguishing between pWW53-like and pWW0-like xyl genes is also important in phylogenetic analysis of the upper pathway operon. The majority of so far characterized TOL plasmids of pseudomonads carry catabolic genes falling in these two groups. The third group is represented by two plasmids only, pSVS11 and pSVS15 (the latter is IncP-9 plasmid). The main hallmark of particular upper pathway operon is the distinct xylA sequence (less than 84% amino acid sequence identity with the XylA of pWW0), which is assumed to have a different origin from that of nearby located xyl genes (Sentchilo et al., 2000).

Plasmid-derived homologous catabolic genes have been also determined from several biodegradative sphingomonads (Alphaproteobacteria), however they are found to be distantly related and likely originate from a distinct genetic source (Nojiri et al., 2009).

Considering the importance of the xylA gene for distinguishing between different types of TOL catabolic operons, it was chosen as a marker gene for screening of plasmid-bearing toluene degrading pseudomonads previously isolated by us from Baltic Sea water (Table 1).

The multiple alignment analysis of amplified 203 bp long xylA gene fragment (Fig. S1) revealed the presence of all three xylA types (pWW0-like, pWW53-like and pSVS11/pSVS15-like) among the isolates. The xylA sequences of plasmids pD67 and pD2RT (93.6% identical to each other) shared only 74 to 77.5% of identity to their homologs, which appeared to be highly similar to those of plasmids pWW0 and pWW53 (97 to 100% of identity). For P. stutzeri strains and Pseudomonas sp. 2B49 it is not defined whether the xyl genes are located on plasmid or chromosome. In addition, the replication type of the plasmids of these strains remains unknown yet, since the PCR-based replicon typing of IncP (Jutkina et al., 2011) and pD2RT (data not shown) backbones yielded no amplification products except for the presence of IncP-9 repA gene in multi-plasmid strains 2A54, 2D47 and 2D49.

To compare the catabolic regions of the two different TOL plasmids, pD2RT and the IncP-9 plasmid pD67, the complete nucleotide sequence of the catabolic region together with the flanking area of the latter plasmid was determined. Data analysis revealed that both plasmids share the same organization of xyl gene clusters with entire nucleotide sequence identity of 99%. Comparison of the flanking regions extended the similarity between these two plasmids, referring to the close relatedness between their MGEs. Thus, the catabolic areas of both plasmids might have originated from common evolutionary ancestor and incorporated into the dissimilar replicons due to the activity of MGEs. Interestingly, the meta-pathway operon of pD67 was found to be 100% identical to that of Pseudomonas sp. strain ST41 (AY599747) and as determined in our previous work (Jutkina et al., 2011) along with plasmid pSVS15 it forms the theta-subgroup of IncP-9 plasmid family.
3.7. Enzyme activity testing and growth kinetics

Plasmid acquisition can induce significant changes in bacterial cell physiology, alteration of bacterial growth in response to plasmid presence has been frequently reported (Diaz Ricci and Hernandez, 2000). In the present study we observed the effect of two different TOL plasmids pD2RT and pD67 on the growth profile of the bacterial strain P. putida PaW340 in comparison to their wild-type hosts P. migulae D2RT and P. migulae D67. All TOL plasmid bearing strains were grown in minimal medium with m-toluate as sole carbon and energy source. Although the bacterial strains P. migulae D2RT and D67 are phylogenetically very closely related (Jutkina et al., 2011), the growth rate of the strain D2RT was almost three times slower than that of the strain D67. The growth profile of the transconjugant strain P. putida PaW340 was influenced by the acquired plasmid, pD2RT or pD67, indicating the probable effect of plasmid specific features on the biodegradative capacity of the culture (Fig. 6).

It has been shown that acquisition of a catabolic plasmid by an unsuitable host can lead to accumulation of highly toxic compounds such as catechol, an intermediate in the carbazole, toluene, naphthalene etc. pathways, which in turn has dramatic impact on both the plasmid and the host (Nojiri, 2013). We examined the expression of catechol 2,3-dioxygenase (C23O) (encoded by xylE), which catalyzes the extradiol ring-cleavage of catechol, in crude extracts of the strains: P. migulae D2RT, P. migulae D67 and both P. putida PaW340 transconjugants, in order to compare the enzymatic activities of the xylE products in particular strains. No remarkable differences in catechol 2,3-dioxygenase (C23O) enzyme activities were detected among these TOL plasmid bearing strains grown either solely on m-toluate or on glucose with m-toluate as a meta-operon inducer (Table S2).

Considering the high conservation level of the meta-pathway of both plasmids (the identical promoter region and the high degree of identity (99–100%) between most of xyl genes) as well as the distinct origin of their backbones it is interesting to clarify the probable causes of observed differences of the bacterial growth profiles in our forthcoming studies.

4. Conclusions

Previous data indicate that genes responsible for degradation of toluene via the meta catabolic pathway are commonly associated with plasmids of IncP-7 and IncP-9 groups. The complete sequence of the plasmid pD2RT considerably broadens the range of carriers for toluene catabolic genes. Although it was previously shown by Sentchilo and colleagues (Sentchilo et al., 2000) that TOL plasmids may be affiliated to different incompatibility groups, our study provides the first complete sequence of a TOL plasmid, which is not related to the well-studied IncP plasmids. Comparative analysis of pD2RT backbone gene products, responsible for plasmid replication, maintenance and conjugative transfer, shows the closest homology to those of plasmids pGRT1 from P. putida strain DOT-T1E and pBVIE04 from B. vietnamiensis strain G4. Interestingly, all three plasmids enable their host bacteria to tolerate toluene, playing crucial role for their survival. However, the mechanisms encoded on these plasmids are different: pD2RT is a TOL plasmid, pBVIE04 enables degradation of toluene via toluene-2-monoxygenase pathway, pGRT1 encodes a toluene efflux pump. Surprisingly, the backbone region of pD2RT shares the highest degree of identity (83–100% amino acid identity) with one of the partial genome sequences of P. syringae pv. glycinea strain B076. Pseudomonas syringae comprises plant-pathogenic species, whose ecological and pathogenic features are commonly associated with pPT23A plasmid family (Ma et al., 2007). The plasmid pD2RT exhibits distant relatedness to this plasmid group as well. Our results confirm that TOL plasmids are not restricted to particular plasmid groups, and dissemination of highly similar toluene-degrading...
genes between distinct replicons like pD2RT and pD67 is likely associated with MGEs, e.g., large class II transposons and/or IS elements. We have also observed differences in the growth profile of the transconjugant host, *P. putida* PAW340, influenced by particular TOL plasmid, referring to the importance of plasmid and host compatibility in the biodegradative capacity of a catabolic plasmid-bearing strain.

**Acknowledgments**

This work was supported by the Institute of Molecular and Cell Biology, University of Tartu, by the Research Grant SF0180026s08 and RLOMR CELMS from the Ministry of Education and Research, and by the Estonian Science Foundation grant 7827. The work presented was partly funded by the Lundbeck Foundation (www.lundbeckfoundation.com) project DK nr R44-A4384.

**Appendix A. Supplementary data**

Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.plasmid.2013.09.003.

**References**


Saumaa, S. et al., 2006. Involvement of DNA mismatch repair in stationary-phase mutagenesis during prolonged starvation of Pseudomonas putida. DNA Repair 5, 505–514.


