A 2,4-dichlorophenoxyacetic acid degradation plasmid pM7012 discloses distribution of an unclassified megaplasmid group across bacterial species

Yoriko Sakai,1 Naoto Ogawa,2 Yumi Shimomura1 and Takeshi Fujii1

Correspondence
Yoriko Sakai
kyoriko@affrc.go.jp

Received 22 October 2013
Accepted 15 January 2014

1Environmental Biofunction Division, National Institute for Agro-Environmental Sciences, 3-1-3 Kannondai, Tsukuba, Ibaraki 305-8604, Japan
2Faculty of Agriculture, Shizuoka University, 836 Ohya, Surugaku, Shizuoka 422-8529, Japan

Analysis of the complete nucleotide sequence of plasmid pM7012 from 2,4-dichlorophenoxyacetic-acid (2,4-D)-degrading bacterium Burkholderia sp. M701 revealed that the plasmid had 582,142 bp, with 541 putative protein-coding sequences and 39 putative tRNA genes for the transport of the standard 20 aa. pM7012 contains sequences homologous to the regions involved in conjugal transfer and plasmid maintenance found in plasmids by 2p from Burkholderia sp. YI23 and pBVIE01 from Burkholderia sp. G4. No relaxase gene was found in any of these plasmids, although genes for a type IV secretion system and type IV coupling proteins were identified. Plasmids with no relaxase gene have been classified as non-mobile plasmids. However, nucleotide sequences with a high level of similarity to the genes for plasmid transfer, plasmid maintenance, 2,4-D degradation and arsenic resistance contained on pM7012 were also detected in eight other megaplasmids (~600 or 900 kb) found in seven Burkholderia strains and a strain of Cupriavidus, which were isolated as 2,4-D-degrading bacteria in Japan and the United States. These results suggested that the 2,4-D degradation megaplasmids related to pM7012 are mobile and distributed across various bacterial species worldwide, and that the plasmid group could be distinguished from known mobile plasmid groups.

INTRODUCTION

Mobile genetic elements such as plasmids and transposons are major tools for generating new genetic characteristics in prokaryotes. Bacteria that can degrade man-made compounds have been isolated and the genes responsible for these activities often exist on mobile genetic elements (Nojiri et al., 2004; Ogawa et al., 2004; van der Meer, 2008). The broadleaf herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) was one of the first herbicides developed and has been used as a model chemical for microbial degradation of xenobiotics (Chinalia et al., 2007). In 1977, a plasmid encoding genes for 2,4-D degradation was reported for the first time (Fisher et al., 1978; Pemberton & Fisher, 1977). Around the same time, Cupriavidus pinatubonensis strain JMP134 (formerly Alcaligenes eutrophus, Ralstonia eutropha and Cupriavidus necator) (Sato et al., 2006) was isolated. This strain, with its conjugal 2,4-D degradation plasmid pJP4, has since been used as a representative strain for the study of microbial 2,4-D degradation and horizontal transfer of degradation genes (Don & Pemberton, 1981; Pérez-Pantoja et al., 2012). pJP4 can be transferred into a broad range of hosts with high efficiency and was classified as belonging to the IncP-1 incompatibility group (Don & Pemberton, 1981).

Although many strains with plasmid-borne genes encoding 2,4-D-degrading enzymes have been isolated, only five of these plasmids [pJP4 (Trefault et al., 2004), pEST4011 from Achromobacter xylosoxidans strain EST4002 (Vedler et al., 2004), plBJ1 from Burkholderia cepacia strain 2a (Poh et al., 2002; Xia et al., 1998), pDB1 from Variovorax sp. strain DB1, and p712 from Pseudomonas picketti strain 710 (Kim et al., 2013)] have been sequenced fully. All of these plasmids, with sizes ranging from 62 to 99 kb, are classified in the IncP-1 incompatibility group.

Mobile plasmids have been classified based on incompatibility tests (Taylor et al., 2004), replicon typing by hybridization with specific probes (Couturier et al., 1988) or PCR with specific primers detecting plasmid maintenance regions (Carattoli et al., 2005; Götz et al., 1996).

Abbreviations: 2,4-D, 2,4-dichlorophenoxyacetic acid; ICE, integrative conjugative element; T4CP, type IV coupling protein; T4SS, type IV secretion system.

Four supplementary tables and four supplementary figures are available with the online version of this paper.
However, it is difficult to classify plasmids newly identified by genome analysis without incompatibility testing or replicon typing. Recently, the amino acid sequences of relaxase, an essential enzyme for plasmid mobilization, have been proposed as a tool for carrying out mobile plasmid classification (Francia et al., 2004; Garcillán-Barcia et al., 2009). Using this classification system, IncP-1 incompatibility group plasmids, including pJP4, pEST4011, pDB1, pDB2 and p712, are classified in the MOB plasmid group (Francia et al., 2004; Garcillán-Barcia et al., 2009; Kim et al., 2013; Poh et al., 2002; Xia et al., 1998).

In our previous study, 14 strains of 2,4-D-degrading bacteria were isolated from paddy fields in various regions of Japan (Sakai et al., 2007). We found that eight of the 14 strains had megaplasmids of similar sizes (~600 kb), with highly similar 2,4-D degradation gene sequences. As seven of the eight strains were identified as Burkholderia spp. and the other as a Cupriavidus sp., based on 16S rDNA sequence analysis, we suggested that these plasmids were transmissible megaplasmids containing 2,4-D degradation genes.

In this work, the complete nucleotide sequence of pM7012 from one of the eight strains, Burkholderia sp. M701 (Sakai et al., 2007), was determined and compared with other related sequences. The findings suggested that 2,4-D degradation megaplasmids related to pM7012 are mobile and that the plasmids can be distinguished from known mobile plasmid groups.

METHODS

Megaplasmid extraction. DNA from pM7012 was separated by PFGE based on a previous method (Takami et al., 1999) with modifications, as described below. Strain M701 was cultured in 0.5 × Luria–Bertani broth (Sambrook & Russell, 2001) at 28 °C until OD600 0.6 and was then embedded in 1% (w/v) agarose gel plugs (each 100 μl plug contained cells from 1 ml culture) using plug moulds (Bio-Rad). The plugs were incubated in lysozyme solution (4 mg ml⁻¹) in TE (Tris/EDTA) buffer at 37 °C for 2 h, washed in TE buffer for 20 min and then washed twice in 2% (w/v) SDS-TE buffer for 20 min. The washed plugs were incubated in 1% (w/v) SDS-TE buffer with 0.2% proteinase K (2 °C) overnight. The plugs were then washed in 1% (w/v) SDS-TE buffer for 1 h and then washed twice with TE buffer for 20 min. The DNA in the plugs was separated using a CHEF-DR II PFGE apparatus (Bio-Rad) in 0.5 × TBE (Tris/borate/EDTA) at 14 °C (parameters: run time, 16 h; gradient, 6.0 V cm⁻¹; included angle, 120°; initial switch time, 2.98 s; final switch time, 57.62 s; ramping factor, linear) and the gel was stained with ethidium bromide. To remove contaminating chromosomal DNA, gel blocks containing pM7012 DNA were cut out from the PFGE gel and embedded in new agarose gel for a second round of PFGE using the same conditions as the first round. The gel blocks containing pM7012 DNA were excised, and the plasmid DNA was electro-eluted into cellulose tubes and then purified with phenol (Sambrook & Russell, 2001).

Sequencing and annotation. The extracted pM7012 DNA was physically fragmented and a library of short DNA pM7012 fragments (2–3 kb) was constructed using pUC118 as a vector (Sambrook & Russell, 2001). The paired-ends of the insert DNA were sequenced by the Sanger method using a 3100 Genetic Analyzer (Applied Biosystems) with a BigDye Terminator v. 3.1 cycle sequencing kit (Applied Biosystems). The nucleotide sequence of pM7012 was also determined using an Illumina Genome Analyzer (Illumina). The sequence data were assembled using ATSQ (Genetix) and Genomics Workbench (CLC bio) software. To confirm the direction and order of the obtained contigs, a large insert (15–40 kb) library was constructed using a CopyControl BAC Cloning Kit (Epicentre). Paired-ends of the large insert DNA were sequenced using the Sanger method as above. Remaining gaps were closed by primer-walking techniques with sequence-derived oligonucleotides. Annotation and comparative analysis of the sequence of pM7012 were carried out using the following software/databases: In-silico Molecular Cloning (IMC; in silico biology), BLAST (Altschul et al., 1997), Pfam (Finn et al., 2010), tRNAscan (Schattner et al., 2005), tRNADB-CE (Abé et al., 2011) and MEGAS (Tamura et al., 2011).

Sequencing of para, traD and arsA in other strains. Partial nucleotide sequences of the genes para (encoding ENTPase for plasmid partitioning), traD (encoding type IV coupling proteins (4TCPs) for plasmid transfer) and arsA (encoding ArsA for arsenic resistance) from other 2,4-D-degrading strains were amplified using an i-Cycler PCR apparatus (Bio-Rad). The 30 μl reaction mixtures contained 10 ng genomic DNA, 0.75 U Ex Taq polymerase (Takara Bio), 1× Ex Taq buffer, 6 pmol dNTPs and 30 pmol of the respective primer of each of the following primer pairs: pM7012_parAsh (5′-TTTCACGCTCGAAGCTACTC-3′) and pM7012_parAr (5′-CGGTCATGGATTTCCGTTG-3′); pM7012_traDf (5′-AAACAGATGTCACACAAACCGG-3′) and pM7012_traDr (5′-CAACGGGCACCCGTTTCAC-3′); pM7012_arsAf (5′-ATCGGTAATCGGCCTACA-3′) and pM7012_arsAr (5′-AGCGGGATGTTTCCTACG-3′). The PCR profile was 3 min at 94 °C; 25 cycles of 15 s at 94 °C, 15 s at 56 °C (for paraA), 55 °C (for traD), or 51 °C (for arsA); 15 s at 72 °C; and 3 min at 72 °C. The PCR products were directly sequenced using an ABI PRISM 3100 Genetic Analyzer as above.

PFGE and hybridization. Plasmids from the 2,4-D-degrading strains from paddy fields in Japan, Burkholderia spp. strains M701, Y212, C308, T201, T301, H801 and F54, and Cupriavidus sp. strain Y103 (Sakai et al., 2007), and Burkholderia tropica strain RASC from the United States (formally strain TF3D, B. cepacia) (Belum et al., 2010; Tomo et al., 1995; Vallaey et al., 1996) (provided by Dr Terence L. Marsh at Michigan State University) were separated by PFGE as described above. Southern hybridization targeting the DNA in the PFGE gel was performed according to a previous report, using DIG-labelled tfICM7012_parAshM7012, traDPM7012 and arsAPM7012 probes (Sakai et al., 2007). Probes were generated using DNA from strain M701 as a template using the primers and PCR parameters described above.

RESULTS AND DISCUSSION

Overview of pM7012

The complete nucleotide sequence of pM7012 from the 2,4-D-degrading bacterium Burkholderia sp. M701 revealed that the plasmid was 582 142 bp in size, with a GC content of 59.7% and 541 putative protein-coding sequences. Among these coding sequences, only 187 were annotated as functional proteins. Coding sequences of unknown function, many of which exhibited no significant homology to any sequences in the databases, were localized specifically in one half of pM7012 (Fig. 1). At both ends of this region, several GC-rich sections with highly repetitive sequences...
were observed. The other half of the plasmid contained coding sequences predicted to encode systems involved in arsenic resistance, 2,4-D degradation and type IV secretion. Between the arsenic resistance genes and 2,4-D degradation genes, coding sequences homologous to those in a chromid (‘secondary chromosome’) (Harrison et al., 2010) region of Burkholderia spp. were found. Thirty-nine putative tRNA genes were also found in pM7012.

Comparative sequence analysis of pM7012 and the other plasmids revealed that the overall genetic structure of pM7012 was similar to that of plasmid byi_2p (356 kb), from the fenitrothion-degrading Burkholderia sp. YI23, (Lim et al., 2012), and that of plasmid pBVIE01 (397 kb), from the trichloroethylene-degrading Burkholderia sp. G4 (GenBank accession number CP000617), although the identity of each corresponding gene between pM7012 and the two plasmids was lower than that between byi_2p and pBVIE01 (Fig. S1, available in the online Supplementary Material). byi_2p and pBVIE01 were found in genome sequence data, and their properties have not been reported. The regions for conjugal transfer and plasmid maintenance were highly similar between byi_2p and pBVIE01, although the transcriptional direction of the region containing genes

---

**Fig. 1.** Map of pM7012 in Burkholderia sp. M701. The first and second circles (from outside to inside) represent genes transcribed in the clockwise and counterclockwise direction, respectively, with colours representing the Clusters of Orthologous Groups classification, except genes annotated as involved in 2,4-D degradation, arsenic resistance, plasmid transfer, gene mobilization and genes related to phage. The third circle represents tRNA genes, and the fourth and fifth circles represent pM7012_repeat A sequences in the clockwise and counterclockwise direction, respectively. The sixth circle shows GC contents from 47 to 62 % separated by different colours, with a mean value of 54 %. The regions in synteny with byi_2p and pBVIE01 are shown by dotted blue lines.
for the type IV secretion system (T4SS) was opposite (Fig. S1). The regions of pM7012 in synteny with byi _2p and pBVIE01 are shown in Fig. 1. Those regions might contain the core genes of the megaplasmids. Some regions in heavy metal resistance plasmid pMOL30 (233 kb), from *Cupriavidus metallidurans* strain CH34, have synteny with those in pBVIE01 (Mergay et al., 2009; Monchy et al., 2007). In addition, some parts of pM7012 also showed synteny with those of pMOL30 (Fig. S1). byi _2p, pBVIE01 and pMOL30 did not have gene clusters for 2,4-D degradation, arsenic resistance or the tRNA gene clusters.

### 2,4-D degradation genes

The genes necessary for conversion of 2,4-D into β-ketoisovalerate (tfdFAKBIEICIDIRI), a central metabolite for pseudomonads (Trefault et al., 2004), form a cluster (tfd gene cluster) in pM7012 (Fig. 2). The organization of this cluster in pM7012 was most similar to that of pEST4011, while the order of the genes, tfdBIEICIDIRI in pM7012, resembles that of the corresponding genes of the tfd II gene cluster of pJP4 (Fig. 2). The organization of the tfd gene cluster in strain RASC, in which the partial nucleotide sequences of tfdA, tfdB and tfdC have ≥98% similarity to those of pM7012 (Sakai et al., 2007), has not been reported. The second tfd gene cluster, tfdCIEI in pM7012, was not a simple duplication of part of the first gene cluster, tfdBIEICIDIRI. In the upstream region of the tfdCIIEI in pM7012, the pseudogene named ΔtfdRI, which has a frame shift in the middle of the gene, was identified. The nucleotide sequence of ΔtfdRI was 91% similar to that of tfdR from pEST4011, while the similarity of the sequences of tfdR and ΔtfdRI in pM7012 was 72%.

Although the nucleotide sequences of tfdCIEI and tfdCIIEI in pM7012 were identical, except for 84 bp of sequence at the end of tfdE, the nucleotide sequence of tfdBII in pM7012 was 77, 78 and 79% similar to those of tfdBII in pM7012, tfdB in pEST4011 and tfdBII in pJP4, respectively. The differences in similarity between the corresponding genes in the first and second tfd gene clusters in pM7012 suggested that the second tfd gene cluster had been constructed with genes from several types of tfd gene clusters, including gene cluster tfdBIEICIDIRI.

IS1071-based transposons have been found in all fully sequenced 2,4-D degradation plasmids, including pEST4011 and pJP4 (Fig. 2) (Kim et al., 2013; Poh et al., 2002; Trefault et al., 2004; Vedler et al., 2004). IS1071 insertion sequences have also been found with the degradation genes of some chemicals and are thought to play an important role in integration of such degradation gene clusters into replicons (Ng & Wyndham, 1993; Sota et al., 2006). However, IS1071-based transposons were not found near the tfd gene cluster in pM7012. A sequence that was mostly identical to the inverted repeat sequence in IS1071, which belongs to the IS3 family (Hasebe et al., 1998), was found in the region downstream of tfdF in pM7012 (Fig. 2). Near the insertion sequence, there was a short coding sequence that had an 85 aa sequence that was homologous to that of the transposase of IS605 found in *Xylella fastidiosa* subsp. sandyi Ann-1 (404 aa; GenBank accession number EAO31770). Downstream of the tfdBII gene in pM7012, there were repeated sequences that were found frequently in pM7012 and were named pM7012_repeat A (Table S1). At ~10 kb downstream of the tfdBII gene in pM7012, an integrase gene coexisted with one pM7012_repeat A. Such integrase genes

---

**Fig. 2.** Gene organization of 2,4-D degradation genes and the surrounding region in pM7012 and other related plasmids. Dark grey arrows represent 2,4-D degradation genes from pM7012 and highly homologous genes from other plasmids. Light grey arrows represent another set of tfd genes in pJP4. Black arrows represent transposases, resolvases or integrase. White arrows represent other genes. A dotted white arrow labelled ΔtfdRI in pM7012 indicates that the middle part of the sequence has a frame shift. Boxes represent insertion sequences. Fine line arrows represent duplicated regions of each gene cluster. The number under each arrow represents nucleic acid sequence identity (%) between the gene and the corresponding gene from the gene cluster tfdFAKBIEICIDIRI in pM7012. GenBank accession numbers: pEST4011, Ay540995; pJP4, Ay365053.
accompanied by repeat sequences are characteristic of some genomic islands [including integrative conjugative elements (ICEs) that contain catabolic genes like Tn4371 carrying genes for the degradation of biphenyl and polychlorinated biphenyls (Toussaint et al., 2003), and clc elements carrying genes for chlorocatechol degradation (Ravatn et al., 1998)]. There is a possibility that a genomic island carrying 2,4-D degradation genes was integrated into the ancestor plasmid of pM7012. However, it is difficult to identify the extremities of the genomic island, because there are so many 'repeat A' in the sequence of pM7012. Therefore, the integrase might be a remnant of a mobile genetic element. byi_2p and pBVIE01 had homologues of the integrase gene in pM7012, and similarity between the amino acid sequence of the integrase of byi_2p or pBVIE01 and that of pM7012 was 42%. Many other repeat sequences and sequences for transposases and integrases were also found in pM7012 (Tables S1 and S2). The existence of many insertion sequences or repeat sequences and the chimeric structure of pM7012 might reflect intensive genetic rearrangements during molecular evolution of pM7012. The tfl gene clusters of pM7012 might have been integrated in one of these rearrangements.

### Arsenic resistance transposon

pM7012 had a Tn3 family transposon with an arsenic resistance gene cluster, arsRCDAB, that had a similar gene organization to that of the arsenic resistance transposon TnLfrars in Leptospirillum ferrophilum strain Fairview, which was isolated from an arsenopyrite bioreoxidation tank containing a high concentration of arsenic (Tuffin et al., 2006) (Fig. S2). The transposon found in pM7012 was named Tn6233, following the recommendations of the transposon nomenclature system (Roberts et al., 2008). The similarities of the amino acid sequences of the corresponding arsenic resistance enzymes in the pM7012 and Fairview gene clusters ranged from 93 to 98%, and the nucleotide sequences of the inverted repeats in their transposons were mostly identical (Fig. S2). ars genes are found widely in microbes (Silver & Phung, 2005) and TnLfrars-related transposons have been found in different replicons (Tuffin et al., 2006). Therefore, it is supposed that the TnLfrars-like transposons have been spread among microbes in the environment.

### Features of genes with chromid origins

The region between the arsenic resistance genes and the 2,4-D degradation genes in pM7012 was 59 kb, and had 51 protein-coding sequences. The nucleotide sequences of the coding sequences of this region were homologous to those of a part of chromids of three Burkholderia strains, namely strain CCGE1002, isolated as a legume-nodulating bacterium (Ormeño-Orrillo et al., 2012), strain PsJN, a plant growth-promoting endophyte (Weilharter et al., 2011), and strain LB400, a polychlorinated biphenyl-degrading bacterium (Chain et al., 2006), as well as other Burkholderia spp. The amino acid sequences of 27 of the 51 coding sequences in pM7012 were homologous to those of the corresponding coding sequences from the chromid of strain CCGE1002, with >80% similarity. In addition, the order of 35 of the 51 coding sequences in pM7012 and the corresponding genes in the chromid of CCGE1002 was identical. The same trend was found between the coding sequences of pM7012 and the corresponding coding sequences of the strains PsJN and LB400. Analysis of the metabolic functions of these 51 coding sequences in pM7012, using the KEGG Automatic Annotation Server, revealed that the region was made up of genes that mediate parts of various metabolic pathways (Fig. 1). Therefore, it was most likely that this region of pM7012 was acquired from a chromid from bacteria of the genus Burkholderia. A possibility that pM7012 possesses essential genes for growth of strain M701 cannot be excluded because curing of pM7012 had not been successful. However, pM7012 has several features different from the definition of chromids. (1) pM7012 had a lower GC content compared with the GC content of chromosomes in Burkholderia spp., depending on the genomic DNA database. Usually, chromids have a nucleotide composition close to the chromosome (Harrison et al., 2010). (2) Mobility of a pM7012-like plasmid including the chromid region among genera had been indicated (see ‘Spread of 2,4-D degradation megaplasmids related to pM7012’). Generally, chromids carry genus-specific genes (Harrison et al., 2010). (3) pBVIE01, whose genetic structure was similar to that of pM7012, is classified as a plasmid (Harrison et al., 2010).

The end of the chromid region in pM7012 encoded phospholipase C and this sequence was disrupted by the Tn6233 resolvase-coding gene (tnpR). However, direct repeat sequences were not found on either side of Tn6233. Therefore, it is most likely that Tn6233 integrated originally into the phospholipase C gene in the chromid of a strain of Burkholderia sp. We propose that an ancestor plasmid of pM7012 invaded the chromid by homologous recombination at or around Tn6233, and then the plasmid was resolved from the chromid with acquisition of Tn6233 and a part of the chromid adjacent to it in a Burkholderia sp. strain. As with the F plasmid, many conjugative plasmids can integrate into host chromosomes and carry part of the chromosome with them when they are resolved (Lawley et al., 2004).

Thirty-nine putative tRNA genes for the standard 20 aa were found in pM7012. pM7012 possessed two tRNA gene clusters, named tRNA gene clusters I and II (Fig. 1, Table S3). tRNA gene cluster I was constructed from three tRNA gene clusters, named tRNA gene clusters I and II (Fig. 1, Table S3). tRNA gene cluster I was constructed from three tRNA gene clusters, named tRNA gene clusters I and II (Fig. 1, Table S3). tRNA gene cluster I was constructed from three tRNA gene clusters, named tRNA gene clusters I and II (Fig. 1, Table S3).
Although their sequences were homologous to those from various other organisms (Table S1), the putative genes encoded 36 tRNAs for the standard 20 aa without duplication of anticodons. Sets of tRNA genes encoding all standard 20 aa, without duplication of anticodons, have also been found in genomes of other microbes such as *Anaplasma marginale* St Maries and *Ehrlichia canis* Jake [Genomic tRNA Database (Chan & Lowe, 2009)]. This suggests that tRNA gene cluster II from pM7012 originated from the genome of an unknown micro-organism.

**Plasmid maintenance region**

The plasmid maintenance region has a plasmid partition region and plasmid replication genes. The plasmid partition region of pM7012 consisted of a plasmid centromere gene (*parS*), and genes encoding a Walker-type ATPase (*parA*) and a centromere-binding protein (*parB*), which are commonly found in the plasmid partition region of other low-copy-number plasmids such as the F plasmid (Fig. 3) (Hayes & Barilla, 2006). The plasmid replication gene (*repA*) was found ~7 kb downstream of the plasmid partition region. The organization of these genes in pM7012 was mostly the same as in byi_2p and pBVIE01, and similar to that of pMOL30 and a plasmid from the naphthalene-degrading *Pseudomonas putida* strain ND6 (pND6-2) (Li et al., 2013) (Fig. 3). A similar gene organization was also found in the chromosome of *P. putida* strain H8234, which was isolated from a hospital as an antibiotic-resistant bacterium (Molina et al., 2013).

Phylogenetic analysis of the amino acid sequences of the ATPases for plasmid partitioning (ParA, SopA, RepA, IncC and YafB) (Fig. S2) showed that the sequences of ParA from pM7012, byi_2p, pBVIE01, pMOL30, pND6-2 and the chromosome of strain H8234 formed a cluster. Although the cluster was closely related to the cluster made up of the sequences from pCAR1 and pND6-1, classified in the IncP-7 incompatibility group (Shintani et al., 2006), the arrangement of genes in the plasmid maintenance region and the palindrome sequences of *parS* were different (Fig. 3). Additionally, coexistence of pND6-1 and pND6-2 in

![Fig. 3. (a) Gene organization in plasmid maintenance regions, and (b) palindrome sequences in *parS*/sopC regions of pM7012 and other related plasmids and a chromosome from strain H8234. Dark grey arrows labelled ‘A’ represent *parA*/sopA genes, grey arrows labelled ‘B’ represent *parB*/sopB genes, light grey arrows labelled ‘rep’ represent genes for replication protein RepA and black squares labelled ‘S’ represent *parS*/sopC regions. Labels ‘h’ and ‘d’ on the arrows indicate genes for hypothetical protein and diguanylate cyclase, respectively. Numbers under the arrows represent amino acid sequence identities (%) between each gene and the corresponding gene in pM7012. GenBank accession numbers: byi_2p, CP003091; pBVIE01, NC_007971; H8234, CP005976; pND6-2, NC_018746; pCAR1, NC_007971; pND6-1, NC_007971; plasmid F, NC_007971; H8234, CP005976; pND6-2, NC_018746; pCAR1, NC_007971; pND6-1, NC_007971; F plasmid, NC_007971.](image-url)
their host, *P. putida* strain ND6 (Li et al., 2012), meant these two plasmids were compatible. These facts indicated that the plasmid maintenance regions of pM7012, byi_2p, pBVIE01, pMOL30 and pND6-2 had a recent common ancestor, and that the plasmids form a group distinct from the IncP-7 incompatibility group.

**Genes for plasmid transfer**

For conjugative transfer, the T4SS, a gene for T4CP and a relaxase (TraI) are required (Lawley et al., 2003). All of these essential genes for F plasmid-like T4SS and T4CP were found in pM7012, except *traI* (Fig. 4). The organization of putative genes for conjugative transfer in pM7012 was very similar to several other plasmids, including byi_2p and pBVIE01, and the chromosome of strain H8234 had similar sequences to T4CP (Fig. 4). It has been reported that the amino acid sequences of T4CPs could be used as a classification tool as the sequences are highly conserved (Smillie et al., 2010). Therefore, the phylogeny of T4CPs in representative plasmids, including pM7012 (Garcillé-Barcia et al., 2009) and related sequences, was analysed (Fig. 5).

One cluster containing the pM7012 sequence was distinct from the clusters of the MOB group plasmids, classified based on the sequence of relaxase (Garcillé-Barcia et al., 2009). The sequences of T4CPs in the pM7012 cluster were found in plasmids from betaproteobacteria and the chromosome of *Pseudomonas* sp. strain H8234.

The chromosome of strain H8234 had gene sets for plasmid maintenance and conjugative transfer that were homologous to those of pM7012 (Figs 3 and 4). The organization of coding sequences in these regions of strain H8234 and pM7012 was highly similar, as were the arrangements of the coding sequences around these regions (Fig. S4). These results suggested that a mobile element related to pM7012 had integrated into the chromosome of H8234. The gene clusters for conjugative transfer in RPME01 and pM7012 were also highly homologous (Fig. 4), although the phylogenetic positions of ParA were different (Fig. S3). When whole structures were compared between these plasmids, approximately half of RPME01 had synteny with pM7012; RPME01 looked like a fusion of a plasmid related to pM7012 with another plasmid with a different plasmid maintenance system.

---

**Fig. 4.** Gene organization for conjugative transfer genes of pM7012 and other related plasmids. Light grey arrows represent genes encoding core T4SS proteins, dark grey arrows represent genes encoding proteins for pilus assembly and mating pair stabilization, and black arrows labelled ‘D’ and ‘I’ represent *traD* (T4CP gene, originally annotated as *traG* in plasmids, except the F plasmid and pM7012) and *traI*, respectively. Graduated arrows labelled ‘hel’ and ‘topo III’ represent putative genes for helicase and topoisomerase III. White arrows represent other genes. Capital letters represent *tra* genes (e.g. ‘G’ represents *traG*). Numbers below arrows indicating *traU, traD, traC* and topoisomerase III represent per cent amino acid sequence identity between the indicated gene and the corresponding sequences of pM7012. Directions of the gene clusters are indicated by 5′ and 3′. Double slash indicates non-contiguous regions omitted in this figure. GenBank accession numbers: byi_2p, CP003091; pBVIE01, NC_009230; RPME01, NC_008826; H8234, CP005976; 1Ebn1, NC_006823; bgla_3p, NC_015378; F plasmid, NC_002483.
It has been reported that a relaxase (TraI) is necessary for conjugal transfer of plasmids and can be used for plasmid classification (García-Barría et al., 2009). However, the DNA sequence of pM7012 contained no coding sequence with any considerable homology to known relaxases. The genes for auxiliary proteins required for DNA processing in preparation for conjugative transfer in IncF plasmids, TraY and TraM (de la Cruz et al., 2010), were also not detected. Smillie et al. (2010) recognized that the T4SSs of pBVIE01, RPE01 and plasmid 1 of strain EbN1 (plasmid 1 EbN1) also do not have any relaxase homologues. Thus, the plasmids and related sequences belonging to the same T4CP cluster as pM7012 form a ‘no relaxase’ group (Fig. 5). Of course, there is a possibility that these plasmids lost unstable relaxases during the stabilization of the plasmids in their hosts, especially under laboratory conditions.

Fig. 5. Phylogenetic tree of T4CP amino acid sequences from pM7012, the related plasmids and plasmids belonging to the MOB families. Related sequences from chromosomes and ICEs were added. Parentheses after the plasmid name represent MOB group, incompatible group, size and accession number of T4CP; ND, no data. The tree was reconstructed by maximum-likelihood analysis using MEGA5 software. Bootstrap values (calculated from 1000 replicates) >50 are shown.
It was suggested that most of the proteobacterial T4SSs with no relaxase were involved in protein transport and not in conjugation (Guglielmini et al., 2011). pBVIE01, RPME01 and 1_18n1 have all been defined as non-mobile plasmids and, because of the lack of relaxase, their T4SSs were suggested to be involved in protein transport (Smillie et al., 2010). However, the highly homologous sequences of 2,4-D degradation genes on plasmids of similar size in Burkholderia strains and a Cupriavidus strain strongly suggested that pM7012-related plasmids are mobile (Sakai et al., 2007). Therefore, it is possible that the 2,4-D degradation plasmids related to pM7012 belong to an unclassified mobile plasmid group.

### Spread of 2,4-D degradation megaplasmids related to pM7012

If plasmids found in different bacterial species have the same structure, it strongly suggests that they are mobile. In a previous report (Sakai et al., 2007), strain M701 carrying pM7012 and seven other strains with largely identical 2,4-D degradation gene sequences were isolated from Japanese paddy fields. These eight strains were defined as several species of Burkholderia and a strain of Cupriavidus, and their 2,4-D degradation genes were detected on megaplasmids of similar size (~600 kb). Therefore, the similar structures of these eight plasmids strongly suggest horizontal transfer of megaplasmids with 2,4-D degradation genes.

To investigate this possibility, in addition to the 2,4-D degradation genes, we analysed the homologous sequences of traD, parA and arsA in the seven other strains. Using PCR products from pM7012 targeting these three genes as probes for Southern hybridization, signals were detected for all three genes on the megaplasmids of all other strains (Fig. 6). Sequencing of PCR products showed >98% similarity in the nucleotide sequences of the three genes between the seven megaplasmids and the corresponding sequences in pM7012 (Table S4). The fact that all eight strains have highly homologous T4CP genes suggests that these 2,4-D degradation plasmids belong to the same T4CP cluster as pM7012. Considering that the plasmids were found in distinct strains of Burkholderia and Cupriavidus, these results indicate that the megaplasmids in our study are mobile (or at least were mobile in the past).

The nucleotide sequences of the 2,4-D degradation genes in pM7012 were >98% similar to those of B. tropica strain RASC (Sakai et al., 2007), isolated from return-activated sludge in the United States (Tonso et al., 1995). The nucleotide sequences of tfdC, traD, parA and arsA in the megaplasmid of strain RASC (~900 kb) were mostly identical (>97%) to the corresponding gene sequences in pM7012 (Fig. 6, Table S4). These results suggest that the megaplasmid from strain RASC is also related to pM7012 and that 2,4-D degradation megaplasmids related to pM7012 are not found only in Japan.

To test the mobility of pM7012 and pM7012-like plasmids, filter-mating experiments were conducted several times using plasmid-free derivatives of Burkholderia sp. strain NK8 (Francisco et al., 2001) and C. necator strain NH9 (Ogawa & Miyashita, 1995) as recipients, and strains M701 and other strains with pM7012-like plasmids obtained in the previous work (Sakai et al., 2007) as donors. However, no transconjugant was detected in this experimental system. The Southern hybridization experiments and sequencing results clearly indicate that the megaplasmids with 2,4-D degradation genes moved between different species in the past. These results suggest that these megaplasmids have lost key genes (maybe relaxases) for conjugative transfer or that they require other mating conditions that were absent in our experimental design. In the case of megaplasmid pSymA having all genes for conjugative transfer in Sinorhizobium meliloti 1021, transfer is undetectable under common laboratory conditions because of the repression of conjugation genes by a transcriptional repressor encoded by the rctA gene (Pérez-Mendoza et al., 2005). There is a possibility that translocated relaxase genes in another replicon in the host cell mediate the transfer of pM7012-related plasmids. Whilst it has been proposed that all mobile plasmids, including non-self-transmissible plasmids, have relaxase...
(Smillie et al., 2010), it was reported recently that a plasmid without any conjugative transfer genes was transferred with the help of a coexisting plasmid carrying homologues of conjugative transfer genes (Li et al., 2013). Another hypothesis is that these 2,4-D degradation megaplasmids or plasmids in the 'no relaxase' group have either a novel relaxase gene or a novel mechanism for plasmid transfer, although expression of the system might be repressed in our filter-mating condition. Comparing the putative coding sequences around the T4SSs of pM7012 with the related replicons containing homologous tra genes revealed a putative gene for DNA topoisomerase III, which belongs to the topoisomerase IA group, in all sequences (Fig. 4). The relaxase domains of TraI in RP4 and the F plasmid can cleave ssDNA and attach their tyrosine residue to the 5' end of DNA at the nicking site, as does topoisomerase IA (Bugreev & Nevinsky, 2009; de la Cruz et al., 2010; Pansegrau et al., 1993). Thus, the putative topoisomerase III in the pM7012-related plasmids might perform the role of relaxase during plasmid transfer. To confirm these hypotheses, further analyses are necessary.

Most well-studied catabolic plasmids belong to the IncP-1, IncP-2, IncP-7 and IncP-9 incompatibility groups (van der Meer, 2008). However, sequence analysis of pM7012 indicates that pM7012 and related plasmids with 2,4-D degradation genes cannot be classified in any incompatibility groups. To the best of our knowledge, this is the first report suggesting the spread of catabolic genes by large (>500 kb) unclassified megaplasmids that apparently do not belong to the Inc groups above.

It is noteworthy that another family of megaplasmids has been described in the genus Cupriavidus, which share a very similar core structure in the group and carry different accessory genes encoding hydrogenotrophy (pHG1), nitrogen fixation (pRALTA) or metal resistance (pMOL28) (Mergeay et al., 2009). The acquisition of accessory genes by way of megaplasmids could be a trait for the family Burkholderiaceae. The present study suggests that the plasmids of the pM7012 group, with accessory genes such as 2,4-D degradation genes, could contribute to the catabolism of Burkholderia and related genera under various environments.

ACKNOWLEDGEMENTS

We are grateful to Dr Terence L. Marsh of Michigan State University for providing B. tropica strain RASC. This work was supported by a Grant-in-Aid from the Ministry of Agriculture, Forestry and Fisheries of Japan (no. eDNA-10-102-1, Soil eDNA).

REFERENCES


Poh, R. P., Smith, A. R. & Bruce, I. J. (2002). Complete characterization of Tn5330 from Burkholderia cepacia strain 2a (pJBJ1) and studies of 2,4-dichlorophenoxyacetate uptake by the organism. Plasmid 48, 1–12.


Edited by: E. Top