ORIGINAL ARTICLE

Isolation and sequence analysis of pCS36-4CPA, a small plasmid from *Citrobacter* sp. 36-4CPA

CrossMark

Natalia V. Zharikova [a,1], Timur R. Iasakov [a,*,1], Boris K. Bumazhkin [b],
Ekaterina O. Patutina [b], Evgeniia I. Zhurenko [a], Vladislav V. Korobov [a],
Alina I. Sagitova [a], Boris B. Kuznetsov [b], Tatiana V. Markusheva [a]

[a] *Ufa Institute of Biology, Russian Academy of Sciences, Russia*
[b] *Bioengineering Center, Russian Academy of Sciences, Russia*

Abstract A small plasmid designated pCS36-4CPA with a size of 5217 base pairs and G-C content of 50.74% was isolated from *Citrobacter* sp. 36-4CPA. The origin of replication (*ori*) of the plasmid was identified as a region of about 800 bp in length with an identity of 67.1% to the ColE1 plasmid at the nucleotide level. The replication region contained typical elements of ColE1-like plasmids: RNA I and RNA II with their corresponding −10 and −35 boxes, a single-strand initiation site (*ssi*), and a lagging-strand termination site (*terH*). As seen in other ColE1-like plasmids, pCS36-4CPA carried mobilisation machinery that include *mobABCD* genes but it did not possess the *rom* gene. Analysis of the multimer resolution site (*mrs*) was performed and XerC and XerD binding sites were identified. Also, the 70-nt transcript Rcd of pCS36-4CPA was predicted and similarity of the transcript's secondary structure with those of the ColE1-family was shown. The cargo module of pCS36-4CPA contained three open reading frames (ORFs). Two of them (ORF5 and ORF6) showed no significant homology to any known gene sequences but contained putative THAP DNA-binding (DBD) and type II restriction endonuclease *Eco*O109I domains. The seventh open reading frame (ORF7) encodes YhdJ-like DNA modification methylase. The region highly homologous to pCS36-4CPA was found in the *Salmonella* phage SE2 genome.

© 2016 The Authors. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

* Corresponding author. Tel./fax: + 7 347 2356247.
E-mail addresses: iasakov@anrb.ru, iasakovtimur@gmail.com (T.R. Iasakov).
[1] Natalia V. Zharikova and Timur R. Iasakov both contributed equally to this article.
Peer review under responsibility of King Saud University.

1. Introduction

Citrobacter is a polyphyletic genus of Gram-negative aerobic rod-shaped bacteria that belongs to the family *Enterobacteriaceae*. This bacterial family is widely spread in the environment, including soil, food, waste water, and human or mammalian intestine (Sakimbaeva, 1985; Mohanty et al., 2007; Petty et al., 2010; Kumar et al., 2013).

It is known that *Citrobacter* can be used in glucosamine (GlcN) production (Kim et al., 2012), shikimic acid overproduction (Ghosh et al., 2012) or bacteriocin synthesis (Shanks et al., 2012). Also, the production of 1,3-propanediol (PD) from biodiesel-derived crude glycerol by a *Citrobacter freundii* was investigated (Anand and Saxena, 2012; Metsoviti et al., 2013).

The xenobiotic degradation is an important feature of a *Citrobacter* spp. *Citrobacter* are known to be able to degrade a range of aromatic compounds, such as phenol (Mohite et al., 2011), benzoate and hydroxyl benzoic acid (Selvakumaran et al., 2011). Also, efficient degradation of polycyclic aromatic hydrocarbons (PAH's) by pure *Citrobacter* strains was discovered, for example, biphenyl and phenanthrene biodegradation under anaerobic conditions (Grishchenkov et al., 2002; Li and Zhu, 2012). The PAH-degrading (acenaphthylene, phenanthrene, bena[a]anthracene, benzo[a]pyrene and others) microbial consortia obtained from petroleum-contaminated soils included *Citrobacter* (Wu et al., 2013). In addition to aromatic compounds, *Citrobacter* are capable of degrading pesticides and biocides such as methyl parathion, chlorpyrifos and tributyltin (Pino and Peñuela, 2011; Sakultantimetha et al., 2009). Also, synthetic dyes pose a serious threat to the environment (Saratale et al., 2011). *Citrobacter* spp. exhibited azo-, anthraquinone, indigo, and triphenylmethane dye decolourisation activity (Wang et al., 2009; Oh et al., 2011; Jang et al., 2005).

In humans, some *Citrobacter* species are hospital-acquired pathogens and can cause diarrhoea (Bai et al., 2012), meningitis (Tan et al., 2010) or neonatal meningitis (Lipsky et al., 1980), perineal ecthyma gangrenosum (Reich et al., 2004). Other species such as *Citrobacter rodentium* can cause colitis in mice (Fredrickson et al., 2010). *Citrobacter koseri* and *C. freundii* are the major human pathogens among *Citrobacter* species (Samonis et al., 2009; Lin et al., 2011).

However, little is known about ColE1-family plasmids from the genus *Citrobacter*, except for the ColA plasmid, which was fully characterised (Morlon et al., 1988). The investigations in that field focused mainly on the plasmid-encoded AmpC β-lactamases (Barlow and Hall, 2002; Sekiguchi et al., 2008). In this paper, we report the sequence analysis of the novel cryptic plasmid pCS36-4CPA which was isolated from *Citrobacter*.

2. Materials and methods

2.1. Media and growth conditions

Luria–Bertani broth medium (LB broth: 10 g/l bactotryptone (Difco, USA), 5 g/l bacto yeast extract (Difco, USA), and 10 g/l NaCl, pH 7.5) was used for the cultivation of *Citrobacter* sp. 36-4CPA. The strain was incubated in 250 ml flasks containing 100 ml medium at 30 °C on an orbital shaker (180 rpm) for 16–18 h.

2.2. Isolation and identification of the strain

Citrobacter sp. 36-4CPA was isolated from hydrocarbon-contaminated soil samples. The isolate was identified by 16S rRNA gene phylogenetic analysis as a member of the polyphyletic genus *Citrobacter*. Genomic DNA was isolated from

bacterial culture, as previously described (Boulygina et al., 2002). The 16S rRNA partial gene (1480 bp) was amplified by using universal primers 27f and 1492r (Edwards et al., 1989). PCR products were purified by LMP (low melting point) agarose and agarase AgarACE™ (Promega). The partial 16S rRNA gene was sequenced using universal primers 27f and 1492r with a BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) on a 3730 DNA Analyser (Applied Biosystems) according to the manufacturer's instructions.

2.3. Isolation, cloning, and sequencing of pCS36-4CPA

Plasmid DNA was isolated by Birnboim–Doly alkaline lysis method (Birnboim and Doly, 1979). Purified through LMP (low melting point) agarose and agarase AgarACE™ (Promega), plasmid DNA was analysed by 1% agarose gel electrophoresis. The plasmid was digested with the restriction endonucleases *Kzo*9I, *Rsa*I and *Alu*I (Promega). Later, *Kzo*9I, *Rsa*I and *Alu*I fragments were cloned into pGEM-3Zf(+) (Promega). The recombinant plasmid DNA was then transformed into *Escherichia coli* DH5α and isolated with a Wizard MaxiPrep kit (Promega). DNA sequencing was performed using universal M13 primers with a BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) on a 3730 DNA Analyser (Applied Biosystems) according to the manufacturer's instructions.

2.4. Sequence annotation and phylogenetic analysis of pCS36-4CPA

Prediction of open reading frames (ORF) and plasmid map visualisation was performed using SnapGene Viewer (SnapGene® software (from GSL Biotech; available at snapgene.com)). Annotation of DNA sequences and protein motifs was performed by the Basic Local Alignment Search Tool (BLAST) software and Conserved Domain Database (CDD) from the NCBI (Altschul et al., 1990; Marchler-Bauer et al., 2013). Multiple sequence alignments and phylogenetic analyses were conducted in MEGA4 (Tamura et al., 2007). The evolutionary history was inferred using the Neighbour-Joining method (Saitou and Nei, 1987). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test is shown next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and are in the units of the number of base substitutions per site. Predicted secondary structure of DNA and RNA molecules was performed by the Mfold web server (Zuker, 2003). Sequence logo was created by WebLogo 3.3 software (Schneider and Stephens, 1990; Crooks et al., 2004).

3. Results and discussion

3.1. Isolation and identification of strain

Bacterial strain 36-4CPA was isolated from hydrocarbon-contaminated soil samples of the Ufa industrial area. The partial sequence of the 16S rRNA gene with a length of 1480 bp was amplified and sequenced to identify the
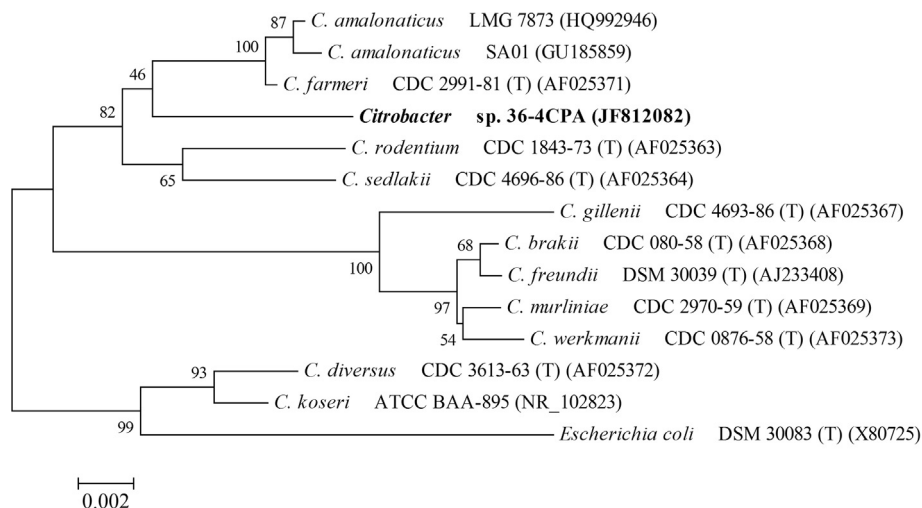
Figure 1 Phylogenetic tree based on 1480 nucleotides of 16S rRNA gene sequences. The bar represents 2 nucleotide substitutions per 1000 nucleotides. Nodal robustness of the tree was assessed using 1000 bootstrap replicates. The NCBI GenBank accession number for each type strain tested is shown in parentheses. 16S rRNA gene of *Escherichia coli* DSM 30083 was used as outgroup.

phylogenetic position of the 36-4CPA strain. Preliminary analysis of the sequence was performed with the rRNA database RDP (Cole et al., 2014) and BLAST (Altschul et al., 1990; Marchler-Bauer et al., 2013). The sequence analysis showed a high level of homology to twelve 16S rRNA gene sequences of the *Citrobacter* genus. Multiple sequence alignments and phylogenetic analyses revealed that 16S rRNA sequence of 36-4CPA strain clustered with type *Citrobacter farmeri* and *Citrobacter amalonaticus* strains (Fig. 1). Identity at the nucleotide level between 36-4CPA and these strains was 98.9% and 98.7%, respectively. According to 16S rRNA gene comparison, strain 36-4CPA belonged to the genus *Citrobacter*. The 16S rRNA gene sequence of *Citrobacter* sp. 36-4CPA determined in this investigation was deposited in GenBank under the accession number JF812082.

3.2. DNA sequence and organisation of pCS36-4CPA

Agarose gel electrophoresis of total genomic DNA revealed the presence of a small plasmid with a length of approximately 5 kb. The plasmid was named pCS36-4CPA. We constructed an incomplete cloned library of pCS36-4CPA plasmid DNA which contains overlapping restriction fragments resulting from digestion by *Rsa*I, *Alu*I and *Kzo*9I. A complete sequence was obtained by assembling individual reads with sequencing by primer walking strategy. The entire plasmid comprised 5217 bp with a G-C content of 50.74%. A map of plasmid pCS36-4CPA is shown in Fig. 2. Also, in the same figure, mobilisation, plasmid stability regions and open reading frames are shown. The complete nucleotide sequence of the pCS36-4CPA plasmid was deposited in GenBank under the accession number KM373703.

3.3. Replication module

BLAST queries showed 88.0% and 87.8% similarity of the region from 134 to 947 nt of pCS36-4CPA with the replication origins of plasmids pRK2 (AY639886) and p29807

(AJ132618), respectively. They are both members of ColE1-family plasmids as well as p15A (V00309), pMB1 (FJ437239), RSF1030 (J01784), CloDF13 (X04466) and ColA (M37402). All plasmids of this family have a replication region (*ori*) similar to ColE1 (Strauch et al., 2000; Štěpánek et al., 2005; Camps, 2010; Morlon et al., 1988). According to those data, we hypothesised that pCS36-4CPA belongs to the ColE1 family. The sequence analysis showed 67.1% similarity between *ori* pCS36-4CPA and ColE1. Based on that result pCS36-4CPA belonged to the ColE1 plasmid family. Also, the comparison of *ori* regions of pCS36-4CPA and well characterised *C. freundii* CA31 plasmid ColA was performed. Their sequence identity was 77.2%. Nevertheless, phylogenetic analysis revealed that the replication origins of two *Citrobacter* plasmids, ColA and pCS36-4CPA, belong to different phylogenetic groups (Fig. 3C).

As is commonly accepted, representatives of the ColE1 plasmid family comprised of the typical *ori* elements. Among them, are two RNAs with regulatory function: pre-primer RNA II and antisense RNA I, *ssi*-site and a *terH* sequence. The function of these partly overlapping RNAs is the regulation of plasmid copy number. Pre-primer RNA II is transcribed from *ori* and forms an *R-loop* after hybridisation with plasmid DNA template. This structure would be further processed by RNaseH to create a 3′-OH end for DNA polymerase I that initiates prime leading-strand synthesis (Camps, 2010; Rozhon et al., 2010). In our work, a preprimer RNA II of pCS36-4CPA with a length of 521 bp was predicted. The promoters (P2) of RNA II synthesis, −35 and −10 boxes, were found (Fig. 3B).

It is known that promoters (P2) are highly conserved (Rozhon et al., 2006). But comparison of the pCS36-4CPA-35 box sequence with other ColE1-family plasmids revealed the replacement of GA with CC at position 170–171. Also, the replacement of a C for an A was found at position 194 of the −10 box sequence of the pCS36-4CPA (Fig. 4). However, the ColA plasmid possessed −35 and −10 boxes that were absolutely identical to ColE1 ones.
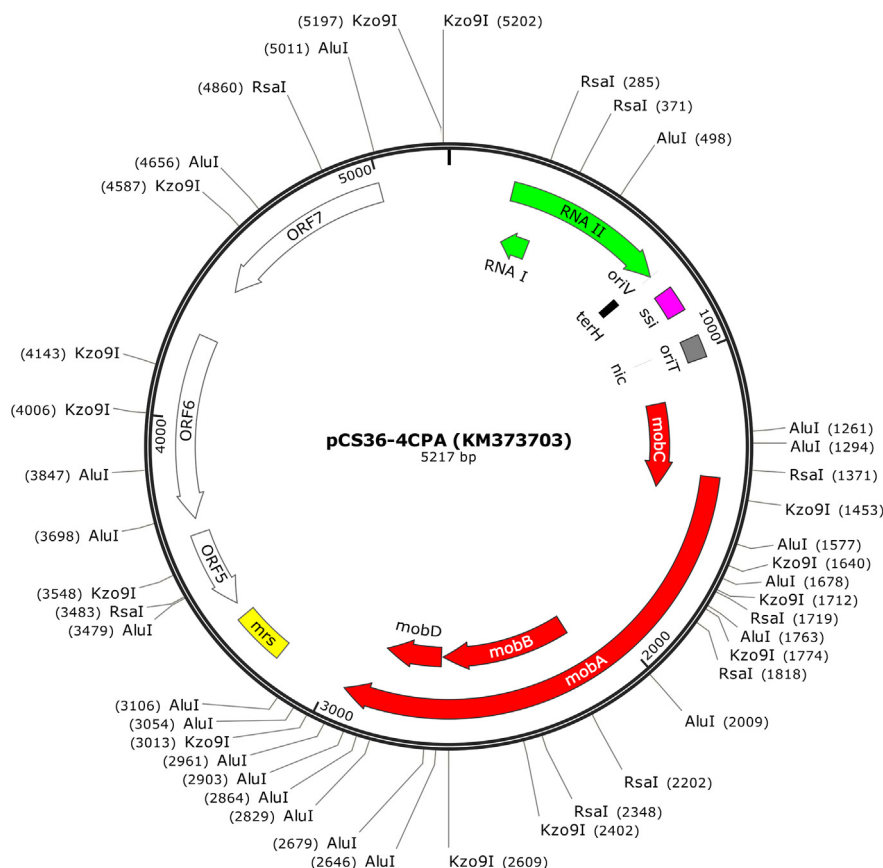
Figure 2 Plasmid map of pCS36-4CPA with restriction sites of the endonucleases *Alu*I, *Kzo*9I and *Rsa*I. The regions transcribed as RNAI and RNAII are shown as green arrows. The single-strand initiation site (*ssi*) and multimer resolution site (*mrs*) are coloured in pink and yellow, respectively. The mobilisation machinery genes are depicted as red arrows. The open reading frames ORF5, ORF6 and ORF7 are shown as unpainted arrows. Abbreviations: *oriV*, origin of replication; *oriT*, origin of transfer; *nic*, sequence-specific cleavage site.

The antisense RNA I in pCS36-4CPA *ori* was predicted in addition to RNA II. The 104 nt-long antisense RNA I of pCS36-4CPA is shown in Fig. 3B. The negative regulation of replication occurred by the antisense RNA I that was transcribed from the opposite DNA strand. RNA I fully overlaps with the 5′-end of RNA II. The interaction between two complementary sequences of RNA transcripts prevents primer-initiated replication (Davison, 1984; Helmer-Citterich et al., 1988; Wu and Liu, 2010). The RNA I promoters (P1) were found and compared with those of other ColE1-like plasmids. The −35 box of the RNA I antisense promoter P1 was identical to each member of the ColE1-family, except for pRK2. The −10 box sequence had one replacement, C to A at position 320, which was identical to p29807, p15A, pRK2, and RSF1030 (Fig. 4).

Both pre-primer RNA II and antisense RNA I forms three stem-loops (SL1, SL2 and SL3) at the 5′-end which mediates their interactions (Camps, 2010). The predicted stem-loops of pCS36-4CPA were partially different from the classical ColE1-type. The loops of SL1, and SL2 were identical to each other and consisted of four nucleotides (G-G-G-G). The third loop was formed by seven nucleotides (G-U-A-C-A-A-G) (Fig. 3D), while all three loops in the ColE1-family plasmids consisted of six nucleotides each (Rozhon et al., 2006). It is surprising that loops of SL1 and SL2 of ColA were identical to those of pCS36-4CPA (G-G-G-G). The SL3 loop differed from both pCS36-4CPA and ColE1. Nevertheless, the −10

and −35 promoter sequences of regulatory RNAs were not identical (except −35 box of RNA I) to each other.

The hybrid RNA II-RNA I duplex is stabilised by the Rop protein (Camps, 2010). According to Di Primo (2008), dissociation of that complex occurs more than two orders of magnitude slower with the Rop protein. The Rop protein is encoded by *rom* (RNA one inhibition modulator gene) (Chan et al., 1985). No significant homology was found between the ColE1 *rom* gene and any predicted ORF of pCS36-4CPA. Our attempts to predict that gene in pCS36-4CPA failed. The absence of functional *rom* gene homologues in several ColE1-family plasmids, such as pHW15 and pRK10 (Rozhon et al., 2006; Ibrahim et al., 2009), was previously shown. The lagging-strand synthesis occurred in the single-strand initiation site (*ssi*) and lagging-strand termination site (*terH*). The 82-nt *ssi* signal was located within the 60 nt upstream of *oriV*. The function of that sequence is binding DnaB helicase and DnaG primase. It is known that the DNA sequence of *ssi* folds into the stem-loop structure (Rozhon et al., 2006). The modelled stem-loop of pCS36-4CPA is shown in Fig. 3E. As generally known, plasmid replication starts in a unique site of the DNA sequence called *oriV*. The *oriV* of plasmid pCS36-4CPA was predicted at position 726 by sequence comparison with these ColE1 and other plasmids of the ColE1-family (p29807, p15A and others). They all have absolutely identical sequence at *oriV*. Thus, the pCS36-4CPA origin of replication included all of the same regulatory
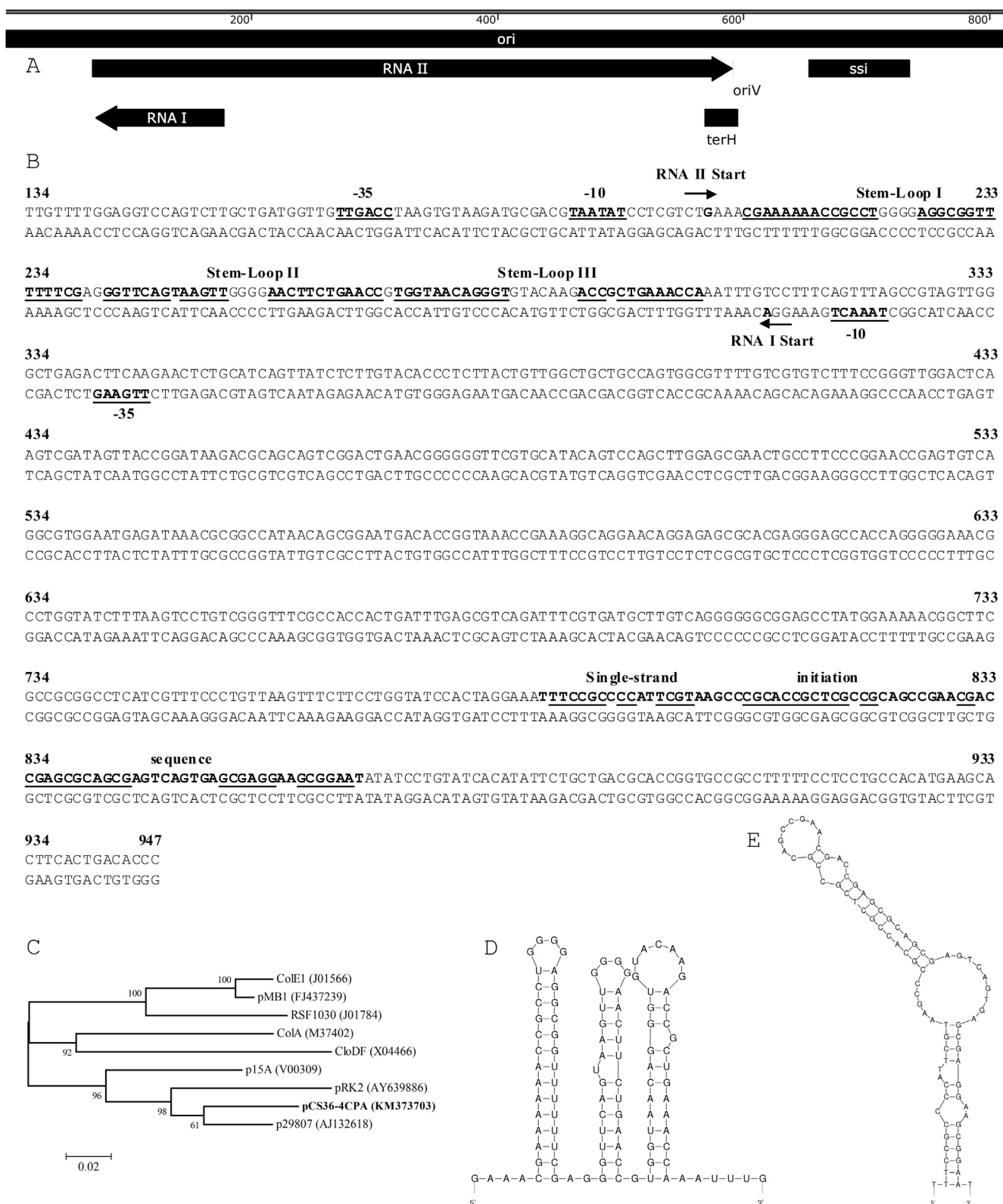
A

200'   400'   600'   800'

ori

RNA II

oriV

ssi

RNA I

terH

B

RNA II Start

134   -35   -10   →   Stem-Loop I   233

TTGTTTTGGAGGTCCAGTCTTGCTGATGGTTG**TTGACC**TAAGTGTAAGATGCGACG**TAATAT**CCTCGTCT**G**AAA**CGAAAAAACCGCCT**GGGG**AGGCGGTT**

AACAAAACCTCCAGGTCAGAACGACTACCAACAACTGGATTCACATTCTACGCTGCATTATAGGAGCAGACTTTGCTTTTTTGGCGGACCCCTCCGCCAA

234   Stem-Loop II   Stem-Loop III   333

TTTTCGAG**GGTTCAGTAAGTT**GGGG**AACTTCTGAACC**G**TGGTAACAGGGT**GTACAAG**ACCGCTGAAACCA**AATTTGTCCTTTCAGTTTAGCCGTAGTTGG

AAAAGCTCCCAAGTCATTCAACCCCTTGAAGACTTGGCACCATTGTCCCACATGTTCTGGCGACTTTGGTTTAAAC**A**GGAAAG**TCAAAT**CGGCATCAACC

RNA I Start

-10

334   433

GCTGAGACTTCAAGAACTCTGCATCAGTTATCTCTTGTACACCCTCTTACTGTTGGCTGCTGCCAGTGGCGTTTTGTCGTGTCTTTCCGGGTTGGACTCA

CGACTCT**GAAGTT**CTTGAGACGTAGTCAATAGAGAACATGTGGGAGAATGACAACCGACGACGGTCACCGCAAAACAGCACAGAAAGGCCCAACCTGAGT

-35

434   533

AGTCGATAGTTACCGGATAAGACGCAGCAGTCGGACTGAACGGGGGGTTCGTGCATACAGTCCAGCTTGGAGCGAACTGCCTTCCCGGAACCGAGTGTCA

TCAGCTATCAATGGCCTATTCTGCGTCGTCAGCCTGACTTGCCCCCCAAGCACGTATGTCAGGTCGAACCTCGCTTGACGGAAGGGCCTTGGCTCACAGT

534   633

GGCGTGGAATGAGATAAACGCGGCCATAACAGCGGAATGACACCGGTAAACCGAAAGGCAGGAACAGGAGAGCGCACGAGGGAGCCACCAGGGGGAAACG

CCGCACCTTACTCTATTTGCGCCGGTATTGTCGCCTTACTGTGGCCATTTGGCTTTCCGTCCTTGTCCTCTCGCGTGCTCCCTCGGTGGTCCCCCTTTGC

634   733

CCTGGTATCTTTAAGTCCTGTCGGGTTTCGCCACCACTGATTTGAGCGTCAGATTTCGTGATGCTTGTCAGGGGGGCGGAGCCTATGGAAAAACGGCTTC

GGACCATAGAAATTCAGGACAGCCCAAAGCGGTGGTGACTAAACTCGCAGTCTAAAGCACTACGAACAGTCCCCCCGCCTCGGATACCTTTTTGCCGAAG

734   Single-strand   initiation   833

GCCGCGGCCTCATCGTTTCCCTGTTAAGTTTCTTCCTGGTATCCACTAGGAAA**TTTCGCCCCATTCGTAAGCCCGCACCGCTCGCCGCAGCCGAACGAC**

CGGCGCCGGAGTAGCAAAGGGACAATTCAAAGAAGGACCATAGGTGATCCTTTAAAGGCGGGGTAAGCATTCGGGCGTGGCGAGCGGCGTCGGCTTGCTG

834   sequence   933

CGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAATATATCCTGTATCACATATTCTGCTGACGCACCGGTGCCGCCTTTTTCCTCCTGCCACATGAAGCA

GCTCGCGTCGCTCAGTCACTCGCTCCTTCGCCTTATATAGGACATAGTGTATAAGACGACTGCGTGGCCACGGCGGAAAAAGGAGGACGGTGTACTTCGT

934   947

CTTCACTGACACCC

GAAGTGACTGTGGG

C

100 — ColE1 (J01566)

100 — pMB1 (FJ437239)

— RSF1030 (J01784)

— ColA (M37402)

92 — CloDF (X04466)

— p15A (V00309)

— pRK2 (AY639886)

96

98 — **pCS36-4CPA (KM373703)**

61 — p29807 (AJ132618)

0.02

D

5' G–A–A–A–C–G–A–G–G–C–G–U–A–A–A–U–U–U–G 3'

E

Figure 3 The replication origin of pCS36-4CPA (KM373703). Structure (A), sequence (B) and important features of the origin of replication of pCS36-4CPA. The -10, -35 signals and nucleotides forming the stems of the structures of (D) and (E) are bold and underlined. Phylogenetic tree of ColE1 family plasmids replication origins (C). Predicted secondary structure of the 5′ part of RNA II (D). Predicted secondary structure of the single-strand initiation site (E).
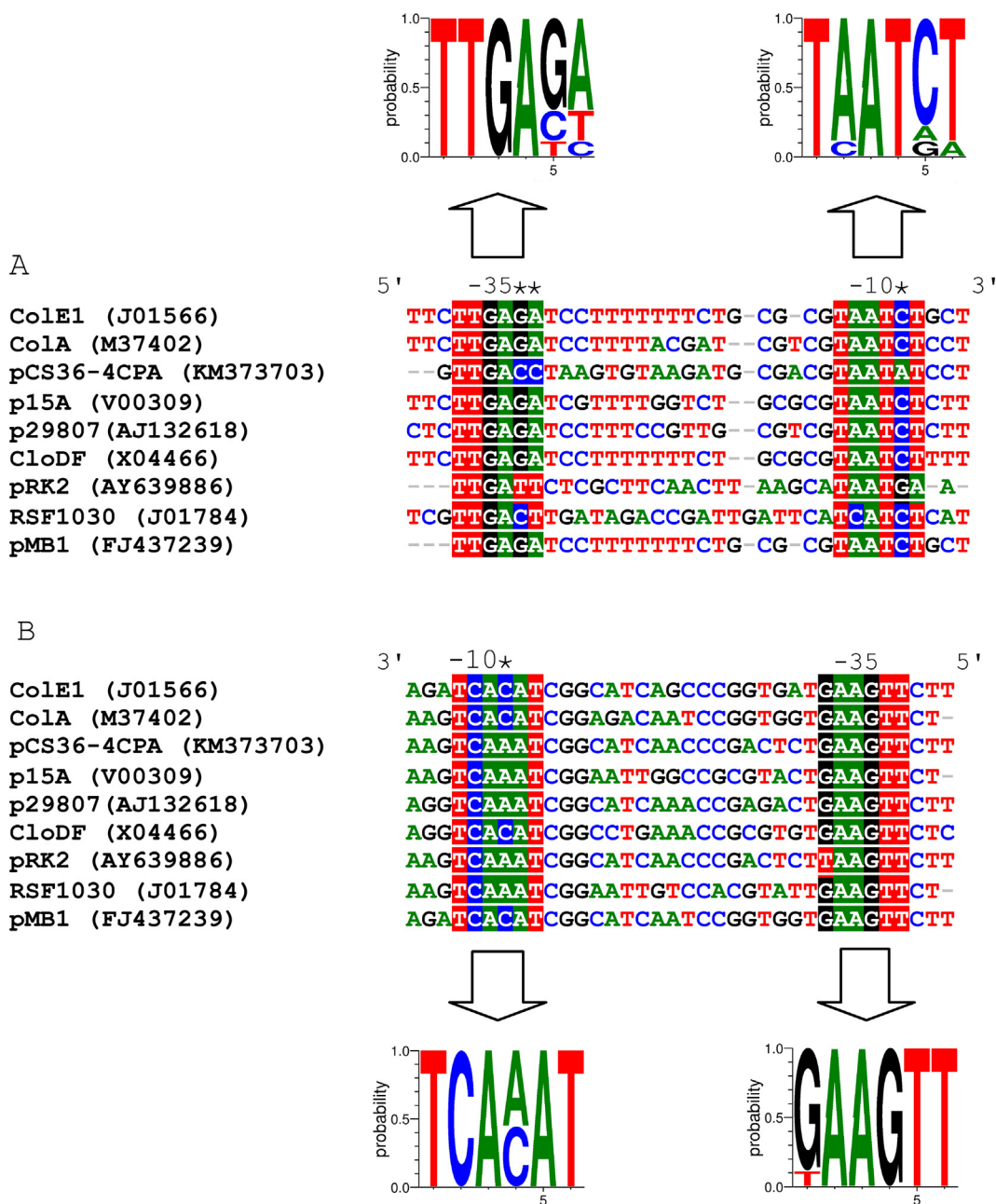
Figure 4 An alignment of ColE1 type plasmids RNA II (A) and RNA I (B) −35 and −10 boxes. The −35 and −10 boxes are highlighted and shown in sequence logo. The asterisks (*) indicate the positions where nucleotides have been replaced.

elements as ColE1: RNA I and RNA II with their corresponding −10 and −35 boxes, a single-strand initiation site (*ssi*), and a *terH* sequence (Fig. 3A). According to the results of the *in silico* analysis of the replication region, plasmid pCS36-4CPA belongs to the ColE1-family as well as ColA. Meanwhile, *ori* of both plasmids pCS36-4CPA and ColA have polyphyletic origin.

3.4. Propagation module (mobilisation machinery of pCS36-4CPA)

The four open reading frames (ORFs) were identified within the region 1139-2948 in pCS36-4CPA. Two of them entirely

and one partly overlapped the ORF by 1548 nt, with an identity level of 97–99% with the *mobA* gene of the plasmids pSP291-3 (CP004094), pIGRW12 (EF088686), pEC01 (AB117929), pS51B (AB583678) and others. Three other ORFs were recognised as *mobC*, *mobB* and *mobD* genes according to the results of the BLAST search. Thus, *mobA–D* genes were classified as the MOB machinery of pCS36-4CPA (Fig. 2). The function of that region is plasmid mobility. Plasmids with a minimal gene set (MOB only or MOB + type IV coupling protein (T4CP)) were named mobilisable, unlike conjugative (self-transmissible) plasmids with a full set (MOB + T4CP + type IV secretion system (T4SS)) of components of a conjugative apparatus (Garcillán-Barcia et al., 2009;

Smillie et al., 2010). Mobility of plasmids closely correlated to their size. Most mobilisable plasmids are small plasmids about 5 kb in size. That size allows them to carry the basic replication module and a few adaptive genes according to the basic trend of mobilisable plasmids (Garcillán-Barcia et al., 2011). Only MOB module components without T4CP + T4SS were predicted in pCS36-4CPA. Thus, the pCS36-4CPA plasmid could be classified as mobilisable.

The MOB machinery codes for Mbe-like proteins that play a key role in the mobilisation process. A full set of *mbe*-genes

allows plasmid mobilisation. Relaxase MbeA (=MobA) and accessory protein MbeC (=MobC) possess the major functions in plasmid mobilisation. MbeB (=MobB) and MbeD (=MobD) proteins are part of relaxosome and putative entry exclusion proteins, respectively (Varsaki et al., 2012). In pCS36-4CPA, we found the presence of *mbe*-like genes (*mobC*, *mobB* and *mobD*) and the absence of an *mbeE*-like gene. These genes play non-essential roles in the mobilisation process. Plasmids with analogous gene organisation of MOB (ColK, ColA and other) form a separate subgroup within the
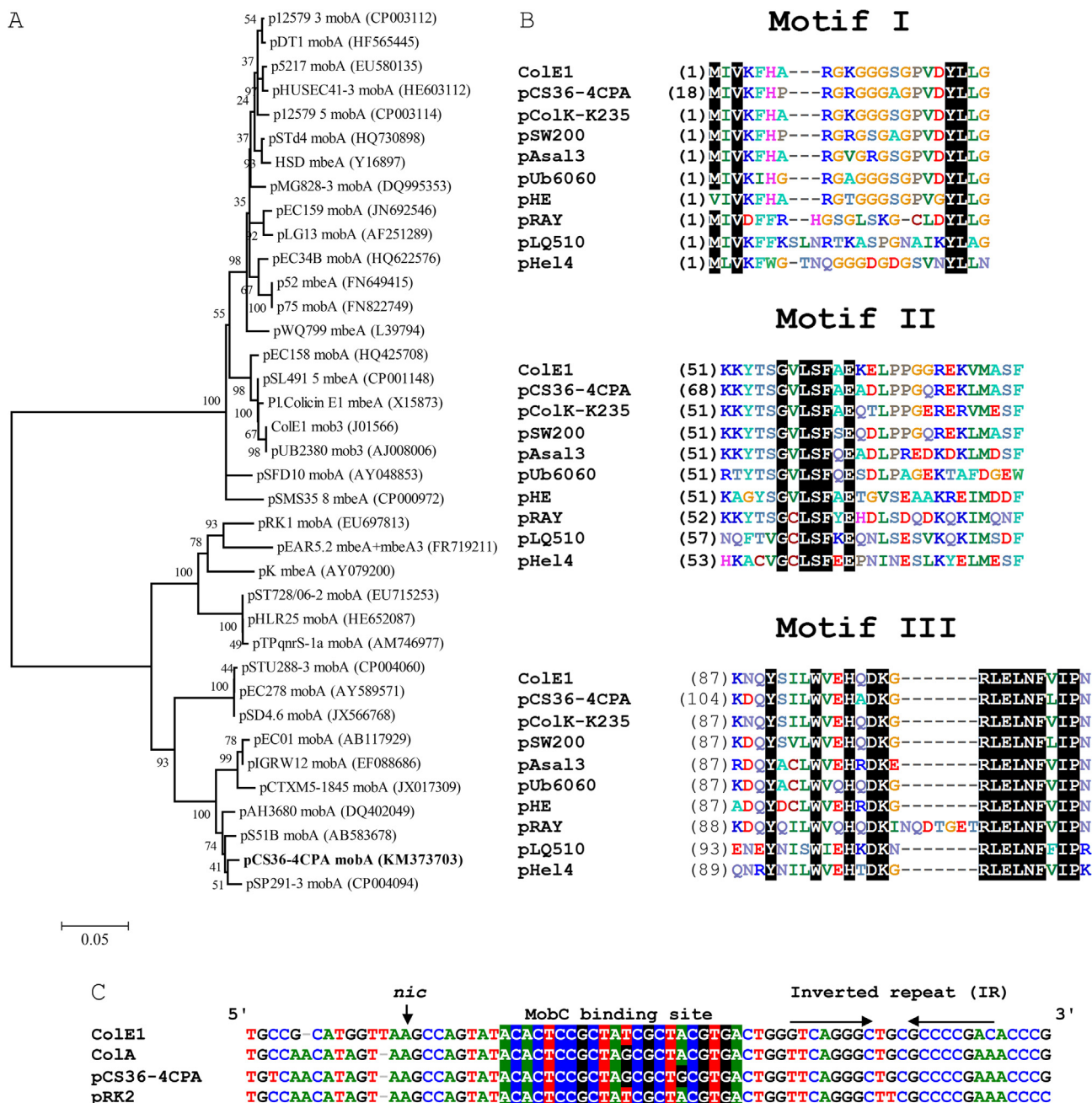

Figure 5 Neighbor-joining dendrogram of aligned relaxases closely related to pCS36-4CPA sequence (A). Alignment of ColE1-superfamily relaxases motifs I–III (B). The conservative motifs are highlighted. An alignment of predicted ColE1-like plasmids *oriT* region (C). The putative MobC binding sites are highlighted. The *nic* position and inverted repeats are indicated by a vertical and horizontal arrow, respectively.

ColE1-superfamily (Francia et al., 2004). It is known that relaxase is encoded by the *mobA* gene. In common, relaxase is a multidomain protein which plays a key role in plasmid mobilisation. Analysis of the relaxase sequences related to pCS36-4CPA revealed that the plasmid is phylogenetically divergent from ColE1 cluster (Fig. 5A). DNA sequence identity between pCS36-4CPA and ColE1 plasmid relaxases was 63.6% only. However, we attempted to analyse the conservative motifs of pCS36-4CPA MobA because relaxase amino acid sequences are used for the plasmid classification (Francia et al., 2004). *In silico* analysis of the pCS36-4CPA translated *mobA* sequence revealed an identical HEN motif of the ColE1-superfamily relaxases MOB$_{HEN}$ family (Fig. 5B). According to updated relaxase classification (Garcillán-Barcia et al., 2009), MOB$_{HEN}$ belongs to the MOB$_{P5}$ clade of the MOB$_P$ cluster. Thus, pCS36-4CPA, like ColE1, belongs to the MOB$_P$ cluster according to relaxase classification. This observation does not preclude to a common universal conjugative process.

Also, the MOB region contains an origin of conjugative transfer (*ori*T) and *nic* site. The high level of similarity between the *ori*T regions of ColA and ColE1 was previously described by Morlon et al. (1988). Multiple alignment of ColE1 *ori*T with pCS36-4CPA, pRK2, ColA and p29807 revealed closely related sequences except p29807 (Fig. 5C). CloDF13 plasmid was excluded from the analysis because Núñez and de la Cruz (2001) showed the position of the CloDF13 *nic* in the sequence 5′-GGGTG/GTCGGG-3′, which differed from ColE1. The putative *nic* site was strongly conservative. Highly similar DNA sequences in the *ori*T region and *nic* site in both *Citrobacter* plasmids explain their crucial role in dissemination between bacteria. MobC binding site and inverted repeat (IR) were predicted in the *ori*T region. The inverted repeat sequences of both pCS36-4CPA and ColA plasmids were identical to each other.

3.5. Stability module (multimer resolution site)

In silico analysis of 3175–3340 region (166 bp length) revealed the presence of a multimer resolution site (*mrs*) in pCS36-4CPA. It is commonly accepted that the site-specific recombination at the *mrs* resolves dimers and multimers to the monomeric state. That process increases the plasmid stability, preventing uncontrolled plasmid multimerisation. Such multimer formation leads to so-called "dimer catastrophe" (Summers et al., 1993). The multimer resolution site (also known as *cer*) of ColE1 was previously well characterised. Analogous regions were found in other plasmids like ColK, CloDF13 and ColA, among others (Summers et al., 1985; Morlon et al., 1988). An alignment of the pCS36-4CPA *mrs* sequence with those of several plasmids of the ColE1 family is shown in Fig. 6A. The process of site-specific recombination is promoted by XerC and XerD proteins together with accessory proteins ArgR and PepA. DNA-binding proteins XerC, XerD and ArgR have conserved sites in the DNA sequence within *mrs* (Hodgman et al., 1998). Sequence analysis of pCS36-4CPA revealed the presence of all these highly conserved sites in the *mrs* sequence. The comparison of

chromosomal XerC and XerD binding sites in 28-nucleotide motif (*dif*) in proteobacteria revealed that the XerD binding site was much more conserved (Carnoy and Roten, 2009). This output correlates well with the results of the *mrs* sequence alignment of our study. XerD binding sites were identical between closely related ColE1-like plasmids. XerC binding sites were much less conserved than XerD binding sites. Interestingly, the consensus sequence of proteobacterial XerD binding sites was identical to pCS36-4CPA and other ColE1-like plasmids, except for G/T exchange (Fig. 6A).

At the same time, in addition to site-specific recombination, the cell cycle regulator Rcd transcribed from the *mrs* sequence is involved in plasmid stability. According to the Rcd checkpoint hypothesis, the division of bacterial cells carrying plasmids in the multimer state is delayed (Sharpe et al., 1999; Balding et al., 2006). This non-coding RNA transcript (Rcd) prevents the division of cells containing plasmid dimmers. Inhibition of cell division occurs through indole production stimulation (Gaimster and Summers, 2015). Analysis of the predicted *cer*-like sequence of pCS36-4CPA revealed the presence of short 70-nucleotide transcript (Rcd). The transcription of Rcd is controlled by two promoters (P$_{cer}$) placed in the −35 and −10 boxes. The spacer, a region between both boxes, plays a crucial role in the activity of P$_{cer}$ (Chatwin and Summers, 2001). The spacer lengths both of ColE1 and pCS36-4CPA were 15 bp. That is why P$_{cer}$ activity of pCS36-4CPA is likely to be similar to ColE1 P$_{cer}$. Predicted minimal free energy (MFE) structures are shown in Fig. 6B. Two alternative structures of the ColA plasmid have equal MFE values. The secondary structures of all of the ColE1-related Rcds, including pCS36-4CPA, formed the single-stem loop. However, the secondary structures of p29807, pRK2 and ColA appeared to be more similar to ColE1 than other plasmids. One of the alternative ColA structures was similar to pCS36-4CPA. The 15-nt conserved sequence in Rcd is a probable candidate for interactions with the target site (Sharpe et al., 1999). We noted that the 15-nt sequence of other ColE1-like plasmids differ much more from ColE1 (Fig. 6A). An identity between the probable 15-nt sequence of p29807 and pCS36-4CPA with ColE1 was only 53.3%; in contrast, these values for the pairs pRK2-ColE1, CloDF13-ColE1, and ColA-ColE1 were 93.3%, 86.7% and 93.3%, respectively.

In addition to the four proteins (XerC, XerD, ArgR and PepA) involved in multimer resolution, Blaby and Summers (2009) proposed the existence of a fifth chromosomally-encoded FIS protein. The role of this sequence-specific protein is proposed to be repression of the Rcd promoter (P$_{cer}$). The analysis of probable FIS binding sites in pCS36-4CPA and other related ColE1-like plasmids revealed the presence of only FIS binding site II (Fig. 6A). The exception is pRK2, which has an FIS binding site I possessing a high level of homology with the analogous nucleotide sequence in *cer* of ColE1. It was not consistent with the results obtained by Blaby and Summers, who showed that FIS binds only one of the binding sites using EMSA results. It is probable that one of the two binding sites is sufficient for the normal activity of FIS protein.
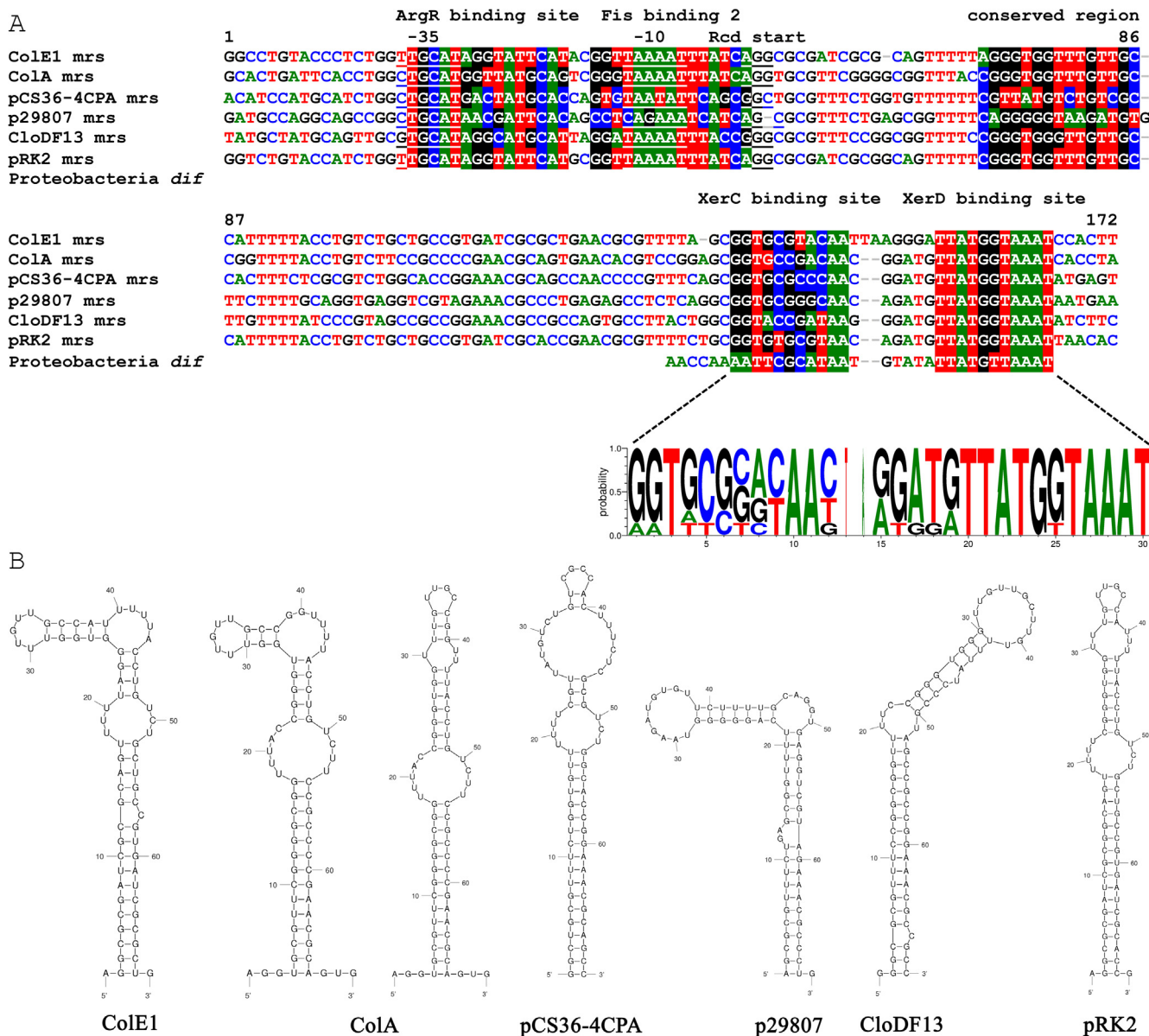
Figure 6 An alignment of ColE1-type plasmids multimer resolution site sequences. The ArgR, XerC and XerD binding sites in *mrs* are highlighted. The −10 and −35 boxes of the ColE1 *cer* (P*cer*) promoter and the putative corresponding boxes of the pCS36-4CPA are underlined. Nucleotides in Rcd start position are bold and underlined. XerC and XerD binding sites shown in sequence logo (A). Predicted secondary structures of all of the ColE1-related Rcds (B).

3.6. Cargo module (open reading frames)

Analysis of ORFs showed that plasmid pCS36-4CPA did not carry any colicinogenic factors, unlike ColA and ColE1. Three open reading frames were found in the pCS36-4CPA sequence possessing 84, 194 and 190 codons. All of the ORFs of the cargo module are encoded in the opposite orientation from those of the MOB genes (Fig. 2). A BLAST search of ORF5 and ORF6 did not reveal any significant similarities to known DNA sequences. Unlike ORF5 and ORF6, a BLAST search against the DNA sequence of ORF7 showed a high level of identity (99%) with the *Salmonella* phage SE2 DNA methylase N-4/N-6 domain protein. *In silico* analysis of amino acid sequences encoded by

ORF5 and ORF6 showed two putative conserved domains. According to these data, ORF5 contains a conserved THAP DNA-binding domain (DBD) and ORF6 contains a typical type II restriction endonuclease *Eco*O109I domain. Analysis of the translated sequence of ORF7 reveals the presence of conserved motifs for YhdJ-like DNA modification methylases family belonging to S-adenosylmethionine-dependent methyltransferase (Mtase) superfamily. These results show that pCS36-4CPA probably can carry a restriction-modification system (RMS). Such possibility for a number of plasmids of the ColE1-family has been previously demonstrated (Gregorova et al., 2002; Mruk et al., 2001). However, it should be noted, that plasmid ColA from *Citrobacter* does not carry RMS and that the *Eco*O109I restriction-

modification system is chromosome-encoded (Kita et al., 1999). *Eco*O109I RMS contains endonuclease, methylase and the positive regulator C.*Eco*O109I (Kita et al., 2002). We have not found *Eco*O109I methylase or positive regulator in pCS36-4CPA. But potentially YhdJ-like methylase, which is encoded by ORF7, can function as part of the plasmid pCS36-4CPA restriction-modification system. Such small plasmids in *Enterobacteriaceae* carrying RMS may have a protective role against phages (Gregorova et al., 2002).

Analysing the ORFs of pCS36-4CPA, we noticed that the plasmid sequence between 4412 and 364 nt (including ORF7 located at position 4433–5002 nt) was highly homologous (99.1%) with region 31,107–32,284 nt of the *Salmonella* phage SE2 genome. These results may be explained by recombination events between the plasmid and phage DNA. Formation hybrids between plasmid and phage DNA or full phage genome integration into plasmid was previously observed (Michel and Ehrlich, 1986; Kanda et al., 1989). We assume that the fragment of the plasmid pCS36-4CPA (4412–364 nt) was integrated into the *Salmonella* phage SE2 and not vice versa. This assumption may be explained by several reasons. At first, *Salmonella* phage SE2 genome's G-C content is 49.6% against 40.2% of the integrated region from the pCS36-4CPA. Secondly, we did not find any significant homology between the integrated region and genomes of the other *Salmonella* phages (such as SS3e, vB_SenS-Ent1, f18SE, SETP3 and fSE1C) which are closely related to SE2 (homologous over 90%). Thus, *Salmonella* phage SE2 genome contains apparently the foreign integrated DNA from plasmid pCS36-4CPA.

4. Conclusion

In this study we focused on detailed sequence analysis of the newly isolated small ColE1-like plasmid pCS36-4CPA. It was found that the replication module had 67.1% identity to ColE1 *ori* and contained all necessary regulatory elements such as RNA I, RNA II, single-strand initiation site (*ssi*), and *terH* sequence. MOB machinery and the *mrs* region of pCS36-4CPA were similar to that of the ColE1 plasmid. Also pCS36-4CPA carried cargo module consisting of three open reading frames. However, plasmid pCS36-4CPA did not carry any colicinogenic factors, unlike ColE1. The sequencing and comparing of pCS36-4CPA with other plasmids of ColE1-family revealed two important points. First, the plasmid replication region (*ori*) of *Citrobacter* plasmids has a polyphyletic origin. Secondly, plasmid and chromosomal-located protein binding sites in *mrs* are related. Thus, pCS36-4CPA is the first fully annotated plasmid from a member of the genus *Citrobacter* that probably carries restriction-modification system and has recombination events with phage genome.

Acknowledgments

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Anand, P., Saxena, R.K., 2012. A comparative study of solvent-assisted pre-treatment of biodiesel derived crude glycerol on growth and 1,3-propanediol production from *Citrobacter freundii*. New Biotechnol. 29, 199–205.

Bai, L., Xia, S., Lan, R., Liu, L., Ye, C., Wang, Y., Jin, D., Cui, Z., Jing, H., Xiong, Y., Bai, X., Sun, H., Zhang, J., Wang, L., Xu, J., 2012. Isolation and characterisation of cytotoxic, aggregative *Citrobacter freundii*. PLoS One. http://dx.doi.org/10.1371/journal.pone.0033054.

Balding, C., Blaby, I., Summers, D., 2006. A mutational analysis of the ColE1-encoded cell cycle regulator Rcd confirms its role in plasmid stability. Plasmid 56, 68–73.

Barlow, M., Hall, B.G., 2002. Origin and evolution of the AmpC β-Lactamases of *Citrobacter freundii*. Antimicrob. Agents Chemother. 46 (5), 1190–1198.

Birnboim, H.C., Doly, J., 1979. A Rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513–1523.

Blaby, I.K., Summers, D.K., 2009. The role of FIS in the Rcd checkpoint and stable maintenance of plasmid ColE1. Microbiology. http://dx.doi.org/10.1099/mic.0.029777-0.

Boulygina, E.S., Kuznetsov, B.B., Marusina, A.I., Tourova, T.P., Kravchenko, I.K., Bykova, S.A., Kolganova, T.V., Galchenko, V.F., 2002. A study of nucleotide sequences of *nifH* genes of some methanotrophic bacteria. Microbiology 71 (4), 425–432.

Camps, M., 2010. Modulation of ColE1-like plasmid replication for recombinant gene expression. Recent Pat. DNA Gene Seq. 4 (1), 58–73.

Carnoy, C., Roten, C.-A., 2009. The *dif*/Xer recombination systems in proteobacteria. PLoS One. http://dx.doi.org/10.1371/journal.pone.0006531.

Chan, P.T., Ohmori, H., Tomizawa, J., Lebowitz, J., 1985. Nucleotide sequence and gene organization of ColE1 DNA. J. Biol. Chem. 260 (15), 8925–8935.

Chatwin, H.M., Summers, D.K., 2001. Monomer-dimer control of the ColE1 P_{cer} promoter. Microbiology 147, 3071–3081.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M., 2014. Ribosomal database project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. http://dx.doi.org/10.1093/nar/gkt1244.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.

Davison, J., 1984. Mechanism of control of DNA replication and incompatibility in ColE1-type plasmids. Gene 28 (1), 1–15.

Di Primo, C., 2008. Real time analysis of the RNAI–RNAII–Rop complex by surface plasmon resonance: from a decaying surface to a standard kinetic analysis. J. Mol. Recognit. 21, 37–45.

Edwards, U., Rogall, T., Blöcker, H., Emde, M.D., Böttger, E.C., 1989. Isolation and direct complete nucleotide determination of entire genes, characterization of gene coding for 16S ribosomal RNA. Nucleic. Acids Res. 17, 7843–7853.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Francia, M.V., Varsaki, A., Garcillán-Barcia, M.P., Latorre, A., Drainas, C., de la Cruz, F., 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. 28, 79–100.

Fredrickson, T.N., Hartley, J.W., Morse III, H.C., 2010. *Citrobacter*-induced colitis in mice with murine acquired immunodeficiency syndrome. Vet. Pathol. 47 (2), 312–317.

Gaimster, H., Summers, D., 2015. Plasmids in the driving seat: the regulatory RNA Rcd gives plasmid ColE1 control over division and growth of its *E. coli* host. Plasmid 78, 59–64.

Garcillán-Barcia, M.P., Francia, M.V., de la Cruz, F., 2009. The diversity of conjugative relaxases and its application in plasmid classification. FEMS Microbiol. Rev. 33, 657–687.

Garcillán-Barcia, M.P., Alvarado, A., de la Cruz, F., 2011. Identification of bacterial plasmids based on mobility and plasmid population biology. FEMS Microbiol. Rev. 35, 936–956.

Ghosh, S., Chisti, Y., Banerjee, U.C., 2012. Production of shikimic acid. Biotechnol. Adv. 30 (6), 1425–1431.

Gregorova, D., Pravcova, M., Karpiskova, R., Rychlik, I., 2002. Plasmid pC present in *Salmonella enterica* serovar Enteritidis PT14b strains encodes a restriction modification system. FEMS Microbiol. Lett. 214 (2), 195–198.

Grishchenkov, V.G., Slepen'kin, A.V., Borodin, A.M., 2002. Anaerobic degradation of biphenyl by the facultative anaerobic strain *Citrobacter freundii* BS2211. Prikl. Biokhim. Mikrobiol. 38 (2), 145–148.

Helmer-Citterich, M., Anceschi, M.M., Banner, D.W., Cesareni, G., 1988. Control of ColE1 replication: low affinity specific binding of Rop (Rom) to RNAI and RNAII. EMBO J. 7 (2), 557–566.

Hodgman, T.C., Griffiths, H., Summers, D.K., 1998. Nucleoprotein architecture and ColE1 dimer resolution: a hypothesis. Mol. Microbiol. 29 (2), 545–558.

Ibrahim, K.S., Bakkiyaraj, D., James, R., Babu, T.G., Pandian, S.T. K., 2009. Isolation and sequence analysis of a small cryptic plasmid pRK10 from a corrosion inhibitor degrading strain *Serratia marcescens* ACE2. Plasmid 62 (3), 183–190.

Jang, M.-S., Lee, Y.-Mi, Kim, C.-Ho, Lee, J.-H., Kang, D.-W., Kim, S.-Jo, Lee, Y.-C., 2005. Triphenylmethane reductase from *Citrobacter* sp. strain KCTC 18061P: purification, characterization, gene cloning, and over expression of a functional protein in *Escherichia coli*. Appl. Environ. Microbiol. 71 (12), 7955–7960.

Kanda, K., Tan, Y., Aizawa, K., 1989. A novel phage genome integrated into a plasmid in *Bacillus thuringiensis* strain AF101. J. Gen. Microbiol. 135, 3035–3041.

Kim, Y.-M., Lee, S.-E., Park, B.-S., Son, Mi-K., Jung, Y.-Mi, Yang, S.-Ok, Choi, H.-K., Hur, S.-Ho, Yum, J.H., 2012. Proteomic analysis on acetate metabolism in *Citrobacter* sp. BL-4. Int. J. Biol. Sci. 8 (1), 66–78.

Kita, K., Tsuda, J., Kato, T., Okamoto, K., Yanase, H., Tanaka, M., 1999. Evidence of horizontal transfer of the *Eco*O109I restriction-modification gene to *Escherichia coli* chromosomal DNA. J. Bacteriol. 181 (21), 6822–6827.

Kita, K., Tsuda, J., Nakai, S.-Y., 2002. C.*Eco*O109I, a regulatory protein for production of *Eco*O109I restriction endonuclease, specifically binds to and bends DNA upstream of its translational start site. Nucleic. Acids Res. 30 (16), 3558–3565.

Kumar, S., Kaur, C., Kimura, K., Takeo, M., Raghava, G.P.S., Mayilraj, S., 2013. Draft genome sequence of the type species of the genus *Citrobacter*, *Citrobacter freundii* MTCC 1658. Genome Announc. http://dx.doi.org/10.1128/genomeA.00120-12.

Li, F., Zhu, L., 2012. Effect of surfactant-induced cell surface modifications on electron transport system and catechol 1,2-dioxygenase activities and phenanthrene biodegradation by *Citrobacter* sp. SA01. Bioresour. Technol. 123, 42–48.

Lin, S.-Yi, Ho, M.-W., Yang, Ya-F., Liu, J.-H., Wang, I-K., Lin, S.-H., Huang, C.-C., 2011. Abscess caused by *Citrobacter koseri* infection: three case reports and a literature review. Intern. Med. 50, 1333–1337.

Lipsky, B.A., Hook 3rd, E.W., Smith, A.A., Plorde, J.J., 1980. *Citrobacter* infections in humans: experience at the Seattle Veterans Administration Medical Center and a review of the literature. Rev. Infect. Dis. 2 (5), 746–760.

Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M.K., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I.,

Lanczycki, C.J., Lu, F., Lu, S., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Bryant, S.H., 2013. CDD: conserved domains and protein three-dimensional structure. Nucleic. Acids Res. http://dx.doi.org/10.1093/nar/gks1243.

Metsoviti, M., Zeng, An-P., Koutinas, A.A., Papanikolaou, S., 2013. Enhanced 1,3-propanediol production by a newly isolated *Citrobacter freundii* strain cultivated on biodiesel-derived waste glycerol through sterile and non-sterile bioprocesses. J. Biotechnol. 163, 408–418.

Michel, B., Ehrlich, 1986. Illegitimate recombination occurs between the replication origin of plasmid pC194 and a progressing replication fork. EMBO J. 5 (13), 3691–3696.

Mohanty, S., Singhal, R., Sood, S., Dhawan, B., Kapil, A., Das, B.K., 2007. *Citrobacter* infections in a tertiary care hospital in Northern India. J. Infect. 54, 58–64.

Mohite, B.V., Pawar, S.P., Morankar, A., 2011. Isolation, selection and biodegradation profile of phenol degrading bacteria from oil contaminated soil. Bull. Environ. Contam. Toxicol. 87, 143–146.

Morlon, J., Chartier, M., Bidaud, M., Lazdunski, C., 1988. The complete nucleotide sequence of the colicinogenic plasmid ColA. High extent of homology with ColE1. Mol. Gen. Genet. 211 (2), 231–243.

Mruk, I., Sektas, M., Kaczorowski, T., 2001. Characterization of pEC156, a ColE1-type plasmid from *Escherichia coli* E1585-68 that carries genes of the *Eco*VIII restriction-modification system. Plasmid 46, 128–139.

Núñez, B., de la Cruz, F., 2001. Two atypical mobilization proteins are involved in plasmid CloDF13 relaxation. Mol. Microbiol. 39 (4), 1088–1099. http://dx.doi.org/10.1046/j.1365-2958.2001.02308.x.

Oh, Y.-K., Seol, E.-H., Park, S., Park, S., 2011. Decolourisation of synthetic dyes by *Citrobacter amalonaticus* Y19. J. Taiwan Inst. Chem. Eng. 42, 492–497.

Petty, N.K., Bulgin, R., Crepin, V.F., Cerdeño-Tárraga, A.M., Schroeder, G.N., Quail, M.A., Lennard, N., Corton, C., Barron, A., Clark, L., Toribio, A.L., Parkhill, J., Dougan, G., Frankel, G., Thomson, N.R., 2010. The *Citrobacter rodentium* genome sequence reveals convergent evolution with human pathogenic *Escherichia coli*. J. Bacteriol. 192 (2), 525–538.

Pino, N., Peñuela, G., 2011. Simultaneous degradation of the pesticides methyl parathion and chlorpyrifos by an isolated bacterial consortium from a contaminated site. Int. Biodeterior. Biodegrad. 65, 827–831.

Reich, H.L., Williams, F.D., Naik, N.S., Honig, P.J., Yan, A.C., 2004. Non-pseudomonal ecthyma gangrenosum. J. Am. Acad. Dermatol. 50 (5), 114–117.

Rozhon, W.M., Petutschnig, E.K., Jonak, C., 2006. Isolation and characterization of pHW15, a small cryptic plasmid from *Rahnella genomospecies* 2. Plasmid 56, 202–215.

Rozhon, W., Petutschnig, E., Khan, M., Summers, D.K., Poppenberger, B., 2010. Frequency and diversity of small cryptic plasmids in the genus *Rahnella*. BMC Microbiol. http://dx.doi.org/10.1186/1471-2180-10-56.

Saitou, N., Nei, M., 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sakimbaeva, S.D., 1985. Dependence of the dynamics of the multiplication of bacteria in the genus *Citrobacter* on the temperature factor and on the product type in experimental contamination. Vopr. Pitan. 4, 56–59.

Sakultantimetha, A., Keenan, H.E., Dyer, M., Beattie, T.K., Bangkedphol, S., Songsasen, A., 2009. Isolation of tributyltin-degrading bacteria *Citrobacter braakii* and *Enterobacter cloacae* from butyltin-polluted sediment. J. ASTM Int. 6 (6), 1–6.

Samonis, G., Karageorgopoulos, D.E., Kofteridis, D.P., Matthaiou, D.K., Sidiropoulou, V., Maraki, S., Falagas, M.E., 2009. *Citrobacter* infections in a general hospital: characteristics and outcomes. Eur. J. Clin. Microbiol. Infect. Dis. 28 (1), 61–68.

Saratale, R.G., Saratale, G.D., Chang, J.S., Govindwar, S.P., 2011. Bacterial decolourisation and degradation of azo dyes: a review. J. Taiwan Inst. Chem. Eng. 42 (1), 138–157.

Schneider, T.D., Stephens, R.M., 1990. Sequence logos: a new way to display consensus sequences. Nucleic. Acids Res. 18, 6097–6100.

Sekiguchi, J., Morita, K., Kitao, T., Watanabe, N., Okazaki, M., Miyoshi-Akiyama, T., Kanamori, M., Kirikae, T., 2008. KHM-1, a novel plasmid-mediated metallo-β-lactamase from a *Citrobacter freundii* clinical isolate. Antimicrob. Agents Chemother. 52 (11), 4194–4197.

Selvakumaran, S., Kapley, A., Kashyap, S.M., Daginawala, H.F., Kalia, V.C., Purohit, H.J., 2011. Diversity of aromatic ring-hydroxylating dioxygenase gene in *Citrobacter*. Bioresour. Technol. 102, 4600–4609.

Shanks, R.M.Q., Dashiff, A., Alster, J.S., Kadouri, D.E., 2012. Isolation and identification of a bacteriocin with antibacterial and anti-biofilm activity from *Citrobacter freundii*. Arch. Microbiol. 194 (7), 575–587. http://dx.doi.org/10.1007/s00203-012-0793-2.

Sharpe, M.E., Chatwin, H.M., Macpherson, C., Withers, H.L., Summers, D.K., 1999. Analysis of the ColEl stability determinant Rcd. Microbiology 145, 2135–2144.

Smillie, C., Garcillán-Barcia, M.P., Francia, M.V., Rocha, E.P.C., de la Cruz, F., 2010. Mobility of plasmids. Microbiol. Mol. Biol. Rev. 74 (3), 434–452.

Štěpánek, V., Valešová, R., Kyslík, P., 2005. Cryptic plasmid pRK2 from *Escherichia coli* W: sequence analysis and segregational stability. Plasmid 54 (1), 86–91.

Strauch, E., Voigt, I., Broll, H., Appel, B., 2000. Use of a plasmid of a *Yersinia enterocolitica* biogroup 1A strain for the construction of cloning vectors. J. Biotechnol. 79, 63–72.

Summers, D., Yaish, S., Archer, J., Sherratt, D., 1985. Multimer resolution systems of ColE1 and ColK: localisation of the crossover site. Mol. Gen. Genet. 201, 334–338.

Summers, D.K., Beton, C.W., Withers, H.L., 1993. Multicopy plasmid instability: the dimer catastrophe hypothesis. Mol. Microbiol. 8 (6), 1031–3108.

Tamura, K., Nei, M., Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbour-joining method. Proc. Natl. Acad. Sci. (USA) 101, 11030–11035.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Tan, C.-K., Lai, C.-C., Lin, S.-H., Liao, C.-H., Huang, Yu-T., Hsueh, Po-R., 2010. Fatal *Citrobacter farmeri* meningitis in a patient with nasopharyngeal cancer. J. Clin. Microbiol. 48 (4), 1499–1500.

Varsaki, A., Lamb, H.K., Eleftheriadou, O., Vandera, E., Thompson, P., Moncalián, G., de la Cruz, F., Hawkins, A.R., Drainas, C., 2012. Interaction between relaxase MbeA and accessory protein MbeC of the conjugally mobilisable plasmid ColE1. FEBS Lett. 586, 675–679.

Wang, H., Su, J.Q., Zheng, X.W., Tian, Y., Xiong, X.J., Zheng, T.L., 2009. Bacterial decolourisation and degradation of the reactive dye Reactive Red180 by *Citrobacter* sp. CK3. Int. Biodeterior. Biodegrad. 63, 395–399.

Wu, Y.-Ch., Liu, Sh.-T., 2010. A sequence that affects the copy number and stability of pSW200 and ColE1. J. Bacteriol. 192 (14), 3654–3660.

Wu, M., Chen, L., Tian, Y., Ding, Yi, Dick, W.A., 2013. Degradation of polycyclic aromatic hydrocarbons by microbial consortia enriched from three soils using two different culture media. Environ. Pollut. 178, 152–158.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic. Acids Res. 31 (13), 3406–3415.