

RESEARCH LETTER – Environmental Microbiology

# PCR-based detection of composite transposons and translocatable units from oral metagenomic DNA

Supathep Tansirichaiya, Peter Mullany and Adam P. Roberts\*

Department of Microbial Diseases, University College London, Eastman Dental Institute, 256 Gray's Inn Road, London WC1×8LD, UK

\*Corresponding author: Department of Microbial Diseases, University College London, UCL Eastman Dental Institute, 256 Gray's Inn Road, London WC1×8LD, UK. Tel: +00-44 (0)20-3456-1044; Fax: +00-44-(0)20-3456-1127; E-mail: [adam.roberts@ucl.ac.uk](mailto:adam.roberts@ucl.ac.uk)

One sentence summary: Using a PCR approach, we have detected composite transposons and TUs directly from human oral metagenomic DNA.

Editor: Marilyn Roberts

## ABSTRACT

A composite transposon is a mobile genetic element consisting of two insertion sequences (ISs) flanking a segment of cargo DNA often containing antibiotic resistance (AR) genes. Composite transposons can move as a discreet unit. There have been recently several reports on a novel mechanism of movement of an IS26-based composite transposon through the formation of a translocatable unit (TU), carrying the internal DNA segment of a composite transposon and one copy of a flanking IS. In this study, we determined the presence of composite transposons and TUs in human oral metagenomic DNA using PCR primers from common IS elements. Analysis of resulting amplicons showed four different IS1216 composite transposons and one IS257 composite transposon in our metagenomic sample. As our PCR strategy would also detect TUs, PCR was carried out to detect circular TUs predicted to originate from these composite transposons. We confirmed the presence of two novel TUs, one containing an experimentally proven antiseptic resistance gene and another containing a putative universal stress response protein (UspA) encoding gene. This is the first report of a PCR strategy to amplify the DNA segment on composite transposons and TUs in metagenomic DNA. This can be used to identify AR genes associated with a variety of mobile genetic elements from metagenomes.

**Keywords:** composite transposon; translocatable unit (TU); oral metagenomic DNA

## INTRODUCTION

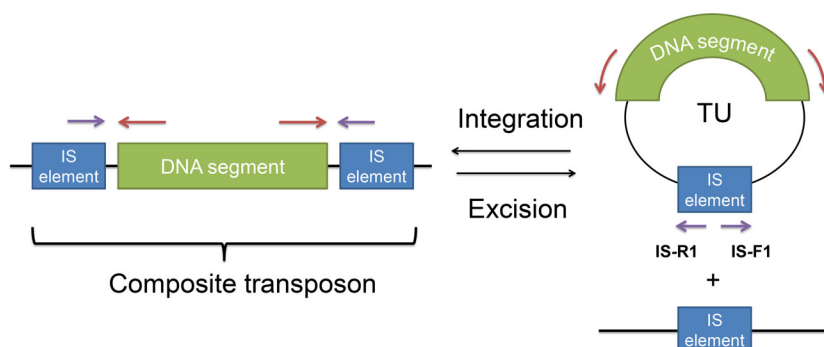
Composite transposons are common mobile genetic elements (MGEs) responsible for the dissemination of genes responsible for bacterial adaptation and survival including those conferring antibiotic resistance and xenobiotic degradation (Nojiri, Shintani and Omori 2004; Bennett 2008). They consist of two insertion sequences (ISs) that flank a segment of DNA, which can translocate as a whole unit (including the two flanking IS elements) and the active IS element alone can also translocate out from the unit.

Recently, another mechanism for the translocation of DNA within an IS26-based composite transposon was described

(Harmer, Moran and Hall 2014; Harmer and Hall 2015). The model is summarized in Fig 1. A single copy of the IS element is excised together with flanking DNA to form a circular molecule. To distinguish this from 'conventional' transposition of IS elements, the transposing region was called a translocatable unit (TU). This molecule can be formed by two different mechanisms: intramolecular replicative transposition and conservative transposition by excising out from the composite transposon (Fig. 1). Most of the DNA segments reported to be found on TUs are antibiotic resistance genes, such as the kanamycin resistance gene in IS26-*aphA1a* TU and tetracycline resistance gene in IS1216-*tet(S)* TU (Ciric et al. 2014; Harmer and Hall 2015).

Received: 1 August 2016; Accepted: 5 August 2016

© FEMS 2016. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)



**Figure 1.** The structures of composite transposons and TUs. The composite transposons consist of two IS elements (blue box) flanking DNA segment (green box). The TU circular molecule can be excised from the composite transposon. The purple and red arrows represent the binding site of the DNA primers for the amplification of composite transposon and TUs, respectively.

Previously, studies on composite transposons and TUs focused on cultivable bacteria. However, most of the bacteria in environmental samples have not yet been cultured in the laboratory. For example, more than 700 bacterial species have been identified in human oral cavity, but less than half of them can be cultivated (Wade 2011). Human oral metagenomic DNA was used to screen for composite transposons and TUs, as several TUs were found in oral bacteria, including *Streptococcus oralis* and *S. infantis* (Ciric, Mullany and Roberts 2011; Ciric et al. 2014).

The aim of this work is to determine whether composite transposons and TUs could be detected through PCR in metagenomic DNA. To do this, PCR primers were designed to amplify DNA flanked by IS elements to check for the presence of composite transposons in oral metagenomic DNA. Then, another set of primers was designed to determine if TUs derived from these putative composite transposons were present. We showed that novel TUs were detectable within human oral metagenomic DNA.

## MATERIALS AND METHODS

### Acquisition of the human oral metagenomic DNA

Saliva samples were collected from 11 healthy males and females from Department of Microbial Diseases, University College London (UCL) Eastman Dental Institute as described previously (Tansirichaiya et al. 2016). All of the volunteers did not receive any antibiotic treatment for at least 3 months, and gave the written consent prior to the sample collection. The saliva collection and processing procedures were approved by the UCL Ethics Committee (project number 5017/001) and are described previously (Tansirichaiya et al. 2016).

### PCR amplification

The amplifications on the human oral metagenome were performed with the primers listed in Table S1 (Supporting Information). The PCR reaction was carried out with an initial denaturation at 94°C for 3 min, following by 35 cycles of (i) denaturation at 94°C for 1 min (ii) annealing at 50°C–65°C for 30 s and (iii) extension at 72°C for 3 min for standard PCR, 10 min for long PCR and 5 min for Q5 PCR, and final extension at 72°C for 10 min. The standard PCR reaction was contained 15 µL of 2X BioMix Red (Bioline, London, UK), 50–100 ng of DNA template, 0.2 µM of each primer and molecular grade water (Sigma, Dorset, UK) up to 30 µL. The long PCR reaction contained 0.25 µL of TaKaRa Ex Taq (5 units/µL) (Takara Bio, Saint-Germain-en-Laye, France), 4 µL of

dNTP mixture (2.5 mM each), 5 µL of 10X Ex Taq buffer, 0.2 µM of each primer, 50–100 ng of DNA template and molecular grade water up to a total volume of 50 µL. Q5 PCR reaction consisted of 12.5 µL Q5 high-fidelity 2X master mix (NEB, Hitchin, UK), 0.2 µM of each primer, 50–100 ng of DNA template and molecular grade water up to a total of 25 µL. PCR amplicons were visualized by gel electrophoresis on 1% agarose gel stained with GelRed nucleic acid stain (Biotium, Cambridge, UK).

### PCR purification, ligation and transformation

PCR products were purified by using either the QIAquick PCR Purification Kit or the QIAquick Gel Extraction Kit (Qiagen), following the manufacturer's protocols. The purified products were subsequently cloned into the pGEM-T Easy vector (Promega, Southampton, UK) and transformed into *Escherichia coli* α-Select Silver Efficiency competent cells (Bioline) by heat shock. Cells were grown on the Luria–Bertani (LB) agar containing 100 µg/mL ampicillin, 40 µg/mL X-Gal and 0.4 mM IPTG, and incubated overnight at 37°C.

### Plasmid isolation and sequencing

By using the blue-white screening, the white colonies (containing inserts) were subcultured into 5 mL of LB broth containing 100 µg/mL ampicillin and incubated overnight in 37°C shaker. The plasmid isolation was performed by using QIAprep Spin Miniprep Kit (Qiagen), as per the manufacturer's instruction. The inserts containing in each plasmid were sequenced using M13 forward and M13 reverse primers at the Beckman Coulter Genomics (Genewiz, Essex, UK) with an ABI 3730XL. If the initial sequencing reaction cannot cover the sequences of the inserts, extra primers were designed and used for additional sequencings.

### Sequence analyses

The analysis of sequencing results was performed with BioEdit software version 7.2.0 (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>). If the insert was sequenced using multiple primers, the sequences were assembled with CAP contig function within the software (Huang 1992). The vector sequences were initially trimmed from the sequences by using VecScreen analysis tool (<http://www.ncbi.nlm.nih.gov/tools/vecscreen>) and identified for the primer binding sites. The sequences were then analyzed by using BlastN and BlastX for the matching with the sequences in the databases, ISFinder for the identification

of IS element and Clustal Omega for the alignment of the sequences (Altschul et al. 1990; Siguier et al. 2006). The sequences of all composite transposons (CTA1 to CTA5) were submitted to the DNA database with the accession numbers from KX305930 to KX305934.

## RESULTS

### Confirmation of IS elements in oral metagenomic DNA

Twelve IS elements were selected for the screening of composite transposons, including those of the IS26 family (IS26, IS240, IS257 and IS1216), IS elements found in *Streptococcus* spp. (the most prevalent bacteria in oral cavity) (IS861, IS1161, IS1167, IS1381 and IS1548), and IS elements commonly associated with plasmids and transposons (IS3, IS256 and IS1485) (Kreth, Merritt and Qi 2009; Clewell et al. 2014). Prior to the detection of composite transposons, the presence of each IS element in oral metagenomic DNA was determined by designing the PCR primers to amplify the ISs. The amplicons of the expected size were sequenced to confirm the results. Among 12 IS elements, the sequencing results showed that 6 of them were confirmed to be present in the extracted oral metagenome, these were IS26, IS257, IS1216, IS1161, IS1167 and IS1485.

### Detection of composite transposons by PCR

A set of PCR primers were designed to amplify outwards from the detected ISs. The amplicons from this PCR amplification could be either the DNA segment flanked by two ISs of a composite transposon or the DNA segment carried by a TU structure.

After the screening, five different putative composite transposon amplicons (CTA1–5) were identified, four were IS1216-based amplicons and another was an IS257-based amplicon (Table 1 and Table S1, Supporting Information). The first amplicon (CTA1) contained two potential ORFs: one predicted to encode a small multidrug resistance protein (Qrg) and the other a hypothetical protein (similar to NAD<sup>+</sup> diphosphatase). It had 99% nucleotide identity to the IS1216 composite transposon found on Tn6087 of *Streptococcus oralis* F.M.I.5 (Circ, Mullany and Roberts 2011). The second amplicon (CTA2) was similar to the first one, but it was missing 229 bp of the gene predicted to encode the hypothetical protein and a region between *qrg* and the flanking IS1216.

The next structure (CTA3) was similar to part of plasmid pIL5 and plasmid pBL1 from *Lactococcus lactis* subsp. *lactis* (Sanchez et al. 2000; Gorecki et al. 2011). The main part of this structure (84% of the query) had 98% nucleotide identity with *orf14*, 15 and partial *orf16* from plasmid pIL5, encoding a transposase, a universal stress-like protein and a Mn<sup>2+</sup>/Fe<sup>2+</sup> transporter-like protein, respectively. The matching part of pBL1 plasmid to this structure was an ISS1-like element with 100% nucleotide identity. By analyzing the *orf14* and the ISS1-like element with ISfinder, they both were matched to IS1216 with 100% similarity. Another structure (CTA4), which was similar to the CTA3, was also identified, but it had an additional 2329 bp compared to CTA3. The extra nucleotides were the rest of *orf16* sequences missing from CTA3, and a transposase gene *orf17* from pIL5.

The fifth structure, CTA5, was the only structure amplified with IS257-based primers. Sequencing analysis showed that the kanamycin resistance gene, *knt*, and a truncated *rep* gene were flanked by transposase genes. It had 99% identity to part of plasmid SAP079A from *Staphylococcus aureus* (McDougal et al. 2010).

By using ISFinder, both partial transposase genes were matched to IS257 (100% and 99% nucleotide identity).

### Amplification of putative TU structures

As each of amplicons described above may have formed from a TU template, another set of primers were designed to amplify outward from the DNA segments between the flanking ISs (Fig. 1). The verification PCR was carried out by using Biomix red and additionally using highly processive Q5 polymerase. This was to make sure that the amplicons were not a result of an early fall-off of the polymerase, leaving partial amplicons that could themselves act as primers in subsequent round of PCR. TUs were detected in PCRs carried out with primers designed from the DNA segments within CTA2 and CTA4; amplicons of the expected size were confirmed as containing a single entire copy of IS1216 by DNA sequencing (Fig. 2).

## DISCUSSION

By designing the primers to read outward from IS elements, five different putative composite transposons were identified from human oral metagenomic DNA, each containing different genes predicted to be involved in environmental adaptation. The *qrg* gene in samples CTA1 and CTA2 was predicted to encode a small multidrug resistance protein that confers resistance to cetyltrimethylammonium bromide (CTAB), a cationic antiseptic (Circ, Mullany and Roberts 2011). For the sample CTA5, it contained kanamycin resistance gene *knt*, encoding kanamycin nucleotidyltransferase protein. This protein was shown to confer resistance by catalyzing the transfer of a nucleotidyl group to the 4'-hydroxyl group of the aminoglycoside resulting in inactivating kanamycin drug (Pedersen, Benning and Holden 1995).

A gene (*uspA*) predicted to encode a universal stress protein was found in samples CTA3 and CTA4. The precise biological function of UspA remains unknown, but it was shown that the levels of UspA are elevated during a variety of stress conditions including heat, oxidant exposure, nutrient starvation and exposure to antibiotics including polymyxins and cycloserine (Kvint et al. 2003; Nachin, Nannmark and Nystrom 2005). It was hypothesized to function in the reprogramming of cells toward defense and escape by enhancing the cell's capacity to withstand stresses and modulating activities related to motility and adhesion (Nachin, Nannmark and Nystrom 2005).

By designing another set of primers to determine the presence of TUs based on each amplicon, we confirmed that CTA2 and CTA4 were likely to be present as small circular molecules as the entire predicted single copy of the expected IS element could be amplified. In order to control for PCR artefacts that may result from short primer extension within a chromosomally located composite transposon, we used the highly processive Q5 DNA polymerase that confirmed the results of the Biomix red PCR. The lack of positive results for the TU PCR of CTA1, CTA3 and CTA5 (which contained similar sized direct repeats) demonstrates that PCR artefacts are unlikely. This is the first time that TUs were detected in metagenomic DNA and also the first time that the stress adaptation gene *uspA* was found on a TU. There is a possibility that the TU amplicon could have arisen if the entire composite transposon was repeated in the host genome; however, we think that this is unlikely due to the inherent instability of large repeated units of mobile DNA.

The studies on TUs are still in an early stage, and there are only a small number of identified examples. Reports on the integration and excision of TUs are based on IS26 composite

**Table 1.** The predicted structure of putative composite transposons amplifying from the human oral metagenomic DNA. The open arrowed boxes represent ORFs, pointing in the probable direction of transcription. The IS elements and ORF in the DNA segment are shown in blue and green, respectively. The dash boxes, arrow boxes and dotted lines represent the regions that are not present compared to the sequences in the database.

Sample	Predicted Structure
CTA1 (1772 bp)	
CTA2 (1321bp)	
CTA3 (984 bp)	
CTA4 (3313 bp)	
CTA5 (1658 bp)	

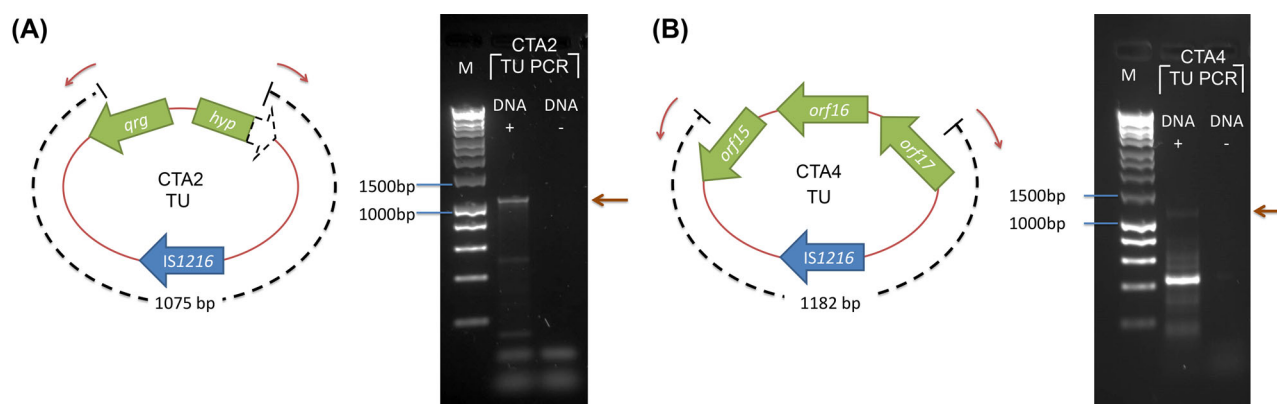
transposons. IS26 is a member of the IS6 family that also contains IS1216 and IS257. The transposase protein encoded by the IS elements in this family has 40%–94% identity. It was shown that IS26 TU preferred to integrate with a preexisting IS26 element via a conservative reaction catalyzed by the Tnp26 transposase, rather than a replicative transposition to a new site (Harmer, Moran and Hall 2014). An intact transposase gene was shown to be important for the integration and excision of the TUs (Harmer, Moran and Hall 2014; Harmer and Hall 2015). Recently, it was shown that RecA-dependent homologous recombination could also mediate the integration of a TU; however, it is not a major factor because it occurs at least two orders of magnitude lower efficiency than the reaction catalyzed by Tnp26 (Harmer and Hall 2016).

TUs are similar to another recently described MGE, called an unconventional circularizable structure (UCS) (Palmieri, Mingoia and Valardo 2013). UCSs can be excised from a replicon that contains two direct repeats (DRs) flanking a DNA segment. After ex-

cision, it results in a non-replicative circular structure containing the DNA segment and one of the DRs. The movement of TUs and UCSs can both occur in RecA-deficient bacteria, suggesting that homologous recombination is not essential for their formation and insertion (Azpiroz, Bascuas and Laviña 2011; Harmer, Moran and Hall 2014). The difference between UCS and TU is that UCS has no recombinase gene to catalyze their insertion reactions (Palmieri, Mingoia and Valardo 2013). This PCR strategy to detect TUs can also be applied to UCSs by designing the primers based on the DRs of the UCSs.

As it was shown that the IS6 family of composite transposons can transpose either as a whole unit of composite transposons or as TUs, the resistance genes associated with them may then have more chance to be transposed and spread in bacterial population (Harmer and Hall 2016). Furthermore, composite transposons are often located on other MGEs such as plasmids and conjugative transposons, which can also facilitate their horizontal gene transfer. If the excised TUs do not integrate into





**Figure 2.** Schematic of the products from the TU confirmation PCR of the sample (A) CTA2 and (B) CTA4. PCR products were separated on 1% agarose and stained with GelRed. Lane M, HyperLadder 1 kb (Bioline, UK). The open arrowed boxes represent ORFs, pointing in the probable direction of transcription. The IS elements and ORF in the DNA segment are shown in blue and green, respectively. The dashed arrow box represents truncated gene. The red arrows represent the binding site of the TU verification primers. The brown arrows indicate the expected PCR products of the TU verification PCR. The black dashed lines indicated the expected size of the amplicons.

a replicon, they will presumably be lost from the population during cellular replication. However, in sample CTA5, we found a truncated *rep* gene on a putative IS257 composite transposon. This raises the possibility that TUs can facilitate the movement of *rep* genes between DNA molecules further adding to the complexity of MGE biology. Indeed, there are some possible structures representing such *rep* containing TU insertion reactions such as  $\Delta repA$ -*repC* that is flanked by IS26 on Tn6029 (Reid, Roy Chowdhury and Djordjevic 2015) and *repB* located next to IS1216 on Tn6079 (de Vries et al. 2011; Reid, Roy Chowdhury and Djordjevic 2015).

In conclusion, we have determined that a metagenomic approach can be used to recover both composite transposons and TUs from oral metagenome by performing PCR amplification with DNA primers based on the IS elements. Due to the fact that the primers were designed based on the IS elements, using this approach could also amplify novel genes carried by those composite transposons. This method might also be a more promising approach for the detection of the TUs in metagenome, as these small circular molecules are likely to be rare and therefore could be missed, or not assembled by metagenomic sequencing.

## SUPPLEMENTARY DATA

Supplementary data are available at FEMSLE online.

**Conflict of interest.** None declared.

## REFERENCES

- Altschul SF, Gish W, Miller W et al. Basic local alignment search tool. *J Mol Biol* 1990;215:403–10.
- Azpiroz MF, Bascuas T, Laviña M. Microcin H47 System: An *Escherichia coli* small genomic Island with novel features. *PLoS One* 2011;6:e26179.
- Bennett PM. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *Brit J Pharmacol* 2008;153:S347–57.
- Ciric L, Brouwer MS, Mullany P et al. Minocycline resistance in an oral *Streptococcus infantis* isolate is encoded by *tet(S)* on a novel small, low copy number plasmid. *FEMS Microbiol Lett* 2014;353:106–15.
- Ciric L, Mullany P, Roberts AP. Antibiotic and antiseptic resistance genes are linked on a novel mobile genetic element: Tn6087. *J Antimicrob Chemoth* 2011;66:2235–9.
- Clewell DB, Weaver KE, Dunny GM et al. *Extrachromosomal and Mobile Elements in Enterococci: Transmission, Maintenance, and Epidemiology*. Boston, MA: Massachusetts Eye and Ear Infirmary, 2014.
- de Vries LE, Valles Y, Agero Y et al. The gut as reservoir of antibiotic resistance: microbial diversity of tetracycline resistance in mother and infant. *PLoS One* 2011;6:e21644.
- Gorecki RK, Koryszewska-Baginska A, Golebiewski M et al. Adaptive potential of the *Lactococcus lactis* IL594 strain encoded in its 7 plasmids. *PLoS One* 2011;6:e22238.
- Harmer CJ, Hall RM. IS26-mediated precise excision of the IS26-aphA1a translocatable unit. *mBio* 2015;6:e01866-15.
- Harmer CJ, Hall RM. IS26-mediated formation of transposons carrying antibiotic resistance genes. *mSphere* 2016;1:e00038-16.
- Harmer CJ, Moran RA, Hall RM. Movement of IS26-associated antibiotic resistance genes occurs via a translocatable unit that includes a single IS26 and preferentially inserts adjacent to another IS26. *mBio* 2014;5:e01801-14.
- Huang X. A contig assembly program based on sensitive detection of fragment overlaps. *Genomics* 1992;14:18–25.
- Kreth J, Merritt J, Qi F. Bacterial and host interactions of oral streptococci. *DNA Cell Biol* 2009;28:397–403.
- Kvint K, Nachin L, Diez A et al. The bacterial universal stress protein: function and regulation. *Curr Opin Microbiol* 2003;6:140–5.
- McDougal LK, Fosheim GE, Nicholson A et al. Emergence of resistance among USA300 methicillin-resistant *Staphylococcus aureus* isolates causing invasive disease in the United States. *Antimicrob Agents Ch* 2010;54:3804–11.
- Nachin L, Nannmark U, Nystrom T. Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion, and motility. *J Bacteriol* 2005;187:6265–72.
- Nojiri H, Shintani M, Omori T. Divergence of mobile genetic elements involved in the distribution of xenobiotic-catabolic capacity. *Appl Microbiol Biot* 2004;64:154–74.
- Palmieri C, Mingoa M, Valardo PE. Unconventional circularizable bacterial genetic structures carrying antibiotic resistance determinants. *Antimicrob Agents Ch* 2013;57:2440–1.

- Pedersen LC, Benning MM, Holden HM. Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase. *Biochemistry* 1995;**34**:13305–11.
- Reid CJ, Roy Chowdhury P, Djordjevic SP. Tn6026 and Tn6029 are found in complex resistance regions mobilised by diverse plasmids and chromosomal islands in multiple antibiotic resistant Enterobacteriaceae. *Plasmid* 2015;**80**:127–37.
- Sanchez C, Hernandez de Rojas A, Martinez B *et al.* Nucleotide sequence and analysis of pBL1, a bacteriocin-producing plasmid from *Lactococcus lactis* IPLA 972. *Plasmid* 2000;**44**:239–49.
- Siguier P, Perochon J, Lestrade L *et al.* ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 2006;**34**:D32–6.
- Tansirichaiya S, Rahman MA, Antepowicz A *et al.* Detection of novel integrons in the metagenome of human saliva. *PLoS One* 2016;**11**:e0157605.
- Wade WG. Has the use of molecular methods for the characterization of the human oral microbiome changed our understanding of the role of bacteria in the pathogenesis of periodontal disease? *J Clin Periodontol* 2011;**38**:7–16.