**Capture of a Novel, Antibiotic Resistance Encoding, Mobile Genetic Element from *Escherichia coli* Using a New Entrapment Vector**

**or**

**A new entrapment vector allows detection of intracellular transposition and capture of mobile genetic elements and antibiotic resistance genes in *Escherichia coli*.**

**Supathep Tansirichaiya1,2,3**, **Sabrina John Moyo3,4, Mohammed Al-Haroni1,2, Adam P. Roberts3\***

**Keywords**: Entrapment vector, Mobile genetic element, Insertion sequence, Translocatable unit, Antimicrobial resistance, Trimethoprim resistance, Clinical isolates

**Running Title**: Capturing mobile DNA in clinical *E. coli*

1Department of Clinical Dentistry, Faculty of Health Sciences, UiT the Arctic University of Norway, Tromsø, Norway

2Centre for New Antimicrobial Strategies, UiT the Arctic University of Norway, Tromsø, Norway

3Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK

4Department of Clinical Science, University of Bergen, 5021, Bergen, Norway.

\*Corresponding author; Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, UK

T: +44(0)151 705 3247. E: Adam.Roberts@lstmed.ac.uk

**Abstract**

**Aims**

Antimicrobial resistance genes (ARGs) are often associated with mobile genetic elements (MGEs), which facilitate their movement within and between bacterial populations. Detection of mobility is therefore important to understand the dynamics of MGE dissemination and their associated genes, especially in resistant clinical isolates that often have multiple ARGs associated with MGEs. Therefore, this study aimed to develop an entrapment vector to capture active MGEs and ARGs in clinical isolates of *Escherichia coli*.

**Methods and Results**

We engineered an entrapment vector, called pBACpAK, to capture MGEs in clinical *E*. *coli* isolates. It contains a *cI*-*tetA* positive selection cartridge in which the *cI* gene encodes a repressor that inhibits the expression of *tetA*. Therefore, any disruption of *cI,* for example, by insertion of a MGE, will allow *tetA* to be expressed and result in a selectable tetracycline-resistant phenotype.

The pBACpAK was introduced into clinical *E. coli* isolates and grown on tetracycline-containing agar to select for clones with the insertion of MGEs into the entrapment vector. Several insertion sequences were detected within pBACpAK, including IS*26*, IS*903B,* and IS*Sbo1*. A novel translocatable unit (TU), containing IS*26* and *dfrA8* was also captured, and *dfrA8* was shown to confer trimethoprim resistance when it was cloned into *E. coli* DH5α.

**Conclusions**

The entrapment vector, pBACpAK was developed and shown to be able to capture MGEs and their associated ARGs from clinical *E. coli* isolates. We have captured, for the first time, a TU encoding antibiotic resistance.

**Significance and Impact of Study**

This is the first time that a TU and associated resistance gene has been captured from clinical *E. coli* isolates using an entrapment vector. The pBACpAK has the potential to be used not only as a tool to capture MGEs in clinical *E. coli* isolates, but also to study dynamics, frequency, and potentiators of mobility for MGEs.

**Introduction**

Antimicrobial resistant (AMR) bacteria are increasing globally, accelerated by the use of antimicrobials that provide selective pressure for resistance development (Ferri et al. 2017; Roope et al. 2019). One of the key phenomena contributing to the spread of AMR is horizontal gene transfer (HGT). This process is often facilitated by mobile genetic elements (MGEs); discrete segments of DNA that can move from one location to another (Roberts et al. 2008; Partridge et al. 2018). Conjugative MGEs (plasmids and transposons) carry genes encoding for protein machinery that can catalyze the transfer between cells (intercellular transposition) through conjugation. For other MGEs, such as insertion sequences (ISs), transposons (Tns), and translocatable units (TUs), they can excise and insert within a genome (intracellular transposition), often from one replicon to another within the same cell. These transposable elements are also frequently located on larger conjugative MGEs which can facilitate their intercellular transposition (Leclercq et al. 2012).

ISs and Tns are often identified through the phenotypic changes seen in the bacterial hosts and through comparative genomics during whole genome sequence analysis. The phenotypic changes can be conferred by the accessory genes carried by these elements; e.g. resistance genes such as *bla*NDM-1 on Tn*3000* and Tn*6360* (Campos et al. 2015; Zhao et al. 2017) and *mcr-1* on Tn*6390* (Liang et al. 2018). The phenotype could also be the result of the insertion itself (Vandecraen et al. 2017), for example, a disruption in *mgrB* by insertion of IS*Kpn13* resulted in colistin resistance in *Klebsiella pneumoniae*, andthe expression of *bla*CTX-M in *Kluyvera ascorbata* is increased when IS*Ecp1B* inserts upstream (Lartigue et al. 2006; Poirel et al. 2015). Most types of MGEs have been shown to be associated and responsible for the translocation of AMR genes, including IS elements, which translocate AMR genes in the form of composite transposons and translocatable units (TUs) (Roberts et al. 2008; Ciric et al. 2011; Harmer et al. 2014; Tansirichaiya et al. 2016). Approximately 75% of the composite transposons listed in The Transposon Registry are associated with genes encoding AMR (Tansirichaiya et al. 2019) and the importance of TUs in the context of clinically important resistance is being increasingly explored (Hubbard et al. 2020).

The use of entrapment vectors is another approach for the identification of MGEs that actively transpose into a pre-determined target site resulting in a selectable phenotype for the cells in which this has occurred. One of these genetic systems is the λ repressor (*cI*) – tetracycline resistance gene (*tetA*)selection cartridge (Fig. 1). The *cI* gene is constitutively expressed by PRM promoter and the encoded λ repressor blocks the expression of the tetracycline resistance gene *tetA* by binding to the PR promoter upstream of *tetA* (Solyga and Bartosik, 2004). Disruption of *cI* by the insertion of MGEs or by mutations, allows positive selection in the presence of tetracycline due to the de-repression of the PR promoter (as the λ repressor is no longer produced). *cI*-*tetA* selection cartridge has been used previously to identify novel MGEs, for example the pMEC1 entrapment vector was used to identify 5 novel ISs and Tn*3434* in *Paracoccus pantotrophus* (Bartosik et al. 2003). Another study used a combination of entrapment vectors, including the *cI*-*tetA*-based vectors (pMEC1 and pMMB2), in Paracoccus spp. and captured 37 IS elements, 1 composite transposon, and 3 non-composite transposons (Dziewit et al. 2012).

In this study, we have developed a single copy entrapment vector called pBACpAK for the detection of MGEs and ARGs in clinical *E. coli* isolates. pBACpAK was introduced into these isolates by electroporation followed by screening for clones with the insertion of MGEs into pBACpAK. Several ISs and a novel translocatable unit (TU) containing a functional trimethoprim resistance gene were found, demonstrating that pBACpAK can be utilised as a tool for the detection of intracellular transposition of MGEs containing AMR genes from one replicon to another in clinical *E*. *coli* isolates.

**Materials and methods**

**Bacterial strains, plasmids, and culture conditions**

All bacterial strains and plasmids used in the study are listed in Table 1. Clinical isolates which have been transformed with pBACpAK and designated with “::pBACpAK”. All bacterialstrains were grown at 37°C in Lysogeny Broth (LB) medium supplemented with appropriate antibiotics (Sigma-Aldrich, UK) with the concentrations as follows; chloramphenicol 12.5 μg ml-1, tetracycline 5 μg ml-1, ampicillin 100 μg ml-1 and trimethoprim 20 μg ml-1.

**Construction of the pBACpAK entrapment vector**

A novel entrapment vector, pBACpAK, was constructed for a direct *in vivo* capture of MGEs in clinical *E. coli* isolates (Fig. 2) (Tansirichaiya 2017). Amplicons containing *cI*-*tetA* selection cartridge (2.9 kb) were amplified from the pAK1 entrapment vector (Department of Bacterial Genetics, University of Warsaw, Poland) by using *cI*-*tetA*(F)-*Xho*I and *cI*-*tetA*(R)-*Xho*I primers (Table S1). The *cI*-*tetA* amplicons were digested with *Xho*I restriction enzyme (NEB, UK) and ligated with pre-digested and dephosphorylated pCC1BAC vector (Fig. S1). The ligation product was desalted and electroporated into *E. coli* EPI300.

**Electrocompetent cell preparation and electroporation**

*E. coli* competent cells were prepared by following the protocol described previously (Dower et al. 1988). An overnight culture of each isolate was prepared in LB medium supplemented with appropriate antibiotics. A fresh 50-ml falcon tube containing LB medium was inoculated with the overnight culture with a starting OD600 of 0.05 and grown until the mid-exponential phase (OD600 of 0.6). The cells were incubated on ice for 10 min and pelleted by centrifugation for 10 min at 4°C and 2500 x *g*. The pellet was washed four times by discarding the supernatant, resuspended in 10 ml of pre-chilled 10% glycerol dissolved in distilled water (Sigma-Aldrich, UK), and pelleted by centrifugation. The pellet was then resuspended in 100 μl of 10% glycerol, transferred to pre-chilled cryotubes, and kept at -80°C.

For electroporation, 50 μl of the competent cells were aliquot into a pre-chilled 1.5 ml microcentrifuge tube and mixed with 10-100 ng of pBACpAK plasmids. The mixture was transferred to a pre-chilled 0.1 cm electroporation cuvette (Bio-Rad, UK) and electroporated with the following conditions: 1.8 kV, 200 Ω, 25 µF. Cells in the cuvette were then resuspended with 950 μl of pre-warmed SOC medium, then transferred to a 50-ml tube and incubated in 37°C shaker for 1 hr. An aliquot of 100 μl of cells was spread on LB agar supplemented with chloramphenicol and incubated overnight. The presence of pBACpAK in the transformants was confirmed by performing *cI*-*tetA* PCR with *cI*-*tetA*-F1 and ERIS primers (Table S1).

**Capturing and screening for clones with the insertion of MGEs in *cI*-*tetA* region on pBACpAK**

*E. coli*::pBACpAK were subcultured into 5 ml of LB broth supplemented with chloramphenicol and incubated for either 4 hrs or overnight in 37°C shaker. Cells were spread on LB agar containing chloramphenicol and tetracycline, and incubated at 37°C for 24-48 hrs. Tetracycline resistance clones were subcultured onto fresh plates and incubated overnight. Clones with the insertion of MGEs into *cI*-*tetA* region of pBACpAK were screened by colony PCR with *cI*-*tetA*-F1 and ERIS primers. Each clone was categorised based on the size of the colony PCR amplicons as i.) No change in the amplicon size: clones with a point mutation(s), small duplication or deletion, ii.) Smaller amplicon size: clones with deletion and iii.) Larger amplicon size: clones with the insertion of MGEs (Fig. 2). All amplicons that were visibly larger than a wild-type *cI*-*tetA* amplicon (1.35 kb) on an agarose gel (increase in amplicon size from 500 bp upward) were sequenced by Sanger sequencing service from Genewiz, UK, and analysed by using BlastN and BlastX to compare the sequences to nucleotide and protein databases, and ISFinder to identify IS elements (Altschul et al. 1990; Siguier et al. 2006). The prediction of plasmid- and chromosome-derived sequences was performed with mlplasmids (Arredondo-Alonso et al. 2018). The whole genome sequencing (WGS) data of *E. coli* EC1444, EC2026, and EC2033 have previously been deposited in the GenBank database under accession numbers PRJNA647256, PRJNA647258 and PRJNA647259, respectively.

**Detection and characterisation of IS*26*-*dfrA8*-containing TU**

The presence of IS*26*-*dfrA8*-containing TU in a wild-type *E. coli* EC2026 was confirmed by PCR, as described previously (Tansirichaiya et al. 2016). TU verification PCR was performed by using Q5 high-fidelity 2X mastermix (NEB, UK) with Response reg-F1 and Response reg-R1 primers (Table S1) to amplify outward from the response regulator gene.

The *cI* region was amplified from a wild-type *E. coli* EC2026 and *E. coli* EC2026::pBAKpAK::*dfrA8*-TU with cI-*tetA*-F1 and ERIS primers by using MyTaq™ Red Mix (Bioline, UK). The amplicons were purified and cloned into a pGEM-T easy vector (Promega, UK) through TA cloning, following the protocol from the manufacturer. The ligation mix was transformed into NEB® 5-alpha Competent *E. coli* with the standard heat-shock transformation protocol. The transformants were grown on LB agar supplemented with ampicillin. Clones with the correct insert size were confirmed by sequencing. To determine trimethoprim resistance phenotype, both *E. coli* containing pGEM-2026-WT-*cI* and pGEM-2026-*dfrA8*-TU were spread and checked for growth on LB agar containing trimethoprim at a concentration of 20 μg ml-1, which is 5 times higher than the EUCAST MIC breakpoint for *E. coli* (EUCAST 2020).

**Results**

**Acquisition of MGEs from *E*. *coli* clinical isolates**

Five chloramphenicol and tetracycline sensitive clinical *E. coli* isolates (EC1168, EC1359A, EC1444, EC2026, and EC2033) (Moyo et al. 2020) were electroporated with pBACpAK and selected on chloramphenicol containing agar. Screening of tetracycline-resistant colonies from all *E. coli*::pBACpAK found 3 different IS elements (IS*26*, IS*903b*, and IS*Sbo1*) from 3 out of 5 isolates (*E. coli* EC1444::pBACpAK, EC2026::pBACpAK, and EC2033::pBACpAK), and also 1 translocatable unit from *E. coli* EC2026::pBACpAK. The number of tetracycline resistance colonies analysed, and the predicted structure and details of captured MGEs, are shown in Fig. 3, Table 2, and Table 3, respectively.

We analysed the genome sequences of the E. coli isolates from where we had successfully captured MGEs in order to predict whether they were originally chromosomal, or plasmid located. Most of the detected MGEs were likely to be plasmid-derived sequences (Table 2). For IS*26* in *E. coli* isolates EC2026, and EC2033, they both contained two IS*26* elements in two separate contigs. Contig 43 of *E. coli* EC2026 and contig 70 of *E. coli* EC2033 were 100% identical and contained IS*26*-*dfrA8* however, 87 bp of IS*26* was missing from both contigs (Fig. 4). The other IS*26*-containing contigs from *E. coli* EC2026, and EC2033 were very similar to each other, and had *bla*TEM-1 resistance gene located next to theIS*26* element, but contig 67 in *E. coli* EC2033 was 18 bp shorter (Fig. 4). Even though both isolates contained almost identical IS*26*-containing contigs, the WGS data shows that they were not clonal strains.

**Characterisation of trimethoprim resistance phenotype conferring by** **IS*26*-*dfrA8* translocatible unit**

The 2-kb translocatable unit captured in *E. coli* EC2026::pBACpAK was shown to contain IS*26*, the trimethoprim resistance gene *drfA8,* and a response regulator gene. The presence of the IS*26*-*dfrA8* TU in *E. coli* EC2026 was confirmed by performing TU verification PCR, described previously (Tansirichaiya et al. 2016), with highly processive Q5 polymerase amplifying outward from the response regulator gene (Fig. 5). The *cI*-*tetA* amplicon containing IS*26*-*dfrA8* TU was cloned into a pGEM-T easy vector and introduced into *E. coli* NEB® 5-alpha, which showed that *dfrA8* genein IS*26*-*dfrA8* TU was functional and conferred trimethoprim resistance, compared to *E.coli*::pGEM-2026-WT-*cI* at a concentration of 20 μg ml-1.

**Discussion**

Entrapment vectors have been used previously to capture MGEs within multiple bacterial species, including *Agrobacterium tumefaciens*, *Rhizobium leguminosarum, Corynebacterium glutamicum,* and *Rhodococcus fascians* (Ulrich and Puhler 1994; Jager et al. 1995; Schneider et al. 2000). By relying on the transposition activity of MGEs for the detection, novel MGEs can be identified, as no prior sequence information is required in the screening. In our study, we developed and demonstrated the use of pBACpAK entrapment vector to capture MGEs through a direct transformation of pBACpAK into the clinical *E. coli* isolates.

We have shown that the pBACpAK vector can be used to capture various MGEs from clinical *E. coli* isolates; IS*26*, IS*903*B, IS*Sbo1* and IS*26*-*dfrA8* TU. The use of entrapment vectors with clinical isolates has been demonstrated previously in *Enterobacter cloacae* and *Citrobacter freundii* by using pRAB1 vector, which found insertions of putative new transposable elements in *E. cloacae* (Raabe et al. 1988). In our study, we designed pBACpAK entrapment vector based on pCC1BAC vector, which incorporates two highly desirable qualities into an entrapment vector; 1, accommodation of large DNA fragments (>100kb), which is important as the size of MGEs can vary from approximately 500 bp for insertion sequences up to and exceeding 100 kb for transposons (Zeng et al. 2019) and 2, as a single copy number vector, it allows capture of genes which could produce proteins that could be toxic to *E. coli* if they are overexpressed (Gubellini et al. 2011; Chen et al. 2018), for example if cloned into a high-copy number vector.

The detection of MGE movement within pBACpAK relies on chloramphenicol and tetracycline resistance phenotypes as selective markers for the vector and the positive selection of MGE insertion within *cI*, respectively. As clinical isolates tend to have multi-drug resistance phenotypes, the pBACpAK vector cannot be used in isolates with either chloramphenicol or tetracycline resistance. Various combinations of resistance genes could be easily included on modified pBACpAK derivatives to allow the capture of MGEs in a wider range of resistant clinical isolates. Another approach is to change from antibiotic selection cartridge to be conditionally lethal genes such as *sacB* gene, encoding for levansucrase (Lessard et al. 1999; Lee et al. 2001).

The pBACpAK entrapment vector contains *E. coli* F-factor single-copy origin of replication, so it may limit the uses of pBACpAK to work in only *E. coli*. Previously, the primary approach to extend the use to other species for entrapment vectors is to add another origin of replication matched with the species of interests, for example, a megaplasmid-specific *repABC*-type replicon in pMMB2 for the detection in *Paracoccus pantotrophus* (Mikosa et al. 2006).

Among the captured MGEs from the three *E. coli* isolates, we identified a novel TU containing *dfrA8* trimethoprim resistance gene, which is the first time that a translocatable unit carrying an antibiotic resistance gene was captured from clinical *E. coli* isolates using an entrapment vector. The ability of IS*26*-*dfrA8* TU to confer trimethoprim resistance was confirmed by the cloning of IS*26*-*dfrA8* TU into *E. coli* DH5α. Previously, resistance genes were captured by entrapment vectors from *Paracoccus spp.*, isolated from the environment, which included *strA-strB* on Tn*5393* and *drrA*-*drrB* on Tn*6097* conferring streptomycin resistance and daunorubicin/doxorubicin resistance (Bartosik et al. 2003; Dziewit et al. 2012).

Translocatable units are recently described as another types of MGE derived from IS*6*/IS*26* family composite transposons in which a single copy of the IS element and the flanking DNA can excise out and form as a circular molecule, leaving another copy of IS element behind (Harmer et al. 2014; Harmer and Hall 2015). Several antimicrobial resistance genes have been reported to be associated with TUs such as IS*1216E*-*optrA*, IS*1216*-*qrg* and IS*1216*-*tet*(S) TUs containing oxazolidinone/phenicol, cetyltrimethylammonium bromide (CTAB) and tetracycline resistance genes, respectively (Ciric et al. 2011; Ciric et al. 2014; He et al. 2016; Tansirichaiya et al. 2016).

Translocatable units can integrate and form cointegrates through three mechanisms: a copy-in mechanism into a random target site, a targeted conservative mechanism with a target site containing another IS*26,* and a homologous recombination between IS*26* on TU and IS*26* on the target site (Harmer and Hall 2016; Harmer and Hall 2020). The formation of cointegrates via copy-in mechanism into random sites was 60-fold lower than the targeted conservative mechanism (Harmer et al. 2014). With pBACpAK, the transposition of the IS*26*-*dfrA8* TU should be through a copy-in mechanism as there is no IS*26* on pBACpAK (Fig. 6a). However, copy-in mechanisms typically result in duplication of IS*26*, forming an IS*26*-based composite transposon; but the IS*26*-*dfrA8* TU captured from *E. coli* EC2026 had only one IS*26* element.

Sequence analysis on the insertion site of the IS*26*-*dfrA8* TU on pBACpAK did not show an 8-bp duplication but rather had a 4-bp deletion, suggesting that its insertion on pBACpAK may have occurred through multiple steps. In order for BACpAK to capture IS*26*-*dfrA8* TU, it could require IS*26* to first insert into pBACpAK which then served as an integration site for the IS*26*-*dfrA8* TU to integrate through either a targeted conservative mechanism or homologous recombination (Fig. 6b), forming an IS*26*-*dfrA8* composite transposon which was followed by a loss of an IS*26*. The multi-events for the insertions of IS*26*-*dfrA8* TU could therefore result in a low probability for the events to occur, which could explain why IS*26*-*dfrA8* TU was not detected in *E. coli* EC2033 despite sharing *IS26-dfrA8* structures in the chromosome.

WGS analysis of the original *E. coli* clinical isolates showed that MGEs captured by pBACpAK were likely to derive from both plasmid and chromosomal DNA with plasmid derived ISs making up the majority (Table 2). Being able to detect transposition between multiple replicons in a cell gives us the possibility to further understand the intracellular dynamics of MGE mobility and investigate factors that may affect this.

To summarise, we have developed a pBACpAK entrapment vector and demonstrated its utility in capturing a variety of MGEs from clinical isolates of *E. coli*.

**Acknowledgments**

We are thankful to Professor Dariusz Bartosik and Dr Magdalena Szuplewska for the providing *of E. coli* strain containing pAK1 entrapment vector to use in this study. Supathep Tansirichaiya is supported by Personal Overseas Research Grant at the Faculty for Health Sciences, UiT The Arctic University of Norway. APR would like to acknowledge funding from the AMR Cross-Council Initiative through a grant from the Medical Research Council, a Council of UK Research and Innovation (Grant Number; MR/S004793/1), and the National Institute for Health Research (Grant number; NIHR200632).

**Conflict of interest**

Nothing to Declare

**References**

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of molecular biology* **215**, 403-410.

Arredondo-Alonso, S., Rogers, M.R.C., Braat, J.C., Verschuuren, T.D., Top, J., Corander, J., Willems, R.J.L. and Schürch, A.C. (2018) mlplasmids: a user-friendly tool to predict plasmid- and chromosome-derived sequences for single species. *Microbial Genomics* **4**.

Bartosik, D., Sochacka, M. and Baj, J. (2003) Identification and characterization of transposable elements of *Paracoccus pantotrophus*. *Journal of bacteriology* **185**, 3753-3763.

Campos, J.C., da Silva, M.J., dos Santos, P.R., Barros, E.M., Pereira Mde, O., Seco, B.M., Magagnin, C.M., Leiroz, L.K., de Oliveira, T.G., de Faria-Júnior, C., Cerdeira, L.T., Barth, A.L., Sampaio, S.C., Zavascki, A.P., Poirel, L. and Sampaio, J.L. (2015) Characterization of Tn*3000*, a Transposon Responsible for *bla*NDM-1 Dissemination among Enterobacteriaceae in Brazil, Nepal, Morocco, and India. *Antimicrobial agents and chemotherapy* **59**, 7387-7395.

Chen, H., Venkat, S., Wilson, J., McGuire, P., Chang, A.L., Gan, Q. and Fan, C. (2018) Genome-Wide Quantification of the Effect of Gene Overexpression on *Escherichia coli* Growth. *Genes (Basel)* 9, 414.

Ciric, L., Brouwer, M.S., Mullany, P. and Roberts, A.P. (2014) Minocycline resistance in an oral *Streptococcus infantis* isolate is encoded by *tet*(S)on a novel small, low copy number plasmid. *FEMS microbiology letters* **353**, 106-115.

Ciric, L., Mullany, P. and Roberts, A.P. (2011) Antibiotic and antiseptic resistance genes are linked on a novel mobile genetic element: Tn*6087*. *The Journal of antimicrobial chemotherapy* **66**, 2235-2239.

Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic acids research* **16**, 6127-6145.

Dziewit, L., Baj, J., Szuplewska, M., Maj, A., Tabin, M., Czyzkowska, A., Skrzypczyk, G., Adamczuk, M., Sitarek, T., Stawinski, P., Tudek, A., Wanasz, K., Wardal, E., Piechucka, E. and Bartosik, D. (2012) Insights into the transposable mobilome of *Paracoccus spp.* (*Alphaproteobacteria*). *PloS one* **7**, e32277-e32277.

EUCAST (2020) European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters, version 10.0. 2020.: Växjö, Sweden.

Ferri, M., Ranucci, E., Romagnoli, P. and Giaccone, V. (2017) Antimicrobial resistance: A global emerging threat to public health systems. *Critical Reviews in Food Science and Nutrition* **57**, 2857-2876.

Gubellini, F., Verdon, G., Karpowich, N.K., Luff, J.D., Boël, G., Gauthier, N., Handelman, S.K., Ades, S.E. and Hunt, J.F. (2011) Physiological response to membrane protein overexpression in *E. coli*. *Mol Cell Proteomics* **10**, M111.007930.

Harmer, C.J. and Hall, R.M. (2015) IS*26*-Mediated Precise Excision of the IS*26*-*aphA1a* Translocatable Unit. *mBio* **6**, e01866-01815.

Harmer, C.J. and Hall, R.M. (2016) IS*26*-Mediated Formation of Transposons Carrying Antibiotic Resistance Genes. *mSphere* **1**.

Harmer, C.J. and Hall, R.M. (2020) IS*26* Family Members IS*257* and IS*1216* Also Form Cointegrates by Copy-In and Targeted Conservative Routes. *mSphere* **5**.

Harmer, C.J., Moran, R.A. and Hall, R.M. (2014) Movement of IS*26*-associated antibiotic resistance genes occurs via a translocatable unit that includes a single IS*26* and preferentially inserts adjacent to another IS*26*. *mBio* **5**, e01801-01814.

He, T., Shen, Y., Schwarz, S., Cai, J., Lv, Y., Li, J., Feßler, A.T., Zhang, R., Wu, C., Shen, J. and Wang, Y. (2016) Genetic environment of the transferable oxazolidinone/phenicol resistance gene *optrA* in *Enterococcus faecalis* isolates of human and animal origin. *Journal of Antimicrobial Chemotherapy* **71**, 1466-1473.

Hubbard, A.T.M., Mason, J., Roberts, P., Parry, C.M., Corless, C., van Aartsen, J., Howard, A., Fraser, A.J., Adams, E.R., Roberts, A.P. and Edwards, T. (2020) Within-patient evolution to piperacillin/tazobactam resistance in a clinical isolate of *Escherichia coli* due to IS*26*-mediated amplification of *bla*TEM-1B. *bioRxiv*, 2020.2002.2028.969360.

Jager, W., Schafer, A., Kalinowski, J. and Puhler, A. (1995) Isolation of insertion elements from Gram-positive *Brevibacterium*, *Corynebacterium* and *Rhodococcus* strains using the *Bacillus subtilis* *sacB* gene as a positive selection marker. *FEMS microbiology letters* **126**, 1-6.

Lartigue, M.F., Poirel, L., Aubert, D. and Nordmann, P. (2006) In vitro analysis of IS*Ecp1B*-mediated mobilization of naturally occurring beta-lactamase gene *bla*CTX-M of *Kluyvera ascorbata*. *Antimicrobial agents and chemotherapy* **50**, 1282-1286.

Leclercq, S., Gilbert, C. and Cordaux, R. (2012) Cargo capacity of phages and plasmids and other factors influencing horizontal transfers of prokaryote transposable elements. *Mobile genetic elements* **2**, 115-118.

Lee, Y.A., Fan, S.C., Chiu, L.Y. and Hsia, K.C. (2001) Isolation of an insertion sequence from *Ralstonia solanacearum* race 1 and its potential use for strain characterization and detection. *Applied and environmental microbiology* **67**, 3943-3950.

Lessard, P.A., O'Brien, X.M., Ahlgren, N.A., Ribich, S.A. and Sinskey, A.J. (1999) Characterization of IS*1676* from *Rhodococcus erythropolis* SQ1. *Applied microbiology and biotechnology* **52**, 811-819.

Liang, B., Roberts, A.P., Xu, X., Yang, C., Yang, X., Wang, J., Yi, S., Li, Y., Ma, Q., Wu, F., Qiu, S. and Song, H. (2018) Transferable Plasmid-Borne *mcr-1* in a Colistin-Resistant *Shigella flexneri* Isolate. *Applied and environmental microbiology* **84**.

Mikosa, M., Sochacka-Pietal, M., Baj, J. and Bartosik, D. (2006) Identification of a transposable genomic island of *Paracoccus pantotrophus* DSM 11072 by its transposition to a novel entrapment vector pMMB2. *Microbiology (Reading, England)* **152**, 1063-1073.

Moyo, S., Manyahi, J., Blomberg, B., Tellevik, M., Masoud, N., Aboud, S., Manji, K., Roberts, A., Hanevik, K., Mørch, K. and Langeland, N. (2020) Bacteraemia, malaria and case fatality among children hospitalised with severe blood stream infection in Dar es Salaam, Tanzania. *Frontiers in Microbiology* **(In press)**.

Partridge, S.R., Kwong, S.M., Firth, N. and Jensen, S.O. (2018) Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clinical microbiology reviews* **31**, e00088-00017.

Poirel, L., Jayol, A., Bontron, S., Villegas, M.V., Ozdamar, M., Turkoglu, S. and Nordmann, P. (2015) The *mgrB* gene as a key target for acquired resistance to colistin in *Klebsiella pneumoniae*. *The Journal of antimicrobial chemotherapy* **70**, 75-80.

Raabe, T., Jenny, E. and Meyer, J. (1988) A selection cartridge for rapid detection and analysis of spontaneous mutations including insertions of transposable elements in Enterobacteriaceae. *Mol Gen Genet* **215**, 176-180.

Roberts, A.P., Chandler, M., Courvalin, P., Guédon, G., Mullany, P., Pembroke, T., Rood, J.I., Jeffery Smith, C., Summers, A.O., Tsuda, M. and Berg, D.E. (2008) Revised nomenclature for transposable genetic elements. *Plasmid* **60**, 167-173.

Roope, L.S.J., Smith, R.D., Pouwels, K.B., Buchanan, J., Abel, L., Eibich, P., Butler, C.C., Tan, P.S., Walker, A.S., Robotham, J.V. and Wordsworth, S. (2019) The challenge of antimicrobial resistance: What economics can contribute. *Science (New York, NY)* **364**, eaau4679.

Schneider, D., Faure, D., Noirclerc-Savoye, M., Barrière, A.-C., Coursange, E. and Blot, M. (2000) A Broad-Host-Range Plasmid for Isolating Mobile Genetic Elements in Gram-Negative Bacteria. *Plasmid* **44**, 201-207.

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J. and Chandler, M. (2006) ISfinder: the reference centre for bacterial insertion sequences. *Nucleic acids research* **34**, D32-36.

Solyga, A. and Bartosik, D. (2004) Entrapment vectors--how to capture a functional transposable element. *Polish journal of microbiology / Polskie Towarzystwo Mikrobiologow = The Polish Society of Microbiologists* **53**, 139-144.

Tansirichaiya, S. (2017) Investigation of Mobile Genetic Elements and Antimicrobial Resistance Genes in Human Oral Metagenomic DNA: UCL (University College London).

Tansirichaiya, S., Mullany, P. and Roberts, A.P. (2016) PCR-based detection of composite transposons and translocatable units from oral metagenomic DNA. *FEMS Microbiology Letters* **363**.

Tansirichaiya, S., Rahman, M.A. and Roberts, A.P. (2019) The Transposon Registry. *Mobile DNA* **10**, 40.

Ulrich, A. and Puhler, A. (1994) The new class II transposon Tn*163* is plasmid-borne in two unrelated *Rhizobium leguminosarum* biovar *viciae* strains. *Mol Gen Genet* **242**, 505-516.

Vandecraen, J., Chandler, M., Aertsen, A. and Van Houdt, R. (2017) The impact of insertion sequences on bacterial genome plasticity and adaptability. *Critical Reviews in Microbiology* **43**, 709-730.

Zeng, L., Zhan, Z., Hu, L., Jiang, X., Zhang, Y., Feng, J., Gao, B., Zhao, Y., Yang, W., Yang, H., Yin, Z. and Zhou, D. (2019) Genetic Characterization of a *bla*VIM–24-Carrying IncP-7β Plasmid p1160-VIM and a *bla*VIM–4-Harboring Integrative and Conjugative Element Tn*6413* From Clinical *Pseudomonas aeruginosa*. *Frontiers in Microbiology* **10**.

Zhao, Y., Wang, L., Zhang, Z., Feng, J., Kang, H., Fang, L., Jiang, X., Zhang, D., Zhan, Z., Zhou, D. and Tong, Y. (2017) Structural genomics of pNDM-BTR harboring In191 and Tn*6360*, and other *bla* NDM-carrying IncN1 plasmids. *Future microbiology* **12**, 1271-1281.

**Table 1** Bacterial strains and plasmids used in this study

|  |  |  |  |
| --- | --- | --- | --- |
| **Strains or Plasmids** | **Characteristics** | **Resistance phenotypes** | **References** |
| Strains |  |  |  |
| *E. coli* |  |  |  |
| EPI300 | - Electrocompetent - Inducible *trfA* gene for high-copy number of pCC1BAC | - | Epicenter, UK |
| NEB® 5-alpha | - Chemically competent | - | New England BioLabs, UK |
| NEB® 5-alpha:: pGEM-2026-WT-*cI* | - *E. coli* NEB® 5-alpha containing pGEM-2026-WT-*cI* | Ampicillin | This study |
| NEB® 5-alpha:: pGEM-2026-*dfrA*8-TU | - *E. coli* NEB® 5-alpha containing pGEM-2026-*dfrA*8-TU | Ampicillin, Trimethoprim | This study |
| EC1168 | - Clinical isolate | N/A | Moyo et al. (2020) |
| EC1168::pBACpAK | - *E. coli* EC1168 containing pBACpAK | Chloramphenicol | This study |
| EC1359A | - Clinical isolate | N/A | Moyo et al. (2020) |
| EC1359A::pBACpAK | - *E. coli* EC1359A containing pBACpAK | Chloramphenicol | This study |
| EC1444 | - Clinical isolate | N/A | PRJNA647256 |
| EC1444::pBACpAK | - *E. coli* EC1444 containing pBACpAK | Chloramphenicol | This study |
| EC2026 | - Clinical isolate | N/A | PRJNA647258 |
| EC2026::pBACpAK | - *E. coli* EC2026 containing pBACpAK | Chloramphenicol | This study |
| EC2026::pBAKpAK::*dfrA8*-TU | - *E. coli* EC2026::pBAKpAK with an insertion *dfrA8* TU | Chloramphenicol, Tetracycline | This study |
| EC2033 | - Clinical isolate | N/A | PRJNA647259 |
| EC2033::pBACpAK | - *E. coli* EC2033 containing pBACpAK | Chloramphenicol | This study |
| Plasmids |  |  |  |
| pCC1BAC | Large insert, Single copy but inducible in *E. coli* EPI300 | Chloramphenicol | Epicenter, UK |
| pAK1 | Small insert, high copy number, *cI*-*tetA* selection cartridge | Kanamycin | Department of Bacterial Genetics, University of Warsaw, Poland |
| pBACpAK | Large insert, Single copy, *cI*-*tetA* selection cartridge | Chloramphenicol | This study |
| pGEM-T easy | Small insert, High copy number, TA cloning | Ampicillin | Promega, UK |
| pGEM-2026-WT-*cI* | pGEM-T easy containing wild-type *cI* amplicon from *E. coli* EC2026::pBACpAK | Ampicillin | This study |
| pGEM-2026-*dfrA*8-TU | pGEM-T easy containing *cI* amplicon from *E. coli* EC2026::pBACpAK-IS*26*-*dfrA8*-TU | Ampicillin, Trimethoprim | This study |

**Table 2** The details of MGEs captured by pBACpAK entrapment vectors

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Bacterial strains  (Accession number) | MGEs | Size (bp) | Posterior  probability  scores\* | Percentage of similarity (ISFinder/NCBI) | | | Direct repeats |
| BlastN | BlastX | Accession number |
| *E. coli* EC1444  (PRJNA647256) | IS*903*B | 1057 | 0.965/0.567 (C) | 97% | 99% | X02527 | GCTCATACG |
| *E. coli* EC2026  (PRJNA647258) | IS*26* | 820 | 0.554/0.527 (P) | 100% | 100% | X00011 | GAAATCAC, GCATTTAA |
| IS*26*-*dfrA8* TU | 2002 | 0.554 (P) | 100% | - | CP044149 | - |
| -IS*26* | 820 | - | 100% | X00011 |
| -*dfrA8* | 510 | - | 100% | WP\_000571065 |
| -Response regulator | 633 | - | 100% | EAA0965179 |
| *E. coli* EC2033  (PRJNA647259) | IS*26* | 820 | 0.854/0.554 (P) | 100% | 100% | X00011 | CGCCTGAC |
| IS*Sbo1* | 1714 | 0.957 (P) | 97% | 97% | CP001062 | - |

**\*** The scores were calculated by using mlplasmid software to predict that the contigs containing each MGE were likely to be either chromosomal-derived (C) or plasmid-derived (P) DNA. The higher the score means the higher chance for the contigs to be derived from each class. Multiple figures means that MGEs were found on multiple contigs in the WGS data.

**Table 3. Number of analysed tetracycline resistance clones by using pBACpAK**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Bacterial strains** | **Number of analysed Tetracycline resistant colonies (%)** | | | |
| **Point mutations** | **Insertion of MGEs** | | **Total** |
| **MGEs** | **Number of colonies** |
| *E. coli* EC1444 | 0 (0) | IS*903*B | 1 (100) | 1 |
| *E. coli* EC2026 | 8 (16.7) | IS*26* | 37 (77.1) | 48 |
| IS*26*-*dfrA8* TU | 3 (6.3) |
| *E. coli* EC2033 | 1 (1.9) | IS*26* | 16 (30.8) | 52 |
| IS*Sbo1* | 35 (67.3) |

**Figure Captions**

**Figure 1.** ***cI*-*tetA* selection cartridge selection system.** The expression of *tetA* tetracycline resistance gene was inhibited by λ repressors. An insertion of MGEs in *cI* gene disrupts the expression of λ repressors, allowing *tetA* to be expressed and conferring tetracycline resistance phenotype. The yellow and blue open arrowed boxes, represent *cI* and *tetA* genes, respectively, which point the direction of transcription. MGE, λ repressors and tetracycline resistance protein were shown as yellow, green and blue rectangles. The black and orange arrows represent promoters and direction with an active translated protein, respectively.

**Figure 2.The structure of pBACpAK entrapment vector and the analysis of tetracycline resistance clones by colony PCR.** pBACpAK was constructed based on pCC1BAC vector by inserting *cI-tetA* selection cartridge at the *Xho*I site. The screening of tetracycline resistance clones was done by colony PCR with *cI*-*tetA*-F1 and ERIS primers (red arrows). The orange, blue, yellow and black open arrowed boxes represent chloramphenicol resistance gene (ChlR), *tetA* resistance gene, *cI* gene and other genes on the vector, pointing in the direction of transcription. The asterisk, dashed box, green box and blue boxes represent small mutations, deletion, mobile genetic element and partial *tetA* in colony PCR amplicons, respectively.

**Figure 3. The structure of *cI-tetA* colony PCR amplicons containing MGEs captured by pBACpAK entrapment vector.** The blue boxes, yellow, green and orange open arrowed boxes represent *tetA*, *cI* MGEs and other genes, respectively.

**Figure 4.** **Genomic comparison of the whole genome sequencing data and the captured MGEs from *E. coli* EC2026and EC2033*.*** The identical DNA regions of MGEs and other genes are shown in yellow and blue, respectively. The green and orange open arrowed boxes represent IS*26* and other genes, respectively.

**Figure 5. Confirmation and characterization of the IS*26*-*dfrA8* TU.** a.) A predicted structure of IS*26*-*dfrA8* TU and the expected PCR product from the TU confirmation PCR. The black arrows and the yellow dashed line indicate primers and the expected size of the PCR. The green and orange open arrowed box represent IS*26* and other genes, respectively. b.) The TU confirmation PCR product on a 1% agarose gel containing HyperLadder 1kb in Lane M. c.) Cultures of *E. coli* NEB® 5-alpha:: pGEM-2026-WT-*cI* and NEB® 5-alpha:: pGEM-2026-*dfrA8*-TU on LB agar supplemented with and without trimethoprim (+TMP and -TMP, respectively).

**Figure 6. An insertion of IS*26*-*dfrA8* TU into pBACpAK.** The IS*26*-*dfrA8* TU could translocate into the *cI* gene directly with a.) copy-in mechanism or with another copy IS*26* through b.) targeted conservative mechanism, forming a composite transposon, which was followed by a loss of one IS*26* from the structure. The blue, yellow, green and orange open arrowed boxes represent *tetA*, *cI* MGEs and other genes, respectively.