Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 29:1-12

Developing a BrkA-based Protein Expression System in *Escherichia coli* Using the Mini-Tn7 Vector System

Dana Lao, Faith Liu, Ava Tiefenbach, Nicholas Viegas

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY Autotransporters are a family of proteins secreted to the cell surface of Gram-negative bacteria that require only the autotransporter protein itself for transport. Many of these autotransporters, such as *Bordetella* resistance to killing (BrkA) protein found in the organism *Bordetella pertussis*, have applications in the realm of protein expression and presentation. Use of autotransporters like BrkA in protein expression systems has largely been confined to the context of plasmid-based systems, which have known limitations including harming bacterial host cell fitness. This has sparked interest in alternatives that bypass these limitations, such as using transposons to mediate insertion of desired genes into the bacterial genome. We leveraged a mini-Tn7 system to introduce a single copy of *brkA* into the *Escherichia coli* genome. 6xHis tagged-*brkA* under the control of the *lac* promoter was amplified and inserted into a linearized mini-Tn7 vector via Gibson assembly. DH10B *E. coli* cells were transformed with our Gibson product together with a helper plasmid to catalyze transposition, with transposition being demonstrated by the presence of gentamicin resistance colonies. *lac*/*brkA*-6xHis was inserted into our transposon vector, and cells transformed with the vector demonstrated gentamicin resistance; however, we were unable to confirm the single insertion of *brkA* into the vector and its specific incorporation into the *E. coli* genome at the expected site through PCR. Sequencing results for our Gibson assembly product also warrant further investigation.

INTRODUCTION

Gram-negative bacteria employ secretion systems to facilitate the transportation of effector molecules across their outer membranes (1-3). Autotransporters are members of the Type V secretion system and a family of proteins that are secreted to the cell surface in Gram-negative bacteria (1, 3). The general structure of autotransporters includes an N-terminal signal peptide which mediates protein transport across the inner membrane, a passenger domain that is secreted and is responsible for the protein's effector functions, and a C-terminal domain enabling transfer across the outer membrane (3). *Bordetella* resistance to killing (BrkA) protein is an autotransporter found in the organism *Bordetella pertussis* that is involved in serum resistance and bacterial adherence (4). The BrkA passenger domain undergoes self-cleavage upon transportation across the bacterial outer membrane but has notably been found to continue adhering tightly to the bacterial surface post-cleavage, a characteristic distinct from most autotransporters (3, 4). The unprocessed BrkA protein is 103 kDa in size, yielding a 73 kDa passenger domain and a 30 kDa transporter domain upon cleavage (3).

Autotransporters like BrkA have a wide range of biotechnological and biomedical applications including the presentation of epitopes for live vaccines and protein secretion (5). Use of autotransporters in these protein expression systems is attractive due to the relative simplicity of the secretion system, as only the autotransporter protein itself is required for transportation, and because of the system's modularity (5). Surface-display-based systems are also advantageous in several ways compared to systems involving intracellular protein expression: purification of the product is greatly simplified, proteins gain stability from being anchored to the host cell membrane, and surface display allows the protein to interact with molecules in the extracellular matrix (5). BrkA is particularly suitable for bacterial surface display of proteins due to its passenger domain remaining strongly associated with the bacterial surface after transportation across the outer membrane (3, 4). Several groups have

Published Online: September 2024

Citation: Lao, Liu, Tiefenbach, Viegas. 2024. Developing a BrkA-based protein expression system in *Escherichia coli* using the mini-Tn7 vector system. UJEMI 29:1-12

Editor: Shruti Sandilya and Ronja Kothe, University of British Columbia

Address correspondence to: https://jemi.microbiology.ubc.ca/

explored the potential of BrkA-based protein expression systems; for example, Sun *et al*. created the BrkAutoDisplay system which leveraged the BrkA autotransporter for membrane display of several proteins including green fluorescent protein, functional enzymes, and a functional single chain antibody fragment in *E. coli* hosts (4).

Plasmid-based protein expression systems such as the BrkAutoDisplay system are widely used for a variety of industrial and research purposes, often using *E. coli* as the host organism due to it being well-characterized and having fast growth kinetics (4, 6, 7). However, the literature notes various obstacles in the application of plasmid-based systems, including the formation of inclusion bodies, heterologous proteins being nonfunctional, and protein production negatively affecting host cell fitness and growth (6, 7). Concurrently, inducible systems such as the T7 expression system which aim to provide greater control over protein expression, tend to suffer from leaky protein expression, often to the detriment of host cell fitness (7). Given these drawbacks, alternatives to plasmid-based protein expression systems are of great interest.

One potential alternative to plasmid-based systems is to use transposon-mediated genomic insertion. Transposons are genetic elements able to move within and across different genomes (8). While many transposon systems generate random insertions in the target genome, certain transposon systems, such as the Tn7 system, specifically insert to a single target site in bacterial genomes with high fidelity (9). Together with a helper plasmid containing a transposase enzyme, such as pTNS2 (10), the Tn7 system integrates DNA fragments in an orientation-specific manner into a single site in bacterial genomes known as the *att*Tn7 site (10-11). Most bacteria, including *E. coli*, contain a single *att*Tn7 site within their genomes that is located downstream of the *glmS* gene that encodes for glucosamine-fructose-6-phosphate aminotransferase (11).

The mini-Tn7 system is an optimized version of the Tn7 system with features that increase its suitability in various applications. This system has a broad host-range, with the only major constraints on bacterial host selection being gene transfer techniques and antibiotic resistance markers on available vectors (9). Addition of transcriptional terminators prevents chromosomal promoters from affecting expression of inserted sequences, and the mini-Tn7 vectors contain multiple cloning sites (MCS) to enable the addition of desired sequences into the vectors (9). The mini-Tn7 system uses suicide delivery vectors such that the vector does not propagate in the intended host cells; for example, certain vectors require a bacterial strain positive for the *pir* gene for propagation of the delivery vector, and strains that are *pir* negative are then used for transposition (9). A variety of different mini-Tn7 vectors containing different antibiotic resistance cassettes have been previously generated, enabling myriad potential applications (9). The use of two antibiotic resistance markers in some mini-Tn7 vectors, one that is transposed and one that is not, allows for specific selection for transposed cells.

There are several advantages to using the mini-Tn7 system for DNA insertion into bacterial genomes. Upon transposition, continued antibiotic selection is no longer required because the system predictably generates a single genomic insertion at the *att*Tn7 site which has been found to be maintained for 100 bacterial generations or more (9). Mini-Tn7-mediated insertions at the *att*Tn7 site in bacterial host cells have also been observed to have minimal effects on bacterial viability and growth (7).

The mini-Tn7 system has been used for single copy number gene insertion in many applications. Jittawuttipoka *et al*. used the mini-Tn7 system for gene expression analysis in *Xanthomonas campestris* and *Xanthomonas oryzae*, as vectors for genomic integration in these species were scarce (8). Romero-Jiménez *et al*. leveraged the stability of the mini-Tn7 system to overexpress c-di-GMP in Gram-negative bacteria to examine c-di-GMP regulatory networks (11). Lambertsen *et al*. also used the mini-Tn7 system for fluorescent tagging of *Pseudomonas putida* (12).

In this study, we aimed to use the mini-Tn7 vector system to integrate the *brkA* gene into the *E. coli* genome. We hypothesized that if *brkA* were inserted into the *E. coli* genome via transposon-mediated genomic insertion, display of the BrkA autotransporter at the bacterial host cell membrane would be observed. We amplified *brkA*-6xHis with the *lac* promoter from the pENS plasmid (Goh *et al*., unpublished manuscript), inserted the gene segments into our transposon vector via Gibson assembly, and transformed *E. coli* cells with our vector

products. Sequencing analysis of our Gibson product highlighted the presence of unexpected constructs, and genomic PCR conducted on the expected *brkA*-containing portion of the transformed *E. coli* genome failed, necessitating further exploration of this system.

METHODS AND MATERIALS

Bacterial strains, plasmids and growth conditions. All bacterial strains and plasmids used in this study are listed in Table 1. Bacterial strains were grown at 37°C, and were streaked on Luria-Bertani (LB) agar plates and/or 5 mL of LB media on a shaker at 200 rpm. If plasmid conferred antibiotic resistance, plates or media were made with 100 μg/mL ampicillin and/or 10 μg/mL or 15 μg/mL gentamicin accordingly.

TABLE. 1 Bacterial strains and plasmids used

	Description	Source
***E. coli* strains**		
EC100D *pir*-116	*pir*+, no selection	Koch *et al.* (16)
EC100D *pir*-116	*pir*+, strain containing pUC18R6K-mini-Tn7T-Gm	Koch *et al.* (16)
DH10B	Commercial electrocompetent cells, no selection	GeneArt™ Gibson Assembly EX Cloning Kit (Invitrogen)
NEB-5α	Strain containing pENS	Goh *et al*., 2023, unpublished manuscript
Pir1	Strain containing pTNS2	Choi *et al.* (10)
Plasmids		
pENS	Plasmid with *brkA* + 6xHis tag, ampicillin resistance.	Goh *et al*., 2023, unpublished manuscript
pUC18R6K-mini-Tn7T-Gm	Mini-Tn7 transposon vector, gentamicin and ampicillin resistance	Koch *et al.* (16)
pTNS2	Helper plasmid, ampicillin resistance	Addgene Plasmid #64968

Plasmid and genomic DNA extraction. Plasmid extraction was performed to extract plasmids pUC19, pENS, pUC18R6K-mini-Tn7T-Gm, pTNS2, and Gibson assembly products pDFAN-1, pDFAN-2.1, and pDFAN2.2 using the Bio Basic EZ-10 Spin Column Plasmid DNA Miniprep Kit Cat #BS614 (Bio Basic). Extraction protocol was performed according to the manufacturer's instructions. Genomic DNA extraction was performed to extract genomic DNA from both untransformed and transformed DH10B *E. coli* cells using the EZ-10 Spin Column Bacterial Genomic DNA Miniprep Kit Cat # BS624 (Bio Basic). Extraction protocol was performed according to the manufacturer's instructions.

PCR primer design and use. Primers were designed using SnapGene and were evaluated for self-dimerization using UniProt's UGENE OligoAnalyzer tool. Information for all PCR primers used in this study, including sequences, are listed in Table 2. Primers were used to test for genomic insertion of the *lac* promoter, *brkA* gene, and 6xHis tag from pENS within the DH10B *E. coli* genome. The *lac* promoter + *brkA* + 6xHis tag in pENS will be referred to as the *lac*/*brkA*-6xHis, *pstS* in the DH10B genome will be referred to as the *pstS* control, and the DH10B *pstS* and hypothetical *brkA* genomic insertion will be referred to as the genomic confirmation segment (Table 2). Using a forward primer complementary to a segment in the genome of DH10B (fwd genomic) and a reverse primer complementary to the hypothetical *brkA* gene insertion (rev genomic), PCR was performed on the genomic DNA of transformed DH10B cells (from grown colonies named A1, A3, A4, A7). These primers were used to determine whether *brkA* was inserted at the *att*Tn7 site of the DH10B genome. The *lac*/*brkA*-6xHis primers (fwd *brkA,* rev *brkA*) were used in the initial amplification of *lac*/*brkA*-6xHis and were used to assess the insertion of *brkA* anywhere in the DH10B genome. pUC19 primers (fwd pUC19 and rev pUC19) that amplified a portion of *lacZα* in pUC19 and primers that amplified *pstS* in the DH10B genome (fwd *pstS* and rev *pstS*) were used as positive

controls. Additionally, combinations of the three sets of primers were used: fwd *pstS* + rev *brkA* and fwd *pstS* + rev genomic.

TABLE. 2 PCR primers.

Region	Short-hand Name of Region	Primer Sequence	Melting Temperature	GC Content
lac promoter + *brkA* + 6xHis tag in pENS	*lac*/*brkA*-6xHis	fwd brkA: 5'-AGCTAATTCGATCATGCATGAGCTCACTAGTGGATCCCCCGATACCGCTCGCCGC-3'	72.2°C	56.4%
		rev *brkA*: 5'-GTTTGGAACTAGATTTCACTTATCTGGTTGGCCTGCAAGGAAGCCCGCTGCTGCTC-3'	70.8°C	51.8%
pstS in the DH10B genome	*pstS* control	fwd *pstS*: 5'-GTACGCCCAGTCGAAGAATT-3'	72.5°C	55.4%
		rev *pstS*: 5'-ATGGGCTGACACTTACCAGA-3'	70.8°C	51.8%
DH10B *pstS* and hypothetical *brkA* genomic insertion	Genomic confirmation segment	fwd genomic: 5'-TGTCACATGGAGTTGGCAGG-3'	57.6°C	55%
		rev genomic: 5'-CGTCAAGGTGGACGCCAACA-3'	60.4°C	60%
Portion of *lacZα* in pUC19	pUC19 control	fwd pUC19: 5'-GTGAAATACCGC ACAGATGC-3'	68.2°C	50%
		rev pUC19: 5'-GGCGTTACCCAA CTTAATCG-3'	68.2°C	50%

Restriction digestion and Gibson assembly. Restriction digestion was performed using the enzymes StuI Cat #15438-013 (Invitrogen) and SmaI Cat #15228-018 (Invitrogen) to linearise the pUC18R6K-mini-Tn7T-Gm vector. Restriction enzyme digest reactions were set up using adapted Invitrogen protocols for each enzyme, and rCutSmart buffer Cat #B6004S (New England Biolabs) was used instead of the buffers listed in the protocol. The digestion was run overnight at 37°C. Gibson assembly was performed following the protocol specified by the GeneArt™ Gibson Assembly EX Cloning Kit Cat #A46634 (Invitrogen).

Transformation. For the EC100D *pir*-116 *E. coli* cells, the cells were made electrocompetent and transformed according to an adapted version of Choi & Schweizer's protocol (9). 1.5 mL of EC100D *pir*-116 *E. coli* overnight cultures were centrifuged at 16000 *g* for 2 minutes at room temperature and resuspended in 1 mL of 300 mM sucrose twice. Cultures were centrifuged once more and resuspended in 300 mM sucrose in preparation for electroporation. 5 µL of 1:5 diluted Gibson product was added to 100 µL of cells. No Gibson product was added to the control. Bio-Rad Gene Pulser/MicroPulser 0.1 cm gap electroporation cuvettes Cat #1652089 (Bio-Rad) and the Bio-Rad Micropulser Electroporator Cat #1652100 (Bio-Rad) at the Ec 1 setting was used to electroporate 100 µL of resuspended cells at a time. 1 mL of LB broth was immediately added to the cells after electroporation, and cells were incubated on a 37°C orbital shaker at 225 rpm for 1 hour. Cells were then plated on LB agar + 100 µg/mL ampicillin and 15 µg/mL gentamicin and left to grow at 37°C over two days. For the DH10B electrocompetent *E. coli* cells included in the GeneArt™ Gibson Assembly EX Cloning Kit Cat #A46634 (Invitrogen), transformation was performed using the manufacturer's protocol. Together with the Gibson assembly product, 50 ng of pTNS2 helper plasmid was added for co-transformation. As a negative control, pUC19 DNA was transformed into electrocompetent DH10B cells. Cells were then plated on LB agar + 15 µg/mL gentamicin and grown overnight at 37°C.

Polymerase chain reaction (PCR). All PCRs were performed in volumes of 20 µL using Platinum SuperFi II DNA Polymerase Cat #12361010 (Invitrogen) in a T100 Thermal Cycler Cat #1861096 (Bio-Rad) according to polymerase user guide instructions. The following parameters were used: an initial denaturation step at 98°C for 30 s followed by 30 cycles of

94°C for 8 s, 60°C for 10 s, and 72°C for 15-23 s/kb varying based on template DNA. The last cycle was followed by a final extension step at 72°C for 5 min.

PCR cleanup. Purification of PCR products was performed using the EZ-10 Spin Column PCR Products Purification Kit Cat #BS664 (Bio Basic) according to manufacturer instructions.

Agarose gel electrophoresis. Agarose gels were prepared using 1X tris-acetate ethylenediaminetetraacetic acid (TAE) buffer to a concentration of 1% agarose. RedSafe Nucleic Acid Staining Solution Cat #21141 (FroggaBio) was added to the agarose mixture to a final concentration of 1X. Samples were dyed using either 6X DNA Gel Loading Dye Cat #R0631 (Fermentas) or 6X Gel Loading Dye, Blue Cat #B7021S (New England Biolabs). O'GeneRuler DNA Ladder Mix Cat #SM1173 (Thermo Scientific), GeneRuler 100 bp Plus DNA Ladder Cat #SM0323 (Thermo Scientific), and/or 100 bp DNA Ladder Cat #566248 (Invitrogen) were loaded on gels. Agarose gels were run at 90V-120 V for approximately 1 hour and visualized using the ChemiDoc Imaging System Cat #1708370 (Bio-Rad).

Plasmid sequencing. Extracted plasmids were shipped to Plasmidsaurus for nanopore sequencing according to company specifications. Sequencing results were analyzed in SnapGene.

RESULTS

To use the mini-Tn7 system to introduce a single copy of *lac*/*brkA*-6xHis into the genome of DH10B *E. coli,* the *lac*/*brkA*-6xHis segment from the pENS plasmid was first amplified, extracted, and inserted into the linearized pUC18R6K-mini-Tn7T-Gm transposon backbone using Gibson assembly to create our mini-Tn7 vector (Fig 1). The vector was then co-transformed into DH10B *E. coli* with the pTNS2 helper plasmid for transposition. Additionally, the Gibson product was transformed into EC100D *pir*-116 *E. coli* for extraction and sequencing (Fig 1).

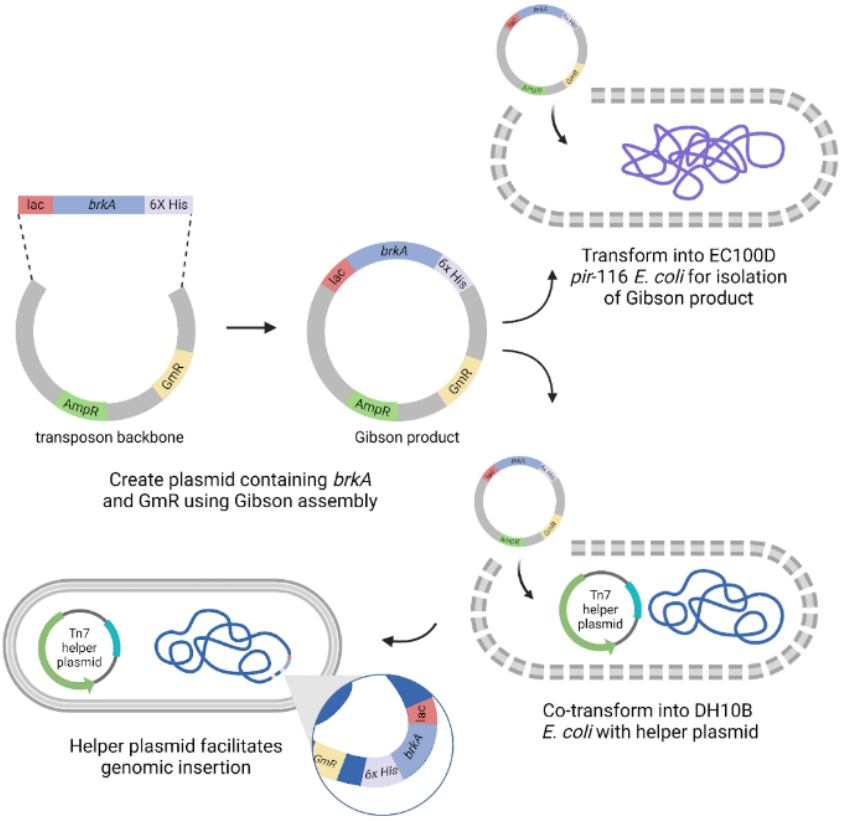

FIG. 1 Schematic of the experimental procedure. The segment containing the *lac* promoter, *brkA*, and a 6xHis tag from the pENS plasmid was amplified using PCR. Restriction enzymes linearized the pUC18R6K-mini-Tn7T-Gm vector to create the transposon backbone, then Gibson assembly was used to insert the amplified segment into the backbone. The Gibson product was co-transformed into DH10B *E. coli* with the helper plasmid for transposition. In parallel, the Gibson product was transformed into EC100D *pir*-116 *E. coli* for extraction and sequencing of the product. Created with BioRender.com.

***lac*/*brkA*-6xHis was amplified from the pENS plasmid.** To amplify *lac*/*brkA*-6xHis from the pENS plasmid in NEB-5α *E. coli*, PCR primers were designed with sticky ends

homologous to the exposed regions in the linearized pUC18R6K-mini-Tn7T-Gm plasmid. PCR was then performed on the pENS plasmid and the products were visualized on an 1% agarose gel (Fig 2). As a negative control, we set up a PCR reaction with the nuclease-free water and *brkA* primers. As a positive control, we set up a PCR reaction with pUC19 plasmid DNA and primers that are known to amplify a 163 bp region on the pUC19 plasmid. PCR products from the amplification of *lac*/*brkA*-6xHis from the pENS plasmid in NEB-5α *E. coli* yielded a ~4.2 kb band, which matches the expected length of this segment (Fig 2). The small product expected from pUC19 amplification with the pUC19 primers was observed. Faint bands around the size of the *lac*/*brkA*-6xHis segment were also observed in the negative control, indicating contamination of the PCR master mix. These bands were less visible than those we postulated to be *lac*/*brkA*-6xHis in the experimental lanes.

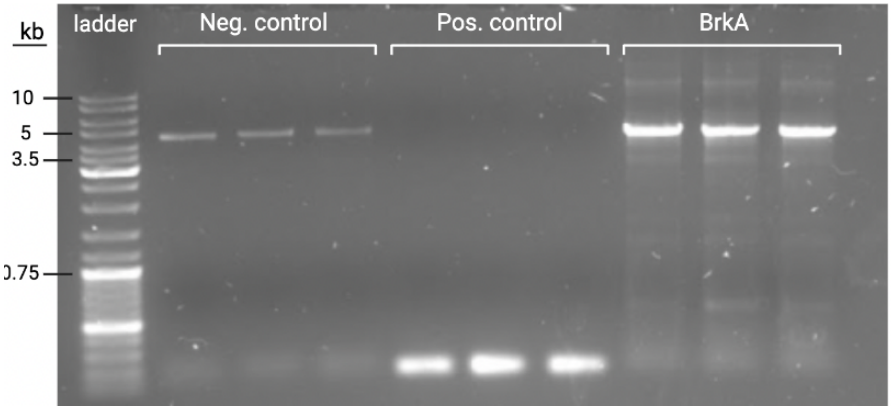

FIG. 2 *lac/brkA*-6xHis from pENS was amplified. fwd *brkA* and rev *brkA* primers were used to amplify the 4.2 kb region containing the *lac* promoter and 6xHis-tagged *brkA* of the pENS plasmid. The negative control used the *brkA* primers with no DNA template. The positive control amplified a 163 bp portion of the pUC19 plasmid using known primers. The O'GeneRuler 1 kb ladder (ThermoFisher) was used as a ladder. Each reaction was completed in triplicate and ran at 90 V on a 1% agarose gel.

Linearization of the pUC18R6K-mini-Tn7T-Gm vector to create a transposon backbone. Restriction enzyme digestion was used to prepare pUC18R6K-mini-Tn7T-Gm for Gibson assembly by linearization. Given the locations of the SmaI and StuI cut sites, the size of the transposon backbone after cleavage at both cut sites would be expected to be 4.3 kb (Supplementary Fig 1). Gel electrophoresis of the restriction digest products cut by both enzymes yielded a distinct band at 4.3 kb, confirming vector linearization (Fig 3). Both single-enzyme digests also resulted in distinct bands, confirming both enzymes were active. The control reaction with no restriction enzyme yielded a larger smeared band, which is likely due to a non-linearized plasmid, thus confirming the band resulted from the restriction enzyme digestion.

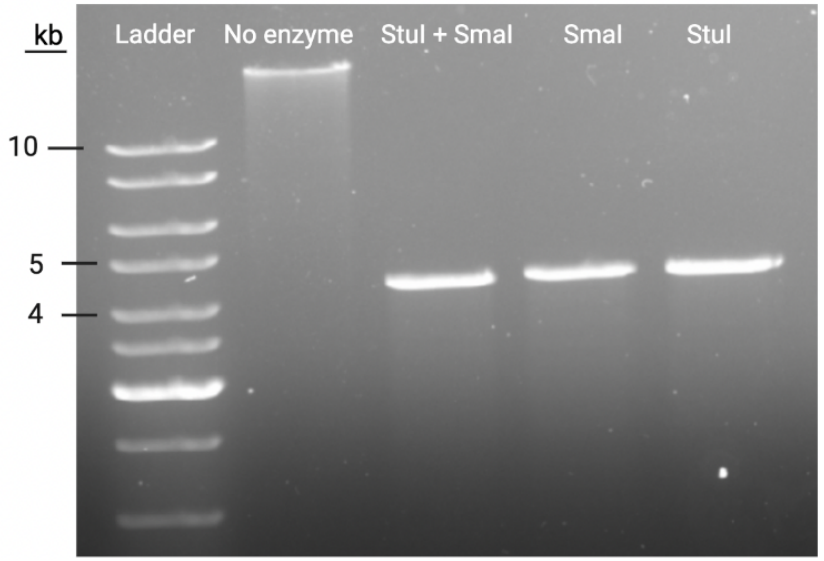

FIG. 3 pUC18R6K-mini-Tn7T-Gm was linearized through digestion with StuI and SmaI. StuI and SmaI (ThermoFisher) were used to linearize pUC18R6K-mini-Tn7T-Gm. Single-enzyme digests and a no-enzyme control were performed to confirm each enzyme was working. The size of the transposon backbone after double enzyme digestion is expected to be 4.3 kb. The O'GeneRuler 1 kb ladder (ThermoFisher) was used as a ladder. Samples were run at 100V on a 1% agarose gel.

Gibson products from insertion of *lac*/*brkA*-6xHis into linearized pUC18R6K-mini-Tn7T-Gm showed unexpectedly long DNA sequencing results. Amplified *lac*/*brkA*-6xHis was cloned into the MCS of the transposon backbone. Gibson assembly was performed, and

the product was transformed into electrocompetent EC100D *pir*-116 *E. coli* cells. Sequencing results (Plasmidsaurus) of the Gibson product extracted from two transformed EC100D *pir*-116 colonies (B1, B2) revealed three plasmids (pDFAN-1, pDFAN-2.1, pDFAN-2.2) with sizes of approximately 17 kb or 25 kb (Fig 4). The sequencing results also showed that *lac*/*brkA*-6xHis was inserted into the vector twice (Fig 4A, 4B, 4C). To verify sequencing results, Gibson products were further digested with SmaI resulting in ~7.5 kb bands when the pDFAN plasmids were digested with SmaI, whereas the undigested controls did not show distinct bands, but very faint products that were large in size, indicating that the plasmids were properly sequenced (Fig 4D).

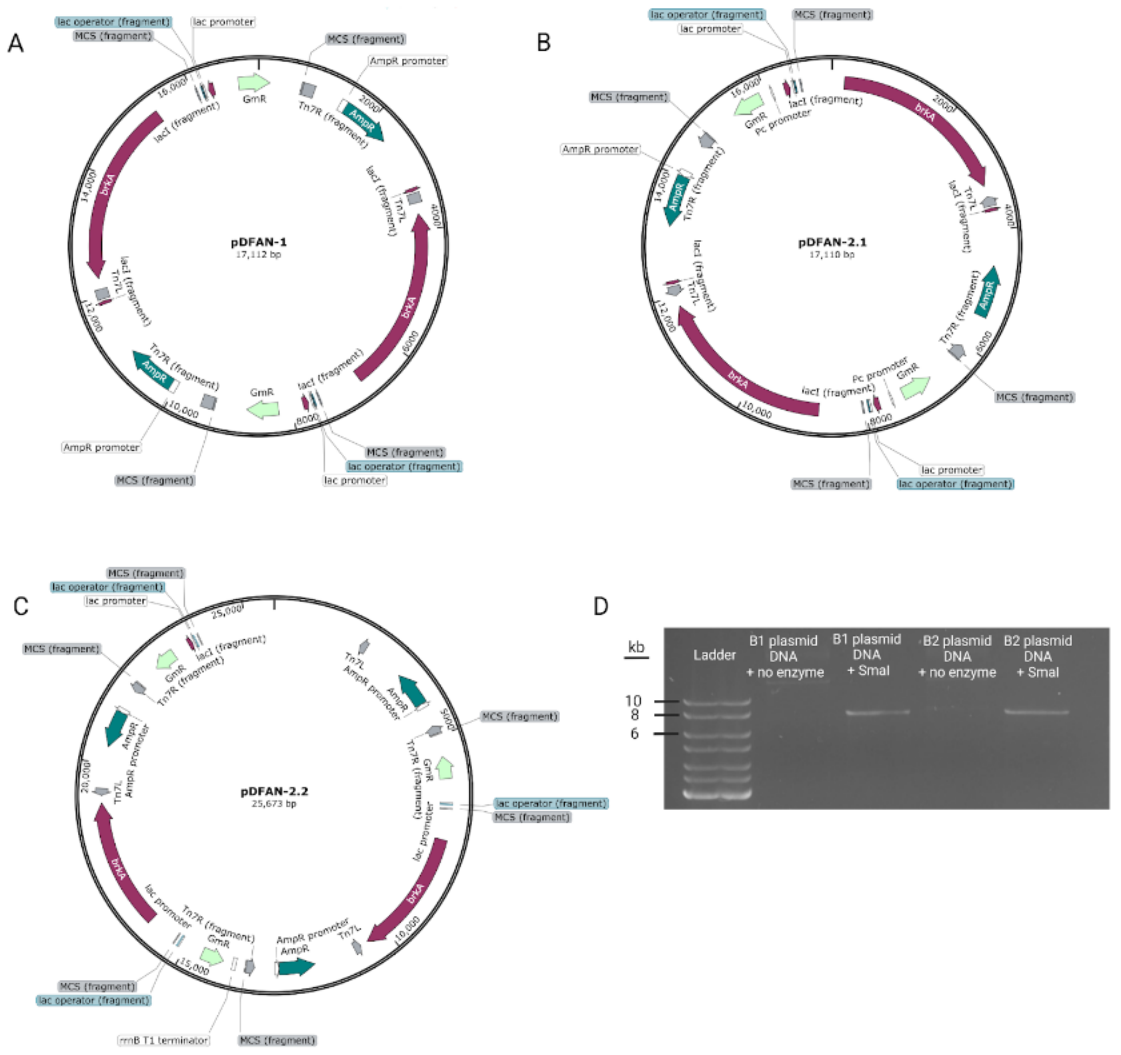

FIG. 4 Gibson assembly product sequencing results showed larger constructs than anticipated. Gibson products were transformed into EC100D *pir*-116 *E. coli* cells, extracted from transformed cells, and plasmids were sequenced. Restriction enzyme digestion with SmaI was performed to verify sequencing results. (A) pDFAN-1 sequence extracted from colony B1. Two *lac/brkA*-6xHis segments and two transposon backbones are present. (B) pDFAN-2.1 sequence extracted from colony B2. Two *lac/brkA*-6xHis segments and two transposon backbones are present. (C) pDFAN-2.2 sequence extracted from colony B2. Two *lac/brkA*-6xHis segments and three transposon backbones are present. (D) SmaI enzyme digest of plasmids B1 and B2. Lanes with enzyme digestion have bands around 7.5 kb, which is the expected size of the Gibson products sequenced.

Genomic insertion of *lac*/*brkA*-6xHis could not be confirmed through PCR. After transforming the direct product of Gibson assembly and pTNS2 helper plasmid into electrocompetent DH10B cells, the genomic DNA of the transformed colonies that grew

overnight (A1, A3, A4, A7) was extracted. PCR reactions were set to test the genomic insertion of *lac*/*brkA*-6xHis into the genome of transformed DH10B cells resolved on 1% agarose gels (Fig 5A). As positive controls, the *pstS* gene in the DH10B genome was amplified (Fig 5B), and a pUC19 positive control was also used to verify the reaction parameters (Fig 5B). As negative controls, H_2O (no template) and untransformed DH10B genomic DNA PCR reactions were performed as well (Fig 5A, 5B). The results shown by the agarose gels indicate that the genomic insertion of *lac*/*brkA*-6xHis at the *att*Tn7 site could not be confirmed. No bands were observed in the lanes with the transformed DH10B genomic DNA and fwd genomic and rev genomic primers, and fwd genomic and rev *brkA* primers (Fig 5A). Furthermore, the lanes with transformed DH10B genomic DNA and fwd *pstS* and rev *brkA* primers also did not have any bands (Fig 5A). Bands are seen in the positive control lanes (fwd *pstS*, rev *pstS*) and the pUC19 control, verifying the template DNA, PCR master mix, and reaction parameters (Fig 5B). ~2.3 kb, ~3.5 kb, and ~4.2 kb bands can be seen in lanes where the initial *brkA* primers were used. This may suggest the presence of *brkA* in the genome, though it cannot be confirmed as the most prominent bands (~2.3 kb) were not the expected size of 4.2 kb (Fig 5B).

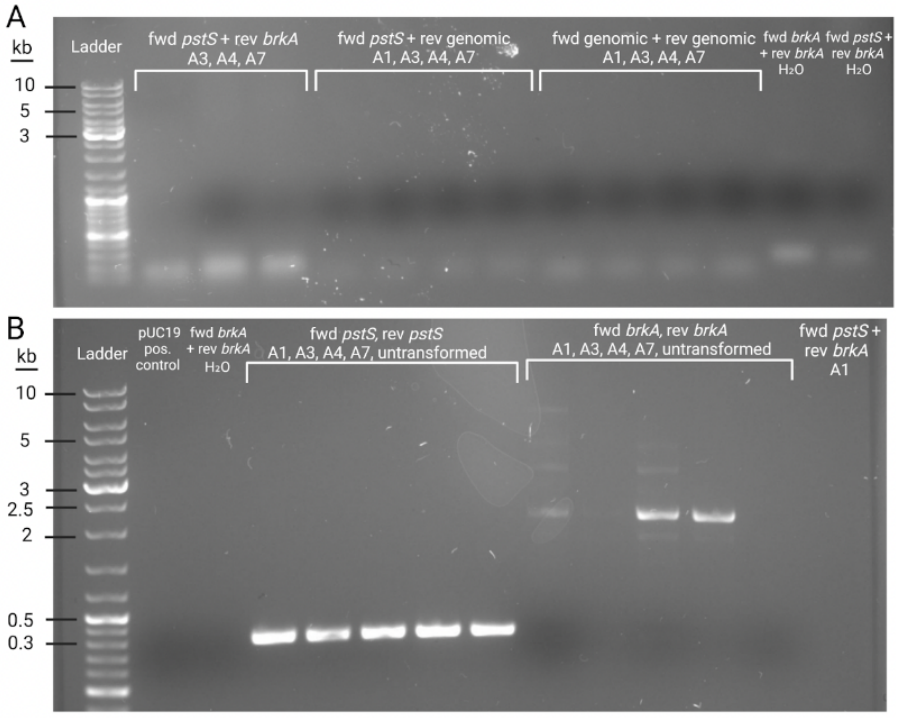

FIG. 5 Genomic insertion of *lac*/*brkA*-6xHis could not be confirmed through PCR. PCR was performed using various primer sets, and the products were run on a 1% agarose gel. A1, A3, A4, and A7 are colonies transformed with the Gibson product and the pTNS2 helper plasmid. No bands were present in the lanes testing for the genomic insertion of *lac*/*brkA*-6xHis in the genomic DNA of the A1, A3, A4, or A7 colonies (A). As controls, the *pstS* gene in the DH10B genome was amplified (B), PCR reactions with primers and the plasmid DNA of untransformed DH10B were performed (B), and a PCR reaction with pUC19 and its associated primers was performed (B). As negative controls, H_2O and untransformed DH10B plasmid DNA PCR reactions were done as well (A, B). Initial *lac*/*brkA*-6xHis primers were also tested with transformed colonies, resulting in band sizes ~2.3 kb, ~3.5 kb, and ~4.2 kb (B).

DISCUSSION

Our study aimed to insert the *lac/brkA*-6xHis region into the genome of DH10B *E. coli* using a mini-Tn7 system. We used plasmid sequencing to investigate our Gibson assembly products and aimed to confirm genomic insertion by amplifying the *lac*/*brkA*-6xHis region from the DH10B genome using genomic PCR.

Sequencing revealed multiple insertions of *lac*/*brkA*-6xHis and backbone in the Gibson product, suggesting recombination may have occurred within the transformants. Sequencing results showed *lac*/*brkA*-6xHis and the mini-Tn7 vector backbone were present more than once in both the B1 and B2 transformants. A restriction digest confirmed that this result was not due to an issue with sequencing (Fig. 4). It appears that both vectors pDFAN-1 and pDFAN-2.1 contain 2 inserts and 2 backbones, also evidenced by their total size being ~17 kb (Fig 4A, 4B). pDFAN-2.2 also contains 2 inserts but seems to contain more than 2 backbones (Fig 4C). We attempted to identify a potential unexpected promoter sequence in our intended mini-Tn7 construct that may have caused over-expression of BrkA, as this may have caused host cell death, leading to only these abnormally large constructs being isolated (6, 7). We theorized that our intended construct may have unintentionally placed a promoter sequence close to *brkA* leading to its overexpression, and that the unexpectedly large

constructs we isolated were obtained because they separated *brkA* from this unintended promoter. However, we could not identify any such promoters. Additionally, the mini-Tn7 vector contains two strong transcriptional terminators to prevent unwanted read-through from genomic promoters upon transposition, and these terminators, which are in the correct orientation, should also prevent any unwanted read-through from any elements on the vector itself (9). As we were able to grow other *E. coli* cells containing the pENS plasmid, which should themselves express similar levels of BrkA as cells containing our Gibson product, we believe it is unlikely that host cell death from overexpression of BrkA caused our Gibson product sequencing results.

Because a high-fidelity polymerase was used during the amplification of the insert, it is also unlikely that polymerase errors caused Gibson assembly failure. Agarose gel electrophoresis confirmed that both the insert and the backbone were the correct size and the sequences in our vectors aligned with the sequences of these segments (Fig 2, 3), so it is unlikely that Gibson assembly failure was due to improper DNA sequences. Because there were multiple variations of the Gibson assembly product in different transformed colonies (Fig 4), we believe the altered plasmids may have resulted from recombination within the EC100D *pir*-116 *E. coli* cells after transformation. If all sequenced plasmids from transformants were identical, then it is possible there was an error in the Gibson assembly. However, this was not the case. Based on these results combined with external research, we believe recombination of the plasmid occurred within the EC100D *pir*-116 *E. coli* cells (13). While we are unsure of the cause of this recombination, it may be exacerbated by the presence of many regions of high GC content (>70%) within both *lac*/*brkA*-6xHis and transposon backbone (14). One other potential reason for our results could be catenation or knotting of the plasmids together, causing multiple copies of the vector and insert to be detected in sequencing, though the efficiency of such DNA topologies arising is relatively unclear (15).

Unexpected bands in negative control raise concern regarding *lac*/*brkA*-6xHis amplification. In our first PCR amplification of *lac*/*brkA*-6xHis, bands are seen in the negative control (Fig 2). However, because the 4.2 kb bands in *lac*/*brkA*-6xHis amplification lanes were much brighter than the bands in the negative control, we assumed the bands observed in the negative control lanes were a result of contamination from either our tubes or from the communal aliquot of nuclease-free water. We observed the expected product in our positive control lanes, which also gave us reason to believe that *lac*/*brkA*-6xHis was truly amplified. It is, however, notable that the laboratory we were operating in was using numerous different DNA constructs containing different versions of the *brkA* gene, any of which may have been a contaminant in our PCR master mix. Therefore, we are unable to fully verify our amplification of *lac*/*brkA*-6xHis.

Genomic insertion of *brkA* at the *att*Tn7 site could not be confirmed. After transposition of the Gibson product, we expected a mini-Tn7-mediated insertion of the product at the *att*Tn7 site of the DH10B *E. coli* genome, which would result in genomic insertion of *lac*/*brkA*-6xHis. Genomic PCR with one primer on the genome (fwd genomic) and the other primer on *brkA* (rev genomic) did not yield bands, indicating that *lac*/*brkA*-6xHis was not present at this expected site. Control PCR reactions amplifying the *pstS* gene confirmed our genomic PCR conditions and the quality and quantity of genomic DNA used, indicating that the results were due to the genomic DNA sequence differing from the expected construct. Interestingly, bands were observed when using the primers originally used to amplify *lac*/*brkA*-6xHis (fwd *brkA* + rev *brkA*), suggesting that *lac*/*brkA*-6xHis was potentially present somewhere in the DH10B genome (Fig. 5B), though not at the intended *att*Tn7 site. However, the brightest bands from using the fwd *brkA* + rev *brkA* primers were ~2.3 kb in size and were thus not the size we were expecting (4.2 kb) based on *in silico* PCR reactions and our initial PCR using those primers. However, we observed very faint bands of expected size (4.2 kb) as well, though they are much less prominent than the ~2.3 kb bands. As such, the expected *lac*/*brkA*-6xHis may have been amplified, but we could not confirm it with our PCR experiment. Therefore, using SnapGene, we tested these same primers to see if they would bind to the DH10B genome anywhere else, and we searched for any potential sites with similar sequence identity to the known *att*Tn7 site at which there may have been transposition. No likely binding sites for the primers were detected, and we could not identify any potential secondary *att*Tn7 sites. This suggests that non-specific amplification was not a

likely reason for these bands. It is also unlikely that these bands resulted from contamination of the PCR master mix, as no bands were observed from the no template control reaction with the *brkA* primers (Fig. 5B). We hypothesize that the issue with the genomic insertion could be due to the unexpected Gibson product results that yielded products with multiple copies of *lac*/*brkA*-6xHis, though we are unsure how these unexpected Gibson constructs may have affected transposition. With sizes of approximately 17 kb and 25 kb, these constructs were much larger than those used in literature, so it is possible that the large size of these products interfered with genomic insertion of *lac*/*brkA*-6xHis. However, the mechanism involved is unknown. Another potential reason for the unexpected bands is that there could have been contamination in our genomic DNA samples. Thus, we are unable to confirm our genomic insertion, and further experiments should be performed to yield more conclusive results.

Limitations Our analysis of the PCR amplicons yielded after the selective amplification of *lac*/*brkA*-6xHis is limited by issues with the negative control. As previously mentioned, bands appeared in the negative control lane, likely a result of contamination during pipetting steps or the incorporation of contaminated water into the PCR master mix. Whilst we proceeded ahead with our experimental protocols, subsequent to consultation with our mentors, it would have been ideal to have proceeded forward with our project with an experimental result that was gained in parallel to a properly performing negative control. This is particularly important, irrespective of the source of contamination, given the usage of our *lac*/*brkA*-6xHis amplicon in downstream assays - more specifically, its insertion into the transposon vector.

The results of our subsequent Gibson assembly and transformation were also inconclusive. Firstly, we did not isolate the pDFAN plasmids before transforming the DH10B cells. We followed the manufacturer's protocol and directly transformed the Gibson products into the DH10B cells right after Gibson assembly. Additionally, while we observed transformed DH10B *E. coli* cells, as evidenced by the growth of Gibson product-containing *E. coli* cells growing on our antibiotic-selective plates, inconsistent growth patterns were observed. These transformants did not survive for a prolonged period, and gentamicin concentration had to be significantly reduced. Unexpectedly, we also observed growth on the part of untransformed cells. These cells would not be expected to produce colonies exhibiting gentamicin resistance. Thus, though we did get transformed colonies on the gentamicin plates, from which we chose specific colonies to be streaked and used in downstream experiments, for experimental robustness, it is not ideal that untransformed cells exhibited growth and resistance properties of their transformed counterparts. Moreover, at a molecular level, sequencing of our Gibson products pointed towards the presence of multiple copies of *lac*/*brkA*-6xHis, amongst other genes, into our designed plasmid. Given our experimental protocols were meant to produce a single insertion event into our linearized mini-Tn7 vector, the possibility of two of such events occurring points towards a shortcoming in our protocol during or before the Gibson assembly step.

With respect to the colonies selected from our likely transformed *E. coli* cells, a limitation to note is our lack of confirmation of transposition location. Insertion of *lac*/*brkA*-6xHis alongside its upstream promoter region via the mini-Tn7 system was meant to be precise, occurring as a single event at a precise location. Unfortunately, our attempts at a genomic PCR did not yield meaningful results, casting uncertainty on the nature of *lac*/*brkA*-6xHis's insertion into the transformed cells' genomes. Moreover, it is unlikely that our Gibson product has lingered in our transformed *E. coli* cells without transposition (thereby conferring gentamicin resistance). This is because our Gibson product is incapable of replicating in instances when it fails to insert into the *E. coli* genome, because the cells require the *pir* gene for plasmid replication (10).

Conclusions In this study, we inserted *lac/brkA*-6xHis into the mini-Tn7 backbone pUC18R6K-mini-Tn7T-Gm using Gibson assembly and transformed this construct into DH10B *E. coli* cells for transposition and EC100D *pir*-116 *E. coli* cells to isolate our Gibson product. We demonstrated amplification of our desired *lac/brkA*-6xHis segment and linearization of our mini-Tn7 vector. Sequencing results of our Gibson assembly product, containing *lac/brkA*-6xHis inserted into our transposon vector, showed two or three copies of the insert and backbone, suggesting the single insertion event did not occur as intended.

DH10B *E. coli* cells co-transformed with the Gibson product and helper plasmid demonstrated growth on gentamicin-containing plates, potentially indicating transposition, though the reason for this growth cannot be confirmed because genomic PCR meant to verify the location of *brkA* into the transformant's genome was unsuccessful, thus warranting further investigation.

Future Directions Several further experiments should be performed to confirm the results of our Gibson assembly and genomic transposition. *lac*/*brkA*-6xHis amplification and its subsequent incorporation into a mini-Tn7 vector by Gibson assembly should be performed again to verify the products and transformants being analyzed for better understanding of the suitability of the mini-Tn7 system for our application. Parallel to genomic PCR, whole genomic sequencing of Gibson transformants should be carried out to verify and confirm the location of transposition, given that a site-specific single insertion is of interest and expected.

Furthermore, it is important to confirm BrkA's surface expression experimentally, which may explain the use of the mini-Tn7 system to facilitate expression of antigens of interest. A western blot on cell lysates should be performed. A 73 kDa protein band would be indicative of BrkA expression. Finally, with the usefulness of the mini-Tn7 system and its facilitation of single transposition events possibly leading to decreased metabolic burden relative to plasmid-based systems (as a result of harboring a single copy of a gene, in this case *brkA*), the parallel comparison of cell growth and viability in putative transformants compared to a plasmid-based system would be useful in directly demonstrating the appropriateness of the mini-Tn7 system in this application.

ACKNOWLEDGEMENTS

We would like to thank Dr. David Oliver and Virginie Jean-Baptiste for their support and guidance in this study. We thank Jade Muileboom for her guidance and technical support. We would like to thank Valentina Santamaría Amorocho and the Murphy lab at UBC for providing us with several *E. coli* strains, the transposon vectors, and the helper plasmid. We thank the Department of Microbiology & Immunology at UBC for providing us with the funding, laboratory space, and resources to perform this work.

CONTRIBUTIONS

All authors contributed to the completion of the experimental work. D.L. contributed to the introduction, methods, and discussion. F.L. contributed to the methods, results, discussion, and figures. A.T. contributed to the methods, discussion, and figures. N.V. contributed to the abstract, introduction, and discussion. All co-authors contributed to the editing and revision of the entire manuscript.

REFERENCES

1. **Green ER, Mecsas J.** 2016. Bacterial Secretion Systems: An Overview. Microbiology Spectrum **4**.
2. **Oliver DC, Huang G, Fernandez RC**. 2003. Identification of Secretion Determinants of the *Bordetella pertussis* BrkA Autotransporter. J Bacteriol **185**:489–495.
3. **Oliver DC, Huang G, Nodel E, Pleasance S, Fernandez RC.** 2003. A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. Molecular Microbiology **47**:1367–1383.
4. **Sun F, Pang X, Xie T, Zhai Y, Wang G, Sun F.** 2015. BrkAutoDisplay: functional display of multiple exogenous proteins on the surface of *Escherichia coli* by using BrkA autotransporter. Microbial Cell Factories **14**:129.
5. **Nicolay T, Vanderleyden J, Spaepen S.** 2015. Autotransporter-based cell surface display in Gram-negative bacteria. Critical Reviews in Microbiology **41**:109–123.
6. **Rosano GL, Ceccarelli EA.** 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol **5**.
7. **Schuster LA, Reisch CR**. Plasmids for Controlled and Tunable High-Level Expression in *E. coli.* Appl Environ Microbiol **88**.
8. **Jittawuttipoka T, Buranajitpakorn S, Fuangthong M, Schweizer HP, Vattanaviboon P, Mongkolsuk S.** 2009. Mini-Tn7 vectors as genetic tools for gene cloning at a single copy number in an industrially important and phytopathogenic bacteria, *Xanthomonas spp*. FEMS Microbiol Lett **298**:111–117.

9. **Choi K-H, Schweizer HP.** 2006. mini-Tn7 insertion in bacteria with single attTn7 sites: example *Pseudomonas aeruginosa*. Nat Protoc **1**:153–161.
10. **Choi K-H, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP.** 2005. A Tn7-based broad-range bacterial cloning and expression system. Nat Methods **2**:443–448.
11. **Romero-Jiménez L, Rodríguez-Carbonell D, Gallegos MT, Sanjuán J, Pérez-Mendoza D.** 2015. Mini-Tn7 vectors for stable expression of diguanylate cyclase PleD* in Gram-negative bacteria. BMC Microbiology **15**:190.
12. **Lambertsen L, Sternberg C, Molin S.** 2004. Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins. Environmental Microbiology **6**:726–732.
13. **Conley EC, Saunders VA, Saunders JR.** 1986. Deletion and rearrangement of plasmid DNA during transformation of *Escherichia coli* with linear plasmid molecules. Nucleic Acids Res **14**:8905–8917.
14. **Marsolier-Kergoat M-C, Yeramian E.** 2009. GC Content and Recombination: Reassessing the Causal Effects for the *Saccharomyces cerevisiae* Genome. Genetics **183**:31–38**.**
15. **Bates AD, Maxwell A.** 1997. DNA topology: Topoisomerases keep it simple. Current Biology **7**:778–781.
16. **Koch D, Chan ACK, Murphy MEP, Lilie H, Grass G, Nies DH**. 2011. Characterization of a Dipartite Iron Uptake System from Uropathogenic *Escherichia coli* Strain F11. Journal of Biological Chemistry **286**:25317–25330.