



# Efficient transposition of Tn4556 by alterations in inverted repeats using a delivery vector carrying a counter-selectable marker for *Streptomyces*

Masahiro Sota<sup>1</sup> · Akiko Sakoda<sup>1</sup> · Haruo Ikeda<sup>2</sup>

Received: 23 September 2018 / Accepted: 26 October 2018 / Published online: 20 November 2018  
© The Author(s) 2018

## Abstract

A 6625-base pair transposon, Tn4556, was initially isolated from a *Streptomyces* strain and a sequence analysis was performed; however, its annotation data remain incomplete. At least three positions were identified as frameshift and base-exchange errors by resequencing. The revised sequence revealed that Tn4556 contains four open reading frames that encode transposase, methyltransferase, isoprenyl diphosphate transferase, and resolvase, respectively. Thirty-eight-base pair inverted repeat (IR) sequences at both ends contained a 1-bp mismatch flanked by a target duplication site, and transposition efficiency was improved by the replacement of imperfectly matched IR-L to perfectly matched IR-L. The detection of Tn4556 transposition was markedly facilitated using a delivery vector carrying a strictly counter-selectable marker for *Streptomyces* strains.

**Keywords** Class-II transposon · Counter selection · *Streptomyces* · Phenylalanyl-t-RNA synthetase

## Introduction

Transposable elements and insertion sequences (ISs) that are capable of moving from a replicon to others are important tools in the study of bacterial genetics and gene expression. The distribution of these mobile elements in the genome is also of interest in the evolution of organisms. Some transposable elements and ISs in *Streptomyces* strains have been isolated [2, 4, 14, 19], and the distribution of ISs in the *Streptomyces* genome was recently reported [10]. Among transposons, the class-II transposable element Tn4556 was identified in neomycin-producing *Streptomyces fradiae* in 1987 [5, 6]. Tn4556 derivatives carrying antibiotic-resistance markers were found to be useful for transposition in

*Streptomyces* strains [16, 22]. The most preferable derivative Tn4560, which contains the resistance marker viomycin phosphotransferase, was used to target genes involved in secondary metabolite biosynthesis [9]. The 6.8-kb sequence of Tn4556 was elucidated in 1990 [18] and at least nine ORFs have been annotated from sequence data; however, some overlapped in both strands. Thus, since the original sequence may contain errors, we resequenced Tn4556 to clarify actual ORFs and examined the efficient transposition of Tn4556.

## Results and discussion

### Resequencing of Tn4556

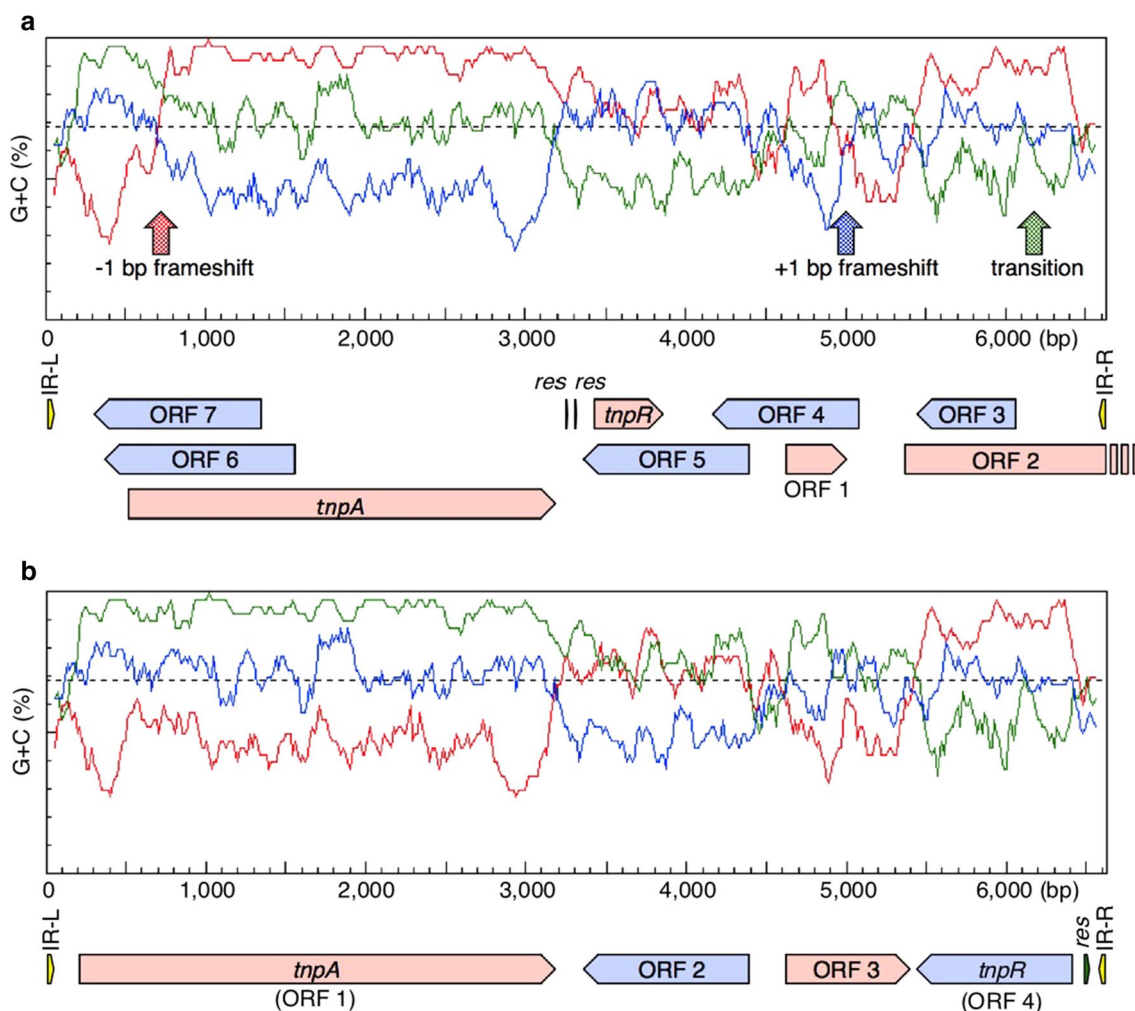
A 6.64-kbp *Bam*HI fragment of Tn4556 in pUC1232 [5], which contains the entire Tn4556 sequence, was subcloned into pUC19 and the sequence was analyzed using a next-generation sequencer (accession # LC417441). Three positions differed from the previous sequence (accession # M29297 [15]). The first position was missing one base (a –1-bp frameshift) between 719 and 720 nt from the left end of the inverted repeat (IR-L) of the original sequence (Fig. 1a). The original ORF1 (892 aa) was annotated as a

This article is part of the Special Issue “Natural Product Discovery and Development in the Genomic Era 2019”.

✉ Haruo Ikeda  
iked@ls.kitasato-u.ac.jp

<sup>1</sup> NAGASE R&D CENTER, NAGASE & CO., LTD., Kobe,  
Hyogo 651-2241, Japan

<sup>2</sup> Kitasato Institute for Life Sciences, Kitasato University,  
Sagamihara, Kanagawa 252-0373, Japan



**Fig. 1** Organization of ORFs and the frame plot of Tn4556 derived from the 6625-bp **a** original sequence and **b** revised sequence. The direction of transcription and relative sizes of the ORFs deduced from an analysis of the nucleotide sequences are indicated. A frame plot was calculated by a window size of 40 codons and step size of 5 codons. Dashed lines indicate average G+C % (68.4%). IR-L, IR-R, and *res* indicate the left end of the inverted repeat, the right end of

the inverted repeat, and the internal resolution site, respectively. The genes *tnpA* and *tnpR* encode transposase and resolvase, respectively. ORF2 and ORF3 in the revised sequence **b** encode predicted methyltransferase and isoprenyl diphosphate transferase, respectively. Vertical arrows in **a** indicate different points between the original and revised sequences of Tn4556

transposase (TnpA), but without the N terminus region found in other TnpAs. Based on the revised sequence (Fig. 1b), the start codon of the revised ORF1 was located upstream of the original ORF1 (from 505 to 3183 nt), and the revised ORF1 was located from 200 (TTG start) to 3184 nt (TGA stop), which encodes 994-aa TnpA (a characteristic motif of PF01526: Tn3 transposase DDE domain was observed), and the deduced polypeptide matched other TnpAs in a BLAST analysis. Furthermore, the revised ORF2 was annotated from 4392 (TGA stop) to 3355 nt (ATG start; encoded in the complementary strand), which encodes a methyltransferase domain-containing protein (345 aa; WP\_085572232 of *Streptomyces* sp. 13-12-16 shows 95% identity and 97% similarity, WP\_086708490 of *S. castelarensis* shows 95% identity and 97% similarity, and WP\_116427250 of *S. spongiicola*

shows 95% identity and 97% similarity), in which characteristic motifs of PF13847 (methyltransferase domain: 127–236 aa), PF08241 (methyltransferase domain: 133–228 aa), and PF13649 (methyltransferase domain: 131–224 aa) were observed, and was identical to ORF5 of the original annotation (Fig. 1a, b). The second position was the insertion of one base (a + 1-bp frameshift) at 5001 nt of the original sequence (Fig. 1a) and this insertion terminated translation because the insertion of an adenine residue made the stop codon TGA. The revised ORF3 was located from 4625 (ATG start) to 5401 nt (TGA stop), which encodes putative isoprenyl diphosphate transferase (258 aa; WP\_085572231 of *Streptomyces* sp. 13-12-16 shows 96% identity and 98% similarity, WP\_037940392 of *S. toyocaensis* shows 96% identity and 98% similarity, and WP\_116427251 of *S. spongiicola*

shows 96% identity and 97% similarity), in which a characteristic motif of PF01255 (putative undecaprenyl diphosphate synthase: 31–257 aa) was observed. The final position was at 6194 nt as a guanine residue (Fig. 1a; the original sequence was adenine). This region was defined as a truncated ORF in the original annotation [18], whereas a gene was located from 5445 (TGA stop) to 6419 nt (ATG start; in the complementary strand) in this resequencing. The revised ORF4 (324 aa) was annotated as a resolvase (TnpR). A previous study reported that the downstream ORF (3420–3851 nt of the original sequence) of TnpA was defined as a TnpR (Fig. 1a; ref 16), whereas the ORF did not contain the conserved motif (PF00589; phage integrase family) found in the TnpRs of transposons; however, the revised ORF4 was defined as a TnpR because the motif for PF00589 was found at 587–971 aa by a Pfam search. Thus, the revised sequence revealed that Tn4556 contains four ORFs. Class-II transposons form a cointegrate intermediate with the target replicon upon transposition and the intermediate is resolved by TnpR at the internal resolution site (*res* site; in many cases, a palindrome-like structure is found in the *res* site). In the original sequence, possible *res* sites were defined between 3243 and 3251 nt, and between 3307 and 3299 nt; however, these two sequences were not found in the palindrome-like structure. On the other hand, since the region between 6497

and 6527 nt in the revised sequence formed a palindrome structure containing an 11-bp stem with a 1-bp mismatch and 9-base loop structure, this region flanking the right end of the inverted repeat (IR-R) may function as a *res* site for resolving the cointegrate intermediate during the transposition process.

### Construction of a new delivery vector carrying a counter-selectable marker for transposition

The transposition of Tn4556 and its derivatives was efficiently performed using a delivery vector replicated in *Streptomyces* strains rather than non-replicative vectors, such as the *Escherichia coli* plasmid pUC19. After transposition, the delivery vector containing the Tn4556 derivative has to be cured from the strain. We used the temperature-sensitive replication vector pKU110, derived from the pIJ101 replicon, for transposition in previous studies [9]; however, the elimination of the vector from the strain was not fully achievable. In the present study, we constructed a new delivery vector carrying a suicide gene for the transposition of Tn4556 derivatives. Although some suicide genes were used for counter selection in *Streptomyces* strains, their use was limited because sensitivity depends on the strain [7, 8]. The *E. coli* phenylalanyl-t-RNA synthetase  $\beta$ -subunit (PheS) was useful as a counter-selectable marker because its A294G variant misincorporates

**Table 1** Susceptibility to 4-chloro-DL-phenylalanine of *Streptomyces* strains and their exoconjugants carrying the mutant type of the *pheS* gene

<i>S. avermitilis</i> MA-4680		4-chloro-DL-phenylalanine ( $\mu\text{g/ml}$ )							
PheS	0	0.156	0.313	0.625	1.25	2.5	5	10	20
Wild type	++	++	++	++	++	++	++	++	++
A339G	++	++	++	++	++	±	–	–	–
A339G/T278S	++	+	±	–	–	–	–	–	–
A339G/T278A	++	+	±	–	–	–	–	–	–
<i>S. lividans</i> TK24		4-chloro-DL-phenylalanine ( $\mu\text{g/ml}$ )							
PheS	0	0.156	0.313	0.625	1.25	2.5	5	10	20
Wild type	++	++	++	++	++	++	++	++	++
A339G	++	++	++	++	++	++	+	–	–
A339G/T278S	++	++	++	+	±	–	–	–	–
A339G/T278A	++	++	++	+	±	–	–	–	–

The PheS state of the wild type indicates strains harboring pGM160 $\Delta$ *aac1::oriT*. A339G, A339G/T278S, and A339G/T278A indicate strains harboring pGM160 $\Delta$ *aac1::oriT::pheS*<sup>A339G</sup>, pGM160 $\Delta$ *aac1::oriT::pheS*<sup>A339G/T278S</sup>, and pGM160 $\Delta$ *aac1::oriT::pheS*<sup>A339G/T278A</sup>, respectively. These plasmids were introduced from *E. coli* GM2929 (*dam dem*) *hsdS::Tn10/pUB307* [12] by conjugation. Exoconjugants were selected from a mixed culture by insensitive to carmonam (100  $\mu\text{g/ml}$ ); counterselect the sensitive *E. coli* donor) and resistance to thiostrepton (20  $\mu\text{g/ml}$ ). Susceptibility to 4-chloro-DL-phenylalanine was examined on YS medium (4 g of yeast extract, 10 g of soluble starch, and 20 g of agar per liter of deionized water and pH was adjusted to 7.0) containing 15  $\mu\text{g/ml}$  of thiostrepton and the various concentrations of 4-chloro-DL-phenylalanine indicated. Synthetic *pheS* genes encoding PheS<sup>A339G</sup>, PheS<sup>A339G/T278A</sup>, and PheS<sup>A339G/T278S</sup> (accession # LC417438, LC417439, and LC417440, respectively) were prepared by Eurofins Genomics Co., Ltd., Tokyo, Japan. Each synthetic DNA was digested with *MluI/XbaI* and the 1.46-kb fragment digested was ligated with the largest *MluI/XbaI*-pGM160 $\Delta$ *aac1::oriT* [12] fragment. pGM160 $\Delta$ *aac1::oriT::pheS* was selected

Abbreviations indicate the following: ++ preferable growth, + growth, ± slight growth (a few colonies occurred), – no growth

4-chlorophenylalanine into cellular proteins during translation, thereby causing cell death [17]. This counter-selectable system was also applied to other bacteria [1, 3, 11, 13, 21]. While wild-type *Streptomyces* is insensitive to 4-chloro-DL-phenylalanine (> 20 mM), we demonstrated for the first time that *Streptomyces* strains carrying an extra copy of the gene encoding the PheS variant are sensitive to 4-chloro-DL-phenylalanine. *S. avermitilis* or *S. lividans* harboring a plasmid containing mutant *pheS* (A339G; corresponding to the A294G variant of *E. coli* PheS) genes (designed from the amino acid sequence of *S. viridochromogenes* DSM 40736; accession # WP\_003989071) was sensitive to ~2.5 mM or 10 mM 4-chloro-DL-phenylalanine, respectively, (Table 1). Since the aa position 251 (Thr) of *E. coli* PheS was involved in substrate recognition, this position was our target for the enhancement of sensitivity to 4-chlorophenylalanine [15]. After the introduction of an aa replacement in the second target at Thr278 of *Streptomyces* PheS (corresponding to Thr251 of *E. coli* PheS), sensitivity to 4-chloro-DL-phenylalanine was enhanced by approximately eightfold by the variant at T278S or T278A in PheS<sup>A339G</sup> (Table 1). We then constructed the new delivery vector pGM160Δ*aac1::oriT::pheS*<sup>A339G/T278A</sup>::Tn4556-*aac3(IV)*, and the transposition of Tn4556-*aac3(IV)* was

efficiently performed because it was easy to isolate progeny carrying the transposon without the delivery vector after a high-temperature incubation at 37 °C and the subsequent selection of 4-chlorophenylalanine resistance.

### Improvements in transposition using perfectly matched IR-L

In bacterial transposons, IRs at both ends of class-II transposons were relatively long (38–110 bp). Tn4556 possessed 38-bp IRs with a 1-bp mismatch (Table 2). The exchange to the perfectly matched IRs of the transposon TnHad2 in the IncP-1β plasmid pUO1 from *Delftia acidovorans* strain B improved transposition frequency [20]. After the cytosine residue of the 5'-end of IR-L of Tn4556 was replaced to a guanine residue, the modified IR-L perfectly matched IR-R (Table 2). Tn4556-*aac3(IV)* carrying perfectly matched IR-L was joined to pGM160Δ*aac1::oriT::pheS*<sup>A339G/T278A</sup> and transposition efficiency was examined. As shown in Table 2, transposition efficiency was approximately five- to tenfold better than that of wild-type Tn4556-*aac3(IV)* and the transposition occurred randomly on *S. avermitilis* chromosome (Fig. 2).

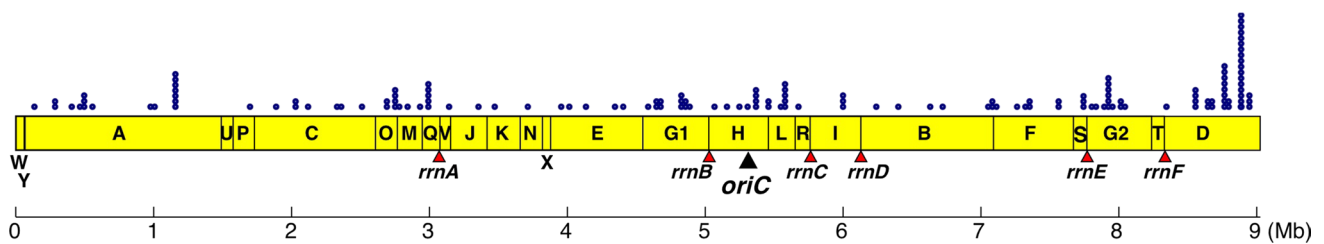
**Table 2** Transposition of wild-type Tn4556-*aac3(IV)* and its derivative consisting of perfectly matched IR-L in *S. avermitilis*

IR-L/IR-R	Transposition efficiency <sup>a</sup>
IR-L cGGGGTTGAGGAACATCCGAACGAAAACCGGCCTAAG 	5.4 × 10 <sup>-4</sup>
IR-R gGGGGTTGAGGAACATCCGAACGAAAACCGGCCTAAG	3.5 × 10 <sup>-4</sup>
	4.2 × 10 <sup>-4</sup>
IR-L' GGGGGTTGAGGAACATCCGAACGAAAACCGGCCTAAG 	1.8 × 10 <sup>-3</sup>
IR-R GGGGGTTGAGGAACATCCGAACGAAAACCGGCCTAAG	3.1 × 10 <sup>-3</sup>
	1.9 × 10 <sup>-3</sup>

pGM160Δ*aac1::oriT::pheS*<sup>A339G/T278A</sup>::Tn4556-*aac3(IV)* and its derivative containing perfectly matched IR-L in *E. coli* GM2929 *hsdS::Tn101* pUB307 were conjugated with spores of *S. avermitilis* SUKA24 [9] on M4 medium [12] supplemented with 20 mM MgCl<sub>2</sub>. After incubation at 30 °C for 18 h, 1 ml of sterile water containing 2 mg/ml of carmonam and 400 μg/ml of thiostrepton was overlaid onto one plate and each plate was incubated at 30 °C for 5 days. Exoconjugants were replicated onto YMS medium [12] containing 15 μg/ml of thiostrepton and 12.5 μg/ml of nalidixic acid. After sporulation, spores on the surface of the agar plate were harvested, and spore numbers were then counted and suspended in 20 v/v % sterile glycerol (defined as thio<sup>f</sup>; stored at -30 °C). Approximately 10<sup>3</sup>–10<sup>4</sup> spores were spread onto one agar plate (antibiotic-free YMS medium) and incubated at 37 °C until sporulation. Each plate was replicated onto YS medium (see Table 1) containing 10 μg/ml of apramycin and 2.5 mM 4-chloro-DL-phenylalanine and the colonies that grew were then counted (defined as thio<sup>s</sup> apm<sup>f</sup>). To confirm curing of the delivery vectors, plates were replicated onto YMS medium containing 15 μg/ml of thiostrepton. After an incubation at 30 °C for 3 days, thiostrepton-resistant colonies were not observed on any plate. Transposition efficiency was calculated by the numbers of thio<sup>s</sup> apm<sup>f</sup>/the numbers of thio<sup>f</sup>. The construction of pGM160Δ*aac1::oriT::pheS*<sup>A339G/T278A</sup>::Tn4556-*aac3(IV)* containing perfectly matched IR-L was as follows: pUC19::Tn4556-*aac3(IV)* (derived from the *Bam*HI fragment from pUC1232; the largest fragment of *Nco*I/*Xho*I-pUC19::Tn4556 was joined with 0.9-kb *Nco*I/*Sal*I-*aac3(IV)* prepared by PCR amplification) was digested with *Eco*RI/*Age*I and the largest 8.6-kb fragment was ligated with a 1.38-kb fragment containing perfectly matched IR-L and the N terminus region of ORF1 that was prepared by PCR amplification with pUC19::Tn4556-*aac3(IV)* as template DNA using the primer pair of forward: 5'-GTgaattcGAGCTCGGTACCCGGGGATCCGGATCCGGG-3' (lowercase characters indicate *Eco*RI site) and reverse: 5'-ACGGTGCCGAaccggtTGATCAGCATGGCCCGCC-3' (lowercase characters indicate the *Age*I site). The largest 7.3-kb *Bam*HI-pUC19::Tn4556-*aac3(IV)* perfectly matched IR-L fragment was joined with the largest 8.4-kb *Bam*HI-digested pGM160Δ*aac1::oriT::pheS*<sup>A339G/T278A</sup> fragment and pGM160Δ*aac1::oriT::pheS*<sup>A339G/T278A</sup>::Tn4556-*aac3(IV)* perfectly matched IR-L was then obtained. The sequence of Tn4556-*aac3(IV)* perfectly matched IR-L has been deposited (accession # LC417442)

<sup>a</sup>Transposition efficiency was calculated from three independent experiments





**Fig. 2** *AseI*-map of *S. avermitilis* chromosome and transposition loci of Tn4556-*aac(3)IV*. The dot indicates transposition locus. The 5-bp target duplication, e.g., GGGTT, TAGAG, TGCTC, GACTG, ACCAT, GGATC, GGAGC, ATGAC, AGGTA and so on, was also

confirmed at the insertion site by the sequence. Abbreviations *oriC* and *rrnA–F* indicate the replication origin and ribosomal RNA (16S–23S–5S rRNA) operons

## Conclusion

We corrected the Tn4556 sequence by resequencing, and the most important ORFs for the class-II transposon, TnpA and TnpR, were accurately annotated. The new delivery vector was useful for the isolation of progeny carrying transpositions because a suicide gene encoding PheS<sup>A339G/T278A</sup> as a counter-selectable marker eliminated the delivery vector by the selection of 4-chlorophenylalanine resistance after transposition. The replacement of the perfectly matched 38-bp IR-L variant enhanced transposition. Tn4556-*aac(3)IV* containing perfectly matched IR-L may be useful for transposon mutagenesis in *Streptomyces* strains.

**Acknowledgements** The original transposon Tn4556 was kindly provided by The UpJohn company (Kalamazoo, MI, USA) in 1990 (to H.I.). This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas FY2016–2020 from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to H.I.) and a grant entitled “Project focused on developing key technologies for discovering and manufacturing drugs for next-generation treatment and diagnosis” from the Japan Agency for Medical Research and Development (to H.I.).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Argov T, Rabinovich L, Sigal N, Herskovits AA (2017) An effective counterselection system for *Listeria monocytogenes* and its use to characterize the monocin genomic region of strain 10403S. *Appl Environ Microbiol* 83:e02927-16
- Bruton CJ, Chater KF (1987) Nucleotide sequence of IS110, an insertion sequence of *Streptomyces coelicolor* A3(2). *Nucleic Acid Res* 15:7053–7065
- Carr JF, Danziger ME, Huang AL, Dahlberg AE, Gregory ST (2015) Engineering the genome of *Thermus thermophilus* using a counterselectable marker. *J Bacteriol* 197:1135–1144
- Chen CW, Yu TW, Chung HM, Chou CF (1992) Discovery and characterization of a new transposable element, Tn4811, in *Streptomyces lividans* 66. *J Bacteriol* 174:7762–7769
- Chung ST (1987) Tn4556, a 6.80-kilobase-pair transposable element of *Streptomyces fradiae*. *J Bacteriol* 169:4436–4441
- Chung ST (1988) Transposition of Tn4556 in *Streptomyces*. *J Ind Microbiol* 29:81–88
- Dubeau MP, Ghinet MG, Jacques PE, Clermont N, Beaulieu C, Brzezinski R (2009) Cytosine deaminase as a negative selection marker for gene disruption and replacement in the genus *Streptomyces* and other Actinobacteria. *Appl Environ Microbiol* 75:1211–1214
- Hosted TJ, Baltz RH (1997) Use of *rpsL* for dominance selection and gene replacement in *Streptomyces roseosporus*. *J Bacteriol* 179:180–186
- Ikeda H, Takada Y, Pang CH, Tanaka H, Omura S (1993) Transposon mutagenesis by Tn4560 and applications with avermectin-producing *Streptomyces avermitilis*. *J Bacteriol* 175:2077–2082
- Ikeda H, Shin-ya K, Omura S (2014) Genome mining of the *Streptomyces avermitilis* genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *J Ind Microbiol Biotechnol* 41:233–250
- Kast P (1994) pKSS—a second-generation general purpose cloning vector for efficient positive selection of recombinant clones. *Gene* 138:109–114
- Komatsu M, Uchiyama T, Omura S, Cane DE, Ikeda H (2011) Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci USA* 107:2646–2651
- Kristich CJ, Chandler JR, Dunny GM (2007) Development of a host genotype-independent counterselectable marker and a high-frequency conjugative delivery system and their use in genetic analysis of *Enterococcus faecalis*. *Plasmid* 57:131–144
- Lydiate FJ, Ikeda H, Hopwood DA (1986) A 2.6 kb DNA sequence of *Streptomyces coelicolor* A3(2) which functions as transposable element. *Mol Gen Genet* 203:79–88
- Mermershtain I, Finarov I, Klipcan L, Kessler N, Rozenberg H, Safro MG (2011) Idiosyncrasy and identity in the prokaryotic phenylalanine system: crystal structure of *E. coli* phenylalanyl-tRNA synthetase complexed with phenylalanine and AMP. *Protein Sci* 20:160–167
- Olson E, Chung ST (1988) Transposon Tn4556 of *Streptomyces fradiae*: nucleotide sequence of the ends and target sites. *J Bacteriol* 170:1955–1967
- Richmond MH (1962) The effect of amino acid analogues on growth and protein synthesis in microorganisms. *Bacteriol Rev* 26:398–420

18. Siemieniak DR, Slightom JL, Chung ST (1990) Nucleotide sequence of *Streptomyces fradiae* transposable element Tn4556: a class-II transposon related to Tn3. *Gene* 86:1–9
19. Solenberg PJ, Baltz RH (1991) Transposition of Tn5096 and other IS493 derivatives in *Streptomyces griseofuscus*. *J Bacteriol* 173:1096–1104
20. Sota M, Kawasaki H, Tsuda M (2003) Structure of haloacetate-catabolic IncP-1 $\beta$  plasmid pUO1 and genetic mobility of its residing haloacetate-catabolic transposon. *J Bacteriol* 185:6741–6745
21. Xie Z, Okinaga T, Qi F, Zhang Z, Merritt J (2011) Cloning-independent and counterselectable markerless mutagenesis system in *Streptococcus mutans*. *Appl Environ Microbiol* 77:8025–8033
22. Yagi Y (1990) Transposition of Tn4560 in *Streptomyces avermitilis*. *J Antibiot* 43:1204–1205