Journal of Energy and Environmental Sustainability, 10 (2020) 59-65

An official publication of the International Society for Energy, Environment and Sustainability (ISEES)

Journal of Energy and Environmental Sustainability

Journal homepage : www.jees.in

Environmental applications of transposable elements: Tn7 and sleeping beauty transposons

SK Varghese[1*] and S Gangamma[2*]

[1]*Panchajanya, Mangalore-575014, India*
[2]*Department of Chemical Engineering, National Institute of Technology Karnataka, Surathkal, Mangalore-575025, India*

ARTICLE INFO

Received : 12 December 2019
Revised : 30 March 2020
Accepted : 04 April 2020

Keywords:
Environmental biotechnology, transposon, sleeping beauty, gene of interest

ABSTRACT

Transposons are DNA elements that can move from one site to another site without sequence homology. The current paper provides a mini-review on Tn7 and Sleeping Beauty (SB) transposons and their applications. The repurposed transposon systems have a wide variety of applications in environmental biotechnology. These systems can accelerate and increase the accuracy of research in environmental biotechnology. Therefore, from the environmental engineering and biotechnology perspective, the current review focused on elaborating working principles and components of Tn7 and SB transposons.Bacterial transposon Tn7 exhibits two elaborate pathways for target integration. One pathway leads to single site integration in a chromosomal target. The second pathway promotes integration into multiple targets in conjugal plasmids. The Tn7 system can be repurposed for the integration of a gene of interest or large DNA cargo into bacterial genomes. The Tn7 system is active in a wide variety of Gram-negative bacteria. Therefore, the Tn7 system offers a great tool for genetic manipulation in bacteria and can be utilized in a variety of applications in the field of environmental engineering and environmental biotechnology. SB transposon is a silent fish transposon. It has been synthetically reactivated and is active in a wide variety of mammalian cells including humans. The SB transposon system is small and contains only a single transposition enzyme. The SB transposon system can be utilized for stable long-term expression of a gene of interest in mammalian cells. The SB system can be used for studying cellular responses to environmental pollutants.

1 Introduction

Mobile genetic element is a term used to refer movement of DNA/RNA elements of considerable size. They jump from a current location in the genome to another location (target location) in the genome or plasmid. Based on the target sequence requirement for jumping, mobile genetic elements can be classified into two types (Watson, 2004). The first type requires the target location should have a specific sequence and the second one is independent of target sequence. This second type is called transpositional recombination (transposons).Transposons are mobile-genetic elements that move DNA without the requirement of such target specificity. In other words, transposons move from one site to other without any or little sequence specificity of the target location. However, some transposons recognize target sequence of other features at the target locations. The Tn7 transposon is very unusual in this regard, and it moves from plasmids to a single site in the chromosomal site in *E.coli* (Lichtenstein and Brenner 1981 and 1982; Ely, 1982).

Due to evolution and host diversification, there is a bewildering variety of transposons (Bourque, et al., 2018). Based on the mechanism used for movement (transposition), transposons can be broadly classified into two categories. One is DNA only transposons and the other is retro-transposons (Bourque et al., 2018, Huang et al., 2012, Feschotte and Pritham, 2007, Wicker et al., 2007). The retro-transposons move by "copy and paste" mechanism. Most of the retro-transposons are viral like elements. For example, HIV-1 retrovirus can integrate into a host chromosome. Such integrated virus into host cell chromosome is called a provirus. A retro-transposon can be visualized as a provirus. However, keep in mind that there are non-viral like retro-transposons. For a more detailed classification of transposons, readers could refer to Bourque, et al., 2018. During movement, these transposons have an RNA intermediate generated by transcription from a promoter located in one of the terminal repeats. This RNA intermediate is converted to cDNA by reverse transcriptase protein. The newly generated cDNA can integrate into a new target location. The DNA only transposons move by the "cut and paste" mechanism (Figure 1). The DNA transposons enzymes cut (excise) transposon at both left and right ends. This excised DNA element can be inserted into another location of the genome with the help of transposon enzymes. All the transposons reviewed in this paper i.e., Tn7 and sleeping beauty (SB) are DNA only transposons. They all move by cut and paste mechanism. More detailed reviews on transposons can be found elsewhere (Craig, 2002; Bourque et al., 2018).

The general structure of a transposon is given in Figure 2A. The major feature of a transposon is the special sequences found at the left and right ends. They are important for defining the physical boundary of a transposon in a genome or plasmid. They are also important for binding and cutting by transposon enzymes. The sequences at the ends have some structure. These structures contain inverted repeats (IR) and direct repeats

* Corresponding Author: *: sgskvnitk@gmail.com, gangamma@nitk.edu.in*

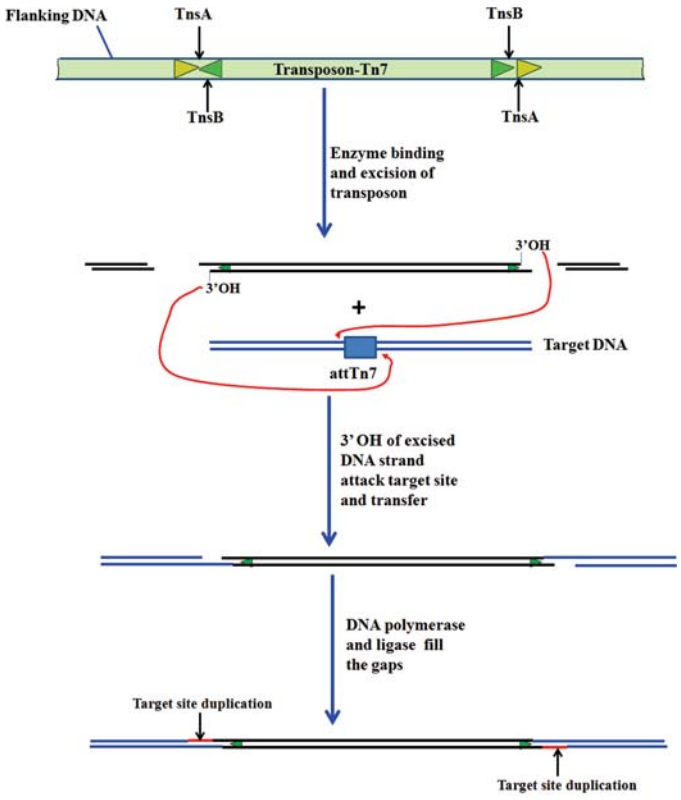

Figure 1. Schematic diagram of the cut and paste mechanism of transposition. Cut and paste mechanism explained by Tn7 transposon. Figure is prepared based on the Peters and Craig, 2001.

(DR). Inverted repeats have similar sequences on the left and the right sides of a transposon. But they are inverted or found in the opposite strands of genomic dsDNA. Inverted repeats have sequences that help transposon enzymes to bind and initiate movement of the transposon. The short direct repeats are found on the left and right sides of the transposon. They are generally originated from a target location during the integration of transposon. However, for some transposons, direct repeat structure could be found embedded in the inverted repeat sequences. The transposons carry few genes within them (Figure 2B). These genes can be classified into three categories, transposase genes, helper genes and host advantage genes. The transposase genes express proteins that help in cutting and joining during the transposition. Helper genes regulate the expression or direct the transposases during the movement or transposition. Host advantage genes give some survival advantage to the host. For example, tn7 transposon carries antibiotic resistance genes. Among the three categories, transposase genes are the most essential and found in all the transposons discussed in the review(Peters and Craig, 2001).

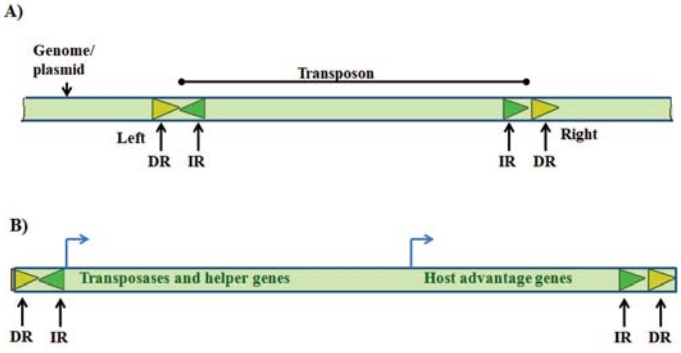

Figure 2. A) General structure of transposon. B) Genes associated with transposon

2. Tn7 transposon

The tn7 transposon was first isolated from a trimethoprim resistant *E. coli* (Barth et al., 1976). However, it was later found that the chromosomal integration of the Tn7 transposon was highly site specific (Lichtenstein and Brenner 1981,1982). Characterization of all transpositions of Tn7 to the *E. coli* genome was found at a single location between PhoS (PstS) and glmS (glutamine synthetase) genes. The unique site in *E.coli* was called attTn7 (Figure 3, McKown, 1988). The Tn7 was also found to integrate with a unique orientation. The orientation of the insertion is also preserved. The left end of Tn7 (Tn7L) is adjacent to phoS, and the right end of Tn7 (Tn7R), is adjacent to glmS (Figure 3). Tn7 was also found to carry resistant genes to trimethoprim, streptomycin and spectinomycin (McKown, 1988). Site specific insertion of Tn7 transposon is observed in many bacterial species (Gay et al., 1986). For example, Tn7 insertions were reported very early in *Caulobacter crescentus* (Ely, 1982), *Vibrio* species (Thompson et al., 1981) and *Pseudomonas aeruginosa* (Fennwald and Shapiro, 1979).

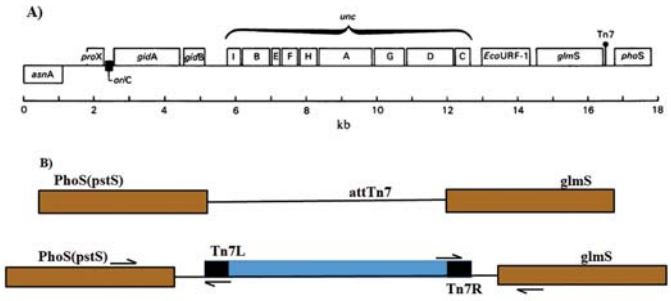

Figure 3. A) Location of Tn7 in *E.coli* chromosome (Walker et al.,1984). B) glmS locus after Tn7 integration.

Apart from drug resistant genes, Tn7 transposon contains TnsABCDE transposon genes. The transposon also contains (cis-acting) specific left flanking sequence called Tn7L and right flanking sequence Tn7R (Figure 4, Peters and Craig 2001). TnsA and TnsB are transposases. These transposases bind to Tn7L and Tn7R sequences and excise the transposon from the current location. TnsD and TnsE are target specificity factors. The tnsC protein is a regulator of transposases. TnsC helps to bridge the interaction between transposases and target specificity factors. Bainton et al., 1993 have shown that TnsC-TnsD likely to form complexes during integration. The TnsC protein also interacts with TnsB protein (Choi et al., 2014), Tn7 proteins and their functions are summarized in Table 1. It is important to note that the Tn7 does not excise from its location unless an attTn7 site is identified in the chromosome (Craig et al., 1997)

Table 1. Tn7 transposon proteins and its functions. (Peters and Craig, 2001)

Protein	Function
TnsA	Transposase subunit. Cutting at the 5' ends of Tn7.
TnsB	Transposase subunit. Cutting at the 3' ends of Tn7.
TnsC	Regulator/connector for TnsA and TnsB proteins.
ATP	dependentDNA-binding
TnsD	Target selection at attTn7 site. Sequence specific DNA binding
TnsE	Target selection on conjugal plasmids or chromosomal sites replication/lagging-strand DNA synthesis. Structure-specific DNA binding.

The left end sequence (Tn7L) is about 150bp and right end sequence (Tn7R) is about 90bp (Peters and Craig, 2001). Both, Tn7L and Tn7R contain several 22 bp transposase binding sites (Figure 4, Tang et al., 1991). There are three non-overlapping transposase binding sequence in Tn7L and there are four overlapping binding sites in Tn7R (Peters, 2014). Tn7R also contain a bacterial constituent promoter region (Gay et al., 1986). This promoter is believed to be helping transcription of TnsAB enzymes.

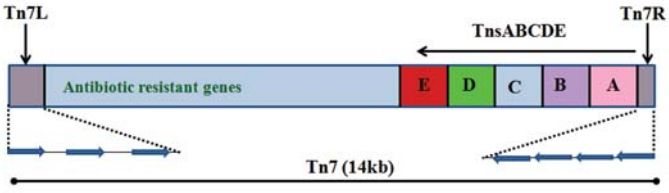

Figure 4. Schematic of Tn7 transposon structure. The transposon is flanked by Tn7L and Tn7R inverted repeats. Tn7L and Tn7R contains 22bp binding site for transposase binding. Tn7L contains three 22bp binding sites (arrows) and TnR contains four binding sites. The transposon encodes five proteins TnsA, TnsB, TnsC, TnsD and TnsE in order. The transposon also carries several antibiotic resistant genes.

Tn7 encodes two different pathways that recognize different types of target site (Figure 5). One pathway is directed by the TnsABC+ TnsD proteins. This pathway directs insertions into a single site, called attTn7, found in *E.coli* genome. Insertion into the attTn7 site does not appear to harm the bacterial host. A second pathway directed by the TnsABC+TnsE proteins. This pathway directs insertions into conjugal plasmids (plasmids capable of mobilizing between bacteria) when they enter the host. Transposition mediated by TnsE pathway was thought to result in random integration onto conjugal plasmids. Peters and Craig, 2001a shows the TnsE based integration is more structure specific and prefers DNA replication sites. Peters and Craig 2000 reports the chromosomal integration of Tn7 mediated pathway near double strand break and where DNA replication ends. Insertion of Tn7 by TnsE pathway into mobile plasmids probably facilitates the horizontal transfer of the element. In contrast TnsD leads to vertical transmission of the element to daughter cells (Peters and Craig, 2001b).

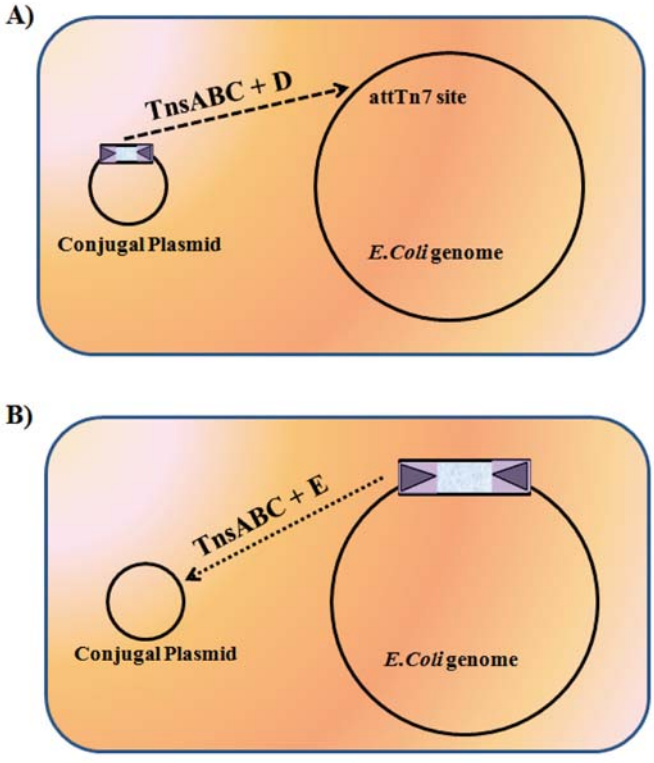

Figure 5. Two pathways of Tn7 transposition. TnsD and TnsE are mutually exclusive target selectors for Tn7 transposition.
A) TnsD pathway targets attTn7 location in the bacterial genome.
B) TnsE targets random integration into conjugal plasmid.
(Based on Peters and Craig, 2001)

The transposases enzymes, TnsA and TnsBbind to the IRs at the both ends of the transposon. They excise Tn7 transposon from the donor backbone by staggered cut at the ends. The 3′ ends of the transposon are cleaved by TnsB. The 5' end of transposon is cleaved by TnsA.Then the TnsB enzyme joins 3′OH ends of the excised transposon to the target DNA. The 3' ends of the transposon is covalently linked to target DNA. But there are short gaps at the 5′ end of the transposon and the target DNA. Host DNA polymerase and ligases repair the gaps to complete Tn7 integration. The "cut and paste" mechanism involved in the process will be a 5-base-pair target-site (CCCGC) duplication at the ends.

McKown et al., 1988 has studied the requirement of additional sequence surrounding the attTn7 site for transposon integration. As previously reported, the study also found that 5bp attTn7 sequence – CCCGC – is duplicated after insertion. McKown et al., 1988 created several plasmids containing left and right flanking attTn7 of various lengths. They evaluated capacity of these plasmids for Tn7 integration. The left and right extend range from -10, 200 to +869 and -3 to + 64. The middle base of the 5bp attTn7 sequence is considered as position 0. The minimal sequence for the orientation specific integration is found to be -3 to 64. The frequency of Tn7 insertion is found between 10^{-3} to 10^{-4}. Studies have reported that TnsD binds mainly between +30 to +50 (Waddell and Craig 1989; Bainton et al., 1993). However, Waddell and Craig 1989 have shown that sequences essential for efficient integration are contained in a bigger range than the observed insertion (attTn7) sequence or TnsD binding site.

The Tn7 transposon exhibits a kind of immunity (Hauer and Shapiro 1984; Arciszewska et al., 1989; Stellwagen and Craig, 1997) against second integration. Already integrated Tn7 transposon can prevent a second integration event. The TnsB and TnsC proteins play an important role in the target immunity (Stellwagen and Craig, 1997). Stellwagen and Craig, 1997 has shown that TnsB might impose target immunity by regulating distribution of TnsC in the target locations. Similar to Tn7 transposon, transposon Tn3 and bacteriophage Mu also exhibit target immunity.Studies have shown that Tn3 and Mu insertions prevent further insertions of the same element (Lee, et al., 1983; Reyes et al., 1987). The presence of Tn3, Tn7 and Mu ends in the target decreases transposition by 100 to 1000 times. Deboy and Craig, 1997 has shown that Tn7 ends prevent integration into genomic sites located 190kb away from the ends. However, integration of Tn7 at large distances (>1.9Mb) away from the Tn7 ends are not affected (Deboy and Craig, 1997).

3. Environmental biotechnology applications of Tn7: Vectors for Tn7 based insertion of DNA sequence into bacterial chromosomes

Transposon Tn7 provides an attractive and convenient tool for insertion of DNA into bacterial chromosomes. It transposes at relatively high frequency in a site-specific and orientation specific manner. In contrast, theTn5-based vectors promote random insertions into the chromosome, and may cause disruption of host genes (deLorenzo et al., 1990). The transposon can integrate into any chromosome or plasmid that has required integration sequence (attTn7, -3 to 64). This requirement is achievable and transposase is active in large number of Gram-negative bacteria. This in fact, helped to create a plasmid system for gene integration into genomes of the Gram-negative bacteria(Barry, 1986; 1988; Grinter, 1983; Bao et al 1991; Shen et al., 1992). Basic design based on the fact that "trans" acting factors (Tn7 enzymes) and "cis" acting factors (Tn7L and Tn7R) can be physically separated into two plasmids.

Moreover, Tn7 transposon is quite large because it requires TnsABCD enzymes for chromosomal integration. Adding to that, the delivery DNA also must be placed between "cis" acting Tn7L and Tn7R sequences. Incorporating all these Tn7 elements into a single plasmid leads to large sized plasmid. This large size can be problematic for cloning. Therefore, several two plasmid systems have been developed for chromosomal cloning. One plasmid contains necessary transposases under a constitutive promoter. This plasmid is called helper plasmid. The second plasmid contains delivery DNA between Tn7L and Tn7R sequences (or transposon or transferable DNA).This plasmid is known as delivery plasmid or transfer plasmid (Figure 6).

However, these vectors were not properly designed to select for transposition in many bacterial species. Bao et al., 1991 solved this issue by using R6K-replicon-based plasmid. Such plasmid can replicate only in bacteria that contain phage lambda-pir gene. This plasmid is therefore maintained in the *Escherichia coli* strains such as SM10 lambda pir; HB101endA: frt lambda pir (Addgene #45472); DH10beta F' DOT sbcC (Addgene #12082). Such conditionally replicating plasmids containing R6K origin of replication are very often used to deliver DNA elements in enteric bacteria. The replication of these plasmids requires

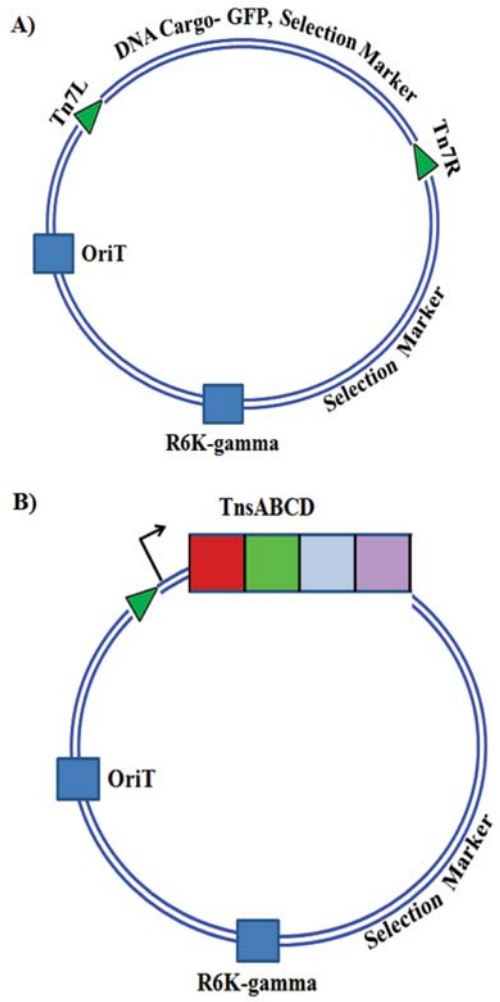

Figure 6. Two plasmid system for Tn7 transposon based genomic engineering in bacteria. A) Tn7 delivery vector B) Tn7 helper vector

the ð protein from pir lambda gene. This gene is usually provided in "trans" in the donor strain. In the absence of pir gene in the recipient strain, these plasmids cannot replicate. Therefore, these plasmids are also called "suicide" vectors. These vectors are rapidly lost from recipient strain but allow the expression or integration of delivery DNA. Currently there are several Tn7 helper and delivery plasmids that uses temperature sensitive origin of replication (Schlechter et al., 2018; Wiles et al., 2018; Jasinska et al., 2020). These plasmids can be cured after Tn7 integration by growing them at higher temperature range.

Most of the Tn7 helper and delivery plasmids contain (origin of transfer) oriT sequence. This allows plasmids to transfer into recipient strain by conjugation. Typical delivery and helper plasmid are given in Figure 6. Conjugation is the most common delivery method for these conditional plasmids. In the early 1980s, several scientists created specialized *Escherichia Coli* donor strains, SM-10 and S17-1. These strains enable mobilization of oriT-containing plasmids to a broad range of recipient strains by conjugation (Simon et al., 1983). These strains are generally called donor strains. Later, pir prophage was incorporated into these strains to enable replication of suicide vectors (Miller and Mekalanos, et al., 1988). A summary of development history of these strains is given in Ferrières et al., 2010. However, recent studies have shown that, bacteriophage Mu in these strains (SM10 and S17-pir) contain an oriT region. This will promote transposition of Mu phage sequence from S17-1 pir and SM10 pir to recipient strains (Ferrières et al., 2010).Apart from conjugation, the plasmids delivery can be achieved through chemical or electroporation-based transformation (Kocch et al., 2001; Choi et al., 2006; Tan et al., 2018). Earlier these methods (chemical transformation and electroporation) were considered as inefficient procedure. However, electroporation method is widely used for inducing Tn7 transposition from conditional plasmids (Kocch et al., 2001; Choi et al., 2006; Tan et al., 2018).

4. RNA guided Tn7 transposition in bacterial genome

The Tn7 transposon integrates into a single chromosomal site called attTn7. Recently, detailed bioinformatics analysis has shown that an association between CRISPR-Cas systems and Tn7-like transposons was identified (Koonin and Makarova, 2017; Peters *et al.*, 2017; Faure *et al.*, 2019). Clustered regularly interspaced short palindromic repeat (CRISPR) and CRISPR-associated (Cas) proteins found in bacteria have been repurposed for genomic engineering in bacteria and eukaryotic cells (Doudna and Charpentier 2014; Knott et al., 2018). The CRISPR-Cas system uses RNA based identification of target DNA or RNA. The system can direct the Cas endonucleases to the sequence specific locations in genomes. The CRISPR-Cas system can be thought of as cargo genes or advantage genes carried by the transposon. Peters et al., 2017 suggested that the CRISPR-Cas system carried by Tn7 transposon might be guiding the transposon for target selection during genome integration. In other words, the Tn7- transposons hijacked the CRISPR system to enable guide RNA (crRNA) directed transposition into the genome. Strecker et al., 2019 tried to repurpose CAST (CRISPR-associated transposase) in bacteria such as *Scynotema hofmanni* (shCAST) and *Aanabena cylindrica* (AcCAST). The researchers have used two plasmid system for this purpose. One is a helper plasmid and the other one is a donor (delivery vector) plasmid. The helper plasmid contains Tn7 proteins TnsB, TnsC, tniQ and Cas12k. The helper plasmid also contains a tracrRNA taken from the shCAST or AcCAST. Both the CAST system was able to integrate transposon DNA cargo into *E.coli* genome unidirectionally. Strecker et al., 2019 have shown that shCast (CRISPR-Cas system in *Cynobacteria Scytonema hofmanni*) can direct RNA guided transposition to 60-66bp downstream of the protospacer. Strecker et al., 2019 found that under over expression conditions that shCast system has off target integration in the *E.coli* genome.

Inspired by Peters et al., 2017, another set of researchers, Klompe et al., 2019 tried to use a CRISPR-Tn7 system in Vibrio Cholera for genetic engineering. The transposon system in *V. cholerae* was called Tn6677 or the *V.cholerae* transposon. The CRISPR-system in the *V.cholerae* Tn7 transposon contains Cas8-Cas5 fusion gene, Cas7, Cas6 and TniQ gene. Analysis has shown that TniQ gene is homologous to TnsD protein. The Tn6677 also contains TnsA, TnsB and TnsC genes. To reuse the *V. cholerae*Tn7 system for genetic engineering, Klompe et al., 2019 used three plasmid system. The first plasmid contains TnsA, TnsB and TnsC genes. The second plasmid contains TniQ, Cas8-Cas5 fusion protein, cas7 and cas6. The second plasmid also contains a CRISPR repeat array that generates guide RNA (or crRNA). The crRNA was designed to target a location in the genome with 5'-CC-3' protospacer adjacent motif (PAM). The third plasmid contains DNA cargo between native terminal repeats of Tn6677 (donor or delivery vector). Klompe et al., 2019 transformed *E.coli* with the three plasmid system. The analysis showed that the DNA cargo in the delivery vector integrated to the desired (guided by CRISPR guide RNA) locations targeted guide RNA (glmS locus, LacZ operon). But the integration was not unidirectional or integration was bidirectional. The integration was 48-50bp downstream of the guide RNA binding sites in the *E.coli* genome. Further experiments have shown that TnsA protein was dispensable for the integration of DNA cargo. Therefore, it is speculated that the transposition was not occurring through the cut and paste mechanism. The DNA cargo (in Donor or delivery plasmid) size was also found to significantly affect the integration efficiency (Klompe et al., 2019). The optimal cargo size was found to be ~0.7kb and the efficacy of integration at a target location in *the E.coli* genome decreased for larger and smaller cargo sizes. Sequencing analysis has shown that three plasmid system based on *V.cholarae* Tn6677 system is very precise in genome integration.

We have seen that the three plasmid system used by Klompe et al., 2019 has low efficiency for larger cargos. Vo et al., 2020 an improved system based on *V.cholerae* Tn6677 Tn7-like transposon system. Vo et al., 2020 converted the three plasmid system into a single plasmid system. All the necessary proteins and crRNA were expressed from this plasmid. The DNA cargo was also placed in the same plasmid. Experiments with *E.coli* have shown that the new system increased integration of the cargo by 2-5 times than the three plasmid system (Vo et al., 2020). The new system also exhibited precise integration near various target locations. Integration efficiency was improved when expression of the proteins was increased.The integration efficiency was found to be higher at 30ÚC compared to 37ÚC. The variation of cargo size from 1-10kb did not affect integration efficiency. The efficiency of the one plasmid system was demonstrated in *Klebsiella oxytoca* and *Pseudomonas putida* bacteria (Vo et al., 2020).

5. Sleeping beauty transposon

Sleeping beauty (SB) is a silent fish transposon and was synthetically reconstructed into an active form (Ivics et al., 1997). Sleeping beauty belongs to Tc1/mariner class of transposon. It contains a single transposase enzyme and the enzyme sequence matches with consensus Tc1/mariner transposon sequences derived from multiple fish species (Ivics et al., 1997). Active transposons in vertebrates are rare or non-existent. Before the sleeping beauty reconstruction in 1997, there was no indication that any DNA-based transposon was active in vertebrates. The transposons in vertebrates were believed to be inactivated by mutations accumulated during evolutionary process (Lohe et al., 1995).

Sleeping beauty belongs to the DNA only transposon class and uses the cut and paste mechanism for transposition. It contains a coding region for a single transposase in the middle of the transposon and is flanked by an inverted terminal repeat (IR) on both sides (Figure 7). The transposase contains an NLS (nuclear localization signal) and a catalytic DDE (Asp, Asp, Glu) domain. The NLS sequence helps to localize the protein to the nucleus. The DDE contains a region that has an RNaseH like domain and DDE provides the essential catalytic active site. The inverted repeats flanking the transposon are 200-250bp in length. Each IRs contain two direct repeats (DRs). These direct repeats within the IRs are transposase binding sites with lengths about 15-20bp (Figure 7). The transposase binds to the terminal repeats (IRs) and catalyzes the excision of the transposon. The transposase also catalyzes the integration of the excised transposon into target locations. The excision and integration of SB transposon take place in four steps (Figure 8, Ivics and Izsvak, 2015). In step one, the transposase dimer binds to the IRs at each end. In the second step, the transposase bends the transposon and brings together the ends of the transposon through interactions of transposase dimers at the ends (Figure 9). This step is known as the tetramerization of transposase and synaptic complex formation step. In the third step, the transposon is excised from the donor DNA. In the fourth step, the excised transposon is integrated with the target site.

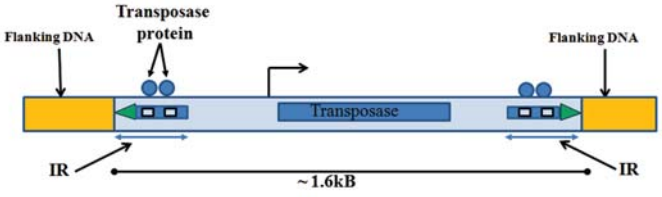

Figure 7. Structure of sleeping beauty transposon. Transposon is flanked by two terminal inverted repeats (IR). Both IRs contain binding sites for SB transposase enzyme (white box, DR). The transposon is about 1.6kB and IRs are about 200-250bp in size. The SB transposon contains single transposase (blue box in middle). Transposase is translated from IRs and 5'UTR region (Ivics, and Izsvák, 2015)

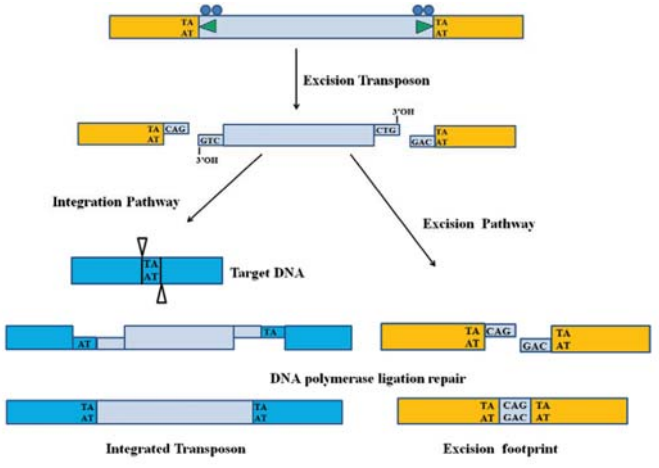

Figure 8. Cut and paste mechanism and excision repair in sleeping beauty transposition. The transposase binds to the IRs at both ends. Events at the integration (target) site is pictured on left and events after transposon excision is pictured on right side. The transposon prefers 5'-TA-3' site for integration. The excision is repaired by NHEJ pathway and may result in significant mutation (Based on Ivics and Izsvak, 2015)

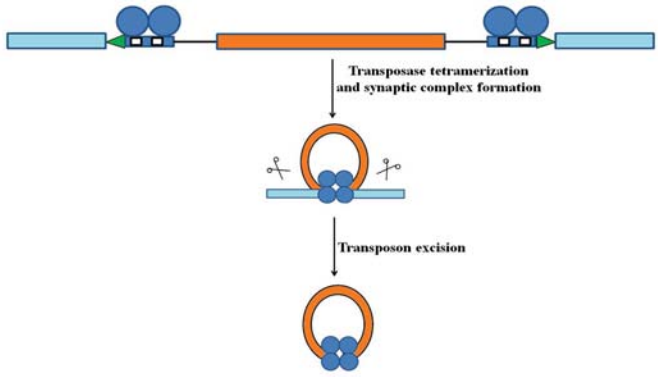

Figure 9. Synaptic complex formation and excision of sleeping beauty transposon. The SB transposase binds to the terminal inverted repeats. The transposase forms a synaptic complex and excise the transposon from the donor site (Ivics, and Izsvák, 2015).

The left IR and the untranslated regions surrounding the transposase (5'UTR) regulate the transcription of the protein and transposition (Moldt, et al., 2007; Ivics and Izsvak, 2015). The reporter assays have shown that the 5'UTR has significant transcription activity (Walisko et al., 2008). The right IR can also induce transcription activity but comparatively less than the left IR and 5'UTR (Ivics and Izsvak, 2015). Experiments have shown that host cell protein HMG2L1 upregulate transcription activity of 5'UTR up to 15 times (Walisko et al, 2008). However, the SB transposase negatively regulates its transcription. Unlike Tn7, the SB transposon integrates randomly in the target chromosome. However, the SB integration prefers 5'-TA-3' sites in the target genome (Figure 8, Ivics et al., 1997; Vigdal et al., 2002). Vigdal et al., 2002 reported that the target site choice of SB is determined by the DNA structure rather than DNA sequence. The TA rich region probably provides flexible DNA structures (Narayanavari et al., 2017). The transposase protein appears in tetrameric form and induces double strand break in opposite strands (Plasterk, 1996; Plasterk et al., 1999).

6. Biotechnology applications: Sleeping beauty transposon

Stable long-term expression of gene of interest in mammalian cells remains a significant challenge for large-scale biotechnological applications. For such a purpose, SB represents an attractive alternative lentiviral vector based transgene expression vector. This is mainly due to its ability to promote relatively easy and efficient genomic integration in larger number of mammalian cell types (Ivics et al., 1997; Luo et al., 1998; Yant et al., 2000; Wilson et al., 2007). To use SB as a tool to integrate transgene into host genome, a two component transposon system is developed. The two plasmid system is very similar to that used in Tn7 transposon (Figure 10). The transposase is expressed in" trans" from a plasmid and the DNA cargo is placed between the inverted terminal repeats in another plasmid. The SB transposase recognizes the IRs in "trans" for its excision and integration activity, it was possible to physically separate the expression of transposase from the recognition motifs for the transposase (DRs and IRs) (Figure 10). This SB based system is suitable to integrate a gene of interest flanked by the IRs. The utility of this tool can be enhanced by combining with other molecular engineering methods such as Cre or FLP recombinases (Narayanavari et al., 2017).

In general, transposition efficiency can be limited by the availability of donor (transposon) and helper (transposase) plasmid in cells. However, the SB transposon system exhibits an overproduction inhibition property. When SB transposase concentration is more than a threshold level, it causes an inhibition on transposition activity. Wilson et al., 2007 examined SB and piggyBac transposon system for transposase over production inhibition. The amount of transposase is controlled by the levels of "helper" plasmid concentrations during transfections. In each experiments the level of the donor (transposon) plasmid is kept constant (50ng, 200ng and 2µg of plasmid DNA) and concentration of helper (transposase) plasmid varied from 200ng to 2µg. In each experiment, the SB transposon showed overproduction inhibition. But the piggyBac system did not exhibit inhibition at any levels of the helper plasmid concentrations. Wu et al., 2006 also observed that SB transposon exhibits over production inhibition.

The SB transposase bind to both inverted repeat elements at the ends of transposon and excise the double strand DNA containing transposon

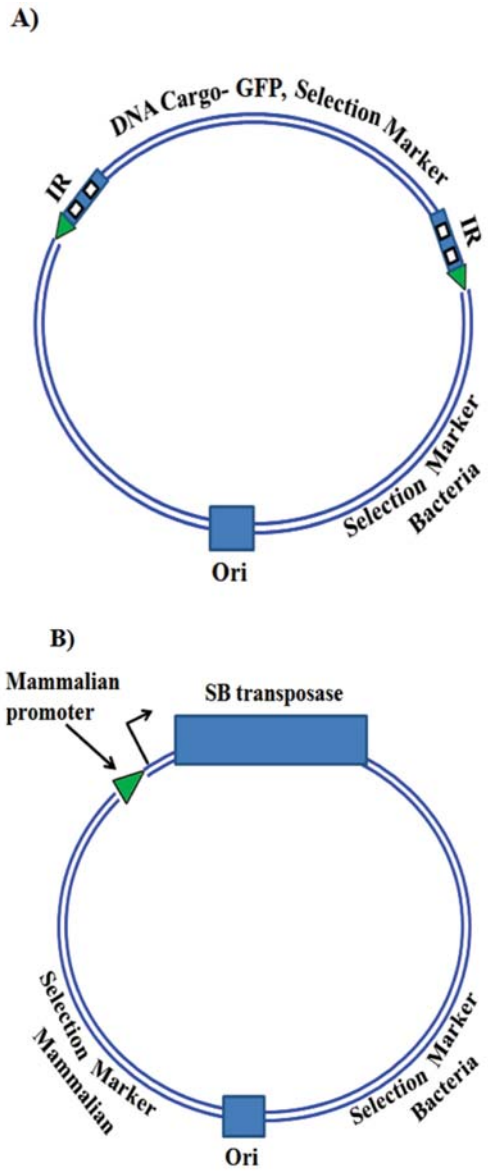

Figure 10. Sleeping beauty transposon system for genetic engineering in mammalian cells. Typically, two plasmid system is used for SB transposition in mammalian cells. A) Transposon delivery vector B) Transposase expressing helper vector

from the donor plasmid (Figure 8).The donor DNA is then repaired by a non-homologous end joining mechanism (NHEJ) (Liu et al., 2005). The NHEJ pathway is known to induce mutations during repair. Repeated excision from mammalian chromosome and integration could take place in case of sustained transposase presence. This is a concern for biotechnology applications in mammalian cells. Liu et al., 2005 developed a method to examine SB excision from donor plasmid. The experiments have shown that SB and transposon excision is only took place in the presence of respective transposon enzymes. In SB system, the repair is associated with significant mutation (insertions) in the donor plasmid.

Sleeping beauty transposon is not a site directed transposon such as Tn7. It can integrate into random locations in genomes. Random integration generally causes perturbation of host gene expression and may lead to diseases such as cancer. Compared to other randomly integrating vectors (such as lentiviral vectors), SB has more chance to integrate into "safe genomic location" (Yant et al., 2000; Wilson et al., 2007). However, this might be a problem to ensure the safety of the biotechnological applications based on SB. Several attempts have been made to add site directed integration ability to SB transposon system (Ivics et al., 2007; Voigt et al., 2012). Recently attempt has been made to direct SB transposon to predetermined sites in human chromosome using CRISPR-Cas9 system (Kovac et al., 2020; Ma et al., 2020). Kovac et al., 2020 used catalytically inactive *S.pyogenes* Cas9 to direct SB transposon to Alu repeats in human genome. The described system will allow a RNA guided integration of SB transposon in human genome.

7. Conclusions and future directions

This paper provides a mini-review on Tn7 and Sleeping Beauty transposons and their applications. Paper is discussing the mechanism and biotechnological application of Tn7 and SB transposons. Paper also discusses the problems associated with these components. Transposon technology provides easily accessible genome manipulating tools for bacterial and mammalian cells. The repurposed transposon systems have a wide variety of applications in environmental biotechnology. These systems can accelerate and increase the accuracy of research in environmental biotechnology. The CRISPR-Cas guided transposon system offers a wide expanding genome engineering tool kit. However, these CRISPR-guided systems still amenable to further optimization and fine tuning for environmental applications. Research and technology development in this area will benefit the environmental biotechnology research community.

References

Arciszewska, L.K., Drake, D. and Craig, N.L., 1989. Transposon Tn7: cis-acting sequences in transposition and transposition immunity. *Journal of MolecularBiology*, *207*(1), pp.35-52.

Bainton, R.J., Kubo, K.M., Feng, J.N. and Craig, N.L., 1993. Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. *Cell*, *72*(6), pp.931-943.

Bao, Y., Lies, D.P., Fu, H. and Roberts, G.P., 1991. An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. *Gene*, *109*(1), pp.167-168.

Barry, G.F., 1986. Permanent insertion of foreign genes into the chromosomes of soil bacteria. *Biotechnology*, *4*(5), pp.446-449.

Barry, G.F., 1988. A broad-host-range shuttle system for gene insertion into the chromosomes of Gram-negative bacteria. *Gene*, *71*(1), pp.75-84.

Bourque, G., Burns, K.H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H.L., Macfarlan, T.S. and Mager, D.L., 2018. Ten things you should know about transposable elements. *Genome Biology*, *19*(1), pp.1-12.

Choi, K.H., Kumar, A. and Schweizer, H.P., 2006. A 10-min method for preparation of highly electrocompetent Pseudomonas aeruginosa cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *Journal of Microbiological Methods*, *64*(3), pp.391-397.

Choi, K.Y., Spencer, J.M. and Craig, N.L., 2014. The Tn7 transposition regulator TnsC interacts with the transposase subunit TnsB and target selector TnsD. *Proceedings of the National Academy of Sciences*, *111*(28), pp.E2858-E2865.

Craig, N.L., 1997. Target site selection in transposition. *Annual Review of Biochemistry*, *66*(1), pp.437-474.

Craig, N.L., 2002. Tn7. In *Mobile DNA II* (pp. 423-456). American Society of Microbiology.

Darzins, A., Kent, N.E., Buckwalter, M.S. and Casadaban, M.J., 1988. Bacteriophage Mu sites required for transposition immunity. *Proceedings of the National Academy of Sciences*, *85*(18), pp.6826-6830.

De Lorenzo, V., Herrero, M., Jakubzik, U. and Timmis, K.N., 1990. Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *Journal of Bacteriology*, *172*(11), pp.6568-6572.

Doudna, J.A. and Charpentier, E., 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213).

Ely, B., 1982. Transposition of Tn7 occurs at a single site on the Caulobacter crescentus chromosome. *Journal of Bacteriology*, *151*(2), pp.1056-1058.

Faure, G., Shmakov, S.A., Yan, W.X., Cheng, D.R., Scott, D.A., Peters, J.E., Makarova, K.S. and Koonin, E.V., 2019. CRISPR–Cas in mobile genetic elements: counter-defence and beyond. *Nature Reviews Microbiology*, *17*(8), pp.513-525.

Fennewald, M.A. and Shapiro, J.A., 1979. Transposition of Tn7 in Pseudomonas aeruginosa and isolation of alk: Tn7 mutations. *Journal of Bacteriology*, *139*(1), pp.264-269.

Ferrières, L., Hémery, G., Nham, T., Guérout, A.M., Mazel, D., Beloin, C. and Ghigo, J.M., 2010. Silent mischief: bacteriophage Mu insertions contaminate products of Escherichia coli random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *Journal of Bacteriology*, *192*(24), pp.6418-6427.

Feschotte, C. and Pritham, E.J., 2007. DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.*, *41*, pp.331-368.

Gay, N.J., Tybulewicz, V.L. and Walker, J.E., 1986. Insertion of transposon Tn7 into the Escherichia coli glm S transcriptional terminator. *Biochemical Journal*, *234*(1), pp.111-117.

Grinter, N.J., 1983. A broad-host-range cloning vector transposable to various replicons. *Gene*, *21*(1-2), pp.133-143.

Hauer, B. and Shapiro, J.A., 1984. Control of Tn 7 transposition. *Molecular and General Genetics MGG*, *194*(1), pp.149-158.

Huang, C.R.L., Burns, K.H. and Boeke, J.D., 2012. Active transposition in genomes. *Annual Review of Genetics*, *46*, pp.651-675.

Ivics, Z. and Izsvák, Z., 2015. Sleeping beauty transposition. *Mobile DNA III*, pp.851-872.

Ivics, Z., Hackett, P.B., Plasterk, R.H. and Izsvák, Z., 1997. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell*, *91*(4), pp.501-510.

Jasinska, W., Manhart, M., Lerner, J., Gauthier, L., Serohijos, A.W. and Bershtein, S., 2020. Chromosomal barcoding of E. coli populations reveals lineage diversity dynamics at high resolution. *Nature Ecology & Evolution*, *4*(3), pp.437-452.

Knott, G.J. and Doudna, J.A., 2018. CRISPR-Cas guides the future of genetic engineering. *Science*, *361*(6405), pp.866-869.

Koch, B., Jensen, L.E. and Nybroe, O., 2001. A panel of Tn7-based vectors for insertion of the gfp marker gene or for delivery of cloned DNA into Gram-negative bacteria at a neutral chromosomal site. *Journal of Microbiological Methods*, *45*(3), pp.187-195.

Koonin, E.V. and Makarova, K.S., 2017. Mobile genetic elements and evolution of CRISPR-Cas systems: all the way there and back. *Genome Biology and Evolution*, *9*(10), pp.2812-2825.

Kovaè, A., Miskey, C., Menzel, M., Grueso, E., Gogol-Döring, A. and Ivics, Z., 2020. RNA-guided retargeting of Sleeping Beauty transposition in human cells. *Elife*, *9*, p.e53868.

Lee, C.H., Bhagwat, A. and Heffron, F., 1983. Identification of a transposon Tn3 sequence required for transposition immunity. *Proceedings of the National Academy of Sciences*, *80*(22), pp.6765-6769.

Lichtenstein, C. and Brenner, S., 1981. Site-specific properties of Tn 7 transposition into the *E. coli* Chromosome. *Molecular and General Genetics MGG*, *183*(2), pp.380-387.

Lichtenstein, C. and Brenner, S., 1982. Unique insertion site of Tn 7 in the *E. coli* chromosome. *Nature*, *297*(5867), pp.601-603.

Liu, G., Aronovich, E.L., Cui, Z., Whitley, C.B. and Hackett, P.B., 2004. Excision of Sleeping Beauty transposons: parameters and applications to gene therapy. *The Journal of Gene Medicine: A cross disciplinary journal for research on the science of gene transfer and its clinical applications*, *6*(5), pp.574-583.

Lohe, A.R., Moriyama, E.N., Lidholm, D.A. and Hartl, D.L., 1995. Horizontal transmission, vertical inactivation, and stochastic loss of mariner-like transposable elements. *Molecular Biology and Evolution*, *12*(1), pp.62-72.

Luo, G., Ivics, Z., Izsvák, Z. and Bradley, A., 1998. Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. *Proceedings of the National Academy of Sciences*, *95*(18), pp.10769-10773.

Ma, S., Wang, X., Hu, Y., Lv, J., Liu, C., Liao, K., Guo, X., Wang, D., Lin, Y. and Rong, Z., 2020. Enhancing site-specific DNA integration by a Cas9 nuclease fused with a DNA donor-binding domain. *Nucleic acids Research*, *48*(18), pp.10590-10601.

McKown, R.L., Orle, K.A., Chen, T. and Craig, N.L., 1988. Sequence requirements of Escherichia coli attTn7, a specific site of transposon Tn7 insertion. *Journal of Bacteriology*, *170*(1), pp.352-358.

Miller, V.L. and Mekalanos, J.J., 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in Vibrio cholerae requires toxR. *Journal of Bacteriology*, *170*(6), pp.2575-2583.

Moldt, B., Yant, S.R., Andersen, P.R., Kay, M.A. and Mikkelsen, J.G., 2007. Cis-acting gene regulatory activities in the terminal regions of sleeping beauty DNA transposon-based vectors. *Human Gene Therapy*, *18*(12), pp.1193-1204.

Narayanavari, S.A., Chilkunda, S.S., Ivics, Z. and Izsvák, Z., 2017. Sleeping Beauty transposition: from biology to applications. *Critical Reviews in Biochemistry and Molecular Biology*, *52*(1), pp.18-44.

Peters, J.E. and Craig, N.L., 2000. Tn7 transposes proximal to DNA double-strand breaks and into regions where chromosomal DNA replication terminates. *Molecular Cell*, *6*(3), pp.573-582.

Peters, J.E. and Craig, N.L., 2001a. Tn7: smarter than we thought. *Nature Reviews Molecular Cell Biology*, *2*(11), pp.806-814.

Peters, J.E. and Craig, N.L., 2001b. Tn7 recognizes transposition target structures associated with DNA replication using the DNA-binding protein TnsE. *Genes & Development*, *15*(6), pp.737-747.

Peters, J.E., Makarova, K.S., Shmakov, S. and Koonin, E.V., 2017. Recruitment of CRISPR-Cas systems by Tn7-like transposons. *Proceedings of the National Academy of Sciences*, *114*(35), pp.E7358-E7366.

Plasterk, R.H., Izsvák, Z. and Ivics, Z., 1999. Resident aliens: The Tc1/mariner superfamily of transposable elements. *Trends in Genetics*, *15*(8), pp.326-332.

Plasterk, R.H.A., 1996. The Tc1/mariner transposon family. *Transposable Elements*, pp.125-143.

Reyes, O., Beyou, A., Mignotte-Vieux, C. and Richaud, F., 1987. Mini-Mu transduction: cis-inhibition of the insertion of Mud transposons. *Plasmid*, *18*(3), pp.183-192.

Schlechter, R.O., Jun, H., Bernach, M., Oso, S., Boyd, E., Muñoz-Lintz, D.A., Dobson, R.C., Remus, D.M. and Remus-Emsermann, M.N., 2018. Chromatic bacteria–A broad host-range plasmid and chromosomal insertion toolbox for fluorescent protein expression in bacteria. *Frontiers in Microbiology*, *9*, p.3052.

Shen, H., Gold, S.E., Tamaki, S.J. and Keen, N.T., 1992. Construction of a Tn7-lux system for gene expression studies in Gram-negative bacteria. *Gene*, *122*(1), pp.27-34.

Simon, R.U.P.A.P., Priefer, U. and Pühler, A., 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Biotechnology*, *1*(9), pp.784-791.

Stellwagen, A.E. and Craig, N.L., 1997. Avoiding self: two Tn7 encoded proteins mediate target immunity in Tn7 transposition. *The EMBO journal*, *16*(22), pp.6823-6834.

Strecker, J., Ladha, A., Gardner, Z., Schmid-Burgk, J.L., Makarova, K.S., Koonin, E.V. and Zhang, F., 2019. RNA-guided DNA insertion with CRISPR-associated transposases. *Science*, *365*(6448), pp.48-53.

Tan, S.Z., Reisch, C.R. and Prather, K.L., 2018. A robust CRISPR interference gene repression system in Pseudomonas. *Journal of Bacteriology*, *200*(7).

Tang, Y., Lichtenstein, C. and Cotterill, S., 1991. Purification and characterisation of the TnsB protein of Tn7: a transposition protein that binds to the ends of Tn7. *Nucleic acids Research*, *19*(12), pp.3395-3402.

Thomson, J.A., Hendson, M. and Magnes, R.M., 1981. Mutagenesis by insertion of drug resistance transposon Tn7 into a vibrio species. *Journal of Bacteriology*, *148*(1), pp.374-378.

Vigdal, T.J., Kaufman, C.D., Izsvák, Z., Voytas, D.F. and Ivics, Z., 2002. Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/mariner transposable elements. *Journal of Molecular Biology*, *323*(3), pp.441-452.

Vo, P.L.H., Ronda, C., Klompe, S.E., Chen, E.E., Acree, C., Wang, H.H. and Sternberg, S.H., 2021. CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nature Biotechnology*, *39*(4), pp.480-489.

Voigt, F., Wiedemann, L., Zuliani, C., Querques, I., Sebe, A., Mátés, L., Izsvák, Z., Ivics, Z. and Barabas, O., 2016. Sleeping Beauty transposase structure allows rational design of hyperactive variants for genetic engineering. *Nature Communications*, *7*(1), pp.1-8.

Waddell, C.S. and Craig, N.L., 1989. Tn7 transposition: recognition of the attTn7 target sequence. *Proceedings of the National Academy of Sciences*, *86*(11), pp.3958-3962.

Walisko, O., Schorn, A., Rolfs, F., Devaraj, A., Miskey, C., Izsvák, Z. and Ivics, Z., 2008. Transcriptional activities of the Sleeping Beauty transposon and shielding its genetic cargo with insulators. *Molecular Therapy*, *16*(2), pp.359-369.

Walker, J.E., Gay, N.J., Saraste, M. and Eberle, A.N., 1984. DNA sequence around the Escherichia coli unc operon. Completion of the sequence of a 17 kilobase segment containing asnA, oriC, unc, glmS and phoS. *Biochemical Journal*, *224*(3), pp.799-815.

Watson, J.D., 2004. *Molecular biology of the gene*. Pearson Education India.

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O. and Paux, E., 2007. A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics*, *8*(12), pp.973-982.

Wiles, T.J., Wall, E.S., Schlomann, B.H., Hay, E.A., Parthasarathy, R. and Guillemin, K., 2018. Modernized tools for streamlined genetic manipulation and comparative study of wild and diverse proteobacterial lineages. *MBio*, *9*(5).

Wilson, M.H., Coates, C.J. and George Jr, A.L., 2007. PiggyBac transposon-mediated gene transfer in human cells. *Molecular Therapy*, *15*(1), pp.139-145.

Wu, S.C.Y., Meir, Y.J.J., Coates, C.J., Handler, A.M., Pelczar, P., Moisyadi, S. and Kaminski, J.M., 2006. piggyBac is a flexible and highly active transposon as compared to sleeping beauty, Tol2, and Mos1 in mammalian cells. *Proceedings of the National Academy of Sciences*, *103*(41), pp.15008-15013.

Yant, S.R., Meuse, L., Chiu, W., Ivics, Z., Izsvak, Z. and Kay, M.A., 2000. Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nature Genetics*, *25*(1), pp.35-41.