Applications of Transposon-Based Gene Delivery System in Bacteria

Choi, Kyoung-Hee* and Kang-Ju Kim

Department of Oral Microbiology, College of Dentistry, Wonkwang University, Iksan 570-749, Korea

Received: November 11, 2008 / Accepted: December 10, 2008

Mobile genetic segments, or transposons, are also referred to as “jumping genes” as they can shift from one position in the genome to another, thus inducing a chromosomal mutation. According to the target site-specificity of the transposon during a transposition event, the result is either the insertion of a gene of interest at a specific chromosomal site, or the creation of knockout mutants. The former situation includes the integration of conjugative transposons via site-specific recombination, several transposons preferring a target site of a conserved AT-rich sequence, and Tn7 being site-specifically inserted at attTn7, the downstream of the essential glmS gene. The latter situation is exploited for random mutagenesis in many prokaryotes, including IS (insertion sequence) elements, mariner, Mu, Tn3 derivatives (Tn4430 and Tn917), Tn5, modified Tn7, Tn10, Tn552, and Ty1, enabling a variety of genetic manipulations. Randomly inserted transposons have been previously employed for a variety of applications such as genetic footprinting, gene transcriptional and translational fusion, signature-tagged mutagenesis (STM), DNA or cDNA sequencing, transposon site hybridization (TraSH), and scanning linker mutagenesis (SLM). Therefore, transposon-mediated genetic engineering is a valuable discipline for the study of bacterial physiology and pathogenesis in living hosts.

Keywords: Transposon, gene integration

Since the advent of bacterial genome sequencing, a great many molecular genetic manipulation tools have been developed and improved, via the application of genetic events that occur naturally in prokaryotes, in order to characterize the genes and their functions in a variety of environments. Many pathogenic bacteria in animals and plants are capable of forming a biofilm in certain environments. Therefore, there is currently more interest in modifying the chromosome itself than in plasmid-based gene manipulation, which necessitates selection pressures such as antibiotic supplementation and is not feasible in such environments.

Bacterial chromosomes can be manipulated either via integration or by excision. Chromosomal DNA manipulation tools can generally be divided into three groups on the basis of mechanism of genetic transfer: (1) homologous recombination; (2) site-specific recombination; (3) transposon-mediated gene integration (Fig. 1). Homologous recombination can generally occur as the result of recombinase RecA-mediated catalytic activity between two long homologous sequences [6]. Site-specific recombination includes enzymes belonging to two major families: (1) the resolvase-invertase family; (2) Int family (FLP-FRT, Cre-loxP). Transposon-mediated gene integration involves phage integration and excision via site-specific recombination between the phage attachment site, attP, and the bacterial attachments site, attB. These events cause either gene insertion or excision on the chromosome. Transposons are mobile genetic materials that move to another genomic position, a process referred to as transposition. Transposition occurs via one of two mechanisms: cut-and-paste transposition (e.g., Tn5 and Tn10) [81], or replicative transposition (e.g., Tn3, Mu, and many IS) [95] by one or both enzymes of transposase and resolvase, leaving one copy on the target DNA or two copies on both donor and target DNA, respectively. In general, transposon-based gene integration does not require homologous sequences for transfer to the chromosome, unlike recombination-associated chromosomal manipulations. Therefore, transposons can be widely utilized for the creation of random mutants, which might be exploited and applied further for the characterization of essentiality and the functions of genes associated with host–pathogen relationship. In addition, a small number of transposons can be integrated at a preferred, neutral, naturally evolved, and defined target site without any deleterious effects or preparatory genetic modification, which involves the construction of a genetically engineered bacterial strain harboring a chromosomal insertion of genes of interest such as reporter genes, green fluorescent protein (gfp), luciferase gene (lux), and β-galactosidase (lacZ)
The gene of interest is inserted into the bacterial chromosome in a variety of ways. Homologous recombination occurs between two homologous nucleotide sequences via the enzymatic activity of RecA, λ-Red, or Homing endonuclease. Conservative site-specific recombination involves enzymes belonging to two major families: the Resolvase-Invertase family (Hin, Gin, Cin, and Pin) and Int family (FLP of yeast and Cre of bacteriophage P1, and the integration and excision of bacteriophage λ). According to the target site-specificity of transposons, the transpositions result in a chromosomal mutation at a specific or random site.

In the text, a variety of useful transposons will be reviewed along with individual examples of their applications.

### Site-Specific Insertion of Transposons

**Transposition via site-specific recombination; conjugative transposons.** Generally, the target site selection of transposon requires sequence similarity between the target and donor DNA segments. The conjugative transposon, Tn4555, found in the Gram-negative *Bacteroides* of the human intestinal tract was determined to integrate into a specific target gene via a site-specific recombination system in an orientation-specific manner [110]. Site-specific integration involves three genes, including an essential int gene encoding for an integrase similar to lambda integrase, and two accessory genes, tnpA and tnpC, which enhance the integration efficiency [111]. However, tnpA deletion caused random insertion, thereby indicating that TnpA performs a function in the facilitation of the site-specific insertion of Tn4555 [111]. Its transposon insertion occurs in *Bacteroides fragilis*, mostly at either end of two direct repeats, PT-1 and PT-2, as a primary target site, whereas random insertions also occur at relatively low frequency rates [110]. Although transposition via Tn4555 is quite similar to lambda integrase, which catalyzes a site-specific recombination between highly homologous attP and attB, Tn4555 transposes via a site-specific recombination of poor homologous sequences of Tn4555 ends to the target site, unlike CTnDOT, SXT, and lambda integration [111].

In addition, the CTnDOT conjugative transposon integrates into the 10-bp sequence of the *B. fragilis* chromosome, which is identical to the transposon end [14]. Another example, the SXT conjugative transposon, which harbors multiple antibiotic resistance genes, inserts into the 5′end of the *prfC* gene of *Vibrio cholerae* by recognizing a 17-bp identical sequence of transposon ends to its target [49].

**Transposition at a conserved AT-rich sequence; IS(s) or Tn(s).** In addition to the target sequence-specific primary target selection, a second pathway was also suggested, which is dependent on the interaction of the integrase protein with the target sequence, with no need for sequence similarity between the target gene and the ends of the transposon [111].

The insertion sequences (ISs), IS605, IS606, and ISHp608 in *Helicobacter pylori* have been shown to transpose at a high frequency into highly selective target sites of *Escherichia coli* [109]. The IS605 integrates into a specific target site, the conserved AT-rich pentanucleotide sequence, TTTAA or TTAAC [52]. Similary, transposons, which lack sequence homology for the selection of the insertion site, but rather integrate in a site-specific manner, include Tn916 and Tn1545 from *Enterococcus faecalis* and *Streptococcus pneumoniae*, respectively [93]. Tn916 and Tn1545, simple conjugative transposons in Gram-positive bacteria, harbor a preferred insertion site, which resides at the position harboring an AT-rich sequence separated by six bases from a T-rich sequence. Therefore, although Tn916 and its relatives have also been found in Gram-negative bacteria including *Neisseria gonorrhoeae* and *Kingella* spp., these transposons insert principally into the chromosomes of Gram-positive bacteria with a high AT content [93, 106]. However, as the transposons

---

**Fig. 1.** Chromosomal gene integration systems used in bacteria. The gene of interest is inserted into the bacterial chromosome in a variety of ways. Homologous recombination occurs between two homologous nucleotide sequences via the enzymatic activity of RecA, λ-Red, or Homing endonuclease. Conservative site-specific recombination involves enzymes belonging to two major families: the Resolvase-Invertase family (Hin, Gin, Cin, and Pin) and Int family (FLP of yeast and Cre of bacteriophage P1, and the integration and excision of bacteriophage λ). According to the target site-specificity of transposons, the transpositions result in a chromosomal mutation at a specific or random site.
may harbor multiple target sites of conserved AT-rich sequences, the use of these transposons is not an appropriate way to introduce any gene of interest at a specific chromosomal site.

**Transposition at a unique attachment site; Tn7.** Tn7 also does not share sequence identity between the target insertion site and the Tn7 ends. Rather, site-specific integration via Tn7 is governed by the TnsD-targeting protein, which is close to TnpA, and directs primary target site insertion in Tn4555 [5]. The unique chromosomal attachment Tn7 site, referred to as atTn7, resides downstream of the essential glmS gene, which encodes for the glucosamine-6-phosphate synthetase involved in cell wall biosynthesis [21]. The site-specific Tn7 insertion site has been identified in many Gram-negative bacteria, including *Desulfovibrio desulfuricans* [113], *E. coli* [66], *Serratia marcescens* [21], *Sphingomonas yanoikuyae* [114], *Vibrio anguillarum* [63], and *Vibrio fischeri* [70] (Table 1).

The mini-Tn7 transposon is a great tool for a single-copy tagging of bacteria in a site-specific manner at a unique and neutral site without any deleterious effects. Heterologous genes including *lacZ* (β-galactosidase), *est*

### Table 1. Determined Tn7 insertion sites (attTn7) in various bacteria.

<table>
<thead>
<tr>
<th>Strains</th>
<th>attTn7 sites</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type I : single glmS-related attTn7 site</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em> K12</td>
<td></td>
<td>[66]</td>
</tr>
<tr>
<td><em>Desulfovibrio desulfuricans</em> G20</td>
<td></td>
<td>[113]</td>
</tr>
<tr>
<td><em>Methylomonas extorquens</em> ATCC55366</td>
<td></td>
<td>[19]</td>
</tr>
<tr>
<td><em>Pseudomonas aeruginosa</em> PA01</td>
<td></td>
<td>[16]</td>
</tr>
<tr>
<td><em>Pseudomonas fluorescens</em> Pf-5</td>
<td></td>
<td>[54]</td>
</tr>
<tr>
<td><em>Pseudomonas putida</em> KT2440</td>
<td></td>
<td>[60]</td>
</tr>
<tr>
<td><em>Pseudomonas syringae</em> pv. tomato str. DC3000</td>
<td></td>
<td>[96]</td>
</tr>
<tr>
<td><em>Sphingomonas yanoikuyae</em> B1</td>
<td></td>
<td>[114]</td>
</tr>
<tr>
<td><em>Vibrio fischeri</em> ES114</td>
<td></td>
<td>[70]</td>
</tr>
<tr>
<td><em>Yersinia pestis</em> KIM</td>
<td></td>
<td>[16]</td>
</tr>
</tbody>
</table>

| **TYPE II : multiple glmS-related attTn7 sites** | | |
| *Burkholderia mallei* ATCC 23344 | | [15] |
| *Burkholderia pseudomallei* K96243 | | [17] |
| *Burkholderia thailandensis* E264 | | [16] |

| **Type III : secondary, glmS-unrelated attTn7 site** | | |
| *Proteus mirabilis* HI4320 | | [18] |

The glmS gene and its downstream gene are shown in closed and open arrows, respectively. The arrows indicate the transcriptional orientation of the genes. The *carA* and *carB* genes in gray arrows encode for a carbamoyl phosphate synthetase that plays an essential role in arginine biosynthesis. The attTn7 sites are designated in closed triangles.
(esterase), and gfp (green fluorescent protein) under the control of the methanol dehydrogenase promoter have been integrated into the intergenic region between glmS and dhaT via the delivery of mini-Tn7 in Methylobacterium extorquens [19]. A gene encoding for green fluorescent protein or luciferase protein under the control of a certain promoter was integrated into the chromosomes of Erwinia chrysanthemi [96], Pseudomonas fluorescens [54], Pseudomonas syringae [96], and Pseudomonas putida [60] by the Tn7-based system. Recently, the mini-Tn7-based gene integration system was improved for gene complementation, gene expression analysis, strain construction, and reporter gene-tagging of Pseudomonas aeruginosa and Yersinia pestis, particularly in biofilm and animal models [16]. More recently, the gene delivery system developed was applied to other organisms, such as Burkholderia spp. and Proteus mirabilis, which were determined to have multiple glmS-linked attTn7 sites and secondary, non-glmS-linked attTn7 site, respectively [15, 18].

As analogs of the glmS gene are found in all organisms, including bacteria, yeast, and even humans, Tn7 transposition appears to be a universal tool for genetic manipulation. For example, two analogs of glmS, glutamine-fructose-6-phosphate-transaminase-1 and -2 (gfpt-1 and gfpt-2) were identified in the human chromosome, and their adjacent positions are a potential target of Tn7 transposition with a higher frequency of insertion neighboring to gfpt-1 than gfpt-2, thereby suggesting that they may prove a useful tool for site-specific DNA delivery in a gene therapy technique [55]. However, although Saccharomyces cerevisiae harbors the glmS analog, gfa-1, at its chromosome XI, does not transpose adjacent to the possible target gfa-1 sequence [55]. A Tn7-like transposon, Tn5468, was found at the intergenic region between glmU encoding for N-acetylg glucosamine-1-uridylyltransferase and the glmS gene in Thiobacillus ferroxidans [75].

Randomly Integrated Transposons

Drosophila-derived transposon; mariner. The mariner family of transposons originally found in Drosophila mauritiana is a DNA element of approximately 1,300 bp in size containing 30-bp short inverted terminal repeats and generates the duplication of a TA dinucleotide target site during transposition [84]. As transposons of the mariner family require no species-specific host factors for proper transposition, they have been extensively utilized for the random mutagenesis of both prokaryotes and eukaryotes [43]. Its use has been genetically manipulated in a variety of bacteria including Gram-negative and Gram-positive bacteria, and Mycobacteria [1, 36, 86]. Furthermore, random insertional mutagenesis via a mariner transposon has also been achieved in Leptospira interrogans, the causative agent of Leptospirosis, which affects both humans and animals – this random insertional mutagenesis was undertaken in an effort to identify several putative virulence factors [8] (Table 2). A mariner transposable element, Himar1, was applied successfully to Gram-positive bacteria such as Bacillus subtilis, in which Tn917 and Tn10 transposons were available for random mutagenesis, but proved less productive [62, 80, 122]. Although the Tn917 integration efficiency was high, actual random mutants are relatively scarce owing to the existence of hot-spots [123]. In the case of Tn10, the requirement for a 6-bp unique target sequence for random transposition reduces the frequency of insertion into the B. subtilis chromosome [40]. Recently, MAR2xT7, a derivative of Himar1, was successfully utilized to create an ordered, nonredundant library of P. aeruginosa PA14 mutants, owing to its minimal insertion-site specificity. This allowed for the identification of 335 putative essential genes in P. aeruginosa via comparison of the Tn5-based genomic library of PA01 with the one established in PA14 [65]. More recently, a Himar1-based random mutagenesis system, designated HimarBP, was developed for the creation of comprehensive transposable libraries of Burkholderia pseudomallei, the etiological agent of melioidosis [82].

Phage-derived transposon; Mu. The genome of the Mu bacteriophage is replicated using its transposition machinery, which entails a transpososome consisting of four Mu bacteriophage is replicated using its transposition machinery, and generates the duplication of a TA dinucleotide target site during transposition [84]. Via transpososome mutagenesis of P. aeruginosa strain PA68 via the electroporation of in vitro assembled Mu complexes with MuA transposase and mini-Mu transposon in order to identify genes involved in twitching motility [94]. Furthermore, it has also been applied in a variety of studies including functional genetic analysis [39], DNA sequencing [38], protein identification via pentapeptide scanning mutagenesis [108], and gene targeting [112] (reviewed in [58]). Its use...
was extended to eukaryotic cells for transfection studies of mammalian cells using Mu proteins fused to Flag-epitope and a SV40-derived nuclear localization signal [91]. For more efficient transposition reaction in mammalian cells, a second Mu ATPase protein, MuB, is required to stimulate the assembly of transpososomes [73]. In addition to the MuA and MuB proteins, bacterial host-encoded Hu factor is also essential for efficient transposition, but mammalian cells can instead use two DNA-binding proteins: HMG (high mobility group) -1 and HMG-2 [61].

Bacteria-derived transposon; Tn(s). This includes various transposons [Tn(s)] such as Tn3, Tn5, modified Tn7, Tn10, Tn552, and so on. Among them, Tn5-based random mutagenesis was utilized in a vast variety of bacterial organisms for the identification of functional genes [26, 31, 59, 64, 98, 103, 107, 117, 121]. Very recently, it was proposed that a minimal genome of bacterial cells can instead use two DNA-binding proteins: HMG (high mobility group) -1 and HMG-2 [61].

### Table 2. Randomly integrated transposons used in bacteria.

<table>
<thead>
<tr>
<th>Sources</th>
<th>Transposons</th>
<th>Properties</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila mauritiana</td>
<td>Himar1 mariner</td>
<td>▶ Cut-and-paste transposition</td>
<td>▶ RM: B. subtilis, B. pseudomallei, L. interrogans, and P. aeruginosa</td>
</tr>
<tr>
<td></td>
<td>Two Himar1 derivatives</td>
<td>▶ TA target site duplication</td>
<td>▶ GF: H. influenzae, P. aeruginosa, and S. pneumoniae</td>
</tr>
<tr>
<td></td>
<td>(MAR2x T7, HimarBP)</td>
<td>▶ No requirement of species-specific host factor</td>
<td>▶ STM: C. jejuni, N. meningococcus, and S. sanguinis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ Broad-host range</td>
<td>▶ TraSH: M. bovis and M. tuberculosis</td>
</tr>
<tr>
<td>Bacteriophage Mu</td>
<td>Mu</td>
<td>▶ Replicative transposition</td>
<td>▶ RM: E. coli, E. carovorata, and P. aeruginosa</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ A 5-bp target site duplication</td>
<td>▶ TLF: L. pneumophila</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ Transposome (4 Mu transposases and 2 transposon right-end DNA segments)</td>
<td>▶ DS: E. coli</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ Requirement of bacterial host factor Hu for transposition</td>
<td>▶ SLM in P. syringae</td>
</tr>
<tr>
<td>Bacteria</td>
<td>Tn3</td>
<td>▶ Replicative transposition</td>
<td>▶ TCF: B. anthracis and S. mutans</td>
</tr>
<tr>
<td></td>
<td>Tn3-like transposon, Tn917</td>
<td>▶ A 5-bp target site duplication</td>
<td>▶ STM: L. monocyctogenes, S. aureus, and S. agalactiae</td>
</tr>
<tr>
<td></td>
<td>Tn3-like transposon, Tn4430</td>
<td>▶ Cut-and-paste transposition</td>
<td>▶ SLM: B. thuringiensis and E. coli</td>
</tr>
<tr>
<td></td>
<td>Tn5</td>
<td>▶ A 9-bp target site duplication</td>
<td>▶ TCF: P. luminescens, P. aeruginosa, P. putida, and V. parahaemolyticus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ Mostly utilized for random mutagenesis</td>
<td>▶ TLF: P. putida, S. enterica serovar Typhimurium, S. melloti, and V. vulnificus</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ Cut-and-paste transposition</td>
<td>▶ GF: E. coli</td>
</tr>
<tr>
<td></td>
<td>Tn10</td>
<td>▶ A 9-bp target site duplication</td>
<td>▶ STM: C. jejuni and X. nematophila</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ Requirement of a 6-bp unique target sequence</td>
<td>▶ TraSH: E. coli and S. enterica serovar Typhimurium</td>
</tr>
<tr>
<td></td>
<td>Tn552</td>
<td>▶ Originated from S. aureus</td>
<td>▶ RM: C. jejuni</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ Replicative transposition</td>
<td>▶ TCF: M. tuberculosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>▶ Short inverted repeat sequences (48-bp)</td>
<td>▶ DS: M. smeagnatis</td>
</tr>
</tbody>
</table>

DS, cDNA sequencing; GF, genetic footprinting; RM, random mutagenesis; SLM, scanning linker mutagenesis; STM, signature-tagged mutagenesis; TCF, transcriptional fusion; TLF, translational fusion; TraSH, transposon site hybridization.

This includes various transposons [Tn(s)] such as Tn3, Tn5, modified Tn7, Tn10, Tn552, and so on. Among them, Tn5-based random mutagenesis was utilized in a vast variety of bacterial organisms for the identification of functional genes [26, 31, 59, 64, 98, 103, 107, 117, 121]. Very recently, it was proposed that a minimal genome of E. coli required only for survival could be created by using the Tn5 system [124].
to create random mutants of Campylobacter jejuni, a common pathogen that causes foodborne disease [20]. Although the Himar1 transposon has been previously applied for the in vivo random mutagenesis of C. jejuni, it appeared to be unsuitable owing to the low efficiency rate of mutation derived from having the high restriction barriers and inefficient expression of transposase proteins in C. jejuni cells [20, 32]. Tn552 is quite similar to Mu, in which it encodes for a transposase TnpA harboring an Asp-Asp-Glu (DDE) active site and a TnpB accessory protein required for transposition, and its transposition occurs via a replicative mechanism. However, as it necessitates the ends of Tn552, which are only 48-bp inverted repeat sequences for efficient transposition, it is simpler than using Mu [34, 72]. The advantages of Tn552 include the fact that it requires only a small transposase protein and an accessory protein for optimal transposition [85].

Although Tn7 displays a high degree of target site specificity in most organisms, it can still be employed for random mutation with the modification of transposase subunits. In general, bacterial Tn7 has a preferred target site, referred to as attTn7, which is downstream of the highly conserved glmS gene encoding for glucosamine-6-phosphate synthetase in the presence of its transposase complex composed of a common core TnsAB, an ATP-dependent DNA-binding protein TnsC, and a targeting protein TnsD. By way of contrast, TnsABC and TnsE, rather than TnsD, preferentially stimulate random transposition into many other sites on conjugating plasmids and proximally into regions of chromosomal DNA replication termination and DNA double-strand breaks [22, 23, 79, 119]. TnsA and TnsB function as a Tn7 transposase, which catalyzes the actual DNA cleavage and rejoining steps [7]. The central regulator TnsC is an ATPase, the function of which is to hydrolyze ATP and achieve the site-specific transposition via binding to attTn7 sites via interactions with TnsD [5, 101]. However, TnsC mutants display gain-of-function phenotypes, which increase transposition frequency in the absence of the target-site determining proteins TnsD or TnsE and engage in altered interactions with ATP and DNA, thereby inducing random transposition events [101]. Tn7-derived large-scale insertional mutagenesis of the S. cerevisiae genome was conducted using the gain-of-function allele of the Tn7-encoded protein TnsC, thereby resulting in the characterization of 25 genes associated with DNA repair, replication, transcription, and chromatin structure [56]. Additionally, random mutants were created successfully in H. influenzae Rd to assess the crucial function of attP as a regulatory gene during natural DNA transformation [37]. Furthermore, this technique was utilized for the performance of in vitro genome-wide mutagenesis, also referred to as transposon-arrayed gene knockout (TAGKO), in filamentous fungi including Magnaporthe grisea and Mycosphaerella graminola, using a modified Tn7 delivery vector containing polyadenylation signals preceded by stop codons in both the Tn7R and Tn7L ends, thereby inducing the premature polyadenylation of mRNA and thus the production of truncated proteins. Thus, it provides a tool for the in vivo study of defective proteins in filamentous fungi [67]. Additionally, many other studies associated with functional analysis including protein–protein interactions, have been predicated on Tn7-mediated random mutagenesis involving the insertion of gene fusions, gene disruptions or deletions, or epitope-tagged sequences into the viral and bacterial genomes [10, 69, 102]. In addition to TnsC mutation, the inactivation of TnsD alone via a site-specific transposition of Tn7 also proved capable of generating random mutations in the sulfate-reducing bacterium Desulfovibrio desulfuricans G20 [113].

Applications of Randomly Integrated Transposons

Genetic footprinting. Transposons have been utilized to determine essential genes in the bacterial genome. This application is referred to as “genetic footprinting,” which is involved in the generation of random transposon mutants followed by the identification of genes that are depleted during outgrowth under certain circumstances. The essential genes are determined via PCR with a pair of primers, one which anneals to the end of the transposon and the other to a nearby gene of interest. This technology was initially exploited for the screening of essential genes in S. cerevisiae under different physiological conditions using the Ty1 transposon [99, 100]. Thus far, there have been several examples in which genetic footprinting systems have been utilized to identify the essentiality of genes in the genome and the potential functions of a large set of genes with unknown function, including a mariner-based mutagenesis in H. influenzae [1, 2], P. aeruginosa [120], and S. pneumoniae [1], and Tn5- or Tn10-based systems in E. coli [4, 29, 41]. For example, the minimal bacterium Mycoplasma genitalium, which possesses the smallest genome and thus a minimal metabolism, was mutated randomly to identify the essentialities of genes using the Tn4001 transposon [30].

Transcriptional and translational fusions. Transposons have been previously utilized for the random generation of transcriptional or translational fusions such as the lacZ, phoA, luxAB, or xylE genes, thus providing an easier way to determine their individual gene expression levels under different conditions, as well as to create mutants [25]. A Tn5-based lux transcriptional reporter transposon was utilized for the study of differentially regulated genes under distinct physiological conditions, thus allowing for further functional studies in Photorhabdus luminescens [117], P. aeruginosa [64], P. putida [26], and Vibrio parahaemolyticus [103]. Moreover, a Tn917 derivative containing a promoterless lacZ reporter gene successfully generated transcriptional
fusions for the screening of environmentally regulated genes in *Streptococcus mutans* [24], or for the identification of regulators of virulence gene transcription in *Bacillus anthracis* [50]. In addition to transcriptional fusions, translational fusions also have been generated to identify the exported and cell envelope-associated proteins and the membrane topology of these proteins [69]. Transposon derivatives with *phoA* fusion are generally used in order that the *phoA* gene devoid of its 5’ signal peptide sequence is fused in-frame to the transposon, and the *phoA* in-frame insertion to the N-terminus of the protein is created via transposition, resulting in a hybrid protein. As PhoA is active in the periplasm, the transposon mutant containing only the exported target protein shows alkaline phosphatase activity, and thus allows for the identification of proteins exported across the cytoplasmic membrane [43]. Tn5-based *phoA* translational fusions have previously been generated in *Klebsiella oxytoca* [77], *P. putida* [31], *Salmonella enterica* serovar Typhimurium [107], *Sinorhizobium melloti* [59], and *Vibrio vulnificus* [121]. In addition to the Tn5 transposon, other transposable elements, Tn10 and Mu, and Tn552 harboring S’-truncated *phoA* were also utilized for the identification of exported proteins in *Legionella pneumophila* and *Mycobacterium tuberculosis* [3, 9, 71].

**Signature-tagged mutagenesis (STM).** The assessment of pathogenic microbial virulence genes is required for the discovery of drug targets and vaccine candidates. Signature-tagged mutagenesis is an improved transposon-based random mutagenesis, which has been designed and used for the identification of virulence factors in animal models, as it has both insertional mutagenesis and negative selection capabilities, and thus obviates the need for the individual tests that were previously the only genetic tool available for the identification of virulence genes [47]. Basically, the transposon containing different signature-tags composed of random sequences generates a tagged mutant library that is tested for attenuated virulence in an animal host, followed by the validation of genes that are not recovered in an output pool following the infection of an input mutant pool, and finally considering them as essential genes for normal cell growth [4, 118]. Thus, the detection and identification of conditionally essential genes will provide a new antimicrobial target in specific environments. In addition, microarray-combined Tn5 transposon

DNA or cDNA sequencing. Although a variety of techniques, including random shotgun subcloning and unidirectional deletions and subcloning, have been developed for DNA sequencing, it still necessitates an efficient DNA sequencing strategy because these techniques are expensive and involve laborious steps [104]. Therefore, the method devised was a transposon-based sequencing approach, in which the sequences generated from random transposon insertions were assembled to yield the entire sequence of interest. The Tn3 transposon efficiently facilitated DNA sequencing in *E. coli* [104], and the *in vitro* transposition of Tn552 was successful for the DNA sequencing of a cosmid containing the *Mycobacterium smegmatis* recBCD gene cluster [34], removing all downstream subcloning steps. Although there exist other well-known strategies of cDNA sequencing strategies, such as primer walking and concatenated cDNA sequencing (CCS) similar to shotgun sequencing, the transposition of the Mu transposon, a simpler approach, could determine entire cDNA sequences, due to its high level of randomness [11]. In addition to the Mu transposition, Tn5 was also successfully utilized for the cDNA sequencing of more than 4,200 mammalian cDNA clones. As transposon-mediated cDNA sequencing involves the use of transposon-specific primers and sequences, which are assembled to obtain the full-length cDNA sequence, it overcomes limitations such as the extensive costs for synthetic oligonucleotides, the many repeated walking steps in larger genomes, and the need for exact annealing of primer and DNA [11, 97].

**Transposon site hybridization (TraSH).** By combining microarray and comprehensive insertional transposon mutagenesis, a new method, transposon site hybridization (TraSH), was developed to identify the differentially expressed specific genes required for optimal growth under different conditions; for example, in minimal medium in comparison with cells growing in rich medium. This powerful tool was exemplified by Tn5- or Tn10-based *E. coli* mutants [4, 118], and mariner-mediated mutants of *Mycobacterium bovis* [89] and *M. tuberculosis* [90]. Putative transcriptional regulators as well as biosynthetic genes could be identified as contributing genes for normal cell growth [4]. Thus, the detection and identification of conditionally essential genes will provide a new antimicrobial target in specific environments. In addition, microarray-combined Tn5 transposon
Scanning linker mutagenesis (SLM). Conventional methods have been widely utilized as a genetic tool to create in-frame insertion mutants, and they include the insertion of oligonucleotide linkers and site-directed mutagenesis, which require considerable efforts to construct many insertion mutants [69]. Scanning linker mutagenesis is a useful technology in which short peptides within transposons are randomly inserted at in-frame positions of a target protein, followed by the removal of unnecessary DNA segments such as an antibiotic resistance gene, thus leaving an in-frame insertion of short peptides only at the target gene. The bulk of the transposon is deleted by two different known methods, which include a restriction enzyme-mediated cleavage and re-ligation, and a site-specific recombination such as Cre-loxP, FLP-FRT (FRT; Flp recombinase target site) and TnpIR-S (IRS; internal resolution site) [43, 88]. SLM strategies vary depending on the peptide sizes, transposons, and deletions of the transposon bulk. Transposable elements including Tn552, Tn3, Tn5, mariner, and IS21 derivatives generate the insertion of differently sized peptides into target proteins, ranging from 4 to 93 amino acids [43]. As an example of the use of the Tn5 transposon, the insertion of the 31 bp in-frame sequence tags harboring protease cleavage sites was utilized for the analysis of the membrane topology of the bacterial membrane-associated proteins in E. coli by which proteins on the periplasmic or cytoplasmic side are far more resistant to proteases than the ones on the membrane-positioned side [69].

Pentapeptide scanning mutagenesis (PSM) is an SLM strategy that involves a 15 bp linker in-frame insertion using a Tn3 derivative, the Tn4430 transposon originating from Bacillus thuringiensis, which was applied successfully both to B. thuringiensis and E. coli [44, 45, 88]. This approach is a powerful tool for structure-function analysis via the dissection of protein functions, such as the interactions between proteins and their transcriptional regulators, or protein–protein interactions of multiprotein subunits. PSM was applied in the assembly of transport proteins for the production of virulence factors. As a few examples, P. aeruginosa XcpZ was identified to be critical in a type II secretion system because domains I, II, and III of XcpZ are involved in the stabilization of other subunits by their interactions [83]. Additionally, the use of PSM using the Tn7 transposon revealed that the plant pathogenic bacterium Xanthomonas campestris necessarily utilizes the conserved C-terminal region of the major pilus component HrpE to oligomerize pilus subunits for a proper protein transport system, which is known to be a type III secretion (TTS) system that confers bacterial pathogenicity in a plant host [115]. In addition, Tn7-based PSM technology has demonstrated that SadB is required for irreversible attachment upon biofilm formation in P. aeruginosa PA14 [12]. In addition to the use of Tn7, a mini-Mu-based PSM tool was used to deduce that the C-terminal sequence of the HrpA protein in P. syringae involved in the type III secretion pathway was an important subunit for pilus assembly [108].

In conclusion, a gene of interest along with a transposon could be integrated into bacterial chromosomes in either a random or site-specific fashion. Thus, although tools for the transposon-mediated genetic manipulation of certain bacteria are either scarce or non-existent, or exhibit a narrow host range, it is a powerful technique for the integration or excision of a gene of interest at a single-copy level on the chromosomal level, which makes it possible to conduct a variety of experiments, including insertional random mutagenesis, gene expression analysis, protein functional studies, or the gene-tagging of bacteria in living organisms.

Acknowledgment

This paper was supported by Wonkwang University in 2008.

References


