Published in final edited form as:

Chem Rev. 2016 October 26; 116(20): 12758–12784. doi:10.1021/acs.chemrev.6b00003.

DNA Transposition at Work

Alison B. Hickman and Fred Dyda*

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, United States

Abstract

DNA transposons are defined segments of DNA that are able to move from one genomic location to another. Movement is facilitated by one or more proteins, called the transposase, typically encoded by the mobile element itself. Here, we first provide an overview of the classification of such mobile elements in a variety of organisms. From a mechanistic perspective, we have focused on one particular group of DNA transposons that encode a transposase with a DD(E/D) catalytic domain that is topologically similar to RNase H. For these, a number of three-dimensional structures of transpososomes (transposase–nucleic acid complexes) are available, and we use these to describe the basics of their mechanisms. The DD(E/D) group, in addition to being the largest and most common among all DNA transposases, is the one whose members have been used for a wide variety of genomic applications. Therefore, a second focus of the article is to provide a nonexhaustive overview of transposon applications. Although several non-transposon-based approaches to site-directed genome modifications have emerged in the past decade, transposon-based applications are highly relevant when integration specificity is not sought. In fact, for many applications, the almost-perfect randomness and high frequency of integration make transposon-based approaches indispensable.

Graphical Abstract

*Corresponding Author: fred.dyda@nih.gov.

Notes

The authors declare no competing financial interest.

1. INTRODUCTION

DNA transposition is one of several enzyme-catalyzed mechanisms by which a discrete segment of DNA is moved from one location in a genome to another. It differs from other pathways for the movement of mobile genetic elements in that all of the intermediates of the reaction are DNA molecules and there is no requirement for extensive homology between DNA sequences at the ends of the mobilized DNA segment (the “transposon”) and the site into which it is moved (the “target” site). In these features, DNA transposons differ from retroelements, which are reverse-transcribed prior to insertion in a target site, and site-specific recombination systems.

DNA transposons have been identified in essentially all genomes that have been sequenced. In prokaryotes and other lower organisms, their movement causes ongoing genomic rearrangement, whereas in higher organisms activity is often severely down-regulated or disabled through mutation. One consequence of transposon inactivation by mutation is that many genomes are riddled with remnants of previously active DNA transposons. Mobile elements are often described as “selfish DNA”, reflecting one perspective that they are self-perpetuating, self-serving systems. However, it is clear that the movement of DNA transposons can cause genomic changes that have functional consequences in their host.^{1–3} The extent to which they may offer an evolutionary advantage to the organisms in whose genomes they reside is an area of active interest.^{4–10}

The transposase enzymes that move DNA transposons are typically multidomain proteins that include a catalytic nuclease domain capable of cleaving one or both DNA strands and subsequently joining the broken DNA end to another DNA strand. Autonomous DNA transposons are those that encode their own transposase (i.e., the transposase gene lies within the transposon itself). Short, truncated transposons can also be mobilized by a transposase encoded elsewhere in the genome as mobility generally only requires two

specific sequences on a single DNA molecule that a transposase can recognize as its ends. Thus, nonautonomous DNA transposons (also called “MITes” for Miniature Inverted-repeat Transposable Elements) have proliferated in many eukaryotic genomes and often far outnumber their autonomous versions.¹¹ This capacity of a transposase to act “in trans” is the basis of many important transposon applications.

The simplest DNA autonomous transposons contain only one open reading frame (ORF) encoding the transposase, and consist solely of the transposase gene flanked by two “ends” that are the binding sites for the transposase. In prokaryotes, such transposons are called Insertion Sequences (ISs). However, transposons can contain multiple genes. Sometimes these are required for mobility and others may just be along for the ride. These latter genes are known as passenger genes, and are particularly important in the spread of antibiotic resistance genes among bacterial populations.

In this review, we discuss what is known about the different ways in which DNA transposons are mobilized by their associated transposase enzymes. These biochemical pathways are placed in the context of the structural information available for DNA transposases and the complexes they form with DNA (“transpososomes”). We will focus specifically on one class of transposases, the so-called DD(E/D) transposases, as they are featured in the second half of the review where we will provide an overview of the wide range of cellular applications for transposons.

It is worth noting that the most recent edition of the textbook *Mobile DNA III*¹² was published in 2015, and individual chapters on many aspects of mobile genetic elements are available through their associated online periodical, *Microbiology Spectrum*. Thus, the reader interested in details on specific types or superfamilies of DNA transposons will find more in-depth reviews in these chapters than we are able to present here. Beyond brief introductions, we have elected to highlight only those contributions to the field that have been published in the past five years or so, with an emphasis on referencing more focused reviews. We also refer the reader to excellent earlier reviews of DNA transposition by Curcio and Derbyshire¹³ and Montañó and Rice.¹⁴

2. OVERVIEW OF DNA TRANSPOSITION: TERMINOLOGY AND BASIC CONCEPTS

DNA transposons were discovered by Barbara McClintock, who was studying mutable genetic loci in maize.¹⁵ The mutations she observed were the consequences of movement of DNA transposons known as *Ac* (for “Activator”) and *Ds* (for “Dissociation”), *Ac* being the autonomous version and *Ds* the nonautonomous forms. *Ac* and *Ds* are members of the *hAT* superfamily (as it is known today) of DNA transposons.

Ac and *Ds* are mobile elements that are moved by being cut out from one site and then inserted into another; they are so-called “cut-and-paste” elements (Figure 1). DNA transposons that move by this pathway have been the most extensively studied to date, both in vivo and in vitro, and are thus the overwhelming focus of this review. However, there are other ways that transposases catalyze the movement of DNA transposons. Two other

characterized pathways require extensive DNA synthesis. One is called “copy-and-paste” or “copy-in” transposition, in which the original version of the transposon remains in place and is not excised but is copied into a new location, as shown in Figure 2. In the other case, “copy-out–paste-in” transposition, the transposon is copied out of its original site and the excised circular dsDNA transposon is then inserted into a new location (Figure 3). Again, one strand of the original transposon remains in its original location.

The cut-and-paste pathway involves introducing double strand breaks (DSBs) at the transposon ends, physically liberating it from its “donor” site, and then inserting it into a new target site. The cutting step is sometimes called “excision” and involves nucleophilic reactions to break each DNA strand. For the first DNA strand break at each end, water is activated as the nucleophile, but the mechanism by which the second strand is broken varies. Depending on the particular transposase (section 3.1), the nucleophile can be water, the 3'-OH of the opposite strand generated during first strand cleavage (in which case a hairpin is formed), and there are examples where a separate enzyme is used to cleave the second strand. Some DNA transposons are excised in linear form; others are excised as circles.

The “pasting” step of cut-and-paste transposition is sometimes called “integration” and again involves nucleophilic reactions, but this time in the context of transesterification. This is also referred to as the “strand transfer” or “joining” step that is the essence of DNA movement; i.e., the connectivity of DNA strands is changed. Most cut-and-paste transposition systems generate target site duplications (TSDs) upon repair of the DNA strand breaks that are introduced by integration (discussed in section 3.2.1). In contrast to the site specificity of their transposon end binding, most transposases are relatively nonspecific in terms of the DNA sequence into which they insert (but see section 5.3).

Transposases generally only catalyze reactions at the ends of their own transposons, as it would be detrimental to their host genomes to have uncontrolled and random DNA breaking and joining. They are thus site-specific endonucleases, recognizing specific sequences that correspond to the two ends of the transposon. By convention, the two ends recognized by the transposase are referred to as the Left End (LE) and the Right End (RE), where the LE is 5' to the direction of transcription of the transposase gene. Some transposases recognize short sequences at each end that are variously referred to as Inverted Repeats (IRs), Terminal Inverted Repeats (TIRs), or Inverted Terminal Repeats (ITRs). For many transposons, the LE and RE sequences are similar but not perfectly identical. Some transposons have far more complicated end sequences that may contain more than one binding site for the transposase and even different numbers of these on each end. In some cases, additional specific sequences are also required that can be hundreds of base pairs interior to the ends (sections 5.4.1 and 5.4.2).

2.1. Classification of DNA Transposon Families and Superfamilies

As new mobile genetic elements are discovered and their transposition pathways worked out, it has become clear that the classification system established when mobile elements were first encountered is too coarse to accommodate the growing menagerie. Although the initial classification^{16,17} into “class I” elements (those that involve RNA intermediates) and “class II” (those that use DNA-only intermediates) is still valid, it is not particularly

informative. Rather, one approach to understanding and classifying mobile elements has been according to their mechanism of strand breakage and integration,¹³ as introduced in the previous section.

An alternate approach to classification is based on the type of the nuclease domain in the transposase. This is usually obvious from the amino acid sequence of the transposase as it appears that there are only a limited number of nuclease domain topologies, or folds, that are used for DNA transposition. This classification system divides DNA transposons into four main classes: DD(E/D) transposons, Y1- and Y2-transposons (reviewed by Chandler et al.¹⁸), serine transposons, and tyrosine transposons. All cut-and-paste transposons studied to date are of the DD(E/D) variety, which refers to three acidic residues in the nuclease active site that coordinate two metal ion cofactors where the third residue is most often glutamic acid, but is sometimes aspartic acid. The fold of a DD(E/D) transposase nuclease is always an RNase H-like fold,¹⁹ although it may contain a variety of insertions relative to the standard or minimal RNase H-like fold.²⁰ The other classes are named according to a conserved tyrosine or serine residue in the active site that serves as the nucleophile for strand cleavage instead of a water molecule; these transposition pathways proceed using covalent intermediates. There are three different nuclease domain folds, corresponding to the tyrosine recombinase fold for tyrosine transposons, the HUH nuclease fold for the Y1- and Y2-transposons, and the serine recombinase fold for serine transposons. Schematic representations of the active sites corresponding to the four classes are shown in Figure 4, and the DD(E/D) and Y1/Y2 transposons that are discussed in this review are listed in Table 1

The conceptual division between prokaryotic and eukaryotic elements that has led to separate approaches to how DNA transposons are named is somewhat artificial given their evident structural and mechanistic similarities.^{21–23} Thus, what might be recognized within eukaryotes as a DNA transposon is called an IS when encountered within a prokaryote. Eukaryotic DNA transposons often have descriptive names linked to the concept of movement—originating in several world languages—such as *piggyBac*, *Transib*, *Ping*, and *Pong*, whereas prokaryotic ISs were systematized early on²⁴ and are instead assigned numbers. This classification and naming inconsistency has been recognized, and efforts have been underway to attempt to establish a universal transposable element classification system.²⁵

The main specialized databases archiving sequence information for DNA transposons include ISfinder²⁴ for ISs and Repbase²⁶ for eukaryotic mobile elements. Unfortunately, most prokaryotic transposons are unaccounted for in these databases.

2.1.1. Overview of Prokaryotic Insertion Sequences and DNA Transposons.

—Historically, the most studied transposons have been Tn5, Tn10, Tn3, and Tn7 (Figure 5), all originally discovered as active mobile elements responsible for the transmission of antibiotic resistance genes in bacteria.^{27–32} Each differs from the others in ways that have revealed important information about transposition pathways, and accumulated information about the mechanisms of transposition of these four is the foundation upon which the field of DNA transposition is built. Two related systems have been intensively studied from the

perspective of understanding the biochemical mechanisms of DNA cleavage and strand transfer, and they have provided invaluable insight into how DNA transposons are mobilized. Although not a DNA transposon by the strictest definition, bacteriophage Mu³³ inserts its genome into its host cell chromosome to establish an infection using an enzyme, MuA, that has a DD(E/D) nuclease domain.³⁴ Similarly, retroviral genomes are integrated into host chromosomes using integrase enzymes that also have DD(E/D) nuclease domains.^{19,35}

As shown in Figure 5, Tn5 and Tn10 are composite transposons, meaning that they are composed of two copies of the same IS (IS50 and IS10, respectively) that bracket several ORFs. Each carries antibiotic resistance passenger genes and moves by a cut-and-paste mechanism.^{36,37} Tn3 from *E. coli* is the archetypal member of a family of transposons that move by replicative transposition (“copy-in” mechanism); it carries a *bla*TEM gene that encodes a β -lactamase.³⁸ Tn7 similarly has antibiotic resistance passenger genes but, in addition to the two genes for its heteromeric transposase (see section 5.3), encodes other genes that are needed for transposition.³⁹ It is also a cut- and-paste transposon.

Prokaryotic ISs are currently classified into ~30 families based on the sequence similarities of their associated transposases and DNA sequences at their ends,²³ and a curated database called ISfinder (<https://www-is.biotoul.fr/>) is available that contains over 4000 IS sequences. The vast majority of these are known, expected, or predicted to have transposases with DD(E/D) nuclease folds.^{22,23} Another class of DNA transposon present in prokaryotes are the Y1- and Y2-transposons, represented by the IS200/IS605 and IS91 families, respectively. In both cases, the catalytic domain is an HUH endonuclease (Figure 4B).^{40,41,18} In the IS200/IS605 case, transposition proceeds through ssDNA intermediates covalently bound to the transposase through 5'-phosphotyrosine linkages.¹⁸ For IS91, experimental information is limited although models for transposition have been proposed.⁴² A schematic of the mechanism used by IS200/IS605 transposons is shown in Figure 6.

Serine and tyrosine transposons are relatively small families of mobile elements. For serine transposons, the predicted resemblance of their transposase catalytic domains to those to serine site-specific recombinases⁴³ such as $\gamma\delta$ resolvase suggests that their transposition mechanisms are likely similar (Figure 4C).⁴⁴ This involves the excision of a double-stranded circular transposon intermediate although there is very little experimental evidence to date. More information is available for a few tyrosine transposons such as CTnDOT and Tn916, which also transpose using an excised circular transposon,⁴⁵ in this case using a mechanism that resembles that used by tyrosine recombinases such as Flp, Cre, and λ integrase (Figure 4D).⁴³

2.1.2. Overview of Eukaryotic DNA Transposons.—There are ~18 superfamilies of DNA transposons, grouped (as with prokaryotic elements) according to similarities in the sequences of their associated transposases and transposon ends.^{21,25,26} Again, the dominant type of transposase catalytic domain is the DD(E/D) nuclease domain;^{22,46} only the *Helitrons* differ with a Y2-type transposase.^{47–49}

A recent analysis⁵⁰ of 23 sequenced vertebrate genomes quantified the wide diversity in not only the proportion of the genome contributed by remnants of mobile elements but also superfamily distribution within individual species. Within this limited data set, the lineage with the greatest mobile element diversity is the teleost fishes, and the most widely distributed superfamilies of DNA transposons are the *Tc1/mariner*, *hAT*, *piggyBac*, and *Helitron* elements.

2.2. Biological Role of DNA Transposition

With the ability to break DNA strands and to then insert more-or-less randomly into another segment of DNA comes the potential to dramatically disrupt genomes. Clearly, this can lead to detrimental effects such as inactivating crucial genes or their promoters, or destroying protein binding sites. On the other hand, disrupting existing cellular pathways—or creating new ones—also has the potential to bring beneficial changes to an organism. Evolutionarily speaking, how the overall balance is tipped—or has been tipped—in different organisms under differing growth conditions is one of the most intriguing aspects of DNA transposition. However, it should be noted that many organisms have evolved mechanisms to suppress the activity of transposons that reside within them,^{51,52} which perhaps partially addresses the question of the balance between beneficial and harmful. In eukaryotes, this down-regulation or “silencing” can be achieved by chromatin modifications such as DNA methylation, histone modifications, or using RNAi pathways.^{51,53–55}

DNA transposition is of clear benefit to bacteria exposed to antibiotics as it provides a way to carry, move, and transmit genes for antibiotic resistance.^{56,57} Hand-in-hand with DNA conjugation, this has led to the emergence of multidrug resistant human pathogens that have become major threats to public health, especially through hospital-acquired multiantibiotic resistant infections. Although associations between specific antibiotic resistance genes and particular DNA transposons have been well-documented,⁵⁷ it is the advent of relatively inexpensive genome sequencing that has only recently allowed in-depth and mechanistic studies of the clinical impact of transposition⁵⁸ to be undertaken.

Integrating bacteriophages such as Mu depend on DNA transposition for their life cycle. In the initial infection or lysogenic phase, upon injection into an *Escherichia coli* cell, the linear Mu genome integrates into the host genome. In Mu's subsequent lytic phase, it moves by intramolecular replicative transposition from one site to another in the genome in order to generate multiple copies of itself. These are excised and packaged into capsids to generate new bacteriophages that are released upon cell lysis to restart the infection cycle.

The beneficial roles of DNA transposons are clear in cases of exaptation, which is the repurposing of the transposon DNA or the transposase enzyme itself to a different function in the cell. Perhaps the best-known example is the RAG1 protein of the vertebrate V(D)J recombination system that originated from an ancient *Transib* transposase.^{59–61} In its new role as part of the adaptive immune system, the RAG1 protein (in combination with other proteins) has maintained its ability to move specific pieces of DNA but does so now in the exquisitely controlled context of rearranging gene segments in order to generate a wide variety of antigen receptors.⁶² Other examples of “domesticated” transposases that have retained some of their catalytic activities include those involved in yeast mating type

switching⁶³ and macro/micronucleus transformations in ciliates.^{64–66} Transposases have also been an important source of DNA-binding and chromatin-associated domains.^{67–70}

The examples of RAG1 and the domesticated *piggyBac* and *Tc1/mariner* transposases in ciliates are experimentally validated examples of former DNA transposases performing new roles. The availability of an ever-increasing number of genome sequences has led to proposals of links between transposons and biological systems that have solid bioinformatic backing but currently lack experimental support. For example, it has recently been proposed⁷¹ that a small group of putative DNA transposons known as “casposons” may be the evolutionary ancestors of the Cas1 proteins of the CRISPR/Cas system found in many bacteria and archaea. It has not yet been confirmed that casposons are, in fact, mobile elements, although it has recently been shown that an archaeal casposase can catalyze the integration of an excised casposon into DNA.⁷²

Similarly, it has been suggested that Polintons, one of the eukaryotic DNA transposon superfamilies, lie at the evolutionary origin of several virus families including the nucleocytoplasmic large DNA viruses (NCLDV)s⁷³ and other eukaryotic DNA viruses such as the Adenoviridae and the Bidnaviridae.^{74,75} Although an intriguing proposal, it is not yet clear if Polintons are indeed transposons: most encode proteins homologous to viral capsid proteins, suggesting that they might instead be viruses or hybrid mobile genetic elements lurking at the boundary between viruses and transposons.⁷⁶

3. CHEMISTRY OF DD(E/D) TRANSPOSASES: DNA CLEAVAGE AND STRAND TRANSFER

Those DD(E/D) transposition systems that have been studied genetically or biochemically have revealed a surprising diversity of mechanisms for accomplishing the task of either moving or copying a transposon DNA sequence into a new location. However, for many DNA transposons, aspects of mechanism have only been deduced from genome analysis such as the detection of “scars” of previous transposition events (i.e., repaired empty sites from which a transposon has moved) or target site duplications. To date, there remain entire prokaryotic and eukaryotic transposon superfamilies for which there is no experimental data regarding their transposition mechanisms.

3.1. How DNA Strands Are Cleaved

As shown in Figure 7, different DD(E/D) transposases have adopted different pathways to liberate their transposon ends from a donor site. However, they all begin with the nucleophilic attack of an activated water molecule at or close to the transposon end.

The chemistry of the initial DNA cleavage reactions⁷⁷ is well-understood in the context of the two metal ion mechanism of phosphodiester bond hydrolysis.⁷⁸ In this type of reaction, protein residues in the enzyme active site coordinate two divalent metal ions (either Mg^{2+} , the physiological ion, or Mn^{2+}) and bind the DNA substrate to correctly orient an attacking water molecule for a nucleophilic S_N2 in-line attack on the phosphorus atom of the scissile phosphate. The reaction proceeds through a pentacovalent intermediate and leads to inversion of configuration at the phosphorus.^{79–81} The phosphodiester bond is broken such

that the products are free DNA ends, one with a 3'-OH group and one with a 5'-phosphate as shown in Figure 4A. Insightful crystallographic⁸² and quantum mechanics/molecular mechanics (QM/MM) simulations studies⁸³ have been carried out with the related enzyme RNase H, and a series of X-ray crystal structures are available representing various stages of the integration pathway of PFV (prototype foamy virus) integrase,^{84–86} closely related to transposases.

DNA cleavage is directed to the appropriate phosphodiester bonds at transposon ends by the specific sequences at the very tips of the transposons that the transposase binds. These, however, are not necessarily the same things as the dominant binding sites are generally just subterminal to the transposon tips. The DNA sequence that flanks the transposon (effectively the target site of the last integration) can also affect the efficiency of cleavage;⁸⁷ this seems to be particularly true for those transposases that integrate into specific sequences (such as the Tc *I/mariners* that insert into TA or *piggyBac* that is targeted to TTAA sequences); the sequences of the TSDs that are generated upon integration are often crucial for subsequent excision.^{88–90}

For DD(E/D) transposases, strand cleavage at each transposon end results in one of the strands with a free 3'-OH group. This is the “bottom strand” at the LE, and the “top strand” at the RE. These strands are designated the “transferred strand” (TS) as the resulting 3'-OH is subsequently the attacking group during integration whereas the other strand is the “nontransferred strand” (NTS) as it does not become connected to the target site strands. Cut-and-paste transposases cleave both the TS and NTS to generate DSBs, while copy-and-paste transposases cleave only the TS. When the transposase generates a DSB, these are typically not blunt; rather, short overhangs (conserved in size for a given transposase) are usually present on either the TS or the NTS. However, cleavage of the TS is always precise, occurring at the phosphodiester bond that represents the end of the transposon, apparently to ensure that genetic information is neither lost nor added to the mobile element during excision (Figure 7A).

Whether the initial cleavage reaction occurs on the TS or the NTS at the transposon end has mechanistic consequences when the liberated 3'-OH group is subsequently used as the nucleophile for cleavage of the second strand. These mechanistically diverging pathways are discussed in sections 3.1.1, 3.1.2, and 3.1.3.

3.1.1. Hairpins and Circles.—How the excision reaction then proceeds depends on the particular transposition system. For cut-and-paste transposons such as Tn5, Tn10, and *piggyBac*, the TS 3'-OH group is used in the second step of the reaction to attack the top strand, resulting in a hairpin on the transposon end (Figure 4A).^{91,36,89} A variation is observed for copy-out–paste-in transposons such as IS911, a member of the large IS3 family, in that the initial 3'-OH formed at one transposon end is used as the nucleophile to cleave the same DNA strand at the *opposite* end of the transposon (Figure 3).⁹² The outcome is that one strand of the transposon forms a covalently closed transposon circle in which the two transposon ends are joined by a ssDNA bridge (also called a “figure 8” intermediate).

If, in contrast, the first cleavage is on the NTS and the resulting 3'-OH group is used to cleave the second strand, then a hairpin is formed on flanking DNA (Figure 7B). This pathway is followed by several superfamilies of eukaryotic transposons such as the *hAT*⁹³ and *Transib*⁶⁰ transposons, and the related RAG1 recombinase,⁶² but has not yet been demonstrated for any prokaryotic transposon.

Hairpins on transposon ends and covalently closed transposon circles must be opened (i.e., a phosphodiester bond must again be broken) if the transposition reaction is to proceed. In the systems that have been studied, this occurs by another transposase-catalyzed nucleophilic attack by a water molecule. In contrast, hairpins on flanking DNA are not opened by the transposase but are dealt with by the DNA repair systems of the host cell.

3.1.2. Pathways Involving Two DNA Strand Cleavage Reactions with Water as the Nucleophile.—

The Tc */mariner* transposons (and, most likely, their prokaryotic cousins from the IS630 family⁹⁴) use another mechanism for generating DSBs: two water-mediated nucleophilic attacks cleave the two DNA strands at each end of the transposon (Figure 7C).⁹⁵ The cleavage site on the NTS is recessed two or three nucleotides within the transposon end and the cleavage site on the TS is precisely at the transposon end.⁹⁶ Recent experimental efforts have provided important insights into the order and regulation of these steps.^{90,97} The eukaryotic P element also generates DSBs at its transposon ends but with a very large staggered offset of 17 nucleotides into the transposon on the NTS.⁹⁸ How this large stagger is achieved is an intriguing outstanding question.

The multimeric Tn 7 transposition system³⁹ uses a completely different approach to generate DSBs at its transposon ends. In this case, two separate proteins cleave each of the strands: ^{99,100} TnsA, a type II restriction enzyme-like protein,¹⁰¹ cleaves the NTS while TnsB, its DDE partner, cleaves the TS. A most elegant experiment was the demonstration that mutating TnsA to prevent top strand cleavage converts Tn 7 from a cut-and-paste transposon into a replicative transposon.¹⁰² Other heteromeric transposases have recently been identified in other Tn 7-like mobile elements.¹⁰³

3.1.3. Pathways Where Only One Strand on Each End Is Cleaved.—

Finally, for replicative transposons such as Tn 3, after the initial cleavage reaction on the bottom strand, the 3'-OH group is used directly for the strand transfer step (see section 3.2.2). This pathway is also used by bacteriophage Mu,¹⁰⁴ and appears restricted to prokaryotes.

3.2. How Transposon Ends Are Inserted into a New DNA Site

3.2.1. Cut-and-Paste Transposases.—

For cut-and-paste DD(E/D) transposases, once transposon ends have been released from their original donor site, the 3'-OH groups on each end serve as the nucleophiles for transesterification reactions at a new site. This is true as well for replicative DD(E/D) transposases, but in addition they drag along the DNA from their original site.

As shown in Figure 8, strand transfer into a new target site involves the coordinated attack of both 3'-OH groups on opposite strands. The reactions generally occur several base pairs apart. The base pair distance is a characteristic of each transposition system, and generally

ranges from 2 to 12 bp.^{1,105} At the molecular level, catalysis of strand transfer uses the same active sites as for the initial step of DNA cleavage at the transposon ends, but now the 3'-OH group at the end of the TS is the activated nucleophile that attacks a phosphodiester bond of the target DNA. A series of structures for the related PFV integrase enzyme have illuminated how the reactions likely proceed.^{84–86}

As transesterification involves the exchange of DNA strand connectivity, strand transfer by a transposon means that breaks are generated in the target DNA strands concomitant with the formation of a new phosphodiester bond. The breaks must be repaired by the host cell to fill in the gaps introduced by the strand transfer offset (Figure 5B). This process generates target site duplications (TSDs)—sometimes known as direct repeats (DR)—on each side of the integrated transposon, and for essentially all DD(E/D) transposons, these are transposition-accompanying modifications to the genome of the host cell. This can be a very useful property as these TSDs are a record of movement, and they allow transposon mobility (or past mobility) to be tracked. In addition to gap repair, any overhanging nucleotides originating from staggered cleavage at transposon ends must be removed.

One notable feature of the DD(E/D) active site arrangement is its pseudosymmetry. This means that the attacking nucleophile (either OH⁻ from water or 3'-OH of a DNA strand) can be on either side (Figure 4A) of the two metal ions, while the scissile phosphate is always central. It is possible that this configurational versatility is exactly why so many transposases and recombinases have DD(E/D) active sites. For many DD(E/D) transposases, the arrangement allows the 3'-OH leaving group created at the initial TS cleavage to stay in the active site, and to act as a nucleophile again as different scissile phosphates come and go, on the way to completing integration.⁸⁰

3.2.2. Replicative Transposases.—Strand transfer catalyzed by prokaryotic replicative transposases also generates breaks in target DNA strands and TSDs, but because of the tagging-along DNA from the original donor site, the reaction intermediate has a more intricate structure (Figure 2). The initial strand transfer product is known as a Shapiro intermediate,¹⁰⁶ and can be viewed as two replication forks. Indeed, replisome assembly^{107,108} and extensive DNA synthesis are required to complete transposition (indicated by the dashed green arrows in Figure 2, inset).

Replicative transposition, in contrast to cut-and-paste transposition, has different outcomes depending on whether the target site is on the same or a different DNA molecule as the transposon donor. If replicative transposition occurs between two different replicons in the cell (i.e., the bacterial chromosome or plasmids), the replicated product in which the two DNA molecules are now fused is known as a cointegrate (Figure 2). Cointegrates can be resolved—separated back into two molecules—by either homologous recombination that splits the cointegrate into two replicons, each with one copy of the transposon, or by a transposon-encoded site-specific resolution system (as with Tn3, Figure 3C). If transposition occurs within a single replicon, replicative transposition has two possible outcomes, depending on the orientation of the attack on target DNA.¹⁰⁹ In one orientation, the result is an inversion of the DNA segment between the two copies of the transposon; in the other, the

segment of DNA between the two copies is deleted. These dramatic genomic rearrangements contribute to the rapid evolution of bacteria, and particularly of the plasmids that they carry.

3.3. Repair of the Donor Site

When a cut-and-paste transposon is excised from the donor site, it leaves a DSB behind (Figure 1). How these are repaired depends on the transposition mechanism, the point in the cell cycle where transposition occurred, and the host organism.^{110,111} At one end of the spectrum are *piggyBac* (section 5.4.2; Figure 9) and a few other transposons such as *mPing*,¹¹² whose particular mechanism of excision allows for seamless repair without DNA synthesis: they leave no trace. Other transposons can be precisely excised and leave behind two copies of the original target site as a footprint; in other cases, excision repair is accompanied by genome modifications that are not always minimal, such as deletions of the flanking sequence or insertion of filler sequence.¹¹¹

In eukaryotes, multiple repair pathways appear able to repair gaps introduced by transposon excision.^{113,114} Gaps can be repaired using a homologous chromosome or a sister chromatid as the template; alternatively, proteins involved in the nonhomologous end-joining (NHEJ) pathway can also play a role.^{115,116} For example, Ku70/80, the protein complex that binds to free DNA ends and provides a platform for the binding of other proteins needed to ligate DNA ends together,¹¹⁷ has been repeatedly implicated.^{116,118} Ku70/80 has also been shown to be involved in the related pathways of V(D)J recombination¹¹⁹ and PiggyMac-mediated¹²⁰ genome rearrangements. For those DNA transposition systems that form flanking hairpins during excision, it seems likely that the Artemis complex¹²¹ will prove to be important.

DNA repair following DNA transposition in prokaryotes has been less intensively studied, but it seems clear that the host cell homologous recombination machinery and the RecBCD complex are involved.^{122–126}

4. INSIGHT INTO DNA TRANSPOSITION MECHANISMS FROM STRUCTURE: OVERALL ARCHITECTURES

The first structural insights into DD(E/D) transposases were provided by crystal structures of the catalytic cores of their related cousin enzymes, retroviral integrases¹⁹ and bacteriophage MuA.³⁴ More recently, far more informative structures of DNA transposases and retroviral integrases bound to their substrates have been emerging, providing intricate and substantive insights into the workings and regulation of transposition. To date, the three-dimensional structures of six DNA transposases have been determined bound to DNA. Four of these correspond to DD(E/D) transposases (those of the *Tn5*, *Mos1*, *afnd* *Hermes* transposons and of bacteriophage Mu; see sections 4.1, 4.2, 4.3, and 4.4) and two are of Y1-transposases^{127–129} from the *IS200/IS605* family. A series of elegant structures have also been determined for PFV integrase bound to its viral ends, to a target site, and after strand transfer.^{84–86} The DNA-bound structure of the accessory protein encoded by the *IS5376* transposon (an *IS21* family member) responsible for target DNA capture has also been determined.¹³⁰

4.1. The Tn5 Transpososome

The Tn5 transposase bound to its cleaved transposon ends is a compact, interwoven protein–DNA complex (Figure 10). Although the protein on its own is a monomer, dimerization occurs upon binding two transposon ends, and the resulting protein–DNA complex has a 2:2 stoichiometry.³⁶ The transpososome structures represent the stage in the transposition reaction where the transposon has been excised from donor DNA, the hairpins on the two ends have been opened, and the complex is ready to bind target DNA.^{131–133}

An important revelation of the structure is that there are protein–DNA contacts along the entire length of the 19-bp transposon inverted repeats (IRs), which are sufficient¹³⁴ for *in vivo* transposition. Furthermore, for a given end, both protein monomers contribute to its binding. Specifically, the tip of each transposon end is bound and processed by the catalytic domain of one monomer (orange in Figure 10), and domains of the transposase that confer site-specific recognition and binding (shown in blue) are contributed by the other monomer. This is called “recognition *in trans*”, and ensures that two transposon ends are bound by a dimeric transposase. Thus, the transposase cannot cleave a single end on its own: only when both ends are located and the dimer is formed will the catalytic domains be appropriately positioned for DNA cleavage and subsequent strand transfer. The DNA binding mode observed for the Tn5 transpososome reflects a rule for DNA transposases for which there is yet no known exception: the active forms of DNA transposases are protein–DNA assemblies that are at least dimeric and two subunits are always required to provide two active sites to act on the two transposon ends.

Another mechanistic aspect of the structure is that a set of protein–DNA interactions stabilize a flipped-out base at the very tip of the transposon end. This is how the transposase makes possible the formation of the tight transposon end hairpin that is needed to generate the DSB.

The usefulness of the Tn5 transposon in applications (see section 5) stems in part from its high activity in many different bacterial species and its lack of target specificity, such that it integrates essentially randomly. Also, unlike several other well-understood transposon systems, it does not require a host factor for activity.¹³⁵ For example, transposition by Tn *10*, *Sleeping Beauty*, and bacteriophage MuA requires or is stimulated by DNA bending proteins.^{37,104,136} In some of these cases, the binding sites for bending proteins are within the transposon ends and are needed to bring specific DNA sequences into appropriate alignment within the synaptic complex, whereas for others, DNA bending proteins bind elsewhere¹³⁷ to assist in bringing the transposon ends together.

It has been possible to increase the transposition activity of the Tn5 transposase through mutation^{138–140} as “hyperactive” transposases are often desirable for application purposes (see section 5). Increasing the *in vivo* activity of Tn5 transposition has also been accomplished by modifying the transposon ends.¹⁴¹ The structure of the transpososome provided a way to rationalize the effects of the activating mutations (which were generated at random), yet the explanations¹³¹ make it difficult to imagine how to successfully reverse the process to design activating mutations on the basis of structure.

4.2. The *Mos1* Transpososome: Paradigm for the Tc 1/*mariner* Superfamily

Mos1, an active eukaryotic transposon originally identified in *Drosophila mauritiana*,¹⁴² is representative of the large and widespread Tc 1/*mariner* superfamily of DNA transposons. Its domain organization (Figure 11) is conserved among other members of the superfamily, which include two of the most commonly used transposases for application purposes: *Sleeping Beauty* and *Himar*. As is true for other Tc 1/*mariner* transposons, *Mos1* uses a cut-and-paste mechanism to transpose.^{143,144} The *Mos1* transposon ends are 28-bp imperfect IRs that differ by 4 bp from each other.

The *Mos1* transposase bound its cleaved transposon ends is a 2:2 protein:DNA complex, but the complex has a remarkably different overall architecture¹⁴⁵ from that of the Tn5 transpososome (Figure 12). In part, this stems from a completely different mode of specific DNA recognition, which involves an N-terminal DNA binding domain with two helix-turn-helix (HTH) motifs in each monomer (shown in blue in Figure 12). The HTH domains recognize sequences interior to the transposon ends and are strung out along the length of the IRs, as had been previously shown¹⁴⁶ for the isolated N-terminal domain of the Tc3 transposase. Interestingly, in contrast to Tn5, *Mos1* is already a dimer without bound ends, and its multimerization state does not change upon transposon end binding. Nevertheless, end binding is in trans, where the catalytic domain that processes one end is part of the same polypeptide protomer that has the two end binding HTH domains that bind the other end. Correspondingly, there is no cleavage activity before the synaptic complex containing both IRL and IRR is assembled.

The structure of the *Mos1* complex with authentic cleaved ends^{145,147} (i.e., a staggered DSB in which the transferred strand is cleaved precisely at the transposon end and recessed by 3 bp on the NTS) represents the same stage of the transposition reaction as that of the Tn5 transpososome: bound to its cleaved transposon end and poised to bind target. A recent structure in which the transferred strand is extended into flank¹⁴⁸ (shown in red bases in Figure 12) provides insight into the role of the flanking TA dinucleotide. The TA dinucleotide represents the TSD generated by *mariner*/Tc 1 transposons (underlined in Figure 12), so it has dual significance during transposition: it is not only the obligate target site, but also consequently the flanking sequence from which the transposon excises when it undergoes the next cycle of transposition. The new structure reveals the structural basis of recognition of the T base, and also suggests a model for how target DNA may be bound.

Nonidentical transposon ends often correlate to differences in transposase binding affinity. It has been shown that modifying *Mos1* (with its asymmetric IRL/IRR ends) such that it has two identical IRR/IRR ends increases the frequency of in vitro transposition ~26-fold,¹⁴⁷ correlating to the reported tighter binding of the *Mos1* transposase to its IRR sequence than to its IRL.¹⁴⁹

4.3. The *Hermes* Transpososome

Hermes is an active *hAT* transposon originally discovered in the housefly.¹⁵⁰ Its transposase is the sole structurally characterized member^{151,152} of the large eukaryotic *hAT* superfamily of DNA transposons that share the same overall protein domain organization, and includes

Ac and *Tol2*.¹⁵³ *Tol2* is one of the very few DNA transposons isolated as an active element from a vertebrate species, the medaka fish.^{154,155}

Even in the absence of DNA, the *Hermes* transposase forms an octamer, arranged as a tetramer of tightly intertwined dimers.¹⁵² It is not yet known if other *hAT* transposases are also octameric, or have different oligomeric states corresponding to different arrangements or combinations of dimers. The structure of the *Hermes* transpososome reveals that the octamer is organized as an interlocked ring of four dimers in which each monomer has one cleaved transposon end bound (Figure 13). This is a nonphysiological situation as a transposon only has two ends, but is a consequence of the short oligonucleotide TIR mimics that were used for the crystallography work; comfortably, the octamer only binds two transposon ends if longer DNA oligonucleotides are used.¹⁵² Once again, the cleaved-ends-ready-to-integrate stage of transposition has been captured.

One observed feature of *hAT* transposons is that very long transposon ends (generally 100 bp or longer) are needed for in vivo transposition.⁹³ Although this requirement for long ends is not unique among eukaryotic transposons, *hAT* transposon ends have also been noted to contain multiple copies of short sequence motifs scattered through the ends, in both orientations, without evident periodicity or patterns.⁹³ One model for *Hermes* transposition suggests that these observations are related, and that the site-specific binding domains of *hAT* transposases (N-terminal “BED” domains, for boundary element domain⁶⁷) bind these subterminal repeated motifs.¹⁵² Thus, an octamer would provide a multitude of DNA binding domains, and this is proposed to be important for affinity or end recognition. Although the construct used for the crystallographic study did not include the BED domain, negatively stained electron microscopy showed that the eight BED domains reside in the center of the octameric ring.¹⁵² How this can be reconciled with a model for DNA binding of longer transposon ends remains to be established.

4.4. The MuA Transpososome

The MuA transpososome¹³⁷ (Figure 14) reveals another splayed-out protein–DNA complex, with multiple specific DNA binding domains studding the ~50 bp transposon ends. Similarly to Tn5, the MuA transposase is a monomer without bound DNA, but the active form of the Mu transposase is a tetramer with bound ends, and tetramerization occurs upon end binding.¹⁵⁶ The tetrameric synaptic complex is then able to capture target DNA. In cells, target DNA is delivered by MuB, an AAA+ ATPase encoded by phage Mu, that is also a nonspecific DNA binder.¹⁰⁴ Each end of the phage genome contains three MuA binding sites although their spacing and orientation differ;¹⁰⁴ on the right end, two of the three MuA binding sites—R1 and R2—are incorporated within the active transpososome (Figure 14). One remarkable feature of the complex is the utilization of the four MuA protomers in different roles and contexts in the assembly. Two MuA protomers in the assembly are mainly associated with R1, and these supply the catalytic subunits whose active sites that are engaged in the cleavage and joining reactions, again in trans. The other two MuA protomers are mainly engaged with the R2 sites, and their catalytic domains play architectural roles. The complex is held together through a plethora of protein–protein and protein–DNA interactions, and while these are perfectly symmetrical between the two ends (as

symmetrical ends were used in crystallization), the interactions formed by the two protomers bound to R1 and R2 differ significantly. Thus, the same amino acid sequences are used in different roles within the transpososome (shown on the right in Figure 14).

One of the most remarkable and telling features of the complex is that the target DNA is severely bent (shown in green, Figure 14). The notion that DNA sequences that are distortable or bendable may be preferred sites for integration has been appreciated from studies on other transposases^{157–160} and retroviral integrases.⁸⁵ Target DNA can also be “pre-bent” by another transposon-encoded protein, as has recently been shown for IS5376, in which the accessory protein IstB bends ~50 bp of target DNA through almost a 180° U-turn.¹³⁰ IstB is also a AAA+ ATPase, and it seems likely that MuB similarly bends target DNA for delivery to MuA.¹⁶¹

These collective results suggest that the phenomenon of target DNA bending might be generalizable to all DD(E/D) transposases, and important either to aid access of the transposase active site to the DNA backbone or to ensure that strand transfer is irreversible.^{130,137,162} In the case of the PFV intasome, bent target DNA is also consistent with its ability to target nucleosomes as sites for integration.¹⁶³

5. TRANSPOSONS AS GENETIC TOOLS: A VERSATILE TOOLSET

The potential of transposons as genetic tools was recognized early on.¹⁶⁴ By its nature, DNA transposition is suited for two main types of applications: to carry out random mutagenesis and to introduce specific genes into a genome. DNA transposition is not the only method available to accomplish either of these tasks, but it has many advantages that have become clear as its applications have been merged with modern techniques such as advanced DNA sequencing technologies and stem cell research.

Applications of transposition for genetic manipulation typically involve bicomponent vector systems: one plasmid encodes a transposon derivative consisting of two transposon ends flanking a gene or genes (such as a reporter gene, a selectable marker, or a corrected gene) and any important regulatory sequences to be carried into the cells, and a second plasmid encodes the transposase under the control of a suitable promoter. This approach therefore mimics the *in vivo* transposition of nonautonomous DNA transposons. There are many different ways to deliver the plasmids, such as transfection, electroporation, or injection, or by packaging them within viral capsids. One variation that has been used to bypass the steps of transposase transcription and translation (and to avoid any possible side effects from random plasmid integration) is to express the protein from injected *in vitro* synthesized mRNA.^{165–167} Alternatively, the transposase protein or preassembled transpososomes can be introduced directly into cells.^{168–170}

The usefulness of a given transposon system for genetic manipulation applications often depends on high transpositional activity in the host of interest, and a lack of selectivity regarding the DNA sequence into which it will insert. Rampant transpositional activity is not a property of most transposons,¹⁷¹ as it seems likely that the most successful coexistence strategy for a DNA transposon regarding its host is to “operate below the radar” by not

disrupting DNA too often. As this can be at odds with the need for efficient transgene delivery, there has been a longstanding interest in generating “hyperactive” versions of the most promising transposon systems. In those cases where this has been successful, this has been achieved by point or cluster mutations in the transposase that increase transposition activity and/or by changing sequences at the transposon ends. One property that might limit transposon activity is the need for a species-specific host factor, and indeed some of the most widely used transposons such as *Tn5*,¹³⁵ *piggyBac*,⁸⁹ and *Tc1/mariners*^{144,172} appear to transpose without the need for other proteins. Presumably other factors in play include how well a transposase is expressed and folded in a cellular environment different than the one from which it was originally isolated, whether it is rapidly degraded, and if it is affected by overproduction inhibition (see section 5.4.3).

Although successful transposon mutagenesis in different species sometimes entails closely matching the original transposon host with the target species, this is not obligatory. Thus, *Ac/Ds* has been successfully used to target plant species other than maize,¹⁷³ ranging from poplar trees¹⁷⁴ and strawberries¹⁷⁵ to rice¹⁷⁶ and soybeans,¹⁷⁷ yet it has also been reported to work effectively in medaka fish.¹⁷⁸ Similarly, *Tol2*, a *hAT* transposon originally isolated from the medaka fish,^{154,155} has not only been successfully applied to manipulate the genome of zebra-fish^{179–181} and frogs,^{182,183} but has been shown¹⁸⁴ to be active in a range of mammalian and human cells.

Whether transposon insertion occurs preferentially into distinct genomic locations also impacts the range of applications of a particular transposon system. Target site selectivity can manifest on several different levels: some transposons integrate into specific DNA sequences (for example, *Tc1/mariners* always integrate into TA) whereas others seem to have “preferred” target sequences.¹⁸⁵ Another feature of target sites that has been reported is a preference for “bendable” sites.^{157–160} Indeed, the MuA transposase preferentially and specifically targets DNA containing mismatched base pairs, most likely due to structural deformability of the target DNA.¹⁸⁶ Genomic context can also be important, as some transposons exhibit a preference for insertion into promoter regions or transcriptional start sites, others for genes or transcriptional units.^{187–191} “Local hopping” refers to the observation that a new site of insertion is often quite close to the site from which the transposon was excised, suggesting that the excised transposon does not spend much time loitering around the nucleus before integrating. This phenomenon has been reported for *Sleeping Beauty*, *Tc1*, the *Pelement*, *Tol2*, and *Ac/Ds*.¹¹⁰

In sections 5.1–5.6, we provide a sampling of some of the exciting research areas involving transposon applications. As the literature from even just the past few years is extensive, we have not attempted to be exhaustive or comprehensive, but rather to direct the reader to more specific reviews and just a few original research reports. We hope that these will provide a flavor of the range of transposon applications, and a sense of how a basic understanding of their mechanisms has impacted the development of new and insightful approaches to scientific questions.

5.1. Gene Deletion Libraries

DNA transposons have been used extensively to generate large, random gene deletion libraries, an approach to gene disruption that offers an alternative to chemical mutagenesis. As the sequences of the transposon are known, they can be used as tags for subsequent DNA sequencing to rapidly identify the precise site of insertion (i.e., to establish what gene or intragenic region has been disrupted).

5.1.1. Bacterial Transposon Libraries.—One of the main applications of transposon insertion libraries has been to identify genes important for bacterial pathogenesis under different host or growth conditions. The main experimental approaches and their uses have been abundantly reviewed.^{192–197}

Transposon libraries have been used in different bacterial species to ask questions such as the following: What genes are required for pathogenicity in a host species or are responsible for different pathogenic properties of different serovars?¹⁹⁸ What genes affect virulence?^{199,200} What genes are involved in specific biochemical, metabolic, or behavioral pathways?^{201,202} What genes mediate pathogen–drug interactions?²⁰³ What genes affect how different bacterial strains coexist under varying environmental conditions?²⁰⁴ As transposon insertions are not targeted to genes per se, the same libraries can be used to interrogate the role of regulatory sequences and noncoding DNA regions.^{198,205}

Transposon mutagenesis in bacteria has gone by many different names and acronyms that are meant to aid in distinguishing between variations in the method and subsequent analysis of the results. Therefore, a variety of descriptors have been used to identify conceptually similar techniques that are distinct in detail, such as transposon gene tagging, genetic footprinting, transposon sequencing (Tn-seq),²⁰⁶ insertion sequencing (INseq),²⁰⁷ high-throughput insertion tracking by deep sequencing (HITS),²⁰⁸ transposon-directed insertion site sequencing (TraDIS),²⁰⁹ transposon site hybridization (TraSH),²¹⁰ and so on.

Generally, genome-wide investigations involving transposon libraries involve first using a transposon system to create a large library of random mutant bacterial strains (Figure 15). The library is then grown under the condition of interest (e.g., in vitro in a specific medium or upon introduction into a host organism in which it causes an infection). After recovery of surviving bacteria and purification of genomic DNA followed by DNA sample processing (which includes fragmentation, ligation of sequencing adaptors, and then PCR enrichment of fragments containing transposon/genome junctions), high-throughput DNA sequencing is used to quantitatively compare the mutant library content before and after the experiment. If mutation of a gene or regulatory region has no effect under the experimental conditions, its frequency of detection before and after should remain unchanged. However, if the mutant is lost or its frequency of detection decreased, then it is presumed to be important under the given growth conditions. In other words, if the particular gene is mutated, then the bacterial strain harboring that mutant can no longer grow, or grow as effectively as other mutants. In this way, it is possible to assign a measure of relative fitness to a given mutation under specific growth conditions. It is worth noting that one of the earliest applications of transposon libraries was to establish the minimal set of genes needed for survival of an organism, or its “essential genome”.^{211–213}

Many different transposon systems have been used to generate prokaryotic transposon libraries, but both Tn5 and (perhaps surprisingly) the eukaryotic transposon *Himar1* have proved particularly useful. Tn5 is highly active, is currently commercially available, and has little or no target specificity. *Himar1*, a *mariner* transposon,¹⁴⁴ is often used for a different added benefit: a mutation in its transposon end at a nonconserved position generates a MmeI restriction site that can be exploited to permit high-throughput DNA sequencing with simplified sample processing relative to other methods.²⁰⁷

The extreme effectiveness of Tn5 and *Himar1* in generating large mutant libraries is reflected in reports of extremely dense levels of insertions. For example, libraries have been created in *Salmonella Typhimurium*¹⁹⁸ and *Caulobacter crescentus*²⁰⁵ using Tn5 that have an insertion every ~8 bp, and a *Yersinia pestis*²⁰⁰ library with an average of one *Himar1* insertion every 25 bp, representing ~70% of all TA dinucleotides in the genome. At this level of coverage, it becomes possible to dissect which domains *within* proteins are important. In addition, transposon insertion site selection preferences start to emerge that may be interesting in their own right. For example, the observation that certain TA sites are not used as *Himar1* target sites has been proposed to reflect the effect of bound H-NS, a nucleoid-associated protein, preventing access of the transpososomes to DNA,¹⁹⁸ a suggestion worth future investigation.

5.1.2. Transposon Libraries in Other Organisms.—Although genetic screens in bacteria—with their short generation time—lend themselves to studies based on selection over time, large-scale genetic screens have also been performed in eukaryotic organisms. Often, this has been done for the purpose of generating a bank of mutant strains in model organisms. For example, it has been possible to disrupt ~70% of protein-coding genes in *Drosophila* using a combination of *Pelement*, *piggyBac*, and, more recently, *Minos* insertions.^{187,214}

In plants, the occurrence of naturally active transposons, not only the *hAT* elements such as *Ac/Ds* or *Tam3* but also those of the *Mutator* (*MuDR*) and *CMC* (*CACTA/En/Spm*) super-families, has provided a palette of tools for genetic manipulation. Large-scale screens provide a way to search for interesting phenotypes and desirable growth properties among important crop species, and transposon mutagenesis offers an alternative when other approaches to genetic manipulation (such as those based on bacteria such as *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes* that can introduce exogenous DNA—"T-DNA" for transfer DNA—into the plant genome) are not practical.^{175,177} Transposon mutagenesis offers a potential advantage over T-DNA insertional mutagenesis, as once a "launching site" (an initial transposon insertion) has been established, new mutants can be continually generated as long as there is a source of the transposase. Clearly, this is a phenomenon not limited to plants, but is particularly helpful when it is difficult or labor-intensive to introduce foreign DNA into cells.

5.2. Transposon-Mediated Insertional Mutagenesis for Cancer Research

In one of their main applications, DNA transposons have been used in model animals and cell lines to address a wide variety of questions relating to the genes involved in cancer development.

Pioneering studies in 2005 showed that the *Sleeping Beauty* transposon system could be used to generate cancer-causing mutations in mice.^{215,216} This allowed the identification of putative cancer genes and correlation to the signaling pathways that might be involved. In this early work,²¹⁵ singly transgenic animals were first generated that harbored a genomic copy of either a hyperactive *Sleeping Beauty* transposase (“SB11”)²¹⁷ or an integrated transposon vector. Upon crossing these singly transgenic animals, double-transgenic mice were created, and those that survived to weaning were monitored for cancer development. By sequencing the *Sleeping Beauty* integration sites in the tumors that developed, it was possible to identify genes that were either disrupted by transposon insertion or activated. The analysis of transposon insertion sites also revealed multiple examples where two genes had been mutated, suggesting cooperation and possibly participation in the same signaling pathways. Similar experiments performed in cancer-predisposed mice deficient for the tumor suppressor p19Arf²¹⁶ led to the identification of a set of genes whose disruption potentially drives tumorigenesis in an initially sensitized background.

Since these ground-breaking studies, *Sleeping Beauty* and *piggyBac* transposition has been used in many large-scale insertional mutagenesis studies, and several excellent recent reviews provide an overview of the scope of transposon-mediated insertional mutagenesis studies in the cancer field.^{218–221} There are certainly limitations to transposon-based experiments of this type, among them that integration site selection is never completely unbiased (so some important genes may be missed), that gene disruption may not faithfully mimic the effect of point mutations, and that—as transposons go through cycles of excision and reintegration—genomic lesions can also be left at excision sites that may have unpredictable consequences.

Despite these issues, transposon-mediated insertional mutational mutagenesis studies have provided invaluable insights. For example, these types of studies have led to the discovery and identification of clinically relevant cancer genes and important cell signaling pathways in various cancer types,^{218–221} insights into how tumors develop depending on different sensitizing genetic backgrounds,²²² determining which genes are important to the development of primary tumors versus metastases,^{223,224} and the identification of genes that mediate resistance to various anticancer drugs^{225,226} such as cisplatin.²²⁷ As transposons can disrupt noncoding regions of the genome, it has also been possible to use transposon mutagenesis to identify new noncoding regions that participate in cancer development.²²⁸

5.3. Single-Copy Tagging Using the Tn7 Transposon

In contrast to all other transposon families, the bacterial Tn7 transposon (Figure 5) possesses the remarkable property of possessing two distinct integration pathways.^{39,229,230} Which one is used depends on whether the active transposition complex incorporates the transposon-encoded TnsD or TnsE protein.

The core of the Tn7 transpososome is a complex containing the TnsA, TnsB, and TnsC proteins.²³¹ TnsD is a site-specific DNA binding protein that binds a highly conserved site within the 3' end of the *glmS* gene in the bacterial chromosome and, in complex with TnsABC, directs integration site—and orientation—specifically to the *glmS* transcriptional terminator.^{157,232,233} Collectively, this region is known as *attTn7* and is considered to be a “safe haven”, as integration just downstream of *glmS* is intergenic. In contrast, the TnsE pathway (TnsABC+E) directs Tn7 integration to nonspecific sites within conjugating plasmids²³⁴ or to DNA containing double-strand breaks.²³⁵ TnsE-directed transposition appears to involve protein–protein interactions with the replication complex via the β -clamp.²³⁶

The site-directed TnsABC+D pathway of Tn7 transposition is an appealing way to introduce exogenous genes into a specific and benign location in bacterial chromosomes.²³⁷ Not only is *glmS* highly conserved across many bacterial species, but it appears to be present in essentially all bacterial genomes sequenced to date and most bacteria have only one *attTn7* site. One use of Tn7 transposition has been to generate bacterial strains that express reporter genes such as GFP or luciferase for in vivo studies.^{238–240} By introducing additional *attTn7* sites into a genome, it is also possible to increase and modulate the expression level of a transgene.²⁴¹ Genes encoding antibiotic resistance can be stably integrated into organisms that are already multidrug-resistant (for example, the hygromycin resistance gene, *hph*, in *Acinetobacter baumannii*²⁴²), or metabolic pathways modulated by adding heterologous genes.²⁴³

Gain-of-function mutants have been identified in the TnsC (an ATPase, similar in function to MuB from the phage Mu system) protein that allow the TnsABC machinery to transposase without target specificity and in the absence of either TnsD or TnsE.²³¹ In this form, the Tn7 transposon has been used for large-scale insertional mutagenesis, much as any other run-of-the-mill transposon.^{193,244}

5.4. Introducing Foreign Genes into a Genome

The notion of introducing a particular foreign gene (or genes) into a genome is the conceptual basis of forward genetic approaches to understanding gene function,^{245–248} manipulating genomes to make transgenic animals to serve as disease models or with new traits,^{249,250} or to provide tissue for xenotransplantation.²⁵¹ It is also the basis of gene therapy in humans as a potential approach to curing a wide range of diseases.¹¹⁰ As seen in section 5.5, it has also been used to generate induced pluripotent stem cells from somatic cells.

For introducing foreign genes into mammalian genomes, the most widely used transposon systems have been the *Sleeping Beauty* and *piggyBac* transposons. Both are versatile methods for transgene delivery, they have proven their utility in a range of mammalian species, and several recent reviews describe their applications, as well as those of other DNA transposons such as *Tol2*.^{110,245–248,252–255} Importantly, all three systems exhibit robust activity in cultured mammalian and human cells.^{256–260}

The use of transposons as a way to permanently introduce a transgene into the genome of a higher organism has an a priori disadvantage due to the generally nonspecific nature of target site selection. Thus, transposons come with the potential for dangerous insertional mutagenesis or even oncogenesis. This is a problem that has impacted other approaches to inserting corrective genes such as the use of retroviruses for transduction,²⁶¹ and can arise not only from the direct disruption of genes or their regulators but also by the introduction of enhancers or promoters from the transgene vector.²⁶² Alternatives that take advantages of targeted approaches to gene introduction include zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR)-associated systems, and these have been reviewed elsewhere.^{263–265}

5.4.1. *Sleeping Beauty* Transposon.—*Sleeping Beauty* is an artificial transposition system that was “awakened from a long evolutionary sleep”²⁶⁶ by determining a consensus sequence for 12 mutated Tc-family transposons found in eight fish species. Both the 340 amino acid transposase and the ~230 bp IRs were then reconstituted, and the so-called “SB10” transposase was shown to transpose in fish, mouse, and human cell lines.²⁶⁶ Since this initial report, *Sleeping Beauty* has proved itself as one of the most important workhorses of transgenesis in higher organisms.^{114,247}

The overall domain organization of the *Sleeping Beauty* transposase, a 340-residue protein, conforms to that of the Tc *I/mariner* superfamily,⁹⁶ and it has the same overall subunit organization as the *Mos1* transposase (Figure 11), although only one HTH subdomain (the “PAI” domain) has been structurally characterized to date.²⁶⁷ In contrast to *Mos1*, however, considerably longer IRs are required for transposition. The reconstructed IRs each have two ~32 bp transposase binding sites, one at the tip of the transposons (called “outer DR”) and an interior site located ~165 bp from the outer DRs (“inner DR”).

Sleeping Beauty moves through a cut-and-paste mechanism and is an unusual Tc *I/mariner*-like element in its reported requirement for a host protein, HMGB1, for transposition¹³⁶ and its interaction with Miz-1, a transcription factor involved in cell cycle regulation.²⁶⁸ As other Tc *I/mariner* transposons, it inserts into TA sites; footprint analysis in mouse embryonic stem cells indicate that cuts at the transposon ends are staggered such that the cut on the nontransferred strand is 3 bp within the transposon.²⁶⁹

To optimize transgenesis experiments, versions of SB10 were sought that were hyperactive in human cells, and the *Sleeping Beauty* transposase has undergone a number of improvements^{270,271} that have culminated in a version, designated SB100X, that is (as the name suggests) 100-fold more active than the original SB10 at mobilizing a chromosomally integrated SB transposon in HeLa cells.²⁷² Relative to SB10, SB100X has five point mutations and one stretch of four amino acids in which all the residues have been mutated (K14R, K33A, R115H, RKEN214–217DAVQ, M243H, T314N). Individually, none of the six individual changes results in more than a 4-fold increase in activity but it was their synergistic effect that resulted in the net overall activity increase. The effects of the mutations defy easy rationalization on the basis of their nature or location, despite the availability of the structure of the homologous *Mos1* transposase bound to DNA (Figure 12).

Rational design of hyperactive transposases appears perilous,²⁷⁰ if not largely unsuccessful. Part of the difficulty may be that increased activity in cells may be unrelated to the chemical steps of transposition, but rather related to other contributing elements such as relief of an inhibition activity,¹³⁹ removal of a negatively regulating phosphorylation site,²⁷³ or improved interaction with a host factor.²⁷⁴ As has been shown for other Tc I/*mariner* transposons,^{275,147} it is also possible to increase the activity of the SB transposition system through rational mutation of the IRs.^{88,247}

There have been many preclinical gene therapy studies using *Sleeping Beauty*, and these have been recently reviewed.^{247,255,276,277} One of the very first studies demonstrated that *Sleeping Beauty* could mediate the in vivo integration of a transgene into the genome of mouse hepatocytes in adult mice.²⁷⁸ *Sleeping Beauty* is also currently being used in four ongoing clinical trials for the treatment of B-cell malignancies such as acute lymphocytic leukemia, chronic lymphocytic leukemia, and non-Hodgkin lymphoma. In these studies, ^{279,280} *Sleeping Beauty* has been used to modify ex vivo either patient- or donor-derived T cells to incorporate a chimeric antigen receptor (CAR) targeted against CD19, a protein specific for malignant B cells.

5.4.2. *piggyBac* Transposon.—The 2.4 kb *piggyBac* transposon (Figure 9) was discovered in 1989 as an active mobile element in the cabbage moth, *Trichoplusia ni*.²⁸¹ *piggyBac* carries a single ORF that encodes a 594 amino acid transposase with a DDD nuclease domain, a variation on the more common DDE motif. The transposon ends are perfect 13 bp TRs, and close to each end but asymmetrically placed are 19 bp internal inverted repeats (designated IR_R and IR_L). *piggyBac* from *T. ni* is one member of what is now recognized to be a superfamily of related elements.^{282–285} An active *piggyBac* transposon has recently been discovered in the genome of the little brown bat, *Myotis lucifugus*, making so-called “*piggyBat*” one of the very few active elements in mammals.²⁸⁶

Two features of the *piggyBac* transposition mechanism set it apart from other transposons: *piggyBac* inserts specifically into the sequence TTAA^{281,287} and it does not leave TSDs behind when it excises: it is excised precisely.^{288,289} *piggyBac* is a cut-and-paste transposon that uses a pathway in which hairpins are formed on transposon ends and are subsequently opened leaving 5′-TTAA overhangs as shown in Figure 9.⁸⁹ The property of excising without leaving a footprint behind is a consequence of an integration–excision cycle in which integration occurs into a palindromic sequence with a 4 bp offset and the hairpins formed during excision are opened in such a way that they have a 4 bp overhang. Only in this way are the gaps left by transposon excision repairable by simple annealing and ligation rather than requiring new DNA synthesis. In contrast, for example, Tc I/*mariners* (and *Sleeping Beauty*) integrate into a palindromic target sequence with a 2 bp offset, but the first nucleophilic attack on the top strand of the transposon is two or three nucleotides into the transposon end whereas the second is precisely at the end on the bottom strand as shown in Figure 7C. After strand transfer, repair cannot be avoided.

As with *Sleeping Beauty*, hyperactive versions of *piggyBac* have been developed for genetic engineering purposes. A first iteration involved codon-optimizing the transposase gene for use in mice, and the resulting increase in activity was shown to be correlated to increased

protein levels.²⁹⁰ In a second iteration, random mutations were introduced into the *piggyBac* transposase using error-prone PCR, followed by screening of activity in yeast and then the combination of several slightly hyperactive mutants to yield a modified *piggyBac* transposase (named hyPBase) that has seven amino acid changes and ~10-fold greater activity in mammalian cells.²⁹¹ Once again, in the absence of three-dimensional structural information, it is next to impossible to work backward to explain why these combinations of mutations lead to a hyperactive phenotype.

Although precise excision is a valuable property of *piggyBac* for genetic experiments, an excised transposon has the potential to reintegrate. As there are circumstances where this is not desirable (e.g., in generating induced pluripotent stem cells; see section 5.5), a version of the *piggyBac* transposase has been generated that is capable of excision but cannot integrate, “Exc⁺Int⁻”.²⁹² A first round of alanine mutation directed at highly conserved basic residues in the catalytic core identified several changes that prevented target DNA binding while leaving transposon end cleavage unaffected. A second round of random mutagenesis then found two amino acid changes that increased excision above wild-type levels.

piggyBac transposes in mammalian cells,^{189,256} and has since been used in a wide variety of genetic modification experiments.^{246,248,254,293} Its ability to seamlessly excise is a particularly valuable property in the context of modifying cells that will be subsequently introduced into a human patient, as it is possible to design vectors that will leave a corrective gene in place but any accompanying gene modifications such as selection markers can be subsequently excised by reintroducing the transposase. *ePiggyBac* is a modified version that has been codon-optimized specifically for use in human embryonic stem cells.²⁹⁴

Attempts have been made to convert *piggyBac* from a nonspecifically integrating transposon into one that is targeted to “safe harbor” sites²⁹⁵ in the genome, for example by fusing a TALE targeting domain to the transposase.^{296,297} Other efforts to target *piggyBac* to specific genomic loci using fusion proteins have been described.²⁹⁸ These efforts have not yet been overwhelmingly successful.

5.4.3. Other Properties of *Sleeping Beauty* and *piggyBac*.—Studies aimed at directly comparing the transposition efficiencies of *Sleeping Beauty* and *piggyBac*^{256–260} have revealed that several transposition phenomena clearly differentially affect the two systems. For example, *piggyBac* appears more capable of effectively transposing relatively large genomic sequences than other transposons. This can be important when it is necessary to introduce more than one gene or large genes accompanied by their regulatory elements. For example, *piggyBac* transposons of up to ~100 kb can transpose in mouse embryonic stem cells²⁹⁹ whereas other transposons such as *Sleeping Beauty* appear much more affected by cargo size.^{217,300} It is not clear what might limit the gene-carrying capacity of a particular DNA transposon, as one obvious possibility—that it becomes more difficult for transposon ends to find each other and form a transpososome as the size of the transposon increases—would appear to apply to all DNA transposases.

Another phenomenon affecting the transposition efficiency of DNA transposons is what is known as overproduction inhibition, or “OPI”.^{301–303,110} This is the property of some

transposases of exhibiting lower activity at higher concentrations, and it has been reported that the problem is more severe for some transposases, including *Sleeping Beauty*, than for others.^{256,258} Whether OPI represents an advantageous inhibitory regulatory mechanism³⁰³ or is a regrettable consequence of protein overproduction in particular cellular backgrounds is not yet clear.

One curious feature of *Sleeping Beauty* transposition is that DNA methylation, normally understood to be a mechanism for down-regulating transposition,^{53,54} appears to stimulate transposition activity in mouse embryonic stem cells.³⁰⁴ This property is also shared with *Frog Prince* (another resurrected Tc *I*-like transposon)³⁰⁵ and *Minos*.³⁰⁶ One proffered explanation for this phenomenon is that the effect may be indirect, and that since CpG methylation induced chromatin condensation, this may stimulate the formation of an active transpososome.³⁰⁶ Why this should be an effect specific to only certain members of the Tc // *mariner*-like family of DNA transposases is not clear, although it could be related to the need to bridge inner and outer IR transposase binding sites within their long transposon ends.

5.5. Application of Transposons to Stem Cell Research: Induced Pluripotent Stem Cells

In mammals, embryonic development is a process during which initially very similar cells go through a program of differentiation to ultimately form all of the different cell types that make up the organism. This property of “pluripotency”—being able to give rise to various cell types—is the crux of new ways of thinking about how to cure or treat human diseases.³⁰⁷ For example, pluripotent cells have the potential to be a source of new cells for transplantation that might treat conditions such as spinal cord injuries, Parkinson’s disease, or multiple sclerosis. For this reason, human embryonic stem cells (hES cells) and cell lines developed from them are being explored as regenerative medicine therapies.³⁰⁸

As an alternative to hES cells, cells derived from patients themselves would bypass any concerns about immune rejection or ethical issues related to the use of human embryonic tissue. They could also be used to model a disease state or to test patient-specific responses to drugs. A major advance toward this goal was achieved when it was shown that the process of differentiation that takes place in embryonic cells can be reversed in adult somatic cells. A ground-breaking discovery in 2006 was that adult cells such as fibroblasts can be reprogrammed back to an embryonic-like state by introducing four genes into the cells and inducing the expression of their gene products.³⁰⁹ The resulting cells were named induced pluripotent stem cells, or iPSCs.

Many methods have been used to introduce genes required for reprogramming into cells, including viral transduction or transfecting with episomal plasmids.³¹⁰ It has been shown that expression of the ectopic factors (typically Oct4, Sox2, c-Myc, and Klf4) is only needed for a brief period of time: once adult cells have become iPSCs, the exogenous proteins are no longer necessary. In fact, prolonged expression of the four transcription factors can change the developmental pattern of the cells³¹¹ or even cause the cells to become oncogenic.³¹²

piggyBac occupies a particular niche in iPSC generation as it provides a way to stably integrate the genes necessary for reprogramming, and then seamlessly remove them once

their job is done. This is accomplished by transiently reexpressing the transposase as first demonstrated in 2009.^{313,314} Clearly, other transposons can also be used for iPSC induction when it is not deemed crucial that the cells be genetically identical to those of the host.^{315,316}

As is possible with other types of stem cells, iPSCs can be genetically modified. Thus, the property of seamless excision by *piggyBac* can also be combined with gene targeting approaches to edit the genomes of iPSCs. For example, a point mutation in the human $\alpha 1$ -antitrypsin gene was corrected ex vivo in patient-derived iPSCs using a ZFN.³¹⁷ *piggyBac* was subsequently used to excise a selection cassette, leaving behind the corrected gene with no foreign DNA sequences. *piggyBac* transposition has also recently been combined with a CRISPR/Cas9 approach to modify iPSCs.³¹⁸ In this case, two mutations in the human hemoglobin β gene were corrected in patient-derived iPSCs. It seems likely that clinical trials will soon be undertaken to test the potential of such therapies.³¹⁹

5.6. Other Applications of Transposons

DNA transposons have been used to provide novel experimental approaches to specific questions. For example, since cut-and-paste transposons move by generating DSBs at the ends of the mobile element, they have been used as a tool to investigate cellular pathways of DSB repair. In one application, a temperature-inducible *Mos1* transposase has been used to trigger the mobilization of a *Mos1* transposon inserted into a *C. elegans* chromosome.³²⁰ Using such a system, processes such as gene conversion using a transgene and the regulation of meiotic crossovers have been studied.

One recent and exciting application is the high resolution mapping of nucleosome positions using transposon integrations into accessible chromatin followed by sequencing. This method, called ATAC-seq,³²¹ takes advantage of an observed property of Tn5 integration that it occurs at high frequency in open chromatin and occurs randomly, while factors bound to DNA such as nucleosomes prevent integration, presumably due to steric reasons. When compared to other nucleosome position mapping techniques, there are several substantial advantages. First, ATAC-seq-based maps can be obtained using only a few cells; when carried out using a microfluidic device,³²² it can be done on the single cell level (scATAC-seq). This opens up the possibility of looking at cell-to-cell variations in the nucleosome landscape. Second, the method is very rapid, so in principle, personal epigenetic profiles on clinically relevant time scales could be generated.

6. OUTLOOK AND CONCLUSION

One surprising aspect of mobile DNA, and of DNA transposition in particular, is the bewildering array of apparently mechanistically different systems that exist (or once existed, with their traces still around) in all kingdoms of life. On the surface, these all evolved to accomplish the single task of moving or copying a defined segment of DNA from one genomic location to another. While many of the systems share some common features such as the ubiquitous DD(E/D) catalytic domain, the overall organization of the transposase protein and the transposon terminal DNA sequences are surprisingly varied. This variation is reflected in the astonishing diversity of transpososome architectures.³²³ If one also takes into

consideration the distinctive retroviral intasome assemblies¹⁶² and the RAG1/2 complex,^{324,325} the conclusion must be drawn that any similarity possibly suggested by the topologically common catalytic core is completely overshadowed by the striking functional and regulatory diversity reflected in these quaternary structures.

Yet what we understand mechanistically is only a small fraction of what is out there. For instance, of the ~30 IS and transposon families currently in ISfinder that are presumably mechanistically different from each other, only two have meaningful structural information available on the transpososome level. Out of the ~18 eukaryotic DNA transposon superfamilies, only two have been structurally studied. Entire mechanistic classes are missing from our current structural and functional understanding.

While many of the transposon groups have been recognized for decades as distinct mobile element families or superfamilies, new superfamilies are being discovered, in part because of the substantial evolutionary impact they apparently have had. One such example is the Polinton superfamily.^{74,75} In this case, there is the complete absence of any experimental insight, let alone any structural insight. Therefore, the field must live with models of their action,³²⁶ models that are entirely reasonable based on the sequence landscape of the participating protein players yet without any current experimental verification and confirmation. Another such example is the group of DNA transposons encoding HUH nucleases. Although the family comprising the IS200/IS605 elements is well-characterized,¹⁸ other families including IS91 and the eukaryotic *Helitron* superfamily are lacking experimental data that would establish their mechanisms. This is even more troubling given that the prokaryotic ISCR family, similar in many respects to IS91, is crucially important in the spread and emergence of antibiotic resistance.³²⁷ These mobile elements, reminiscent of *Helitrons*, have the apparent ability to capture and mobilize genes and gene fragments from the “neighborhood”, implying perhaps very large cargo capacity if they could ever be tamed for gene transfer applications.

The persistent and stubborn paucity of experimental information on transposition systems can certainly be blamed on biochemical difficulties of dealing with such systems and their often inherently low transpositional activities that require sensitive and difficult experiments. Nevertheless, the potential benefit of doing so is clear. There are only a handful of currently used transposition systems, and it is not surprising that they come from those families that have been well-characterized: Tn5, *Sleeping Beauty* of the *mariner* family, *Tol2* of the *hAT* family, and *piggyBac* are all accompanied by substantial biochemical and mechanistic information (if not necessarily at the level of transpososomal three-dimensional structure). Therefore, one must conclude that if such levels of understanding supported by experimental data were available on other currently obscure families, the range of their utilities in applications might expand to reaches that we currently cannot imagine.

ACKNOWLEDGMENTS

We apologize to all of those colleagues whose work we have not cited. We thank Drs. Ivana Grabundzija, Susu He, and Shweta Kailasan for helpful comments on the manuscript. This work was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH.

Biographies

Alison B. Hickman received her B.Sc. degree in chemistry from McGill University in 1983, and her Ph.D. degree in chemistry from the Massachusetts Institute of Technology in 1990. She has been at the National Institutes of Health in Bethesda, MD, since 1991, first as a postdoctoral fellow in the Laboratory of Molecular Biology (LMB) at the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), and then as a senior staff fellow at the National Institute of Child Health and Development (NICHD). She has been in her current position as a staff scientist in the LMB at NIDDK since 1999.

Fred Dyda is a senior investigator in the Laboratory of Molecular Biology at the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), NIH. He obtained his Ph.D. from the Department of Crystallography at the University of Pittsburgh in 1992, and he has been at the NIH since 1993. His research group focuses on understanding how mobile elements move within cells and between cells, using X-ray crystallography as a main tool.

REFERENCES

- (1). Feschotte C; Pritham EJ DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu. Rev. Genet* 2007, 41, 331–368 and references therein. [PubMed: 18076328]
- (2). Chénais B; Caruso A; Hiard S; Casse N The Impact of Transposable Elements on Eukaryotic Genomes: From Genome Size Increase to Genetic Adaptation to Stressful Environments. *Gene* 2012, 509, 7–15 and references therein. [PubMed: 22921893]
- (3). Hayashi-Tsugane M; Maekawa M; Tsugane K A Gain-of-Function Bushy dwarf tiller 1 Mutation in Rice MicroRNA Gene miR156d Caused by Insertion of the DNA Transposon nDart1. *Sci. Rep* 2015, 5, 14357. [PubMed: 26403301]
- (4). McClintock B The Significance of Responses of the Genome to Challenge. *Science* 1984, 226, 792–801. [PubMed: 15739260]
- (5). Wagner A Transposable Elements as Genomic Diseases. *Mol. BioSyst* 2009, 5, 32–35. [PubMed: 19081928]
- (6). Chou HH; Berthet J; Marx CJ Fast Growth Increases the Selective Advantage of a Mutation Arising Recurrently during Evolution under Metal Limitation. *PLoS Genet* 2009, 5, e1000652. [PubMed: 19763169]
- (7). Stoebel DM; Dorman CJ The Effect of Mobile Element IS10 on Experimental Regulatory Evolution in *Escherichia coli*. *Mol. Biol. Evol* 2010, 27, 2105–2112. [PubMed: 20400481]
- (8). Fehér T; Bogos B; Méhi O; Fekete G; Csörgő B; Kovács K; Pósfai G; Papp B; Hurst LD; Pál C Competition between Transposable Elements and Mutator Genes in Bacteria. *Mol. Biol. Evol* 2012, 29, 3153–3159. [PubMed: 22527906]
- (9). Oliver KR; Greene WK Transposable Elements and Viruses as Factors in Adaptation and Evolution: An Expansion and Strengthening of the TE-Thrust Hypothesis. *Ecol. Evol* 2012, 2, 2912–2933. [PubMed: 23170223]
- (10). Platt RN, II; Vandeweghe MW; Kern C; Schmidt CJ; Hoffmann FG; Ray DA Large Numbers of Novel miRNAs Originate from DNA Transposons and Are Coincident with a Large Species Radiation in Bats. *Mol. Biol. Evol* 2014, 31, 1536–1545. [PubMed: 24692655]
- (11). Fattash I; Rooke R; Wong A; Hui C; Luu T; Bhardwaj P; Yang G Miniature Inverted-Repeat Transposable Elements: Discovery, Distribution, and Activity. *Genome* 2013, 56, 475–486 and references therein. [PubMed: 24168668]
- (12). Mobile DNA III; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; American Society for Microbiology: Washington, DC, 2015.
- (13). Curcio MJ; Derbyshire KM The Outs and Ins of Transposition: From Mu to Kangaroo. *Nat. Rev. Mol. Cell Biol* 2003, 4, 865–877. [PubMed: 14682279]

- (14). Montañó SP; Rice PA Moving DNA Around: DNA Transposition and Retroviral Integration. *Curr. Opin. Struct. Biol* 2011, 21, 370–378. [PubMed: 21439812]
- (15). McClintock B The Origin and Behavior of Mutable Loci in Maize. *Proc. Natl. Acad. Sci. U. S. A* 1950, 36, 344–355. [PubMed: 15430309]
- (16). Finnegan DJ Eukaryotic Transposable Elements and Genome Evolution. *Trends Genet.* 1989, 5, 103–107. [PubMed: 2543105]
- (17). Finnegan DJ Transposable Elements. *Curr. Opin. Genet. Dev* 1992, 2, 861–867. [PubMed: 1335807]
- (18). Chandler M; de la Cruz F; Dyda F; Hickman AB; Moncalian G; Ton-Hoang B Breaking and Joining Single-Stranded DNA: The HUH Endonuclease Superfamily. *Nat. Rev. Microbiol* 2013, 11, 525–538 and references therein. [PubMed: 23832240]
- (19). Dyda F; Hickman AB; Jenkins TM; Engelman A; Craigie R; Davies DR Crystal Structure of the Catalytic Domain of HIV-1 Integrase: Similarity to Other Polynucleotidyl Transferases. *Science* 1994, 266, 1981–1986. [PubMed: 7801124]
- (20). Majorek KA; Dunin-Horkawicz S; Steczkiewicz K; Muszewska A; Nowotny M; Ginalski K; Bujnicki JM The RNase H-Like Superfamily: New Members, Comparative Structural Analysis and Evolutionary Classification. *Nucleic Acids Res.* 2014, 42, 4160–4179 and references therein. [PubMed: 24464998]
- (21). Wicker T; Sabot F; Hua-Van A; Bennetzen JL; Capi P; Chalhoub B; Flavell A; Leroy P; Morgante M; Panaud O; et al. A Unified Classification System for Eukaryotic Transposable Elements. *Nat. Rev. Genet* 2007, 8, 973–982. [PubMed: 17984973]
- (22). Hickman AB; Chandler M; Dyda F Integrating Prokaryotes and Eukaryotes: DNA Transposases in Light of Structure. *Crit. Rev. Biochem. Mol. Biol* 2010, 45, 50–69. [PubMed: 20067338]
- (23). Siguier P; Gournay E; Chandler M Bacterial Insertion Sequences: Their Genomic Impact and Diversity. *FEMS Microbiol. Rev* 2014, 38, 865–891 and references therein. [PubMed: 24499397]
- (24). Siguier P; Perochon J; Lestrade L; Mahillon J; Chandler M ISfinder: The Reference Centre for Bacterial Insertion Sequences. *Nucleic Acids Res* 2006, 34, D32–D36. [PubMed: 16381877]
- (25). Piégu B; Bire S; Arensburger P; Bigot Y A Survey of Transposable Element Classification Systems - A Call for a Fundamental Update to Meet the Challenge of Their Diversity and Complexity. *Mol. Phylogenet. Evol* 2015, 86, 90–109. [PubMed: 25797922]
- (26). Jurka J; Kapitonov VV; Pavlicek A; Klonowski P; Kohany O; Walichiewicz J Repbase Update, a Database of Eukaryotic Repetitive Elements. *Cytogenet. Genome Res* 2005, 110, 462–467. [PubMed: 16093699]
- (27). Berg DE; Davies J; Allet B; Rochaix JD Transposition of R Factor Genes to Bacteriophage λ . *Proc. Natl. Acad. Sci. U. S. A* 1975, 72, 3628–3632. [PubMed: 1059152]
- (28). Kleckner N; Chan RK; Tye BK; Botstein D Mutagenesis by Insertion of a Drug-Resistance Element Carrying an Inverted Repetition. *J. Mol. Biol* 1975, 97, 561–575. [PubMed: 1102715]
- (29). Heffron F; Rubens C; Falkow S Translocation of a Plasmid DNA Sequence which Mediates Ampicillin Resistance: Molecular Nature and Specificity of Insertion. *Proc. Natl. Acad. Sci. U. S. A* 1975, 72, 3623–3627. [PubMed: 1103150]
- (30). Barth PT; Datta N; Hedges RW; Grinter NJ Transposition of a Deoxyribonucleic Acid Sequence Encoding Trimethoprim and Streptomycin Resistances from R483 to Other Replicons. *J. Bacteriol* 1976, 125, 800–810. [PubMed: 767328]
- (31). Barth PT; Datta N Two Naturally Occurring Transposons Indistinguishable from Tn7. *J. Gen. Microbiol* 1977, 102, 129–134. [PubMed: 915473]
- (32). Kleckner N Translocatable Elements in Prokaryotes. *Cell* 1977, 11, 11–23 and references therein. [PubMed: 326413]
- (33). Harshey RM The Mu Story: How a Maverick Phage Moved the Field Forward. *Mobile DNA* 2012, 3, 21 and references therein. [PubMed: 23217166]
- (34). Rice P; Mizuuchi K Structure of the Bacteriophage Mu Transposase Core: A Common Structural Motif for DNA Transposition and Retroviral Integration. *Cell* 1995, 82, 209–220. [PubMed: 7628012]
- (35). Krishnan L; Engelman A Retroviral Integrase Proteins and HIV-1 DNA Integration. *J. Biol. Chem* 2012, 287, 40858–40866 and references therein. [PubMed: 23043109]

- (36). Reznikoff WS Transposon Tn5. *Annu. Rev. Genet* 2008, 42, 269–286 and references therein. [PubMed: 18680433]
- (37). Haniford DB; Ellis MJ Transposons Tn10 and Tn5 In *Mobile DNA III*; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; ASM Press: Washington, DC, 2015; pp 631–645 and references therein, DOI: 10.1128/microbiolspec.MDNA3-0002-2014.
- (38). Nicolas E; Lambin M; Dandoy D; Galloy C; Nguyen N; Oger CA; Hallet B The Tn3-Family of Replicative Transposons In *Mobile DNA III*; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; ASM Press: Washington, DC, 2015; pp 693–726 and references therein, DOI: 10.1128/micro-biolspec.MDNA3-0060-2014.
- (39). Peters JE; Craig NL Tn7: Smarter Than We Thought. *Nat. Rev. Mol. Cell Biol.* 2001, 2, 806–814. [PubMed: 11715047]
- (40). Ilyina TV; Koonin EV Conserved Sequence Motifs in the Initiator Proteins for Rolling Circle DNA Replication Encoded by Diverse Replicons from Eubacteria, Eucaryotes and Archaeobacteria. *Nucleic Acids Res.* 1992, 20, 3279–3285. [PubMed: 1630899]
- (41). Koonin EV; Ilyina TV Computer-Assisted Dissection of Rolling Circle DNA Replication. *BioSystems* 1993, 30, 241–268. [PubMed: 8374079]
- (42). Garcillán-Barcia MP; Bernales I; Mendiola MV; de la Cruz F IS91 Rolling-Circle Transposition In *Mobile DNA II*; Craig NL, Craigie R, Gellert M, Lambowitz AM, Eds.; ASM Press: Washington, DC, 2002; pp 891–904 and references therein.
- (43). Grindley NDF; Whiteson KL; Rice PA Mechanisms of Site-Specific Recombination. *Annu. Rev. Biochem* 2006, 75, 567–605 and references therein. [PubMed: 16756503]
- (44). Boocock MR; Rice PA A Proposed Mechanism for IS607-Family Serine Transposases. *Mobile DNA* 2013, 4, 24. [PubMed: 24195768]
- (45). Rajeev L; Malanowska K; Gardner JF Challenging a Paradigm: The Role of DNA Homology in Tyrosine Recombinase Reactions. *Microbiol. Mol. Biol. Rev* 2009, 73, 300–309. [PubMed: 19487729]
- (46). Yuan YW; Wessler SR The Catalytic Domain of All Eukaryotic Cut-and-Paste Transposase Superfamilies. *Proc. Natl. Acad. Sci. U. S. A* 2011, 108, 7884–7889. [PubMed: 21518873]
- (47). Kapitonov VV; Jurka J Rolling-Circle Transposons in Eukaryotes. *Proc. Natl. Acad. Sci. U. S. A* 2001, 98, 8714–8719. [PubMed: 11447285]
- (48). Kapitonov VV; Jurka J Helitrons on a Roll: Eukaryotic Rolling-Circle Transposons. *Trends Genet* 2007, 23, 521–529. [PubMed: 17850916]
- (49). Pritham EJ; Feschotte C Massive Amplification of Rolling-Circle Transposons in the Lineage of the Bat *Myotis lucifugus*. *Proc. Natl. Acad. Sci. U. S. A* 2007, 104, 1895–1900. [PubMed: 17261799]
- (50). Chalopin D; Naville M; Plard F; Galiana D; Volff JN Comparative Analysis of Transposable Elements Highlights Mobilome Diversity and Evolution in Vertebrates. *Genome Biol. Evol* 2015, 7, 567–580. [PubMed: 25577199]
- (51). Levin HL; Moran JV Dynamic Interactions between Transposable Elements and Their Hosts. *Nat. Rev. Genet* 2011, 12, 615–627 and references therein. [PubMed: 21850042]
- (52). Burns KH; Boeke JD Human Transposon Tectonics. *Cell* 2012, 149, 740–752. [PubMed: 22579280]
- (53). Zemach A; Zilberman D Evolution of Eukaryotic DNA Methylation and the Pursuit of Safer Sex. *Curr. Biol* 2010, 20, R780–R785 and references therein. [PubMed: 20833323]
- (54). Saze H; Tsugane K; Kanno T; Nishimura T DNA Methylation in Plants: Relationship to Small RNAs and Histone Modifications, and Functions in Transposon Inactivation. *Plant Cell Physiol.* 2012, 53, 766–784 and references therein. [PubMed: 22302712]
- (55). Castel SE; Martienssen RA RNA Interference in the Nucleus: Roles for Small RNAs in Transcription, Epigenetics and Beyond. *Nat. Rev. Genet* 2013, 14, 100–112 and references therein. [PubMed: 23329111]
- (56). Alekshun MN; Levy SB Molecular Mechanisms of Antibacterial Multidrug Resistance. *Cell* 2007, 128, 1037–1050 and references therein. [PubMed: 17382878]
- (57). Partridge SR Resistance Mechanisms in Enterobacteriaceae. *Pathology* 2015, 47, 276–284 and references therein. [PubMed: 25764207]

- (58). He S; Hickman AB; Varani AM; Siguier P; Chandler M; Dekker JP; Dyda F Insertion Sequence IS26 Reorganizes Plasmids in Clinically Isolated Multidrug-Resistant Bacteria by Replicative Transposition. *mBio* 2015, 6, e00762–15. [PubMed: 26060276]
- (59). Kapitonov VV; Jurka J RAG1 Core and V(D)J Recombination Signal Sequences Were Derived from *Transib* Transposons. *PLoS Biol* 2005, 3, e181. [PubMed: 15898832]
- (60). Hencken CG; Li X; Craig NL Functional Characterization of an Active Rag-Like Transposase. *Nat. Struct. Mol. Biol* 2012, 19, 834–836. [PubMed: 22773102]
- (61). Kapitonov VV; Koonin EV Evolution of the RAG1-RAG2 Locus: Both Proteins Came from the Same Transposon. *Biol. Direct* 2015, 10, 20. [PubMed: 25928409]
- (62). Schatz DG; Swanson PC V(D)J Recombination: Mechanisms of Initiation. *Annu. Rev. Genet* 2011, 45, 167–202 and references therein. [PubMed: 21854230]
- (63). Rajaei N; Chiruvella KK; Lin F; Åström SU Domesticated Transposase Kat1 and Its Fossil Imprints Induce Sexual Differentiation in Yeast. *Proc. Natl. Acad. Sci. U. S. A* 2014, 111, 15491–15496. [PubMed: 25313032]
- (64). Baudry C; Malinsky S; Restituito M; Kapusta A; Rosa S; Meyer E; Bétermier M PiggyMac, a Domesticated *piggyBac* Transposase Involved in Programmed Genome Rearrangements in the Ciliate *Paramecium tetraurelia*. *Genes Dev.* 2009, 23, 2478–2483. [PubMed: 19884254]
- (65). Nowacki M; Higgins BP; Maquilan GM; Swart EC; Doak TG; Landweber LF A Functional Role for Transposases in a Large Eukaryotic Genome. *Science* 2009, 324, 935–938. [PubMed: 19372392]
- (66). Cheng CY; Vogt A; Mochizuki K; Yao MC A Domesticated *piggyBac* Transposase Plays Key Roles in Heterochromatin Dynamics and DNA Cleavage during Programmed DNA Deletion in *Tetrahymena thermophila*. *Mol. Biol. Cell* 2010, 21, 1753–1762. [PubMed: 20357003]
- (67). Aravind L The BED Finger, a Novel DNA-Binding Domain in Chromatin-Boundary-Element-Binding Proteins and Transposases. *Trends Biochem. Sci* 2000, 25, 421–423. [PubMed: 10973053]
- (68). Cordaux R; Udit S; Batzer MA; Feschotte C Birth of a Chimeric Primate Gene by Capture of the Transposase Gene from a Mobile Element. *Proc. Natl. Acad. Sci. U. S. A* 2006, 103, 8101–8106. [PubMed: 16672366]
- (69). Liu D; Bischerour J; Siddique A; Buisine N; Bigot Y; Chalmers R The Human SETMAR Protein Preserves Most of the Activities of the Ancestral Hsmar1 Transposase. *Mol. Cell. Biol* 2007, 27, 1125–1132. [PubMed: 17130240]
- (70). Feschotte C Transposable Elements and the Evolution of Regulatory Networks. *Nat. Rev. Genet* 2008, 9, 397–405. [PubMed: 18368054]
- (71). Krupovic M; Makarova KS; Forterre P; Prangishvili D; Koonin EV Casposons: A New Superfamily of Self-Synthesizing DNA Transposons at the Origin of Prokaryotic CRISPR-Cas Immunity. *BMC Biol* 2014, 12, 36. [PubMed: 24884953]
- (72). Hickman AB; Dyda F The Casposon-Encoded Cas1 Protein from *Aciduliprofundum boonei* is a DNA Integrase that Generates Target Site Duplications. *Nucleic Acids Res* 2015, 43, 10576–10587. [PubMed: 26573596]
- (73). Abergel C; Legendre M; Claverie JM The Rapidly Expanding Universe of Giant Viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiol. Rev* 2015, 39, 779–796. [PubMed: 26391910]
- (74). Krupovic M; Koonin EV Polintons: A Hotbed of Eukaryotic Virus, Transposon and Plasmid Evolution. *Nat. Rev. Microbiol* 2015, 13, 105–115. [PubMed: 25534808]
- (75). Koonin EV; Dolja VV; Krupovic M Origins and Evolution of Viruses of Eukaryotes: The Ultimate Modularity. *Virology* 2015, 479–480, 2–25.
- (76). Krupovic M; Bamford DH; Koonin EV Conservation of Major and Minor Jelly-Roll Capsid Proteins in Polinton (Maverick) Transposons Suggests that They are bona fide Viruses. *Biol. Direct* 2014, 9, 6. [PubMed: 24773695]
- (77). Hickman AB; Dyda F Mechanisms of DNA Transposition In Mobile DNA III; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; ASM Press: Washington, DC, 2015; pp 531–553 and references therein, DOI: 10.1128/microbiolspec.MDNA3-0034-2014.

- (78). Steitz TA; Steitz JA A General Two-Metal-Ion Mechanism for Catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A* 1993, 90, 6498–6502. [PubMed: 8341661]
- (79). Mizuuchi K; Nobbs TJ; Halford SE; Adzuma K; Qin J A New Method for Determining the Stereochemistry of DNA Cleavage Reactions: Application to the *SfiI* and *HpaII* Restriction Endonucleases and to the MuA Transposase. *Biochemistry* 1999, 38, 4640–4648. [PubMed: 10194386]
- (80). Kennedy AK; Haniford DB; Mizuuchi K Single Active Site Catalysis of the Successive Phosphoryl Transfer Steps by DNA Transposases: Insights from Phosphorothioate Stereoselectivity. *Cell* 2000, 101, 295–305. [PubMed: 10847684]
- (81). Yang W; Lee JY; Nowotny M Making and Breaking Nucleic Acids: Two-Mg²⁺-Ion Catalysis and Substrate Specificity. *Mol. Cell* 2006, 22, 5–13 and references therein. [PubMed: 16600865]
- (82). Nowotny M; Gaidamakov SA; Crouch RJ; Yang W Crystal Structures of RNase H Bound to an RNA/DNA Hybrid: Substrate Specificity and Metal-Dependent Catalysis. *Cell* 2005, 121, 1005–1016. [PubMed: 15989951]
- (83). Rosta E; Nowotny M; Yang W; Hummer G Catalytic Mechanism of RNA Backbone Cleavage by Ribonuclease H from Quantum Mechanics/Molecular Mechanics Simulations. *J. Am. Chem. Soc* 2011, 133, 8934–8941. [PubMed: 21539371]
- (84). Hare S; Gupta SS; Valkov E; Engelman A; Cherepanov P Retroviral Intasome Assembly and Inhibition of DNA Strand Transfer. *Nature* 2010, 464, 232–236. [PubMed: 20118915]
- (85). Maertens GN; Hare S; Cherepanov P The Mechanism of Retroviral Integration from X-ray Structures of Its Key Intermediates. *Nature* 2010, 468, 326–329. [PubMed: 21068843]
- (86). Hare S; Maertens GN; Cherepanov P 3'-Processing and Strand Transfer Catalysed by Retroviral Integrase in Crystallo. *EMBO J* 2012, 31, 3020–3028. [PubMed: 22580823]
- (87). Wu Z; Chaconas G Flanking Host Sequences Can Exert an Inhibitory Effect on the Cleavage Step of the in Vitro Mu DNA Strand Transfer Reaction. *J. Biol. Chem* 1992, 267, 9552–9558. [PubMed: 1315758]
- (88). Cui Z; Geurts AM; Liu G; Kaufman CD; Hackett PB Structure-Function Analysis of the Inverted Terminal Repeats of the Sleeping Beauty Transposon. *J. Mol. Biol* 2002, 318, 1221–1235. [PubMed: 12083513]
- (89). Mitra R; Fain-Thornton J; Craig NL *piggyBac* Can Bypass DNA Synthesis during Cut and Paste Transposition. *EMBO J* 2008, 27, 1097–1109. [PubMed: 18354502]
- (90). Claeys Bouuaert C; Chalmers R Transposition of the Human *Hsmar1* Transposon: Rate-Limiting Steps and the Importance of the Flanking TA Dinucleotide in Second Strand Cleavage. *Nucleic Acids Res* 2010, 38, 190–202. [PubMed: 19858101]
- (91). Haniford DB Transpososome Dynamics and Regulation in Tn10 Transposition. *Crit. Rev. Biochem. Mol. Biol* 2006, 41, 407–424 and references therein. [PubMed: 17092825]
- (92). Chandler M; Fayet O; Rousseau P; Ton-Hoang B; Duval-Valentin G Copy-Out-Paste-In Transposition of *IS911*: A Major Transposition Pathway In Mobile DNA III; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; ASM Press: Washington, DC, 2015; pp 591–607 and references therein, DOI: 10.1128/microbiolspec.MDNA3-0031-2014.
- (93). Atkinson PW *hAT* Transposable Elements In Mobile DNA III; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; ASM Press: Washington, DC, 2015; pp 775–802 and references therein, DOI: 10.1128/microbiolspec.MDNA3-0054-2014.
- (94). Feng X; Colloms SD In Vitro Transposition of *ISY100*, a Bacterial Insertion Sequence Belonging to the *Tc1/mariner* Family. *Mol. Microbiol* 2007, 65, 1432–1443. [PubMed: 17680987]
- (95). Dawson A; Finnegan DJ Excision of the *Drosophila* Mariner Transposon *Mos1*: Comparison with Bacterial Transposition and V(D)J Recombination. *Mol. Cell* 2003, 11, 225–235. [PubMed: 12535535]
- (96). Plasterk RHA; Izsvák Z; Ivics Z Resident Aliens: The *Tc1/mariner* Superfamily of Transposable Elements. *Trends Genet* 1999, 15, 326–332. [PubMed: 10431195]
- (97). Claeys Bouuaert C; Walker N; Liu D; Chalmers R Crosstalk between Transposase Subunits during Cleavage of the *mariner* Transposon. *Nucleic Acids Res.* 2014, 42, 5799–5808. [PubMed: 24623810]

- (98). Beall EL; Rio DC *Drosophila* P-element Transposase is a Novel Site-Specific Endonuclease. *Genes Dev.* 1997, 11, 2137–2151. [PubMed: 9284052]
- (99). Sarnovsky RJ; May EW; Craig NL The Tn7 Transposase is a Heteromeric Complex in which DNA Breakage and Joining Activities are Distributed between Different Gene Products. *EMBO J.* 1996, 15, 6348–6361. [PubMed: 8947057]
- (100). Biery MC; Lopata M; Craig NL A Minimal System for Tn7 Transposition: The Transposon-Encoded Proteins TnsA and TnsB Can Execute DNA Breakage and Joining Reactions that Generate Circularized Tn7 Species. *J. Mol. Biol.* 2000, 297, 25–37. [PubMed: 10704304]
- (101). Hickman AB; Li Y; Mathew SV; May EW; Craig NL; Dyda F Unexpected Structural Diversity in DNA Recombination: The Restriction Endonuclease Connection. *Mol. Cell* 2000, 5, 1025–1034. [PubMed: 10911996]
- (102). May EW; Craig NL Switching from Cut-and-Paste to Replicative Tn7 Transposition. *Science* 1996, 272, 401–404. [PubMed: 8602527]
- (103). Peters JE; Fricker AD; Kapili BJ; Petassi MT Heteromeric Transposase Elements: Generators of Genomic Islands across Diverse Bacteria. *Mol. Microbiol* 2014, 93, 1084–1092. [PubMed: 25091064]
- (104). Mizuuchi K Transpositional Recombination: Mechanistic Insights from Studies of Mu and Other Elements. *Annu. Rev. Biochem* 1992, 61, 1011–1051 and references therein. [PubMed: 1323232]
- (105). Siguier P; Filée J; Chandler M Insertion Sequences in Prokaryotic Genomes. *Curr. Opin. Microbiol* 2006, 9, 526–531. [PubMed: 16935554]
- (106). Shapiro JA Molecular Model for the Transposition and Replication of Bacteriophage Mu and Other Transposable Elements. *Proc. Natl. Acad. Sci. U. S. A* 1979, 76, 1933–1937. [PubMed: 287033]
- (107). Nakai H; Doseeva V; Jones JM Handoff from Recombinase to Replisome: Insights from Transposition. *Proc. Natl. Acad. Sci. U. S. A* 2001, 98, 8247–8254. [PubMed: 11459960]
- (108). Burton BM; Baker TA Mu Transpososome Architecture Ensures that Unfolding by ClpX or Proteolysis by ClpXP Remodels but Does Not Destroy the Complex. *Chem. Biol* 2003, 10, 463–472. [PubMed: 12770828]
- (109). Arthur A; Sherratt D Dissection of the Transposition Process: A Transposon-Encoded Site-Specific Recombination System. *Mol. Gen. Genet* 1979, 175, 267–274. [PubMed: 392228]
- (110). Skipper KA; Andersen PR; Sharma N; Mikkelsen JG DNA Transposon-Based Gene Vehicles - Scenes from an Evolutionary Drive. *J. Biomed. Sci* 2013, 20, 92 and references therein. [PubMed: 24320156]
- (111). Roffler S; Wicker T Genome-Wide Comparison of Asian and African Rice Reveals High Recent Activity of DNA Transposons. *Mobile DNA* 2015, 6, 8 and references therein. [PubMed: 25954322]
- (112). Gilbert DM; Bridges MC; Strother AE; Burckhalter CE; Burnette JM, III; Hancock CN Precise Repair of *mPing* Excision Sites is Facilitated by Target Site Duplication Derived Microhomology. *Mobile DNA* 2015, 6, 15. [PubMed: 26347803]
- (113). Majumdar S; Rio DC P Transposable Elements in *Drosophila* and Other Eukaryotic Organisms In *Mobile DNA III*; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; ASM Press: Washington, DC, 2015; pp 727–752 and references therein, DOI: 10.1128/microbiolspec.MDNA3-0004-2014.
- (114). Ivics Z; Izsvák Z *Sleeping Beauty* Transposition In *Mobile DNA III*; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; ASM Press: Washington, DC, 2015; pp 853–874 and references therein, DOI: 10.1128/microbiolspec.MDNA3-0042-2014.
- (115). Gloor GB; Moretti J; Mouyal J; Keeler KJ Distinct *P*-Element Excision Products in Somatic and Germline Cells of *Drosophila melanogaster*. *Genetics* 2000, 155, 1821–1830. [PubMed: 10924477]
- (116). Izsvák Z; Stüwe EE; Fiedler D; Katzer A; Jeggo PA; Ivics Z Healing the Wounds Inflicted by *Sleeping Beauty* Transposition by Double-Strand Break Repair in Mammalian Somatic Cells. *Mol. Cell* 2004, 13, 279–290. [PubMed: 14759372]

- (117). Grundy GJ; Moulding HA; Caldecott KW; Rulten SL One Ring to Bring Them All - The Role of Ku in Mammalian Non-Homologous End Joining. *DNA Repair* 2014, 17, 30–38 and references therein. [PubMed: 24680220]
- (118). Beall EL; Rio DC *Drosophila* IRBP/Ku p70 Corresponds to the Mutagen-Sensitive *mus309* Gene and is Involved in P-element Excision In Vivo. *Genes Dev.* 1996, 10, 921–933. [PubMed: 8608940]
- (119). Malu S; Malshetty V; Francis D; Cortes P Role of Non-Homologous End Joining in V(D)J Recombination. *Immunol. Res* 2012, 54, 233–246 and references therein. [PubMed: 22569912]
- (120). Marmignon A; Bischerour J; Silve A; Fojcik C; Dubois E; Arnaiz O; Kapusta A; Malinsky S; Bétermier M Ku-Mediated Coupling of DNA Cleavage and Repair during Programmed Genome Rearrangements in the Ciliate *Paramecium tetraurelia*. *PLoS Genet.* 2014, 10, e1004552. [PubMed: 25166013]
- (121). Ma Y; Pannicke U; Schwarz K; Lieber MR Hairpin Opening and Overhang Processing by an Artemis/DNA-Dependent Protein Kinase Complex in Nonhomologous End Joining and V(D)J Recombination. *Cell* 2002, 108, 781–794. [PubMed: 11955432]
- (122). Hagemann AT; Craig NL Tn7 Transposition Creates a Hotspot for Homologous Recombination at the Transposon Donor Site. *Genetics* 1993, 133, 9–16. [PubMed: 8380272]
- (123). Jang S; Sandler SJ; Harshey RM Mu Insertions Are Repaired by the Double-Strand Break Repair Pathway of *Escherichia coli*. *PLoS Genet* 2012, 8, e1002642. [PubMed: 22511883]
- (124). Choi W; Jang S; Harshey RM Mu Transpososome and RecBCD Nuclease Collaborate in the Repair of Simple Mu Insertions. *Proc. Natl. Acad. Sci. U. S. A* 2014, 111, 14112–14117. [PubMed: 25197059]
- (125). Jang S; Harshey RM Repair of Transposable Phage Mu DNA Insertions Begins Only when the *E. coli* Replisome Collides with the Transpososome. *Mol. Microbiol* 2015, 97, 746–758. [PubMed: 25983038]
- (126). Dillingham MS; Kowalczykowski SC RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks. *Microbiol. Mol. Biol. Rev* 2008, 72, 642–671 and references therein. [PubMed: 19052323]
- (127). Ronning DR; Guynet C; Ton-Hoang B; Perez ZN; Ghirlando R; Chandler M; Dyda F Active Site Sharing and Subterminal Hairpin Recognition in a New Class of DNA Transposases. *Mol. Cell* 2005, 20, 143–154. [PubMed: 16209952]
- (128). Barabas O; Ronning DR; Guynet C; Hickman AB; Ton-Hoang B; Chandler M; Dyda F Mechanism of IS200/IS605 Family DNA Transposases: Activation and Transposon-Directed Target Site Selection. *Cell* 2008, 132, 208–220. [PubMed: 18243097]
- (129). Hickman AB; James JA; Barabas O; Pasternak C; Ton-Hoang B; Chandler M; Sommer S; Dyda F DNA Recognition and the Precleavage State during Single-Stranded DNA Transposition in *D. radiodurans*. *EMBO J.* 2010, 29, 3840–3852. [PubMed: 20890269]
- (130). Arias-Palomo E; Berger JM An Atypical AAA+ ATPase Assembly Controls Efficient Transposition through DNA Remodeling and Transposase Recruitment. *Cell* 2015, 162, 860–871. [PubMed: 26276634]
- (131). Davies DR; Goryshin IY; Reznikoff WS; Rayment I Three-Dimensional Structure of the Tn5 Synaptic Complex Transposition Intermediate. *Science* 2000, 289, 77–85. [PubMed: 10884228]
- (132). Lovell S; Goryshin IY; Reznikoff WS; Rayment I Two-Metal Active Site Binding of a Tn5 Transposase Synaptic Complex. *Nat. Struct. Biol* 2002, 9, 278–281. [PubMed: 11896402]
- (133). Klenchin VA; Czyz A; Goryshin IY; Gradman R; Lovell S; Rayment I; Reznikoff WS Phosphate Coordination and Movement of DNA in the Tn5 Synaptic Complex: Role of the (R)YREK Motif. *Nucleic Acids Res.* 2008, 36, 5855–5862. [PubMed: 18790806]
- (134). Sasakawa C; Carle GF; Berg DE Sequences Essential for Transposition at the Termini of IS50. *Proc. Natl. Acad. Sci. U. S. A* 1983, 80, 7293–7297. [PubMed: 6316355]
- (135). Goryshin IY; Reznikoff WS Tn5 *in Vitro* Transposition. *J. Biol. Chem* 1998, 273, 7367–7374. [PubMed: 9516433]
- (136). Zayed H; Izsvák Z; Khare D; Heinemann U; Ivics Z The DNA-Bending Protein HMGB1 is a Cellular Cofactor of *Sleeping Beauty* Transposition. *Nucleic Acids Res.* 2003, 31, 2313–2322. [PubMed: 12711676]

- (137). Montañó SP; Pigli YZ; Rice PA The Mu Transpososome Structure Sheds Light on DDE Recombinase Evolution. *Nature* 2012, 491, 413–417. [PubMed: 23135398]
- (138). Wiegand TW; Reznikoff WS Characterization of Two Hypertransposing Tn5 Mutants. *J. Bacteriol* 1992, 174, 1229–1239. [PubMed: 1310499]
- (139). Weinreich MD; Gasch A; Reznikoff WS Evidence that the *cis* Preference of the Tn5 Transposase is Caused by Nonproductive Multimerization. *Genes Dev.* 1994, 8, 2363–2374. [PubMed: 7958902]
- (140). Zhou M; Reznikoff WS Tn5 Transposase Mutants that Alter DNA Binding Specificity. *J. Mol. Biol* 1997, 271, 362–373. [PubMed: 9268665]
- (141). Zhou M; Bhasin A; Reznikoff WS Molecular Genetic Analysis of Transposase-End DNA Sequence Recognition: Cooperativity of Three Adjacent Base-Pairs in Specific Interaction with a Mutant Tn5 Transposase. *J. Mol. Biol* 1998, 276, 913–925. [PubMed: 9566196]
- (142). Bryan GJ; Jacobson JW; Hartl DL Heritable Somatic Excision of a *Drosophila* Transposon. *Science* 1987, 235, 1636–1638. [PubMed: 3029874]
- (143). van Luenen HGAM; Colloms SD; Plasterk RHA The Mechanism of Transposition of Tc3 in *C. elegans*. *Cell* 1994, 79, 293–301. [PubMed: 7954797]
- (144). Lampe DJ; Churchill MEA; Robertson HM A Purified *mariner* Transposase is Sufficient to Mediate Transposition *in Vitro*. *EMBO J.* 1996, 15, 5470–5479. [PubMed: 8895590]
- (145). Richardson JM; Colloms SD; Finnegan DJ; Walkinshaw MD Molecular Architecture of the Mos1 Paired-End Complex: The Structural Basis of DNA Transposition in a Eukaryote. *Cell* 2009, 138, 1096–1108. [PubMed: 19766564]
- (146). Watkins S; van Pouderooyen G; Sixma TK Structural Analysis of the Bipartite DNA-Binding Domain of Tc3 Transposase Bound to Transposon DNA. *Nucleic Acids Res.* 2004, 32, 4306–4312. [PubMed: 15304566]
- (147). Trubitsyna M; Grey H; Houston DR; Finnegan DJ; Richardson JM Structural Basis for the Inverted Repeat Preferences of *mariner* Transposases. *J. Biol. Chem* 2015, 290, 13531–13540. [PubMed: 25869132]
- (148). Dornan J; Grey H; Richardson JM Structural Role of the Flanking DNA in *mariner* Transposon Excision. *Nucleic Acids Res.* 2015, 43, 2424–2432. [PubMed: 25662605]
- (149). Zhang L; Dawson A; Finnegan DJ DNA-Binding Activity and Subunit Interaction of the *mariner* Transposase. *Nucleic Acids Res.* 2001, 29, 3566–3575. [PubMed: 11522826]
- (150). Warren WD; Atkinson PW; O'Brochta DA The Hermes Transposable Element from the House Fly, *Musca domestica*, is a Short Inverted Repeat-Type Element of the *hobo*, *Ac*, and *Tam3 (hAT)* Element Family. *Genet. Res* 1994, 64, 87–97. [PubMed: 7813905]
- (151). Hickman AB; Perez ZN; Zhou L; Musingarimi P; Ghirlando R; Hinshaw JE; Craig NL; Dyda F Molecular Architecture of a Eukaryotic DNA Transposase. *Nat. Struct. Mol. Biol* 2005, 12, 715–721. [PubMed: 16041385]
- (152). Hickman AB; Ewis HE; Li X; Knapp JA; Laver T; Doss AL; Tolun G; Steven AC; Grishaev A; Bax A; et al. Structural Basis of *hAT* Transposon End Recognition by Hermes, an Octameric DNA Transposase from *Musca domestica*. *Cell* 2014, 158, 353–367. [PubMed: 25036632]
- (153). Calvi BR; Hong TJ; Findley SD; Gelbart WM Evidence for a Common Evolutionary Origin of Inverted Repeat Transposons in *Drosophila* and Plants: *hobo*, *Activator*, and *Tam3*. *Cell* 1991, 66, 465–471. [PubMed: 1651170]
- (154). Koga A; Inagaki H; Bessho Y; Hori H Insertion of a Novel Transposable Element in the Tyrosinase Gene is Responsible for an Albino Mutation in the Medaka Fish, *Oryzias latipes*. *Mol. Gen. Genet* 1995, 249, 400–405. [PubMed: 8552044]
- (155). Koga A; Suzuki M; Inagaki H; Bessho Y; Hori H Transposable Element in Fish. *Nature* 1996, 383, 30.
- (156). Lavoie BD; Chan BS; Allison RG; Chaconas G Structural Aspects of a Higher Order Nucleoprotein Complex: Induction of an Altered DNA Structure at the Mu-Host Junction of the Mu Type 1 Transpososome. *EMBO J.* 1991, 10, 3051–3059. [PubMed: 1655409]
- (157). Kuduvali PN; Rao JE; Craig NL Target DNA Structure Plays a Critical Role in Tn7 Transposition. *EMBO J.* 2001, 20, 924–932. [PubMed: 11179236]

- (158). Pribil PA; Haniford DB Target DNA Bending is an Important Specificity Determinant in Target Site Selection in Tn10 Transposition. *J. Mol. Biol* 2003, 330, 247–259. [PubMed: 12823965]
- (159). Hackett CS; Geurts AM; Hackett PB Predicting Preferential DNA Vector Insertion Sites: Implications for Functional Genomics and Gene Therapy. *Genome Biol.* 2007, 8, S12. [PubMed: 18047689]
- (160). Pflieger A; Jaillet J; Petit A; Augé-Gouillou C; Renault S Target Capture during Mos1 Transposition. *J. Biol. Chem* 2014, 289, 100–111. [PubMed: 24269942]
- (161). Mizuno N; Drami anin M; Mizuuchi M; Adam J; Wang Y; Han YW; Yang W; Steven AC; Mizuuchi K; Ramón-Maiques S MuB is an AAA+ ATPase that Forms Helical Filaments to Control Target Selection for DNA Transposition. *Proc. Natl. Acad. Sci. U. S. A* 2013, 110, E2441–E2450. [PubMed: 23776210]
- (162). Cherepanov P; Maertens GN; Hare S Structural Insights into the Retroviral DNA Integration Apparatus. *Curr. Opin. Struct. Biol* 2011, 21, 249–256. [PubMed: 21277766]
- (163). Maskell DP; Renault L; Serrao E; Lesbats P; Matadeen R; Hare S; Lindemann D; Engelman AN; Costa A; Cherepanov P Structural Basis for Retroviral Integration into Nucleosomes. *Nature* 2015, 523, 366–369. [PubMed: 26061770]
- (164). Kleckner N; Roth J; Botstein D Genetic Engineering in Vivo Using Translocatable Drug-Resistance Elements. *J. Mol. Biol* 1977, 116, 125–159. [PubMed: 338917]
- (165). Wilber A; Wangenstein KJ; Chen Y; Zhuo L; Frandsen JL; Bell JB; Chen ZJ; Ekker SC; McIvor RS; Wang X Messenger RNA as a Source of Transposase for Sleeping Beauty Transposon-Mediated Correction of Hereditary Tyrosinemia Type I. *Mol. Ther* 2007, 15, 1280–1287. [PubMed: 17440442]
- (166). Sumiyama K; Kawakami K; Yagita K A Simple and Highly Efficient Transgenesis Method in Mice with the *Tol2* Transposon System and Cytoplasmic Microinjection. *Genomics* 2010, 95, 306–311. [PubMed: 20219670]
- (167). Ivics Z; Mátés L; Yau TY; Landa V; Zidek V; Bashir S; Hoffmann OI; Hiripi L; Garrels W; Kues WA; et al. Germline Transgenesis in Rodents by Pronuclear Microinjection of *Sleeping Beauty* Transposons. *Nat. Protoc* 2014, 9, 773–793. [PubMed: 24625778]
- (168). Shibano T; Takeda M; Suetake I; Kawakami K; Asashima M; Tajima S; Taira M Recombinant Tol2 Transposase with Activity in Xenopus Embryos. *FEBS Lett.* 2007, 581, 4333–4336. [PubMed: 17716667]
- (169). Lamberg A; Nieminen S; Qiao M; Savilahti H Efficient Insertion Mutagenesis Strategy for Bacterial Genomes Involving Electroporation of In Vitro-Assembled DNA Transposition Complexes of Bacteriophage Mu. *Appl. Environ. Microbiol* 2002, 68, 705–712. [PubMed: 11823210]
- (170). Goryshin IY; Jendrisak J; Hoffman LM; Meis R; Reznikoff WS Insertional Transposon Mutagenesis by Electroporation of Released Tn5 Transposition Complexes. *Nat. Biotechnol* 2000, 18, 97–100. [PubMed: 10625401]
- (171). Lampe DJ Bacterial Genetic Methods to Explore the Biology of *mariner* Transposons. *Genetica* 2010, 138, 499–508. [PubMed: 19711186]
- (172). Vos JC; De Baere I; Plasterk RHA Transposase is the Only Nematode Protein Required for in Vitro Transposition of Tc1. *Genes Dev.* 1996, 10, 755–761. [PubMed: 8598301]
- (173). Osborne BI; Baker B Movers and Shakers: Maize Transposons as Tools for Analyzing Other Plant Genomes. *Curr. Opin. Cell Biol* 1995, 7, 406–413. [PubMed: 7662372]
- (174). Fladung M; Polak O *Ac/Ds*-Transposon Activation Tagging in Poplar: A Powerful Tool for Gene Discovery. *BMC Genomics* 2012, 13, 61. [PubMed: 22309468]
- (175). Lu N; Carter JD; Medina TB; Holt SH; Manrique-Carpintero NC; Upham KT; Pereira A; Shulaev V; Veilleux RE Transposon Based Activation Tagging in Diploid Strawberry and Monoploid Derivatives of Potato. *Plant Cell Rep.* 2014, 33, 1203–1216. [PubMed: 24728112]
- (176). Wang N; Long T; Yao W; Xiong L; Zhang Q; Wu C Mutant Resources for the Functional Analysis of the Rice Genome. *Mol. Plant* 2013, 6, 596–604. [PubMed: 23204502]
- (177). Mathieu M; Winters EK; Kong F; Wan J; Wang S; Eckert H; Luth D; Paz M; Donovan C; Zhang Z; et al. Establishment of a Soybean (*Glycine max* Merr. L) Transposon-Based Mutagenesis Repository. *Planta* 2009, 229, 279–289. [PubMed: 18855007]

- (178). Boon Ng GH; Gong Z Maize *Ac/Ds* Transposon System Leads to Highly Efficient Germline Transmission of Transgenes in Medaka (*Oryzias latipes*). *Biochimie* 2011, 93, 1858–1864. [PubMed: 21777650]
- (179). Asakawa K; Suster ML; Mizusawa K; Nagayoshi S; Kotani T; Urasaki A; Kishimoto Y; Hibi M; Kawakami K Genetic Dissection of Neural Circuits by *Tol2* Transposon-Mediated Gal4 Gene and Enhancer Trapping in Zebrafish. *Proc. Natl. Acad. Sci. U. S. A* 2008, 105, 1255–1260. [PubMed: 18202183]
- (180). Kawakami K; Abe G; Asada T; Asakawa K; Fukuda R; Ito A; Lal P; Mouri N; Muto A; Suster ML; et al. zTrap: Zebrafish Gene Trap and Enhancer Trap Database. *BMC Dev. Biol* 2010, 10, 105. [PubMed: 20950494]
- (181). Seiler C; Gebhart N; Zhang Y; Shinton SA; Li YS; Ross NL; Liu X; Li Q; Bilbee AN; Varshney GK; et al. Mutagenesis Screen Identifies *agtpbp1* and *eps15L1* as Essential for T Lymphocyte Development in Zebrafish. *PLoS One* 2015, 10, e0131908. [PubMed: 26161877]
- (182). Hamlet MRJ; Yergeau DA; Kulyev E; Takeda M; Taira M; Kawakami K; Mead PE *Tol2* Transposon-Mediated Transgenesis in *Xenopus tropicalis*. *Genesis* 2006, 44, 438–445. [PubMed: 16906529]
- (183). Lane MA; Kimber M; Khokha MK Breeding Based Remobilization of *Tol2* Transposon in *Xenopus tropicalis*. *PLoS One* 2013, 8, e76807. [PubMed: 24116167]
- (184). Kawakami K *Tol2*: A Versatile Gene Transfer Vector in Vertebrates. *Genome Biol* 2007, 8 (Suppl 1), S7 and references therein. [PubMed: 18047699]
- (185). Linheiro RS; Bergman CM Testing the Palindromic Target Site Model for DNA Transposon Insertion using the *Drosophila melanogaster* P-Element. *Nucleic Acids Res.* 2008, 36, 6199–6208 and references therein. [PubMed: 18829720]
- (186). Yanagihara K; Mizuuchi K Mismatch-Targeted Transposition of Mu: A New Strategy to Map Genetic Polymorphism. *Proc. Natl. Acad. Sci. U. S. A* 2002, 99, 11317–11321. [PubMed: 12177413]
- (187). Bellen HJ; Levis RW; He Y; Carlson JW; Evans-Holm M; Bae E; Kim J; Metaxakis A; Savakis C; Schulze KL; et al. The *Drosophila* Gene Disruption Project Using Transposons With Distinctive Site Specificities. *Genetics* 2011, 188, 731–743. [PubMed: 21515576]
- (188). Woodard LE; Li X; Malani N; Kaja A; Hice RH; Atkinson PW; Bushman FD; Craig NL; Wilson MH Comparative Analysis of the Recently Discovered *hAT* Transposon *TcBuster* in Human Cells. *PLoS One* 2012, 7, e42666. [PubMed: 23166581]
- (189). Ding S; Wu X; Li G; Han M; Zhuang Y; Xu T Efficient Transposition of the *piggyBac* (*PB*) Transposon in Mammalian Cells and Mice. *Cell* 2005, 122, 473–483. [PubMed: 16096065]
- (190). Liang Q; Kong J; Stalker J; Bradley A Chromosomal Mobilization and Reintegration of *Sleeping Beauty* and *PiggyBac* Transposons. *Genesis* 2009, 47, 404–408. [PubMed: 19391106]
- (191). Huang X; Guo H; Tammanna S; Jung YC; Mellgren E; Bassi P; Cao Q; Tu ZJ; Kim YC; Ekker SC; et al. Gene Transfer Efficiency and Genome-Wide Integration Profiling of *Sleeping Beauty*, *Tol2*, and *PiggyBac* Transposons in Human Primary T cells. *Mol. Ther* 2010, 18, 1803–1813. [PubMed: 20606646]
- (192). Mazurkiewicz P; Tang CM; Boone C; Holden DW Signature-Tagged Mutagenesis: Barcoding Mutants for Genome-Wide Screens. *Nat. Rev. Genet* 2006, 7, 929–939 and references therein. [PubMed: 17139324]
- (193). Choi KH; Kim KJ Applications of Transposon-Based Gene Delivery System in Bacteria. *J. Microbiol. Biotechnol* 2009, 19, 217–228 and references therein. [PubMed: 19349746]
- (194). Barquist L; Boinett CJ; Cain AK Approaches to Querying Bacterial Genomes with Transposon-Insertion Sequencing. *RNA Biol.* 2013, 10, 1161–1169 and references therein. [PubMed: 23635712]
- (195). Brochado AR; Typas A High-Throughput Approaches to Understanding Gene Function and Mapping Network Architecture in Bacteria. *Curr. Opin. Microbiol* 2013, 16, 199–206. [PubMed: 23403119]
- (196). van Opijnen T; Camilli A Transposon Insertion Sequencing: A New Tool for Systems-Level Analysis of Microorganisms. *Nat. Rev. Microbiol* 2013, 11, 435–442 and references therein. [PubMed: 23712350]

- (197). Biswas I Genetic Tools for Manipulating *Acinetobacter baumannii* Genome: An Overview. *J. Med. Microbiol* 2015, 64, 657–669. [PubMed: 25948809]
- (198). Barquist L; Langridge GC; Turner DJ; Phan MD; Turner AK; Bateman A; Parkhill J; Wain J; Gardner PP A Comparison of Dense Transposon Insertion Libraries in the *Salmonella* Serovars Typhi and Typhimurium. *Nucleic Acids Res.* 2013, 41, 4549–4564. [PubMed: 23470992]
- (199). Hensel M; Shea JE; Gleeson C; Jones MD; Dalton E; Holden DW Simultaneous Identification of Bacterial Virulence Genes by Negative Selection. *Science* 1995, 269, 400–403. [PubMed: 7618105]
- (200). Palace SG; Proulx MK; Lu S; Baker RE; Goguen JD Genome-Wide Mutant Fitness Profiling Identifies Nutritional Requirements for Optimal Growth of *Yersinia pestis* in Deep Tissue. *mBio* 2014, 5, e01385–14. [PubMed: 25139902]
- (201). Girgis HS; Liu Y; Ryu WS; Tavazoie S A Comprehensive Genetic Characterization of Bacterial Motility. *PLoS Genet.* 2007, 3, e154.
- (202). Griffin JE; Gawronski JD; DeJesus MA; Ioerger TR; Akerley BJ; Sassetti CM High-Resolution Phenotypic Profiling Defines Genes Essential for Mycobacterial Growth and Cholesterol Catabolism. *PLoS Pathog* 2011, 7, e1002251. [PubMed: 21980284]
- (203). Gallagher LA; Shendure J; Manoil C Genome-Scale Identification of Resistance Functions in *Pseudomonas aeruginosa* Using Tn-seq. *mBio* 2011, 2, e00315–10. [PubMed: 21253457]
- (204). Wu M; McNulty NP; Rodionov DA; Khoroshkin MS; Griffin NW; Cheng J; Latreille P; Kerstetter RA; Terrapon N; Henrissat B; et al. Genetic Determinants of In Vivo Fitness and Diet Responsiveness in Multiple Human Gut *Bacteroides*. *Science* 2015, 350, aac5992. [PubMed: 26472912]
- (205). Christen B; Abeliuk E; Collier JM; Kalogeraki VS; Passarelli B; Collier JA; Fero MJ; McAdams HH; Shapiro L The Essential Genome of a Bacterium. *Mol. Syst. Biol* 2011, 7, 528. [PubMed: 21878915]
- (206). van Opijnen T; Bodi KL; Camilli A Tn-seq: High-Throughput Parallel Sequencing for Fitness and Genetic Interaction Studies in Microorganisms. *Nat. Methods* 2009, 6, 767–772. [PubMed: 19767758]
- (207). Goodman AL; McNulty NP; Zhao Y; Leip D; Mitra RD; Lozupone CA; Knight R; Gordon JI Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. *Cell Host Microbe* 2009, 6, 279–289. [PubMed: 19748469]
- (208). Gawronski JD; Wong SMS; Giannoukos G; Ward DV; Akerley BJ Tracking Insertion Mutants Within Libraries by Deep Sequencing and a Genome-Wide Screen for *Haemophilus* Genes Required in the Lung. *Proc. Natl. Acad. Sci. U. S. A* 2009, 106, 16422–16427. [PubMed: 19805314]
- (209). Langridge GC; Phan MD; Turner DJ; Perkins TT; Parts L; Haase J; Charles I; Maskell DJ; Peters SE; Dougan G; et al. Simultaneous Assay of Every *Salmonella* Typhi Gene Using One Million Transposon Mutants. *Genome Res.* 2009, 19, 2308–2316. [PubMed: 19826075]
- (210). Sassetti CM; Boyd DH; Rubin EJ Comprehensive Identification of Conditionally Essential Genes in Mycobacteria. *Proc. Natl. Acad. Sci. U. S. A* 2001, 98, 12712–12717. [PubMed: 11606763]
- (211). Akerley BJ; Rubin EJ; Camilli A; Lampe DJ; Robertson HM; Mekalanos JJ Systematic Identification of Essential Genes by *In Vitro mariner* Mutagenesis. *Proc. Natl. Acad. Sci. U. S. A* 1998, 95, 8927–8932. [PubMed: 9671781]
- (212). Hutchison CA, III; Peterson SN; Gill SR; Cline RT; White O; Fraser CM; Smith HO; Venter JC Global Transposon Mutagenesis and a Minimal Mycoplasma Genome. *Science* 1999, 286, 2165–2169. [PubMed: 10591650]
- (213). Salama NR; Shepherd B; Falkow S Global Transposon Mutagenesis and Essential Gene Analysis of *Helicobacter pylori*. *J. Bacteriol* 2004, 186, 7926–7935. [PubMed: 15547264]
- (214). Nagarkar-Jaiswal S; Lee PT; Campbell ME; Chen K; Anguiano-Zarate S; Gutierrez MC; Busby T; Lin WW; He Y; Schulze KL; et al. A Library of MiMICs Allows Tagging of Genes and Reversible, Spatial and Temporal Knockdown of Proteins in *Drosophila*. *eLife* 2015, 4, e05338.

- (215). Dupuy AJ; Akagi K; Largaespada DA; Copeland NG; Jenkins NA Mammalian Mutagenesis Using a Highly Mobile Somatic Sleeping Beauty Transposon System. *Nature* 2005, 436, 221–226. [PubMed: 16015321]
- (216). Collier LS; Carlson CM; Ravimohan S; Dupuy AJ; Largaespada DA Cancer Gene Discovery in Solid Tumours Using Transposon-Based Somatic Mutagenesis in the Mouse. *Nature* 2005, 436, 272–276. [PubMed: 16015333]
- (217). Geurts AM; Yang Y; Clark KJ; Liu G; Cui Z; Dupuy AJ; Bell JB; Largaespada DA; Hackett PB Gene Transfer into Genomes of Human Cells by the Sleeping Beauty Transposon System. *Mol. Ther* 2003, 8, 108–117. [PubMed: 12842434]
- (218). Mann MB; Jenkins NA; Copeland NG; Mann KM *Sleeping Beauty* Mutagenesis: Exploiting Forward Genetic Screens for Cancer Gene Discovery. *Curr. Opin. Genet. Dev* 2014, 24, 16–22. [PubMed: 24657532]
- (219). Bermejo-Rodríguez C; Pérez-Mancera PA Use of DNA Transposons for Functional Genetic Screens in Mouse Models of Cancer. *Curr. Opin. Biotechnol* 2015, 35, 103–110. [PubMed: 26073851]
- (220). Chénais B Transposable Elements in Cancer and Other Human Diseases. *Curr. Cancer Drug Targets* 2015, 15, 227–242 and references therein. [PubMed: 25808076]
- (221). DeNicola GM; Karreth FA; Adams DJ; Wong CC The Utility of Transposon Mutagenesis for Cancer Studies in the Era of Genome Editing. *Genome Biol.* 2015, 16, 229. [PubMed: 26481584]
- (222). Takeda H; Wei Z; Koso H; Rust AG; Yew CCK; Mann MB; Ward JM; Adams DJ; Copeland NG; Jenkins NA Transposon Mutagenesis Identifies Genes and Evolutionary Forces Driving Gastrointestinal Tract Tumor Progression. *Nat. Genet* 2015, 47, 142–150. [PubMed: 25559195]
- (223). Wu X; Northcott PA; Dubuc A; Dupuy AJ; Shih DJH; Witt H; Croul S; Bouffet E; Fults DW; Eberhart CG; et al. Clonal Selection Drives Genetic Divergence of Metastatic Medulloblastoma. *Nature* 2012, 482, 529–533. [PubMed: 22343890]
- (224). Moriarity BS; Otto GM; Rahrmann EP; Rathe SK; Wolf NK; Weg MT; Manlove LA; LaRue RS; Temiz NA; Molyneux SD; et al. A *Sleeping Beauty* Forward Genetic Screen Identifies New Genes and Pathways Driving Osteosarcoma Development and Metastasis. *Nat. Genet* 2015, 47, 615–624. [PubMed: 25961939]
- (225). Choi J; Landrette SF; Wang T; Evans P; Bacchiocchi A; Bjornson R; Cheng E; Stiegler AL; Gathiaka S; Acevedo O; et al. Identification of PLX4032-Resistance Mechanisms and Implications for Novel RAF Inhibitors. *Pigm. Cell Melanoma Res.* 2014, 27, 253–262.
- (226). Perna D; Karreth FA; Rust AG; Perez-Mancera PA; Rashid M; Iorio F; Alifrangis C; Arends MJ; Bosenberg MW; Bollag G; et al. BRAF Inhibitor Resistance Mediated by the AKT Pathway in an Oncogenic BRAF Mouse Melanoma Model. *Proc. Natl. Acad. Sci. U. S. A* 2015, 112, E536–E545. [PubMed: 25624498]
- (227). Tsutsui M; Kawakubo H; Hayashida T; Fukuda K; Nakamura R; Takahashi T; Wada N; Saikawa Y; Omori T; Takeuchi H; Kitagawa Y Comprehensive Screening of Genes Resistant to an Anticancer Drug in Esophageal Squamous Cell Carcinoma. *Int. J. Oncol* 2015, 47, 867–874. [PubMed: 26202837]
- (228). Mann MB; Black MA; Jones DJ; Ward JM; Yew CCK; Newberg JY; Dupuy AJ; Rust AG; Bosenberg MW; McMahon M; et al. Transposon Mutagenesis Identifies Genetic Drivers of *Braf*^{V600E} Melanoma. *Nat. Genet* 2015, 47, 486–495. [PubMed: 25848750]
- (229). Waddell CS; Craig NL Tn7 Transposition: Two Transposition Pathways Directed by Five Tn7-Encoded Genes. *Genes Dev.* 1988, 2, 137–149. [PubMed: 2834269]
- (230). Parks AR; Peters JE Tn7 Elements: Engendering Diversity from Chromosomes to Episomes. *Plasmid* 2009, 61, 1–14. [PubMed: 18951916]
- (231). Stellwagen AE; Craig NL Gain-of-Function Mutations in TnsC, an ATP-Dependent Transposition Protein That Activates the Bacterial Transposon Tn7. *Genetics* 1997, 145, 573–585. [PubMed: 9055068]
- (232). Bainton RJ; Kubo KM; Feng JN; Craig NL Tn7 Transposition: Target DNA Recognition is Mediated by Multiple Tn7-Encoded Proteins in a Purified In Vitro System. *Cell* 1993, 72, 931–943. [PubMed: 8384534]

- (233). Mitra R; McKenzie GJ; Yi L; Lee CA; Craig NL Characterization of the TnsD-*attTn7* Complex That Promotes Site-Specific Insertion of *Tn7*. Mobile DNA 2010, 1, 18. [PubMed: 20653944]
- (234). Wolkow CA; DeBoy RT; Craig NL Conjugating Plasmids are Preferred Targets for Tn7. Genes Dev. 1996, 10, 2145–2157. [PubMed: 8804309]
- (235). Peters JE; Craig NL Tn7 Transposes Proximal to DNA Double-Strand Breaks and into Regions Where Chromosomal DNA Replication Terminates. Mol. Cell 2000, 6, 573–582. [PubMed: 11030337]
- (236). Parks AR; Li Z; Shi Q; Owens RM; Jin MM; Peters JE Transposition into Replicating DNA Occurs through Interaction with the Processivity Factor. Cell 2009, 138, 685–695. [PubMed: 19703395]
- (237). Choi KH; Gaynor JB; White KG; Lopez C; Bosio CM; Karkhoff-Schweizer RR; Schweizer HP A Tn7-Based Broad-Range Bacterial Cloning and Expression System. Nat. Methods 2005, 2, 443–448. [PubMed: 15908923]
- (238). Su S; Saldanha R; Pemberton A; Bangar H; Kawamoto SA; Aronow B; Hassett DJ; Lamkin TJ Characterization of Stable, Constitutively Expressed, Chromosomal Green and Red Fluorescent Transcriptional Fusions in the Select Agent Bacterium, *Francisella tularensis* Schu S4 and the Surrogate Type B Live Vaccine Strain (LVS). Appl. Microbiol. Biotechnol 2013, 97, 9029–9041. [PubMed: 23852642]
- (239). Damron FH; McKenney ES; Barbier M; Liechti GW; Schweizer HP; Goldberg JB Construction of Mobilizable Mini-Tn7 Vectors for Bioluminescent Detection of Gram-Negative Bacteria and Single-Copy Promoter *lux* Reporter Analysis. Appl. Environ. Microbiol 2013, 79, 4149–4153. [PubMed: 23584769]
- (240). Bruckbauer ST; Kvitko BH; Karkhoff-Schweizer RR; Schweizer HP Tn5/7-*lux*: A Versatile Tool for the Identification and Capture of Promoters in Gram-Negative Bacteria. BMC Microbiol. 2015, 15, 17. [PubMed: 25648327]
- (241). Roos K; Werner E; Loessner H Multicopy Integration of Mini-Tn7 Transposons into Selected Chromosomal Sites of a *Salmonella* Vaccine Strain. Microb. Biotechnol 2015, 8, 177–187. [PubMed: 25488129]
- (242). Jacobs AC; Thompson MG; Black CC; Kessler JL; Clark LP; McQueary CN; Gancz HY; Corey BW; Moon JK; Si Y; et al. AB5075, a Highly Virulent Isolate of *Acinetobacter baumannii*, as a Model Strain for the Evaluation of Pathogenesis and Antimicrobial Treatments. mBio 2014, 5, e01076–14. [PubMed: 24865555]
- (243). Majerczyk C; Kinman L; Han T; Bunt R; Greenberg EP Virulence of *Burkholderia mallei* Quorum-Sensing Mutants. Infect. Immun 2013, 81, 1471–1478. [PubMed: 23429539]
- (244). Biery MC; Stewart FJ; Stellwagen AE; Raleigh EA; Craig NL A Simple *In Vitro* Tn7-Based Transposition System with Low Target Site Selectivity for Genome and Gene Analysis. Nucleic Acids. Res 2000, 28, 1067–1077. [PubMed: 10666445]
- (245). Ivics Z; Li MA; Mátés L; Boeke JD; Nagy A; Bradley A; Izsvák Z Transposon-Mediated Genome Manipulation in Vertebrates. Nat. Methods 2009, 6, 415–422. [PubMed: 19478801]
- (246). Kim A; Pytko I Size Matters: Versatile Use of *PiggyBac* Transposons as a Genetic Manipulation Tool. Mol. Cell. Biochem 2011, 354, 301–309. [PubMed: 21516337]
- (247). Ammar I; Izsvák Z; Ivics Z The *Sleeping Beauty* Transposon Toolbox. Methods Mol. Biol 2012, 859, 229–240. [PubMed: 22367875]
- (248). Woodard LE; Wilson MH *piggyBac*-ing Models and New Therapeutic Strategies. Trends Biotechnol. 2015, 33, 525–533. [PubMed: 26211958]
- (249). Garrels W; Ivics Z; Kues WA Precision Genetic Engineering in Large Mammals. Trends Biotechnol. 2012, 30, 386–393 and references therein. [PubMed: 22521716]
- (250). Tan W; Carlson DF; Walton MW; Fahrenkrug SC; Hackett PB Precision Editing of Large Animal Genomes. Adv. Genet 2012, 80, 37–97 and references therein. [PubMed: 23084873]
- (251). Clark KJ; Carlson DF; Fahrenkrug SC Pigs Taking Wing With Transposons and Recombinases. Genome Biol. 2007, 8, S13. [PubMed: 18047690]
- (252). Belay E; Dastidar S; VandenDriessche T; Chuah MKL Transposon-Mediated Gene Transfer into Adult and Induced Pluripotent Stem Cells. Curr. Gene Ther 2011, 11, 406–413. [PubMed: 21864290]

- (253). Di Matteo M; Mátrai J; Belay E; Firdissa T; VandenDriessche T; Chuah MKL PiggyBac Toolbox. *Methods Mol. Biol* 2012, 859, 241–254. [PubMed: 22367876]
- (254). Di Matteo M; Belay E; Chuah MK; VandenDriessche T Recent Developments in Transposon-Mediated Gene Therapy. *Expert Opin. Biol. Ther* 2012, 12, 841–858. [PubMed: 22679910]
- (255). Vand Rajabpour F; Raoofian R; Habibi L; Akrami SM; Tabrizi M Novel Trends in Genetics: Transposable Elements and Their Application in Medicine. *Arch. Iran. Med* 2014, 17, 702–712. [PubMed: 25305771]
- (256). Wu SCY; Meir YJJ; Coates CJ; Handler AM; Pelczar P; Moisyadi S; Kaminski JM *piggyBac* is a Flexible and Highly Active Transposon as Compared to *Sleeping Beauty*, *Tol2*, and *Mos1* in Mammalian Cells. *Proc. Natl. Acad. Sci. U. S. A* 2006, 103, 15008–15013. [PubMed: 17005721]
- (257). Wilson MH; Coates CJ; George AL, Jr. *PiggyBac* Transposon-Mediated Gene Transfer in Human Cells. *Mol. Ther* 2007, 15, 139–145. [PubMed: 17164785]
- (258). Grabundzija I; Irgang M; Mátés L; Belay E; Matrai J; Gogol-Döring A; Kawakami K; Chen W; Ruiz P; Chuah MKL; et al. Comparative Analysis of Transposable Element Vector Systems in Human Cells. *Mol. Ther* 2010, 18, 1200–1209. [PubMed: 20372108]
- (259). Meir YJJ; Weirauch MT; Yang HS; Chung PC; Yu RK; Wu SC Genome-Wide Target Profiling of *piggyBac* and *Tol2* in HEK 293: Pros and Cons for Gene Discovery and Gene Therapy. *BMC Biotechnol* 2011, 11, 28. [PubMed: 21447194]
- (260). Doherty JE; Huye LE; Yusa K; Zhou L; Craig NL; Wilson MH Hyperactive *piggyBac* Gene Transfer in Human Cells and *In Vivo*. *Hum. Gene Ther* 2012, 23, 311–320. [PubMed: 21992617]
- (261). Hacein-Bey-Abina S; Von Kalle C; Schmidt M; McCormack MP; Wulffraat N; Leboulch P; Lim A; Osborne CS; Pawliuk R; Morillon E; et al. *LMO2*-Associated Clonal T Cell Proliferation in Two Patients after Gene Therapy for SCID-X1. *Science* 2003, 302, 415–419. [PubMed: 14564000]
- (262). Hacein-Bey-Abina S; Pai S-Y; Gaspar HB; Armant M; Berry CC; Blanche S; Bleesing J; Blondeau J; de Boer H; Buckland KF; et al. A Modified γ -Retrovirus Vector for X-Linked Severe Combined Immunodeficiency. *N. Engl. J. Med* 2014, 371, 1407–1417. [PubMed: 25295500]
- (263). Kim H; Kim JS A Guide to Genome Engineering With Programmable Nucleases. *Nat. Rev. Genet* 2014, 15, 321–334. [PubMed: 24690881]
- (264). LaFontaine JS; Fathe K; Smyth HDC Delivery and Therapeutic Applications of Gene Editing Technologies ZFNs, TALENs, and CRISPR/Cas9. *Int. J. Pharm* 2015, 494, 180–194. [PubMed: 26278489]
- (265). Aravalli RN; Belcher JD; Steer CJ Liver-Target Gene Therapy: Approaches and Challenges. *Liver Transpl* 2015, 21, 718–737. [PubMed: 25824605]
- (266). Ivics Z; Hackett PB; Plasterk RH; Izsvák Z Molecular Reconstruction of *Sleeping Beauty*, a *Tc1*-Like Transposon from Fish, and Its Transposition in Human Cells. *Cell* 1997, 91, 501–510. [PubMed: 9390559]
- (267). Carpentier CE; Schreifels JM; Aronovich EL; Carlson DF; Hackett PB; Nesmelova IV NMR Structural Analysis of *Sleeping Beauty* Transposase Binding to DNA. *Protein Sci.* 2014, 23, 23–33. [PubMed: 24243759]
- (268). Walisko O; Izsvák Z; Szabó K; Kaufman CD; Herold S; Ivics Z *Sleeping Beauty* Transposase Modulates Cell-Cycle Progression through Interaction with Miz-1. *Proc. Natl. Acad. Sci. U. S. A* 2006, 103, 4062–4067. [PubMed: 16537485]
- (269). Luo G; Ivics Z; Izsvák Z; Bradley A Chromosomal Transposition of a *Tc1*/mariner-Like Element in Mouse Embryonic Stem Cells. *Proc. Natl. Acad. Sci. U. S. A* 1998, 95, 10769–10773. [PubMed: 9724779]
- (270). Zayed H; Izsvák Z; Walisko O; Ivics Z Development of Hyperactive *Sleeping Beauty* Transposon Vectors by Mutational Analysis. *Mol. Ther* 2004, 9, 292–304. [PubMed: 14759813]
- (271). Yant SR; Park J; Huang Y; Mikkelsen JG; Kay MA Mutational Analysis of the N-Terminal DNA-Binding Domain of *Sleeping Beauty* Transposase: Critical Residues for DNA Binding and Hyperactivity in Mammalian Cells. *Mol. Cell. Biol* 2004, 24, 9239–9247. [PubMed: 15456893]
- (272). Mátés L; Chuah MKL; Belay E; Jerchow B; Manoj N; Acosta-Sanchez A; Grzela DP; Schmitt A; Becker K; Matrai J; et al. Molecular Evolution of a Novel Hyperactive *Sleeping Beauty*

Transposase Enables Robust Stable Gene Transfer in Vertebrates. *Nat. Genet* 2009, 41, 753–761. [PubMed: 19412179]

- (273). Beall EL; Mahoney MB; Rio DC Identification and Analysis of a Hyperactive Mutant Form of *Drosophila P*-Element Transposase. *Genetics* 2002, 162, 217–227. [PubMed: 12242235]
- (274). Díaz-Maldonado H; Gómez MJ; Moreno-Paz M; San Martín-Úriz P; Amils R; Parro V; López de Saro FJ Transposase Interaction with the β Sliding Clamp: Effects on Insertion Sequence Proliferation and Transposition Rate. *Sci. Rep* 2015, 5, 13329. [PubMed: 26306550]
- (275). Augé-Gouillou C; Hamelin M-H; Demattei M-V; Periquet M; Bigot Y The Wild-Type Conformation of the *Mos-1* Inverted Terminal Repeats is Suboptimal for Transposition in Bacteria. *Mol. Genet. Genomics* 2001, 265, 51–57. [PubMed: 11370872]
- (276). Swierczek M; Izsvák Z; Ivics Z The *Sleeping Beauty* Transposon System for Clinical Applications. *Expert Opin. Biol. Ther* 2012, 12, 139–153. [PubMed: 22176302]
- (277). Boehme P; Doerner J; Solanki M; Jing L; Zhang W; Ehrhardt A The Sleeping Beauty Transposon Vector System for Treatment of Rare Genetic Diseases: An Unrealized Hope? *Curr. Gene Ther.* 2015, 15, 255–265. [PubMed: 25619886]
- (278). Yant SR; Meuse L; Chiu W; Ivics Z; Izsvák Z; Kay MA Somatic Integration and Long-Term Transgene Expression in Normal and Haemophilic Mice Using a DNA Transposon System. *Nat. Genet* 2000, 25, 35–41. [PubMed: 10802653]
- (279). Hackett PB; Largaespada DA; Cooper LJN A Transposon and Transposase System for Human Application. *Mol. Ther* 2010, 18, 674–683. [PubMed: 20104209]
- (280). Singh H; Moyes JSE; Huls MH; Cooper LJN Manufacture of T Cells Using the Sleeping Beauty System to Enforce Expression of a CD19-Specific Chimeric Antigen Receptor. *Cancer Gene Ther.* 2015, 22, 95–100. [PubMed: 25591810]
- (281). Cary LC; Goebel M; Corsaro BG; Wang HG; Rosen E; Fraser MJ Transposon Mutagenesis of Baculoviruses: Analysis of *Trichoplusia ni* Transposon IFP2 Insertions within the FP-Locus of Nuclear Polyhedrosis Viruses. *Virology* 1989, 172, 156–169. [PubMed: 2549707]
- (282). Handler AM; McCombs SD The *piggyBac* Transposon Mediates Germ-Line Transformation in the Oriental Fruit Fly and Closely Related Elements Exist in Its Genome. *Insect Mol. Biol* 2000, 9, 605–612. [PubMed: 11122469]
- (283). Sarkar A; Sim C; Hong YS; Hogan JR; Fraser MJ; Robertson HM; Collins FH Molecular Evolutionary Analysis of the Widespread *piggyBac* Transposon Family and Related “Domesticated” Sequences. *Mol. Genet. Genomics* 2003, 270, 173–180. [PubMed: 12955498]
- (284). Ray DA; Feschotte C; Pagan HJT; Smith JD; Pritham EJ; Arensburger P; Atkinson PW; Craig NL Multiple Waves of Recent DNA Transposon Activity in the Bat *Myotis lucifugus*. *Genome Res.* 2008, 18, 717–728. [PubMed: 18340040]
- (285). Yusa K *piggyBac* Transposon *In Mobile DNA III*; Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB, Eds.; ASM Press: Washington, DC, 2015; pp 875–892 and references therein, DOI: 10.1128/microbiolspec.MDNA3-0028-2014.
- (286). Mitra R; Li X; Kapusta A; Mayhew D; Mitra RD; Feschotte C; Craig NL Functional Characterization of *piggyBac* from the Bat *Myotis lucifugus* Unveils an Active Mammalian DNA Transposon. *Proc. Natl. Acad. Sci. U. S. A* 2013, 110, 234–239. [PubMed: 23248290]
- (287). Fraser MJ; Cary L; Boonvisudhi K; Wang HGH Assay for Movement of Lepidopteran Transposon IFP2 in Insect Cells Using a Baculovirus Genome as a Target DNA. *Virology* 1995, 211, 397–407. [PubMed: 7645244]
- (288). Fraser MJ; Cliszczon T; Elick T; Bauser C Precise Excision of TTAA-Specific Lepidopteran Transposons *piggyBac* (IFP2) and *tagalong* (TFP3) from the Baculovirus Genome in Cell Lines from Two Species of Lepidoptera. *Insect Mol. Biol* 1996, 5, 141–151. [PubMed: 8673264]
- (289). Elick TA; Bauser CA; Fraser MJ Excision of the *piggyBac* Transposable Element *in Vitro* is a Precise Event that is Enhanced by the Expression of its Encoded Transposase. *Genetica* 1996, 98, 33–41. [PubMed: 8765680]
- (290). Cadiñanos J; Bradley A Generation of an Inducible and Optimized *piggyBac* Transposon System. *Nucleic Acids Res.* 2007, 35, e87. [PubMed: 17576687]

- (291). Yusa K; Zhou L; Li MA; Bradley A; Craig NL A Hyperactive *piggyBac* Transposase for Mammalian Applications. *Proc. Natl. Acad. Sci. U. S. A* 2011, 108, 1531–1536. [PubMed: 21205896]
- (292). Li X; Burnight ER; Cooney AL; Malani N; Brady T; Sander JD; Staber J; Wheelan SJ; Joung JK; McCray PB, Jr.; et al. *piggyBac* Transposase Tools for Genome Engineering. *Proc. Natl. Acad. Sci. U. S. A* 2013, 110, E2279–E2287. [PubMed: 23723351]
- (293). Saha S; Woodard LE; Charron EM; Welch RC; Rooney CM; Wilson MH Evaluating the Potential for Undesired Genomic Effects of the *piggyBac* Transposon System in Human Cells. *Nucleic Acids Res.* 2015, 43, 1770–1782. [PubMed: 25605795]
- (294). Lacoste A; Berenshteyn F; Brivanlou AH An Efficient and Reversible Transposable System for Gene Delivery and Lineage-Specific Differentiation in Human Embryonic Stem Cells. *Cell Stem Cell* 2009, 5, 332–342. [PubMed: 19733544]
- (295). Sadelain M; Papapetrou EP; Bushman FD Safe Harbours for the Integration of New DNA in the Human Genome. *Nat. Rev. Cancer* 2012, 12, 51–58.
- (296). Owens JB; Urschitz J; Stoytchev I; Dang NC; Stoytcheva Z; Belcaid M; Maragathavally KJ; Coates CJ; Segal DJ; Moisyadi S Chimeric *piggyBac* Transposases for Genomic Targeting in Human Cells. *Nucleic Acids Res.* 2012, 40, 6978–6991. [PubMed: 22492708]
- (297). Ye L; You Z; Qian Q; Zhang Y; Che J; Song J; Zhong B TAL Effectors Mediate High-Efficiency Transposition of the *piggyBac* Transposon in Silkworm *Bombyx mori* L. *Sci. Rep* 2015, 5, 17172. [PubMed: 26608076]
- (298). Hong JB; Chou FJ; Ku AT; Fan HH; Lee TL; Huang YH; Yang TL; Su IC; Yu IS; Lin SW; et al. A Nucleolus-Predominant *piggyBac* Transposase, NP-mPB, Mediates Elevated Transposition Efficiency in Mammalian Cells. *PLoS One* 2014, 9, e89396 and references therein. [PubMed: 24586748]
- (299). Li MA; Turner DJ; Ning Z; Yusa K; Liang Q; Eckert S; Rad L; Fitzgerald TW; Craig NL; Bradley A Mobilization of Giant *piggyBac* Transposons in the Mouse Genome. *Nucleic Acids Res.* 2011, 39, e148. [PubMed: 21948799]
- (300). Izsvák Z; Ivics Z; Plasterk RH *Sleeping Beauty*, a Wide Host-Range Transposon Vector for Genetic Transformation in Vertebrates. *J. Mol. Biol* 2000, 302, 93–102. [PubMed: 10964563]
- (301). Chalmers RM; Kleckner N Tn10/IS10 Transposase Purification, Activation, and *in Vitro* Reaction. *J. Biol. Chem* 1994, 269, 8029–8035. [PubMed: 8132525]
- (302). Lohe AR; Hartl DL Autoregulation of *mariner* Transposase Activity by Overproduction and Dominant-Negative Complementation. *Mol. Biol. Evol* 1996, 13, 549–555. [PubMed: 8882498]
- (303). Bire S; Casteret S; Arnaoty A; Piégu B; Lecomte T; Bigot Y Transposase Concentration Controls Transposition Activity: Myth or Reality? *Gene* 2013, 530, 165–171. [PubMed: 23994686]
- (304). Yusa K; Takeda J; Horie K Enhancement of *Sleeping Beauty* Transposition by CpG Methylation: Possible Role of Heterochromatin Formation. *Mol. Cell. Biol* 2004, 24, 4004–4018. [PubMed: 15082793]
- (305). Miskey C; Izsvák Z; Plasterk RH; Ivics Z The *Frog Prince*: A Reconstructed Transposon from *Rana pipiens* with High Transpositional Activity in Vertebrate Cells. *Nucleic Acids Res.* 2003, 31, 6873–6881. [PubMed: 14627820]
- (306). Jursch T; Miskey C; Izsvák Z; Ivics Z Regulation of DNA Transposition by CpG Methylation and Chromatin Structure in Human Cells. *Mobile DNA* 2013, 4, 15. [PubMed: 23676100]
- (307). Okano H; Yamanaka S iPS Cell Technologies: Significance and Applications to CNS Regeneration and Disease. *Mol. Brain* 2014, 7, 22 and references therein. [PubMed: 24685317]
- (308). Whiting P; Kerby J; Coffey P; da Cruz L; McKernan R Progressing a Human Embryonic Stem-Cell-Based Regenerative Medicine Therapy Towards the Clinic. *Philos. Trans. R. Soc., B* 2015, 370, 20140375 and references therein.
- (309). Takahashi K; Yamanaka S Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* 2006, 126, 663–676. [PubMed: 16904174]
- (310). Hu K Vectorology and Factor Delivery in Induced Pluripotent Stem Cell Reprogramming. *Stem Cells Dev.* 2014, 23, 1301–1315 and references therein. [PubMed: 24625220]

- (311). Igawa K; Kokubu C; Yusa K; Horie K; Yoshimura Y; Yamauchi K; Suemori H; Yokozeki H; Toyoda M; Kiyokawa N; et al. Removal of Reprogramming Transgenes Improves the Tissue Reconstitution Potential of Keratinocytes Generated From Human Induced Pluripotent Stem Cells. *Stem Cells Transl. Med* 2014, 3, 992–1001. [PubMed: 25024429]
- (312). Okita K; Ichisaka T; Yamanaka S Generation of Germline-Competent Induced Pluripotent Stem Cells. *Nature* 2007, 448, 313–317. [PubMed: 17554338]
- (313). Woltjen K; Michael IP; Mohseni P; Desai R; Mileikovsky M; Hämäläinen R; Cowling R; Wang W; Liu P; Gertsenstein M; et al. *piggyBac* Transposition Reprograms Fibroblasts to Induced Pluripotent Stem Cells. *Nature* 2009, 458, 766–770. [PubMed: 19252478]
- (314). Yusa K; Rad R; Takeda J; Bradley A Generation of Transgene-Free Induced Pluripotent Mouse Stem Cells by the *piggyBac* Transposon. *Nat. Methods* 2009, 6, 363–369. [PubMed: 19337237]
- (315). Grabundzija I; Wang J; Sebe A; Erdei Z; Kajdi R; Devaraj A; Steinemann D; Szuhai K; Stein U; Cantz T; et al. *Sleeping Beauty* Transposon-Based System for Cellular Reprogramming and Targeted Gene Insertion in Induced Pluripotent Stem Cells. *Nucleic Acids Res.* 2013, 41, 1829–1847. [PubMed: 23275558]
- (316). Kues WA; Herrmann D; Barg-Kues B; Haridoss S; Nowak-Imialek M; Buchholz T; Streeck M; Grebe A; Grabundzija I; Merkert S; et al. Derivation and Characterization of Sleeping Beauty Transposon-Mediated Porcine Induced Pluripotent Stem Cells. *Stem Cells Dev.* 2013, 22, 124–135. [PubMed: 22989381]
- (317). Yusa K; Rashid ST; Strick-Marchand H; Varela I; Liu PQ; Paschon DE; Miranda E; Ordóñez A; Hannan NRF; Rouhani FJ; et al. Targeted Gene Correction of α_1 -Antitrypsin Deficiency in Induced Pluripotent Stem Cells. *Nature* 2011, 478, 391–394. [PubMed: 21993621]
- (318). Xie F; Ye L; Chang JC; Beyer AI; Wang J; Muench MO; Kan YW Seamless Gene Correction of β -Thalassemia Mutations in Patient-Specific iPSCs Using CRISPR/Cas9 and *piggyBac*. *Genome Res.* 2014, 24, 1526–1533. [PubMed: 25096406]
- (319). Parker SL; Perlingeiro RCR Are We There Yet? Navigating Roadblocks to Induced Pluripotent Stem Cell Therapy Translation. *Regener. Med* 2013, 8, 389–391.
- (320). Rosu S; Libuda DE; Villeneuve AM Robust Crossover Assurance and Regulated Interhomolog Access Maintain Meiotic Crossover Number. *Science* 2011, 334, 1286–1289 and references therein. [PubMed: 22144627]
- (321). Buenrostro JD; Giresi PG; Zaba LC; Chang HY; Greenleaf WJ Transposition of Native Chromatin For Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position. *Nat. Methods* 2013, 10, 1213–1218. [PubMed: 24097267]
- (322). Buenrostro JD; Wu B; Litzenburger UM; Ruff D; Gonzales ML; Snyder MP; Chang HY; Greenleaf WJ Single-Cell Chromatin Accessibility Reveals Principles of Regulatory Variation. *Nature* 2015, 523, 486–490. [PubMed: 26083756]
- (323). Dyda F; Chandler M; Hickman AB The Emerging Diversity of Transpososome Architectures. *Q. Rev. Biophys* 2012, 45, 493–521. [PubMed: 23217365]
- (324). Kim MS; Lapkouski M; Yang W; Gellert M Crystal Structure of the V(D)J Recombinase RAG1-RAG2. *Nature* 2015, 518, 507–511. [PubMed: 25707801]
- (325). Ru H; Chambers MG; Fu TM; Tong AB; Liao M; Wu H Molecular Mechanism of V(D)J Recombination from Synaptic RAG1-RAG2 Complex Structures. *Cell* 2015, 163, 1138–1152. [PubMed: 26548953]
- (326). Kapitonov VV; Jurka J Self-Synthesizing DNA Transposons in Eukaryotes. *Proc. Natl. Acad. Sci. U. S. A* 2006, 103, 4540–4545. [PubMed: 16537396]
- (327). Toleman MA; Bennett PM; Walsh TR ISCR Elements: Novel Gene-Capturing Systems of the 21st Century? *Microbiol. Mol. Biol. Rev* 2006, 70, 296–316. [PubMed: 16760305]
- (328). Grindley NDF; Whiteson KL; Rice PA Mechanisms of Site-Specific Recombination. *Annu. Rev. Biochem* 2006, 75, 567–605 and references therein. [PubMed: 16756503]

Figure 1. Schematic of cut-and-paste transposition. The color scheme (orange = transposon DNA; black = transposon ends; blue = target DNA; pale green = target site duplication, TSD) is conserved throughout.

Figure 2. Schematic for copy-and-paste transposition between two replicons. In the center is shown the Shapiro intermediate that forms during replicative transposition. The green dashed arrows (left inset) indicate new DNA synthesis. Cointegrate resolution regenerates the original replicon and a copy of the transposon in the target plasmid flanked by TSDs.

Figure 3.
Schematic for copy-out–paste-in transposition.

Figure 4.

Chemical reactions catalyzed by DNA transposases. (A) DDE active site, based on structures of PFV intasomes.^{84–86} The green DNA represents the cleaved dinucleotide, and orange is the target strand. Spheres indicate bound metal ions. (B) HUH nuclease active site acting on single-stranded DNA.¹²⁹ Shown is the reaction for the Y1-transposase of IS*Dra2* at the transposon Left End (LE). The strand transfer joining reaction results in a sealed donor backbone. (C) The active site serine of a serine recombinase³²⁸ is surrounded by many Arg residues. Upon 180° rotation of one dimer within a tetramer, one strand rotates out of the active site (green) while another rotates in (orange). (D) Crucial residues within the active site of a tyrosine recombinase³²⁸ include a conserved Arg-His-Arg triad. Reproduced with permission from ref 77. Copyright 2015 American Society for Microbiology.

Figure 5.

Genetic organization of several bacterial transposons carrying antibiotic resistance genes. The *tnpA* genes encoding the transposase are shown in purple; for Tn 7, the heterotransposase is composed of the proteins encoded by both *tnpA* and *tnpB*. Tn.5 carries passenger genes encoding resistance to kanamycin (*kan*), bleomycin (*ble*), and streptomycin (*str*). Tn.10 encodes *jemA*, *jemB*, and *jemC* (not labeled) in addition to genes that encode resistance to tetracycline. The IS 10L *tnpA* gene encodes a nonfunctional truncated transposase. Tn.3 encodes its own resolvase encoded by *tnpR* and carries a passenger gene encoding a β -lactamase (*bla*). Five Tn 7 gene products participate in transposition (*tnpA/B/C/D/E*), and the transposon also encodes resistance to trimethoprim (*dhfr*), streptothricin (*sat*), and streptomycin/spectinomycin (*aadA*).

Figure 6. Schematic for single-stranded DNA (ssDNA) transposition catalyzed by *IS200/IS605* transposases. Reactions occur exclusively on ssDNA.¹⁸ The cleavage site at the Left End of the transposon is adjacent to a conserved tetra- or pentanucleotide sequence (light green). Subsequent integration of the circular single-stranded transposon intermediate occurs into a ssDNA target that contains the same conserved sequence.

Figure 7. Schematic of pathways used by different DNA transposases to generate double-strand breaks at transposon ends. NTS, nontransferred strand; TS, transferred strand. The cleavage site on the TS is always precisely at the transposon end; the position of the cleavage site on the NTS varies.

Figure 8. Generation of target site duplications (TSDs) upon staggered strand transfer. Shown is an example where the two sites of strand transfer occur five base pairs apart.

Figure 9.
piggyBac excision and integration.

Figure 10.

Tn5 transpososome. The DDE nuclease domain is shown in orange, the N-terminal domain is in blue, and the C-terminal dimerization domain is in purple. There is a four- β -strand insertion into the DDE domain, shown in red. The active site residues (D97, D188, and E326) are shown as balls and sticks. The 20-mer DNA representing the IR is in pale gray, and the sequence used is shown in the upper left, along with a schematic representation of the Tn5 transposon. PDB ID: 1MUH.

Figure 11.

Comparison of the domain organization of the *Mos1*, *Sleeping Beauty*, and *Himar1* transposases. The coloring scheme corresponds to that of the structure of the *Mos1* transpososome in Figure 12: the DD(E/D) domain is colored orange, the HTH domains (called “PAI” and “RED” in *Sleeping Beauty*) are in blue. Residues observed to be involved in dimerization of the *Mos1* transposase are colored purple.

Figure 12.

Mos1 transpososome. The DDD nuclease domain is shown in orange and the N-terminal domain (two HTH domains connected by a linker) is in blue. Residues involved in multimerization are in purple. The 28-mer DNA representing the IR is in pale gray, and the sequence used is shown on top, along with a schematic representation of the *Mos1* transposon. The bases shown in red correspond to flanking DNA, and the underlined TA is the TSD. PDB ID: 4U7B.

Figure 13.

Hermes transpososome. The DDE nuclease domains are shown in orange, and the insertion domain is in red. Intertwined dimerization domains are in purple. Each monomer in the *Hermes* octamer is bound to one TIR (sequence is shown at the upper left). The *Hermes* transposon Left End (LE) is 449 bp and the Right End (RE) is 464 bp, although the identical TIRs on each end are only 17 bp.¹⁵⁰ One dimer is indicated by the dashed oval, and is shown on the right. PDB ID: 4D1Q.

Figure 14.

Mu transpososome. The crystallized complex on the left contains four MuA monomers (encoded by the *A* gene), and two copies of the bacteriophage Right End (each of which carries two MuA binding sites, R1 and R2) that have inserted into a target 35-mer DNA (shown in green). On the right, two different ways a Mu monomer participates in transpososome organization are shown. The domain organization of MuA is shown at the bottom; the first N-terminal domain and the final C-terminal domain are missing in the crystallized construct. PDB ID: 4FCY.

Figure 15.

Overview of bacterial insertion library process. After DNA sequencing of the recovered DNA, mapping against the genome sequence distinguishes between those regions of the genome that are important for survival under the experimental condition (e.g., *gene1*) and those that are not (*gene2*). Adapted with permission from ref 194. Copyright 2013 Landes Bioscience.

Table 1.

DD(E/D) and Y1/Y2 Transposons and Transposases Discussed in This Review

	transposon class	
	DD(E/D)	Y1 or Y2
Prokaryotic		
	Tn5	IS200/IS605
	Tn10	IS91
	Tn3	
	Tn7	
	IS5376	
	bacteriophage MuA	
Related:	viral integrases	
Eukaryotic superfamily ^{21,25,46}		
<i>hAT</i>	<i>Ac/Ds</i>	
	<i>Tam3</i>	
	<i>Hermes</i>	
	<i>Tol2</i>	
<i>piggyBac</i>	<i>piggyBac</i>	
	<i>piggyBat</i>	
Related:	<i>PiggyMac</i>	
<i>Transib</i>	<i>Transib</i>	
Related:	RAG1	
Tc1/ <i>mariner</i>	<i>Mos1</i>	
	<i>Sleeping Beauty</i>	
	<i>Himar</i>	
	<i>Frog Prince</i>	
	<i>Minos</i>	
<i>P</i>	<i>Pelement</i>	
<i>Mutator</i> or <i>MuDr</i>	<i>Mutator (MuDR)</i>	
<i>CMC</i> or <i>CACTA</i> or <i>En/Spm</i>	<i>CACTA</i>	
	<i>En/Spm</i>	
<i>Helitron</i>		<i>Helitron</i>