Turun yliopisto
University of Turku

ESTABLISHMENT AND ENHANCEMENT OF Mu *IN VITRO* DNA TRANSPOSITION APPLICATIONS: NOVEL PROCEDURES AND NEW TOOLS

Elsi Pulkkinen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. A II osa - tom. 324 | Biologica - Geographica - Geologica | Turku 2016

University of Turku

Faculty of Mathematics and Natural Sciences
Department of Biology
Laboratory of Genetics
DNA recombination group

Supervised by

Professor Harri Savilahti
Docent Saija Haapa-Paananen
Department of Biology
University of Turku
Finland

Reviewed by

Professor Maia Kivisaar
Department of Genetics
University of Tartu
Estonia

Docent Mikko Frilander
Institute of Biotechnology
University of Helsinki
Finland

Opponent

Docent Marko Virta
Department of Food and Environmental Sciences
University of Helsinki
Finland

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-6591-5 (PRINT)
ISBN 978-951-29-6592-2 (PDF)
ISSN 0082-6979 (PRINT)
ISSN 2343-3183 (ONLINE)
Painosalama Oy - Turku, Finland 2016

TABLE OF CONTENTS

ABSTRACT

Transposable elements, or transposons, are DNA segments that have the ability to move from one location on the genome to another. They have been widely utilized as molecular tools in the study of prokaryotes, plants, invertebrates, and recently also in vertebrates. Custom-designed transposons are easy to generate, as any DNA fragment can be inserted between the transposon ends, and this flexibility makes them suitable for a number of different purposes, such as gene delivery and genome modification. One of the best-characterized DNA transposition systems is that of the Mu bacteriophage. The simplified *in vitro* reaction of Mu is widely used as a tool for genetic studies. Mu transposition is mediated by a DNA transposition complex, the Mu transpososome, which is assembled when four transposase proteins bind to sequence-specific binding sites at each transposon end. In this thesis, two novel Mu *in vitro* transposition-based applications were established, the first for measuring the transpositional activity of transposition complexes, and the second for the cloning of circular DNA. In addition, hyperactive MuA transposase variants were generated and their properties were characterized in the context of the function of the Mu transpososome.

The first part of this thesis presents a new assay for monitoring the activity of DNA transposition complexes. The assay is based on an *in vitro* transposition reaction with a special target plasmid, which allows positive selection of transposon insertions regardless of the nature of transposon used. Therefore, the assay is applicable to all types of transposons, and provides, for the first time, a common quality-control measure which will be essential for the evaluation of existing and future transposon-based applications. In this study we demonstrate the functionality of the assay with a set of Mu-based transposons. In addition, we show that the assay yields a linear response as a function of transposon DNA concentration, and also as a function of *Escherichia coli* competency.

The second part of this thesis introduces a new method that will potentially allow any circular DNA, such as uncharacterized plasmids and viral genomes, to be maintained in *E. coli*, thereby simplifying functional and structural studies. This method utilizes the Mu *in vitro* transposition system to deliver both the *E. coli* origin of replication and a selectable marker into DNA-circles, which are then transformed into *E. coli*. This study demonstrates the utility of the assay using a replication-deficient plasmid, and presents altogether ten novel transposons that can be used for this procedure, thus enabling multiple approaches for the study of DNA-circles of interest. In addition, this study presents a novel means to reduce intramolecular transposition, a phenomenon in which a transposon inserts into itself instead of an external target DNA. Here, we show that the amount of intramolecular transposition is affected by the ratio of transposon to target DNA. In addition, we show that the use of MuB, and the simultaneous use of two transposons, one containing the replication origin and other the selectable marker, can reduce this phenomenon.

Lastly, this thesis shows that it is possible to generate hyperactive variants of MuA transposase. Activity-enhancing substitutions were mapped on the Mu transpososome structure, which allowed us to identify the specific locations involved. Purified versions of the MuA variants improved the transposition efficiency of different Mu *in vitro* applications.

The methods established in this thesis are useful additions to the transposon toolbox. In particular, the hyperactive MuA variants that were generated here will substantially improve the use and development of *in vitro* applications based on the Mu system, which could potentially include the development of gene therapy applications. Furthermore, the results with hyperactive MuA variants can be used for structure-function studies of other similar transposases or close relatives, such as HIV integrase, creating opportunities for accelerated progress in this field of research.

TIIVISTELMÄ

Transposonit ovat DNA-jaksoja, joilla on luontainen kyky liikkua kromosomissa paikasta toiseen ja niitä on löydetty kaikista eliöryhmistä, bakteereista ihmisiin. Transposoneja on helppo muokata ja niitä voidaan hyödyntää monipuolisesti eri tarkoituksiin, kuten geeninsiirtoon ja genomien muokkaukseen. Transposoneihin perustuvia työkaluja käytetään sekä bakteereilla, kasveilla, selkärangattomilla, että selkärankaisilla. Bakteriofagi Mu hyödyntää elinkierrossaan transpositiota. Mu:n transpositiomekanismia on tutkittu paljon, ja koeputkessa tapahtuvaa yksinkertaistettua Mu:n transpositioreaktiota käytetään laajasti geeniteknologian työkaluna. Yksinkertaistetussa reaktiossa Mu:n transpositiokompleksi koostuu Mu-DNA:sta, jonka päihin on sitoutuneena neljä MuA-proteiinia. Väitöskirjatyöni tavoitteena oli kahden erilaisen Mu:n transpositioreaktioon perustuvan menetelmän kehittäminen. Ensimmäinen menetelmä oli tarkoitettu transpositiokompleksien tehokkuuden mittaamiseen ja toinen DNA-renkaiden monistamiseen *Escherichia coli* -soluissa transposonien avulla. Lisäksi työssäni tutkittiin Mu:n transpositioreaktion tehokkuuden parantamista muokkaamalla transpositionreaktiossa tarvittavasta MuA proteiinista eri versioita.

Ensimmäiseksi tässä työssä esitellään uusi menetelmä transpositiokompleksien aktiivisuuden mittaamiseen. Kehitetty menetelmä perustuu koeputkessa tapahtuvaan transpositioreaktioon, jossa hyödynnetään kohde-DNA:na erityistä plasmidia. Tämä plasmidi mahdollistaa transpositiotuotteiden valinnan riippumatta siitä mitä ominaisuuksia transposoni sisältää. Täten kehitetty menetelmä soveltuu kaikille transposoneille ja on merkittävä parannus jo olemassa olevien ja kehitteillä olevien transpositioon perustuvien menetelmien arvioimiseen. Työssä todistetaan menetelmän toimivuus erilaisten Mu-transposonien avulla. Lisäksi työssä osoitetaan että menetelmä antaa vertailukelpoisia tuloksia riippumatta transposoni-DNA:n määrästä, tai transformaatiossa käytettävien *Escherichia coli* -solujen transformaatiotehokkuudesta.

Toinen työssä esiteltävä uusi menetelmä mahdollistaa periaatteessa minkä tahansa rengasmaisen DNA:n, kuten vielä tuntemattoman plasmidin tai virus-DNA:n monistamisen ja säilyttämisen *E. coli* -soluissa. Kehittämässämme menetelmässä hyödynnetään Mu:n koeputkessa tapahtuvaa transpositioreaktiota viemään DNA-renkaisiin replikaation aloituskohta ja antibioottivalinnan mahdollistava geeni, jonka jälkeen reaktiotuotteet viedään *E. coli* -soluihin. Tässä työssä osoitetaan menetelmän toimivuus käyttäen plasmidia, joka ei itse pysty replikoitumaan käytetyissä *E. coli* soluissa ja esitellään kymmenen uutta transposonia, jotka mahdollistavat DNA-renkaiden monipuolisen tutkimisen. Johtuen transposonien sisältämästä replikaation aloituskohdasta, menetelmä tuottaa myös jonkun verran taustapesäkkeitä, jotka eivät sisällä haluttua DNA-rengasta vaan aiheutuvat siitä, että transposoni on ulkopuolisen kohde DNA:n sijasta liittynyt itseensä (intramolekulaarinen transpositio). Tässä työssä osoitetaan että transposonin ja DNA-renkaiden välinen suhde vaikuttaa intramolekulaarisen transposition määrään. Lisäksi osoitetaan että käyttämällä MuB-proteiinia tai yhtäaikaisesti kahta transposonia (joista toinen sisältää replikaation aloituskohdan ja toinen antibioottivalinta geenin) voidaan vähentää intramolekulaarisen transposition määrää.

Väitöskirjatyössä osoitetaan lisäksi, että MuA transposaasista voidaan kehittää hyperaktiivisia muotoja. Työssä selvitetään myös transpositiotehokkuutta lisäävien muutosten sijainti suhteessa Mu transpositiokompleksin rakenteeseen ja osoitetaan, että muokattujen MuA-proteiinien avulla voidaan parantaa Mu:n transpositioon perustuvien menetelmien tehokkuutta.

Työssä kehitetyt menetelmät ovat tarpeellinen lisä transposoneja hyödyntävien menetelmien joukkoon. Lisäksi kehitetyt hyperaktiiviset MuA-muodot parantavat merkittävästi Mu:n transpositioon perustuvien menetelmien käytön tehokkuutta, mistä on hyötyä muun muassa geeniterapiaan liittyvien sovellusten kehityksessä. Muokattujen MuA-proteiinien avulla saatuja tuloksia voidaan hyödyntää myös tutkiessa muita samankaltaisia proteiineja, joihin kuuluu myös muun muassa HI-viruksen integraasi.

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred in the text by their roman numerals:

I. Pulkkinen E., Haapa-Paananen S., Savilahti H. An assay to monitor the activity of DNA transposition complexes yields a general quality control measure for transpositional recombination reactions. Mobile genetic elements 2014; 4:1-8.

II. Pulkkinen E., Haapa-Paananen S., Savilahti H. MuA-mediated in vitro cloning of circular DNA: transpositional autointegration and the effect of MuB. Molecular genetics and genomics 2016; 291: 1181-1191.

III. Pulkkinen E., Haapa-Paananen S., Turakainen H., Savilahti H. A set of mini-Mu transposons for versatile cloning of circular DNA and novel dual-transposon strategy for increased efficiency. Plasmid 2016; 86 46-53.

IV. Rasila T.S., Pulkkinen E., Haapa-Paananen S., Kiljunen S., Vihinen M., Paulin L., Rice P.A., Savilahti H. Activity-probing of Mu transposition recombination machinery and generation of hyperactive MuA transposase variants. Manuscript.

Article II has been reprinted with permission of Springer and article III with permission of Elsevier.

Unpublished data will also be presented.

ABBREVIATIONS

aa	amino acid
Ap	ampicillin
ATP	adenosine triphosphate
BAF	barrier-to-autointegration factor
bp	base pair(s)
BSA	bovine serum albumin
CDC	cleaved donor complex
cDNA	complementary deoxyribonucleic acid
Cm	chloramphenicol
C-terminal	carboxy-terminal
CRISPR	Clustered regularly-interspaced short palindromic repeats
DD(E/D)	domain which contain a triad of highly conserved amino acids [aspartate (D), aspartate, and glutamate (E), or third aspartate]
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DSB	double stranded break
ep	error prone
ERV	endogenous retrovirus
HIV	human immunodeficiency virus
HPLC	high-pressure liquid chromatography
HTH	heliz-turn-helix
HUH-transposase	tyrosine-histidine-hydrophobic-histidine transposase
IAS	internal activating sequence
ICE	integrative and conjugative element
IHF	*E.coli* integration host factor
IPTG	isopropyl-β-D-1-thiogalactopyranoside
IS	insertion sequence

LE	left end
LER	left end-enhancer-right end; synaptic complex in Mu transposition
LINE	long interspersed repeated element
LTR	long terminal repeat
MITE	miniature inverted-repeat transposable element
MoMLV	Moloney Murine Leukemia Virus
MuA	bacteriophage Mu transposase protein A
MuB	bacteriophage Mu transposase protein B
NMR	nuclear magnetic resonance
N-terminal	amino-terminal
OH	hydroxide
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PLE	Penelope-like element
RE	right end
RNA	ribonucleic acid
RNaseH	ribonuclease H
SDS	sodium dodecyl sulphate
SINE	short interspersed repeated element
SNP	single-nucleotide polymorphism
SSC	stable synaptic complex
TE	transposable element
TIR	terminal inverted repeat
TriNex	trinucleotide change
YR –element	tyrosine recombinase encoding element
V(D)J	variable (diversity) joining; antigen receptor gene segments
wt	wild type
X-gal	5-bromo-4-chloro-3-inodyl-β-D-galactopyranoside

1 INTRODUCTION

1.1 Transposable elements are major components of genomes

Transposable elements (TEs), or transposons, are defined segments of DNA that have the ability to move or copy themselves into another location within and between genomes, even if those locations are non-homologous (reviewed by Craig et al. 2002; Craig et al. 2015). Transposons were first identified in the 1940s by Barbara McClintock through experimentation with maize (reviewed by McClintock 1987). She proposed that genomes were dynamic and subject to rearrangements and alterations. These radical suggestions were met with skepticism and transposons were long thought of as useless or junk DNA. Since then, however, knowledge of transposons has vastly increased and, far from being considered junk DNA, they are now viewed as major components of genomes which play an important role in evolution, particularly by controlling gene activity (reviewed by Biemont and Vieira 2006).

A variety of TEs have been found in all three domains of life (Aziz et al. 2010; Feschotte and Pritham 2007; Filee et al. 2007; Siguier et al. 2014), and the number of identified elements is continuously growing due to the advent of high-throughput genome sequencing. TEs can make up a large proportion of a genome, for instance, 45% of the human genome (Lander et al. 2001), 85% of the maize genome (Schnable et al. 2009) and 37.5% of the mouse genome (Waterston et al. 2002). TEs tend to be more abundant in larger versus smaller genomes; however, no straightforward relationship between genome size and TE diversity (number of TE types) has been found (reviewed by Elliott and Gregory 2015).

Transposons can influence the evolution of their host by altering gene function; this can occur through insertion, by inducing chromosomal rearrangements, and by providing novel genetic material (reviewed by Feschotte and Pritham 2007; Rebollo et al. 2012). A prime example of beneficial transposon-derived sequences that have been co-opted by the host is found in the *RAG* genes, which are involved in immunoglobulin V(D)J joining (reviewed by Fugmann 2010; Gellert 2002). Most transposon insertions in genomes are molecular fossils that have been inactivated by rearrangements or mutations, and are incapable of further transposition. However, some transposons are actively mobile and cause new insertions. New insertions are most often benign, extremely rarely beneficial, and occasionally harmful, with the potential to cause various diseases, particularly cancers (Lee et al. 2012; reviewed by Goodier and Kazazian 2008).

1.2 Classification of transposable elements

Today, knowledge regarding transposable elements, such as their transposition mechanisms and diversity, has increased dramatically. However, somewhat paradoxically, the classification of these elements has become more challenging. Elements have been divided into classes based on their transposition mechanisms, sequence similarities, and structural relationships, and these classes have been further be divided into various subgroups, such as subclasses, orders, superfamilies, families, and subfamilies (reviewed by Piegu et al. 2015; Wicker et al. 2007).

The first TE classification system separated TEs into two main classes based on their transposition intermediate: RNA (class I or retrotransposons) or DNA (class II or DNA transposons) (reviewed by Finnegan 1989). In Wicker's proposal, presented in Table 1, these two main classes are further subdivided into subclasses, orders, and superfamilies (reviewed by Wicker et al. 2007). Wicker's proposal (reviewed by Wicker et al. 2007), and the similar Repbase proposal (reviewed by Jurka et al. 2005; Kapitonov and Jurka 2008), have been widely used in TE studies. However, critics have identified certain weaknesses in these approaches, such as the fact that they focus only on eukaryotic elements, excluding prokaryotic transposons, and that the classification fails to reflect phylogeny and evolutionary relationships consistently (reviewed by Piegu et al. 2015; Seberg and Petersen 2009).

Table 1. Wicker's proposal for the classification of eukaryotic TEs (Wicker et al. 2007)

	Order	**Superfamilies**
Class I (retrotransposons)		
	LTR	*Ty, Copia, Gypsy, Bel-Pao, Retrovirus, ERV*
	DIRS	*DIRS, Ngaro, VIPER*
	PLE	*Penelope*
	LINE	*R2, RTE, Jockey, L1, I*
	SINE	*tRNA, 7SL, 5S*
Class II (DNA transposons) –subclass I		
	TIR	*Tc1-Mariner, hAT, Mutator, Merlin, Transib, P, PiggyBac, PIF-Harbinger, CACTA*
	Crypton	*Crypton*
Class II (DNA transposons) –subclass II		
	Helitron	*Helitron*
	Maverick	*Maverick-Polinton*

1.2.1 Retrotransposons

Retrotransposons (class I elements) act through an RNA intermediate, moving via a "copy and paste" mechanism. This involves transcription of an RNA intermediate from

a genomic copy, reverse transcription of the RNA intermediate by an element-encoded reverse transcriptase enzyme, and insertion of the cDNA copy at a new genome site (reviewed by Curcio and Derbyshire 2003; Goodier and Kazazian 2008). Retrotransposons are often highly abundant in eukaryotic genomes, especially in mammals (reviewed by Goodier and Kazazian 2008). Class I elements can be divided into long-terminal-repeat (LTR) elements (for example, endogenous retrovirus (ERV) elements and Ty elements of *Saccharomyces cerevisiae*) and non-LTR retrotransposons (for example, long interspersed nuclear elements (LINEs), or short interspersed nuclear elements (SINEs) such as *Alu* elements and bacterial and organellar group II introns) (reviewed by Beauregard et al. 2008; Wicker et al. 2007). In addition to the traditional class I elements, there are also two novel groups of retrotransposons: tyrosine recombinase (YR) -encoding elements (for example, *DIRS*-like elements) and Penelope-like elements (PLEs) (reviewed by Wicker et al. 2007).

As their name suggests, LTR retrotransposons have long terminal repeats at their ends, and they resemble retroviruses both structurally and mechanistically, differing mainly in the fact that retroviruses can form infectious virions that allow them to move horizontally from cell to cell, while LTR retrotransposons can only move between different sites within the genome of a single cell (reviewed by Beauregard et al. 2008). Instead, non-LTR retrotransposons do not contain LTRs and integrate into the genome using target-primed reverse transcription, in which cleaved DNA targets are used to prime reverse transcription of the element's RNA intermediate into cDNA (reviewed by Beauregard et al. 2008). For integration into the genome, LTR retrotransposons employ integrases, YR elements utilize tyrosine recombinases, and PLEs and autonomous non-LTR retrotransposons use endonucleases. Non-autonomous elements, of which a prominent class is SINEs, require LINEs for their propagation (reviewed by Okada et al. 1997; Richardson et al. 2015).

1.2.2 DNA transposons

1.2.2.1 Eukaryotic DNA transposons

DNA transposons (Class II elements) act through either a double- or single-stranded DNA intermediate (reviewed by Curcio and Derbyshire 2003; Hickman and Dyda 2015). Based on their transposition mechanism, eukaryotic DNA transposons can be divided into four groups: i) classic "cut-and-paste" transposons (for example, *hAT*, *Tc1/mariner*, *P*, and *piggyBac* elements), ii) rolling circle transposons (*Helitrons*), iii) self-replicating transposons, which use a self-encoded DNA polymerase (*Mavericks*/*Polintons*), and iv) tyrosine recombinase transposons (*Cryptons*) (reviewed by Feschotte and Pritham 2007; Poulter and Goodwin 2005). Both *Helitrons* and *Mavericks*/*Polintons* probably transpose through a replicative copy-and-paste process. Eukaryotic DNA transposons also include non-autonomous elements, such as miniature inverted repeats (reviewed by Fattash et al. 2013; Feschotte et al. 2002).

1.2.2.2 Prokaryotic DNA transposons

Although found in all kingdoms of life, DNA transposons are particularly prevalent in bacteria. Examples of prokaryotic transposons and their end organization are presented in Figure 1. The simplest autonomous prokaryotic DNA elements are insertion sequences (ISs), which contain a transposase gene flanked by inverted terminal repeats (IRs) (reviewed by Chandler and Mahillon 2002; Siguier et al. 2014; Siguier et al. 2015). Composite transposons (for example, Tn*5* and Tn*10*) contain a pair of insertion elements (IS*50* and IS*10*, respectively) which flank genetic information for antibiotic resistance or other properties (reviewed by Haniford 2002; Haniford and Ellis 2015). Complex transposons (for example Tn*3*) are not terminated with complete IS elements and may encode for several gene products (reviewed by Grindley 2002; Nicolas et al. 2015). Integrative and conjugative elements (for example, Tn*916*) can integrate into the chromosome of their host cell and eventually self-transfer to other bacteria via conjugation (reviewed by Wozniak and Waldor 2010). The most elaborate prokaryotic elements are transposing bacteriophages, such as phage Mu, which use transposition in their lifecycles (reviewed by Chaconas and Harshey 2002; Harshey 2014). A novel class of prokaryotic transposons, casposons, has been recently proposed (Krupovic et al. 2014). Casposons resemble eukaryotic self-synthesizing *Mavericks*/*Polintons* elements; their name derives from the transposase (integrase) protein that they use, which is related to Cas1, a protein component of the bacterial adaptive immunity system CRISPR-Cas.

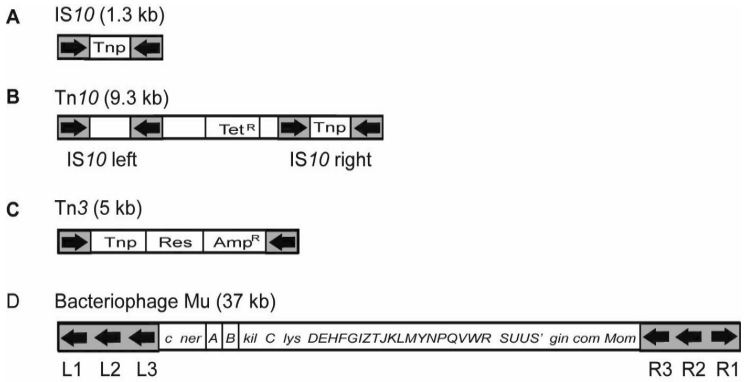

Figure 1. Examples of prokaryotic transposons. Terminal repeats are denoted by arrows. **A.** Insertions sequence IS*10* has one open reading frame which encodes the transposase (Tnp). **B.** Composite transposon Tn*10* has IS*10* elements at both ends flanking a gene that encodes for tetracycline resistance. Only one copy of IS*10* encodes an active transposase. **C.** Complex transposon Tn*3* has three open reading frames, which encode for transposase, site-specific resolvase and ampicillin resistance. **D.** Bacteriophage Mu is both a transposon and a bacteriophage. It encodes for two transposition proteins, A and B. Other genes are involved in regulation of transposition, lysis, head and tail genes, and host range. Drawn according to (Braid et al. 2004; Leach 1996).

1.3 Transposases

Transposases move DNA transposons to new locations, without the need for sequence homology between the transposon and the target site. Transposases contain a catalytic domain that mediates the breakage and joining of DNA; based on the type of the catalytic domain present, they can be divided in four major classes: aspartate-aspartate-glutamate (DDE) transposases, tyrosine-histidine-hydrophobic-histidine (HUH) transposases, tyrosine transposases, and serine transposases (reviewed by Curcio and Derbyshire 2003; Hickman and Dyda 2015). DDE transposases integrate a transposon to a new site via direct transesterification, whereas the other transposase classes use covalent-protein DNA intermediates. Transposases belonging to the novel Cas1 class use the same chemical mechanism for DNA joining as DDE transposases do, but have no structural homology with the latter class (Krupovic et al. 2014).

1.3.1 DDE transposases

Although all the above-mentioned major classes of transposases can be found in all kingdoms of life, the largest class is that of the DDE transposases. Members of this group contain the DD(E/D) domain, also called the RNase H-like fold, which is the most common catalytic nuclease domain fold (reviewed by Hickman and Dyda 2015; Montano and Rice 2011). The DD(E/D) domain contains a triad of highly conserved amino acids [aspartate (D), aspartate, and either glutamate (E) or a third aspartate], which are crucial for the coordination of metal ions required for catalysis (Kulkosky et al. 1992; reviewed by Hickman et al. 2010). These metal ions (most likely Mg^{2+} *in vivo*) are crucial cofactors for two chemical reactions catalyzed by DDE transposases: DNA strand cleavage and strand transfer (reviewed by Hickman and Dyda 2015). In the first chemical step, hydrolysis of the phosphodiester backbone at the transposon ends generates free 3′-OH groups. In the second step, these liberated 3′-OH groups are joined to the target DNA in a transesterification reaction. The nucleophilic attack on the target DNA normally occurs at staggered positions, generally separated by 2-9 nucleotides, with the result that short single-stranded segments of host DNA flank the inserted transposon (reviewed by Curcio and Derbyshire 2003). Repair of these single-stranded gaps by the host enzymes results in target duplication, the length of which is characteristic for each transposon.

1.3.2 Replicative and non-replicative transposition pathways

All DDE transposases catalyze the same chemical reactions which are needed for DNA strand cleavage and strand transfer. However, the detailed mechanisms of the transposition pathways are transposase-specific. The transposition strategy has traditionally been divided into two types: replicative or non-replicative (“cut and paste”) (reviewed by Hickman and Dyda 2015; Turlan and Chandler 2000). Transposition pathways of DDE transposases are presented in Figure 2.

1.3.2.1 Replicative pathway

In the replicative pathway, a new copy of the transposon is generated, and as an end result, one transposon copy appears at the new site and one copy remains at the old site. The replicative pathway can be further divided into two different types, nick-paste-copy and nick-copy-out-paste (reviewed by Curcio and Derbyshire 2003; Hickman and Dyda 2015) (Figure 2.). In the nick-paste-copy pathway, only the 3′-ends of the transposon are nicked and joined to the target, while the transposon 5′-ends remain attached to the donor molecule. Fusion of the cleaved 3′-ends to a target molecule forms a strand transfer intermediate, also called a Shapiro intermediate (Shapiro 1979). Replication from the 3′-OHs in the flanking target DNA generates a cointegrate molecule, which consists of donor and target DNA joined by two copies of the transposon. The cointegrate molecule can be resolved by a recombination reaction, either through homologous recombination between the transposon copies or by a site-specific recombination system encoded by the element, which yields a donor and recipient replicon which each have a copy of the transposon. Examples of transposable elements that utilize the nick-paste-copy pathway are bacteriophage Mu (reviewed by Chaconas and Harshey 2002; Harshey 2014) and members of the Tn*3* family (reviewed by Grindley 2002; Nicolas et al. 2015). In the nick-copy-out-paste pathway, the transposase asymmetrically nicks only one 3′-end of the transposon and the released 3′-OH group joins intramolecularly to just outside of its own 5′-end, thereby circularizing one strand of the transposon (reviewed by Curcio and Derbyshire 2003; Hickman and Dyda 2015). Replication of the 3′-OH of the flanking target DNA copies both transposon strands, resulting in the release of the double-stranded transposon circle and the repair of the donor site. The transposase cleaves the transposon ends, generating two 3′-OH which are then used to attack the target DNA. Examples of elements utilizing the nick-copy-out-paste pathway are IS*911* and other members of the IS*3* family (Sekine et al. 1994; Ton-Hoang et al. 1997; reviewed by Chandler et al. 2015). Retroviral integrases of LTR-retrotransposons are also DDE transposases, and they utilize a copy-out-paste pathway (reviewed by Curcio and Derbyshire 2003) (not shown in Figure 2.). LTR-retrotransposons generate a copy of their genome by transcription, which is followed by the generation of cDNA by reverse transcription. Either the 3′-ends of this cDNA contain a terminal CA dinucleotide (the 3′-CA dinucleotide is conserved among retrotransposons and retroviruses), or the DDE transposase introduces nicks at each end of the proviral DNA, exposing 3′-CA dinucleotides that are joined to the target (Engelman et al. 1991; Vink et al. 1991).

1.3.2.2 Non-replicative pathway

In the non-replicative (cut-and-paste) pathway, the transposon is excised from the original site and integrated into the new site (reviewed by Hickman et al. 2010; Hickman and Dyda 2015; Turlan and Chandler 2000) (Figure 2.). As the initial hydrolysis reaction of DDE transposases catalyzes the cleavage of only one strand at the transposon ends, elements utilizing non-replicative transposition have developed a

variety of strategies to cleave the second strand and thereby introduce a double-strand break (DSB) in order to excise themselves from the donor DNA (reviewed by Hickman et al. 2010; Turlan and Chandler 2000). Perhaps the simplest strategy to excise the transposon is to introduce DSBs by two sequential strand cleavages on opposite strands, a strategy used, for example, by members of the *Tc1/mariner* family (Dawson and Finnegan 2003). Another strategy is found in the bacterial insertion sequences IS*10* and IS*50*, which use a hairpin formation on the transposon ends for excision (Bhasin et al. 1999; Kennedy et al. 1998). The hairpin is formed when the cleaved 3′-OHs at the transposon ends attack the complementary 5′-ends of the transposon. Before insertion into the target DNA, the hairpin is linearized by hydrolysis at the transposon 3′-ends. If the transposon is first nicked at the 5′-ends, the hairpin formation occurs at the flanking donor DNA. This type of hairpin formation is utilized, for example, by members of the *hAT* family of eukaryotic transposons (for example the *Hermes* transposon) and the V(D)J recombination RAG1/2 recombinase (Zhou et al. 2004; reviewed by Schatz and Swanson 2011). Instead, transposon Tn*7* utilizes two proteins to liberate itself from the donor DNA: the TnsB protein for 3′-cleavage and TnsA for 5′-cleavage (Sarnovsky et al. 1996; reviewed by Peters 2014). The Tn*7* transposition pathway also demonstrates the requirement for DSBs in non-replicative transposition: if 5′-cleavage is prevented by mutations in TnsA, Tn*7* is turned into a replicative transposon (May and Craig 1996).

Excision of the non-replicative element from the donor DNA, and possibly the insertion of the element into target DNA, generates gaps which must be repaired. For most transposons, these repair steps are still unclear. Among prokaryotes, repair of the donor site after transposon excision can be performed utilizing homologous recombination, provided that a second copy of the chromosome exists (Hagemann and Craig 1993). It was recently shown that for the non-replicative pathway of Mu, the gaps created following Mu insertion are repaired by the *E. coli* machinery for double-strand break repair, which involve both the replication restart proteins and homologous recombination proteins (Jang et al. 2012).

Among eukaryotes, the double-stranded break caused by excision from the donor site can be repaired by non-homologous end joining (Robert et al. 2008; Staveley et al. 1995; reviewed by Plasterk and van Luenen 2002). However, particularly if transposition occurs during replication, the double-stranded break can also be repaired by homologous recombination with the sister chromatid (Bender et al. 1991; Engels et al. 1990). For both prokaryotes and eukaryotes, although the transposon per se may be non-replicative, repair via homologous recombination can lead to an increase in transposon copy number in the genome.

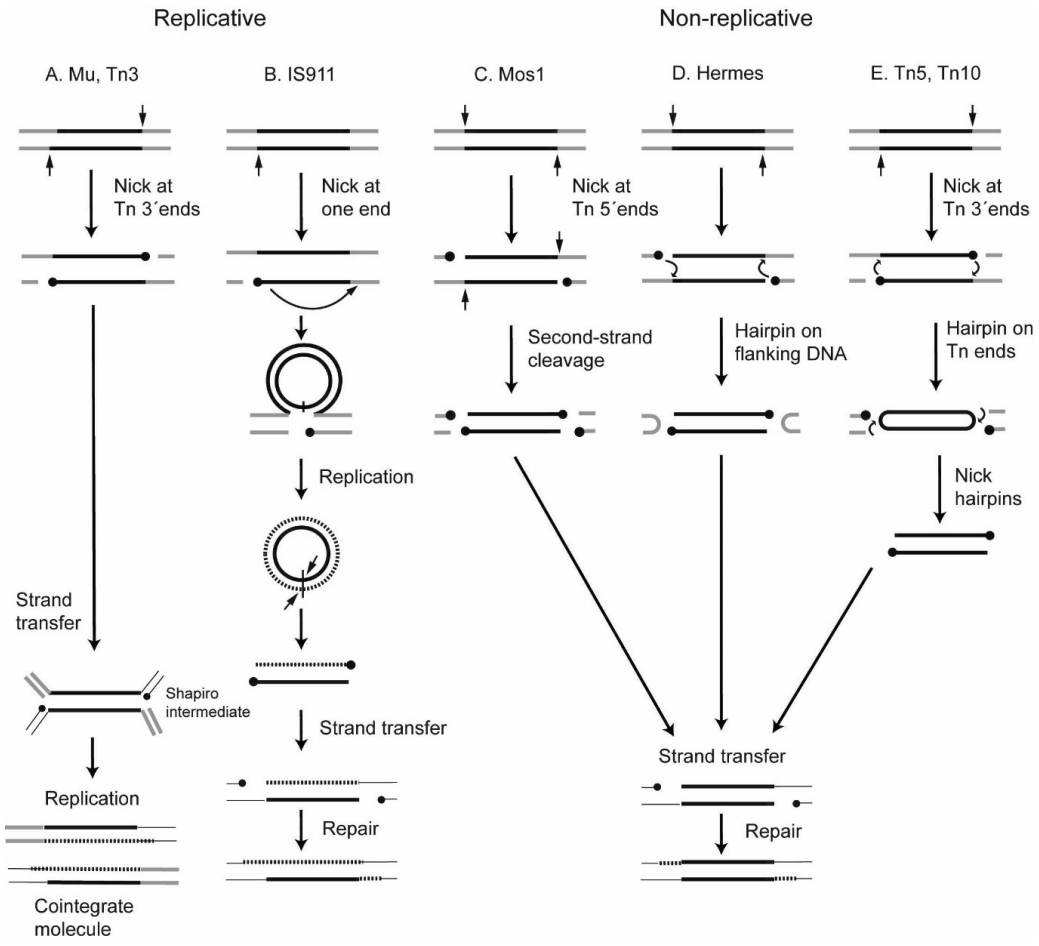

Figure 2. Transposition pathways of DDE transposases. The mechanisms that are used to generate a DNA substrate for strand transfer differ between the DDE transposases. **A.** Mu and Tn3-like elements utilize transposases that nick and join the 3´ends of the transposon to the target, which forms a strand-transfer, or Shapiro, intermediate. Replication from the 3´OHs in the flanking target DNA generates a cointegrate molecule. **B.** IS*911* and other IS3-like elements family members utilize transposases that nick only one 3´end of the transposon and the released 3´OH group joins intramolecularly to just outside of its own 5´end, thereby circularizing one strand of the transposons. Replication results to a release of double stranded transposon circle and repaired donor site. Second round of transposase cleavage generates a linear transposon. **C-E.** Non-replicative transposons utilize transposases that excise a double-stranded copy of the transposon and result in simple insertion. All DDE-transposases attack the target DNA at staggered positions, and when the cell repairs the resulting short single stranded segments of host DNA flanking the inserted transposon it generates target site duplications at both transposon ends. Thick black lines represent transposon DNA; thick grey lines represent flanking host DNA; thin black lines represent target DNA; dashed lines represent newly replicated DNA; circles at the ends of DNA represent exposed 3´OH groups. Drawn according to (Curcio and Derbyshire 2003; Hickman and Dyda 2015).

1.4 Transposable bacteriophage Mu

Mu is both a transposable element and a temperate bacteriophage of *Escherichia coli* and other Gram-negative bacteria. Mu was the first mobile element found in *E. coli*, and it has played a major role in the development of the study of mobile DNA elements (Taylor 1963; reviewed by Harshey 2012). Soon after its discovery, Taylor (1963) noted that it could induce mutations at a high rate and therefore named it Mu for mutator. Mu was the first TE for which an *in vitro* transposition reaction was established (Mizuuchi 1983). This has been invaluable to the study of its reaction chemistry and of the function of the participating proteins (reviewed by Chaconas and Harshey 2002; Harshey 2014).

1.4.1 Transposition mechanism of phage Mu

Mu uses transposition at two different stages during its lifecycle (reviewed by Chaconas and Harshey 2002; Harshey 2014). Upon lysogenization, Mu uses non-replicative transposition to integrate randomly into the bacterial chromosome; then, during the lytic stage, it uses replicative transposition to replicate its DNA. In both stages the transposition mechanism is the same, but the pathway for product resolution differs. First, the transposase protein (MuA) pairs the Mu genome ends together to form a transpososome. At both ends, MuA catalyzes hydrolytic nicking at the Mu-host junction and, during the subsequent strand-transfer reaction, the element's freed 3′-OHs attack the target DNA. The attack at the target DNA occurs at nearly random sites and at staggered positions, separated by five nucleotides. The strand-transfer reaction produces branched intermediates which must be resolved. During the lytic (replicative) stage, the intermediate is resolved by target-primed replication, leading to the duplication of the entire Mu genome. During the initial lysogen formation (cut-and-paste transposition), both strands at each genome-end are cleaved, leading to a simple insertion without duplication. Because the target DNA is cut at staggered positions, single-stranded gaps are created in the host DNA which flank the inserted transposon. Repair of these gaps by the host results in 5-bp target site duplications, the presence of which is a hallmark of Mu transposition (Allet 1979; Kahmann and Kamp 1979; Mizuuchi 1992).

1.4.2 DNA and protein requirements in Mu transposition

Bacteriophage Mu has a double-stranded DNA genome 36,717 base pairs in length (Morgan et al. 2002). The Mu DNA sites and proteins involved in Mu transposition are presented in Figure 3. The left (L) and right (R) ends each contain three MuA-binding sites (L1–L3 and R1–R3, respectively) (Craigie et al. 1984; Groenen and van de Putte 1986; Zou et al. 1991). Three additional MuA-binding sites are also found about 1 kb away from the left end, in the enhancer (E) segment (also called internal activating sequence, or IAS). This region was originally identified as the operator sequence, and

hence the binding sites are named O1-O3. Later it was found that this sequence also enhances transposition (Leung et al. 1989; Mizuuchi and Mizuuchi 1989; Surette et al. 1989). The L and R ends are asymmetric and the binding sites differ in orientation and spacing. Between L1 and L2 exists a binding site for the *E. coli* protein HU, which bends Mu DNA to allow the assembly of transpososomes (Lavoie et al. 1996). Between O1 and O2 there is a binding site for the *E. coli* integration host factor (IHF) protein, which optimizes the assembly of transpososomes (Leung et al. 1989; Mizuuchi and Mizuuchi 1989; Surette et al. 1989). These components are sufficient for executing the DNA cleavage reaction in the presence of the divalent metal ion Mg^{2+}, but the strand-transfer reaction additionally requires phage-encoded MuB. MuB is an ATP-dependent protein that binds DNA non-specifically (Adzuma and Mizuuchi 1988; Adzuma and Mizuuchi 1989; Maxwell et al. 1987). It directs target-site selection and also interacts with MuA to optimize all stages of transpososome assembly (Coros et al. 2003; Mizuno et al. 2013).

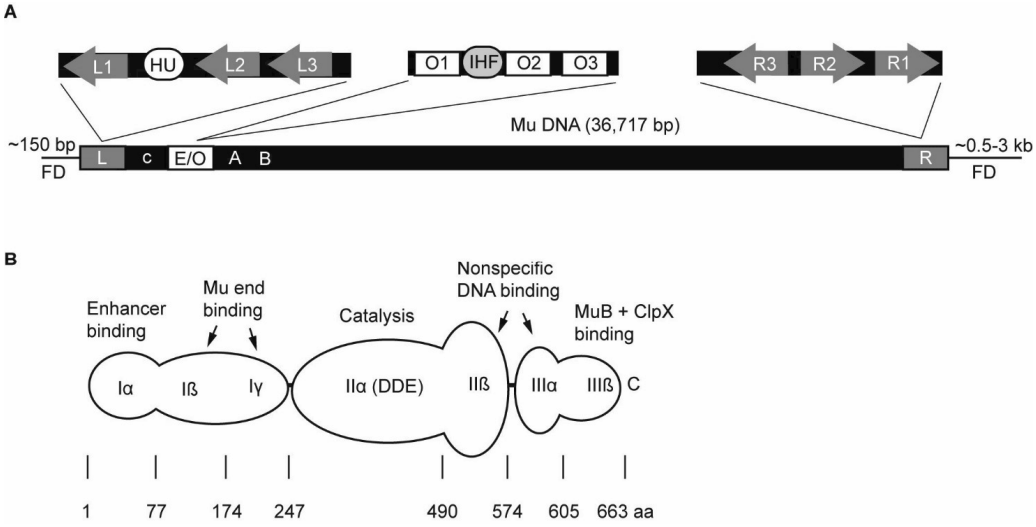

Figure 3. DNA and protein requirements for Mu transposition. **A.** Arrangement of three MuA binding sites at both ends (L1-L3 and R1-R3) of the Mu genome, and also at enhancer (E) segment (O1-O3). Enhancer segment is also labeled O because enhancer and operator (O) region overlap. Binding sites for *E.coli* bending protein HU and *E.coli* integration host factor (IHF) within L and E are shown. The genes regulating transposition (*c*, *A*, *B*) are indicated. Flanking DNA (FD) on both sides of the Mu genome is packaged into virions. Drawn according to (Akhverdyan et al. 2011; Harshey 2014). **B.** Domain and subdomain organization of MuA. The function of each subdomain is indicated. Drawn according to (Harshey 2014; Montano et al. 2012).

1.4.2.1 Minimal requirements for Mu in vitro transposition

While transpososome assembly is normally controlled by several cofactors, the required components can be substantially simplified using modified reaction conditions, particularly with the inclusion of dimethylsulfoxide (DMSO) and the utilization of pre-cut donor DNA (Baker and Mizuuchi 1992; Craigie and Mizuuchi

1986; Mizuuchi and Mizuuchi 1989; Savilahti et al. 1995). Under relaxed reaction conditions, transposition is possible without DNA supercoiling, the enhancer sequence, the MuA domain Iα (N-terminal enhancer-binding domain), HU, IHF, MuB, and the MuA domain IIIβ (C-terminal MuB/ClpX-binding domain) (Savilahti et al. 1995) (Domain organization of MuA is shown in Figure 3). In a minimal setup, the only requirements are a simple buffer, MuA transposase, transposon DNA in a precut linear form that contains only the R1 and R2 MuA binding sites at both ends, and target DNA (Haapa et al. 1999b; Savilahti et al. 1995).

1.4.3 Structure of MuA transposase

MuA is a large (75 kDa, 663 amino acids) protein with a modular organization (Nakayama et al. 1987). It can be divided into three major domains (I, II, III) and seven subdomains (Iα, Iβ, Iγ, IIα, IIβ, IIIα, IIIβ) based on different structural and functional properties (Figure 3B). MuA domain and subdomain organization was originally determined by partial proteolysis (Nakayama et al. 1987). Later, structures of the individual subdomains (except IIIβ) were defined by nuclear magnetic resonance (NMR) or by X-ray crystallography (Clubb et al. 1994; Clubb et al. 1997; Montano et al. 2012; Rice and Mizuuchi 1995; Schumacher et al. 1997). Recently, this latter technique was also used to determine the structure of a Mu transpososome joined to target DNA that contained nearly full-length MuA, lacking only the N- and C-terminal domains Iα and IIIβ (Montano et al. 2012).

The N-terminal I domain contains the Iα subdomain, which binds to the transposon enhancer sequence, and subdomains Iβ and Iγ, which bind to the specific MuA-binding sites at the transposon ends (Clubb et al. 1994; Clubb et al. 1997; Kim and Harshey 1995; Leung et al. 1989; Schumacher et al. 1997). The central II domain contains the catalytic IIα subdomain and subdomain IIβ. Subdomain IIα contains a DDE residue triad (at positions Asp269, Asp336, and Glu392), which is crucial in the catalysis of strand cleavage and strand transfer (Baker and Luo 1994; Kim et al. 1995; Krementsova et al. 1998; Rice and Mizuuchi 1995). Subdomain IIβ has a large, positively charged region exposed on its surface (Rice and Mizuuchi 1995). This participates in binding nonspecific DNA (including DNA that flank the cleavage site and target DNA) (Krementsova et al. 1998; Montano et al. 2012), is important at the transition between the strand-cleavage and strand-transfer steps of transposition (Krementsova et al. 1998), and has been implicated to in the metal-assisted assembly of the MuA tetramer and in intramolecular DNA strand transfer (Namgoong et al. 1998).

The C-terminal III domain contains two subdomains, IIIα and IIIβ. The IIIα subdomain has a role in transpososome assembly (Mariconda et al. 2000; Naigamwalla et al. 1998); it binds DNA nonspecifically and contributes to the capture of target DNA (Montano et al. 2012). The IIIα subdomain also has cryptic nuclease activity which

might be involved in removing the flanking host DNA after initial integration (Choi and Harshey 2010; Montano et al. 2012; Wu and Chaconas 1995). Alternatively, this removal might also be accomplished by IIIα delivering the uncleaved strand into the catalytic site (subdomain IIα) for hydrolysis (Montano et al. 2012). The IIIβ subdomain is essential for interaction with the MuB and ClpX proteins (Levchenko et al. 1997; Wu and Chaconas 1994). Phage-encoded MuB is required for target capture during intermolecular transposition (Maxwell et al. 1987; Mizuno et al. 2013), and it also interacts with MuA to optimize all stages of transpososome assembly (Baker et al. 1991; Mizuuchi et al. 1995; Surette et al. 1991). The host encoded ClpX protein is needed for the remodeling or disassembly of the transpososome after strand transfer and before replication (Levchenko et al. 1997).

1.4.4 Mu transposition complexes in different stages of transposition

Bacteriophage Mu uses higher-order nucleoprotein complexes, or transpososomes, for transposition. Transpososomes contain the phage Mu genome synapsed by the MuA transposase proteins, and transpososome configuration varies through the different stages of Mu transposition (reviewed by Harshey 2014; Mizuuchi 1992).

MuA is inactive as a monomer, and becomes chemically active only when bound to Mu ends and assembled into a tetrameric form (Baker and Mizuuchi 1992; Lavoie et al. 1991; Mizuuchi et al. 1992). Within the tetramer, the catalytic steps are conducted by only two of the MuA subunits, those bound to the L1 and R1 binding sites of Mu, while the other two subunits of the tetramer appear to be essential for the structural integrity of the transpososome (Mariconda et al. 2000; Namgoong and Harshey 1998; Williams et al. 1999). The catalysis of cleavage and strand transfer occurs in *trans*, i.e. with MuA bound to the left end and the catalysis reaction chemistry at the right end, and vice versa (Aldaz et al. 1996; Savilahti and Mizuuchi 1996).

At an early stage in transpososome assembly, a three-site synaptic complex, LER, is formed between the left end, right end, and the enhancer (Watson and Chaconas 1996). In the LER complex, six MuA monomers are bound to L1-L3 and R1-R3 through their Iβμ domains, and interact with E through their Iα domain. The HU and IHF proteins both bend DNA and have important roles in the formation of LER. The interactions between MuA subunits in LER lead to the tetramerization of MuA and the formation of the stable functional core of subsequent complexes (Kobryn et al. 2002; Lavoie et al. 1991). The paired Mu-end structure is held together by the MuA tetramer, which is bound only to the L1, R1, and R2 MuA-binding sites at the Mu ends (Kuo et al. 1991; Lavoie et al. 1991; Mizuuchi et al. 1991). Although the enhancer is not required for these chemical steps, it remains weakly associated with MuA also in tetrameric complexes.

Interactions within the LER complex lead to the formation of more stable tetrameric complexes. The first type of tetrameric complex is type 0, also called a stable synaptic

complex (SSC) (Wang et al. 1996). In the type 0 complex, MuA assumes its active tetrameric configuration, but no strand cleavage has yet occurred. In the presence of Mg^{2+}, single-strand cleavage occurs at both Mu ends, and the type 0 complex is converted into a type 1 complex, also called the cleaved donor complex (CDC) (Craigie and Mizuuchi 1987; Surette et al. 1987). The 3′-OH groups of the cleaved donor complex attack the target DNA in a direct transesterification reaction, which results in the formation of a type 2 complex, also called the strand-transfer complex (STC) (Surette et al. 1987). The type 2 complex is the most stable Mu transpososome. The strand-transfer reaction is assisted by MuB, which interacts with the C-terminus of MuA and affects the target-site selection (Harshey and Cuneo 1986; Leung and Harshey 1991; Maxwell et al. 1987). The resulting strand transfer product is a branched molecule, called the Shapiro intermediate or θ structure (Shapiro 1979), in which MuA is still intact. The MuA tetramer is disassembled by a host protein, ClpX (Kruklitis et al. 1996); in the replicative pathway, this action is followed by replication of Mu and leads to the formation of a cointegrate molecule.

1.4.5 Structure of Mu transposition complex

Although the organization of almost all the individual MuA subdomains has been resolved by NMR or X-ray crystallization, the crystallization of the entire Mu transposition complex proved difficult and was accomplished only recently (Montano et al. 2012). The X-ray crystal structure contained a Mu strand-transfer transpososome, with a pair of Mu-ends derived from the Mu genome's right end, each carrying two MuA binding sites, R1 and R2, a tetramer of truncated MuA proteins (residues 77-605), and target DNA. An earlier electron-microscopy (EM) structure of the Mu transpososome (without the target DNA) had suggested a V-shape for the Mu R-ends and the MuA tetramer (Yuan et al. 2005), and the crystal structure was consistent with those findings. Specifically, it revealed a scissor-like configuration, in which the Mu end DNAs form the handles and the sharply bent target DNA the blades (Montano et al. 2012) (Figure 4.).

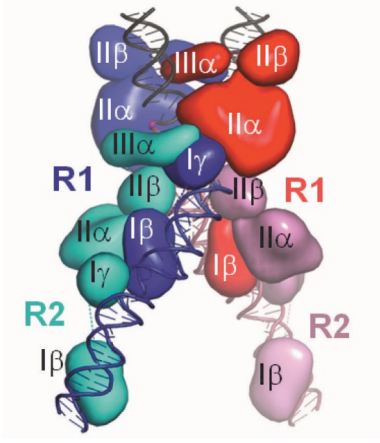

Figure 4. X-ray crystal structure of Mu transpososome joined to target DNA. The structure contains a pair of Mu-ends derived from the Mu genome's right end, each carrying two MuA binding sites, R1 and R2 (pink and blue), tetramer of truncated version of MuA proteins (residues 77-605), missing the N (Iα)- and C (IIIβ)-terminal domains, and target DNA (black). Magenta spheres depict scissile phosphates (only one can be seen in the picture). The figure is generated with PyMOL.

The crystal structure of the Mu transpososome confirmed many earlier results and predictions from biochemical studies and also provided new insights. For example, it was known that the MuA Iβ and Iγ subdomains are responsible for site-specific DNA binding at the Mu ends (Kim and Harshey 1995). The crystal structure revealed that, besides recognizing the specific sites, the R1-bound DNA-binding subdomains are also engaged in protein-protein contacts (Montano et al. 2012). The Iβ subdomain interacts with the IIα region of the subunit at the R2 site at the same Mu end, while the Iγ subdomain has contacts to the IIα region of the subunit at the R1 site at the other Mu end, as well as with subdomain IIIα of the subunit at the R2 site at the same Mu end (Montano et al. 2012).

The crystal structure likewise shed light on the function of the IIα subdomain, which contains the catalytic site. Biochemical studies had shown that with Mu ends that contain the R1-R2 MuA binding sites, only the subunit at the R1 site provides the active site for catalysis, which happens in *trans* (Namgoong and Harshey 1998). The other subdomain of domain II, subdomain IIβ, was predicted to interact with the target DNA (Krementsova et al. 1998). The crystal structure of the Mu transpososome confirmed the results and predictions of these earlier biochemical studies (Montano et al. 2012), and, in addition, showed that the interaction between the IIβ subdomain and the target occurs only in the R1 subunit (Montano et al. 2012). Subdomains IIα and IIβ of the subunit at the R2 site bridge the Mu ends and have a mainly structural role (Montano et al. 2012).

The X-ray structure also resolved the organization of the previously uncharacterized MuA IIIα subdomain (Montano et al. 2012). Like the other subdomains, IIIα has different roles depending on whether the subunit is bound to the R1 or R2 site. In the R1-bound subunit, IIIα stabilizes the strongly bent target DNA. Target bending may help reduce reversion of the strand-transfer complex by straining the DNA conformation in such a way as to increase the distance between the Mu ends and the active site after strand transfer (Montano et al. 2012). Rearrangements at domain III of the R1 subunit after target capture may also allow the ClpX remodeling machinery to recognize the strand-transfer complex for disassembly (Montano et al. 2012). Instead, IIIα of the R2-bound subunit is required for initial transpososome assembly (Mariconda et al. 2000). The crystal structure suggests that this subdomain wraps around the other subunits near the catalytic site and stabilizes the complex (Montano et al. 2012). The IIIα subdomain of the R2-bound subunit may also interact with the flanking host DNA and contribute to its cleavage after insertion into the target DNA (Montano et al. 2012).

1.4.6 Regulation of Mu transposition

1.4.6.1 Control of lysogenic repression

In the lytic state, phage Mu replicates highly efficiently through replicative transposition. However, during the lysogenic state, transposition can be shut down completely (reviewed by Toussaint et al. 1994). The key regulator in the establishment

of lysogeny is the Rep protein (product of the *c* gene), which can shut down Mu transposition functions. This repressor binds to an operator region (which overlaps the enhancer segment) and blocks the expression of early lytic genes, such as MuA and MuB (reviewed by Goosen and van de Putte 1987). Wild-type Mu lysogens are quite stable and are not induced by UV or any other known physical or chemical treatment. However, derivatives of Mu that have repressors sensitive to either high temperature or proteolytic degradation can be induced by high temperature or proteolysis by cellular proteases, respectively (Lamrani et al. 1999; Ranquet et al. 2005; Vogel et al. 1991). The inactivation of the repressor allows the expression of early lytic genes and replicative transposition of Mu.

1.4.6.2 Quality control in transposition

The assembly of Mu transpososomes is regulated at several levels (reviewed by Harshey 2014; Mizuuchi 1992). An essential regulatory element is the enhancer segment, which has an important role in the regulation of early transpososome assembly prior to any irreversible chemical steps (Mizuuchi and Mizuuchi 2001; Watson and Chaconas 1996). The formation of the first transposition complex requires that the enhancer segment (which overlaps the operator sequence) is available and not bound by the Mu repressor protein Rep (Mizuuchi and Mizuuchi 1989; Mizuuchi et al. 1992). Rep shares homology in its DNA-binding domain with the enhancer-binding domain of MuA, and when bound to the operator site, Rep blocks transposition by blocking both MuA binding to the enhancer site and the transcription of early lytic genes (Craigie et al. 1984; Harshey et al. 1985; Leung et al. 1989). Thus, Rep down-regulates transposition at two levels. The enhancer segment is also suggested to play a regulatory role in keeping the MuA monomer catalytically silent, in facilitating and stabilizing synapsis between ends, and in ensuring that the ends of the same Mu genome are paired, instead of the ends from neighboring genomes (reviewed by Harshey and Jayaram 2006).

Transpososome assembly is also regulated by two other important features: “domain sharing”, i.e. a functional active site built from domains of different MuA monomers (Namgoong and Harshey 1998), and *trans* catalysis, with MuA bound to the left end and the catalysis reaction chemistry at the right end, and vice versa (Aldaz et al. 1996; Savilahti and Mizuuchi 1996). These two features ensure that the active site is functional only after multimerization of the MuA monomers, and that donor cleavage happens only after the two Mu ends have been synapsed and have formed a stable complex.

1.4.7 Mu target site selection

Even though transposons can insert into various sites in the target DNA, they all appear to have at least some degree of target-site selectivity (reviewed by Craig 1997). Among transposons, the target-site selectivity of Mu is relatively low, with a target recognition site that is only 5bp long (5′-C-Py-(G/C)-Pu-G-3′, where Py is C or T, and Pu is A or

G) (Haapa-Paananen et al. 2002; Mizuuchi and Mizuuchi 1993). This sequence-level target site preference is encoded by MuA alone, independent of the MuB protein.

MuB is a regulatory protein of the Mu transposon which assists in target site selection and in transposition immunity (Maxwell et al. 1987). It is an ATP-dependent nonspecific DNA-binding protein with ATPase activity. Upon ATP binding, MuB polymerizes and forms helical filaments around DNA (Greene and Mizuuchi 2004; Mizuno et al. 2013), with a modest tendency to form larger filaments on DNA with high A/T content (Greene and Mizuuchi 2004). MuA that is bound to the Mu ends induces ATP hydrolysis by MuB, which dissociates the MuB from DNA, thus generating regions without MuB, which are poor targets for new transposon insertions (Figure 5) (discussed in more detail in section 1.4.8.1) (Adzuma and Mizuuchi 1988; Maxwell et al. 1987). Conversely, the nucleotide-bound MuB stimulates the activity of MuA at several steps, including the formation of the MuA-Mu DNA complex and the nicking of Mu ends (Schweidenback and Baker 2008). The helical parameters of the MuB filament do not match those of the coated DNA, and MuB does not deform the DNA to which it is bound (Mizuno et al. 2013). However, the MuB-imposed symmetry model proposes that MuA and MuB together may induce a symmetry match between DNA and MuB at the MuB filament end and thus deformation of DNA, which results in target capture, as the bent DNA conformation is favored by MuA for transposition (Figure 5.) (Mizuno et al. 2013).

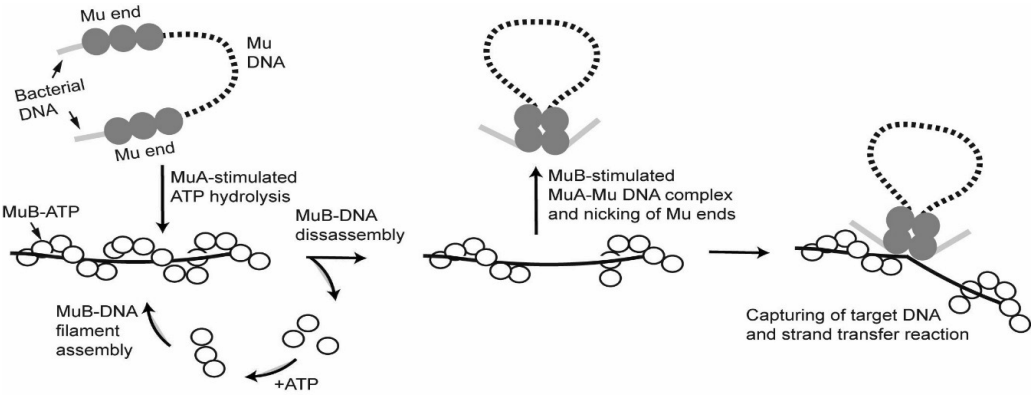

Figure 5. Model for MuB function in target capture and *cis* immunity. Upon ATP binding MuB polymerizes and forms helical filaments around DNA. MuA bound to Mu ends induce the ATP hydrolysis by MuB, which dissociates the MuB from DNA and generates regions without MuB. Reciprocally, the nucleotide-bound MuB stimulate MuA to pair and nick Mu ends. According to Mizuno et al. 2013, MuA and MuB together may induce deformation of DNA at the boundary of MuB filament, which results to target capture. Drawn according to (Mizuno et al. 2013).

1.4.8 Mu transposition immunity

Bacteriophage Mu and other transposons, including Tn*7* and members of the Tn3 family, avoid inserting near or into their own genome through a phenomenon called transposition immunity (reviewed by Craig 1997). For this purpose, Mu utilizes two

mechanisms, *cis* immunity and genome immunity, which operate outside and inside the Mu ends, respectively (Ge et al. 2010).

1.4.8.1 cis immunity

The *cis*-immunity mechanism prevents regions immediately outside the Mu ends from being used as targets by the element. This immunity can extend over large distances from the chromosomal site where the transposon is located, but does not provide protection to the whole bacterial genome harboring the transposon (Adzuma and Mizuuchi 1988; Adzuma and Mizuuchi 1989). A key component of the immunity mechanism involves MuB-MuA interactions. MuB-bound DNA is an efficient transposition target, but MuA bound to Mu ends interacts with nearby MuB-ATP complexes and induces the ATPase activity of MuB. This dissociates MuB from the nearby DNA before target-site commitment, making the region near the Mu ends a poor target for new insertions (Adzuma and Mizuuchi 1991; Greene and Mizuuchi 2002a). The MuA-MuB interaction that removes MuB requires DNA looping between the binding sites of MuA and MuB, and as the loop increase in size over time it drives the transpososomes to distal target sites (Greene and Mizuuchi 2002b; Han and Mizuuchi 2010). MuA and MuB reciprocally stimulate each other's enzymatic activity, such that their interactions can result in either *cis* immunity or target-site selection in *trans*. It is not clear what factors control the temporal order of these two outcomes, but they may be determined by the oligomeric state of MuA, e.g., whether MuA is only bound to Mu ends or already assembled into an active transpososome (Greene and Mizuuchi 2002b).

1.4.8.2 Mu genome immunity

Mu amplifies its DNA 100-fold during the lytic phase, and with each round there is a risk of Mu transposing into itself. The *cis*-immunity mechanism has been shown to be strongest around 5 kb outside the Mu ends (Manna and Higgins 1999), and thus is not expected to be capable of protecting the 37-kb Mu genome completely. Instead, a relatively new Mu genome-immunity mechanism has been described which functions inside the ends and protects Mu from self-integration (Ge et al. 2010). In contrast to the *cis*-immunity mechanism, in the genome-immunity mechanism MuB stays strongly bound within the Mu genome. According to the model for Mu genome immunity, the Mu genome is segregated into an independent chromosomal domain, in which polymerization of MuB on the genome forms a barrier against self-integration (Ge et al. 2010). Mu genome immunity could theoretically be functionally similar to the immunity mechanism of the Moloney Murine Leukemia Virus (MoMLV) which utilizes the barrier-to-autointegration factor (BAF) to protect itself from self-integration (Lee and Craigie 1994; Lee and Craigie 1998). BAF prevents autointegration by compacting the viral DNA and making it inaccessible as a target for integration (Bradley et al. 2005; Zheng et al. 2000).

1.5 Transposon tools

Transposons are powerful tools for investigating and manipulating genes and genomes, and they have been utilized widely in many applications for advanced genetic studies (reviewed by Boeke 2002; Hayes 2003). A variety of transposon tools has been successfully used in different organisms, including prokaryotes, plants, invertebrates, and recently also vertebrates (reviewed by Hayes 2003; Ivics et al. 2009; Ivics and Izsvak 2010; Ni et al. 2008). In the transposition process, the transposon catalyzes its own insertion into a target site, which make transposons natural DNA delivery vehicles. They are extremely suitable for a number of different purposes, as the DNA between the transposon ends can be of any origin and can include useful features such as selectable markers, unique primer binding sites, reporter functions (such as the gene for β-galactosidase), plasmid origins of replication, and controlling elements (such as regulated promoters).

Traditionally transposition applications were performed *in vivo*, which requires either mobilization of an endogenous transposon or the introduction into the host cell, via transformation or bacterial mating, of a plasmid (or plasmids) containing the transposon and transposase. The use of transposons in *in vivo* reactions can be complicated by different factors, such as the restricted host ranges of different elements.

Host range limitations can be avoided by the use of *in vitro* transposition reactions. A number of efficient *in vitro* systems have been established both for prokaryotic and eukaryotic transposons. Examples of prokaryotic transposons include the *in vitro* reactions of Tn*3* (Maekawa et al. 1996), Tn*5* (Goryshin and Reznikoff 1998; Goryshin et al. 2000), Tn*7* (Bainton et al. 1993), Tn*10* (Chalmers and Kleckner 1994), Tn*552* (Leschziner et al. 1998), IS*911* (Polard et al. 1996), and bacteriophage Mu (Haapa et al. 1999b; Mizuuchi 1983). Examples of eukaryotic transposons include the *in vitro* transposition system of yeast, Ty*1* (Devine and Boeke 1994); the *mariner/Tc1* family transposons Tc1 (Vos et al. 1996), *Himar1* (Lampe et al. 1996), and *Mos1* (Tosi and Beverley 2000); *piggyBac* (Mitra et al. 2008); and the hAT superfamily transposon *Hermes* (Zhou et al. 2004).

Increased understanding of transposition mechanisms has facilitated the use of transposable elements for research purposes and enabled the establishment of efficient transposon tools for a variety of applications, including insertional mutagenesis (Dupuy et al. 2005; Lamberg et al. 2002), the generation of induced pluripotent stem cells (Woltjen et al. 2009), and functional genomics studies (Kekarainen et al. 2002; Miskey et al. 2005). In addition, transposon tools are increasingly being developed for gene therapy (Ivics and Izsvák 2006; VandenDriessche et al. 2009) and transgenesis applications (Ivics et al. 2009; Ivics and Izsvak 2010).

1.5.1 Mu as a genetic tool

The minimal Mu transposition system displays high transposition frequency and relatively low target-site selectivity (Haapa et al. 1999b; Haapa-Paananen et al. 2002), and these features make it ideal for a variety of applications. In addition, the Mu transposon ends can be designed to contain nucleotide changes outside of the MuA recognition sequence (Jones 2005; Poussu et al. 2004), which can be exploited when designing novel Mu tools. Mu *in vitro* transposition technology has been used for DNA sequencing (Haapa et al. 1999a), protein engineering for structure/function studies (Baldwin et al. 2008; Edwards et al. 2008; Jones 2005; Poussu et al. 2004; Poussu et al. 2005), genome-wide functional mapping of virus genomes (Kekarainen et al. 2002; Kiljunen et al. 2005; Krupovic et al. 2006; Vilen et al. 2003), construction of gene targeting vectors (Jukkola et al. 2005; Turakainen et al. 2009; Vilen et al. 2003; Zhang et al. 2005), insertional mutagenesis of archaea (Kiljunen et al. 2014), and SNP discovery (Orsini et al. 2007; Yanagihara and Mizuuchi 2002). In addition, functional transpososomes can be pre-assembled *in vitro* and subsequently transformed into host cells, where DNA of interest can be transposed *in vivo* into the genome of the recipient cell (Lamberg et al. 2002; Pajunen et al. 2005). This combination of *in vitro* and *in vivo* systems can be used for highly efficient, species-non-specific gene delivery and insertional mutagenesis, as demonstrated with a variety of Gram-negative and Gram-positive bacteria, yeast, and mammalian cells (Lamberg et al. 2002; Paatero et al. 2008; Pajunen et al. 2005; Tu Quoc et al. 2007; Wu et al. 2009). The current applications for Mu *in vitro* transposition technology are presented in Table 2.

1.5.1.1 DNA Sequencing

The Mu-mediated sequencing approach enables the efficient and easy creation of sequencing templates (Haapa et al. 1999a). In this protocol, the transposon is integrated *in vitro* into the target plasmid DNA to be sequenced. Following re-introduction of the transposition products into bacterial cells, the cells are selected using an antibiotic-resistance marker carried by the transposon to identify clones with transposon insertion. From the selected clones, sequencing is carried out bi-directionally using primers specific for the ends of the transposon. Mu-based sequencing product is commercially available from Thermo Fisher Scientific (TGS, Template Generation System). Recently, Mu transposition technology was also adapted for the generation of genomic DNA libraries using next-generation sequencing systems (Thermo Fisher Scientific, MuSeek library preparation kit). This method utilizes MuA transposase and Mu ends to simultaneously fragment the target DNA and tag the fragment ends with transposon DNA. In a subsequent reaction, fragments with transposon sequence tags at the ends are attached to platform-specific adapter sequences and PCR amplified, then subjected to high-throughput sequencing. A similar Mu-mediated strategy has also been used for monitoring outcome in gene therapy; to directly map the integration sites of therapeutic vectors using DNA barcoding and pyrosequencing (Brady et al. 2011).

1.5.1.2 Functional and structural studies of proteins

Mu-based protein engineering applications can be used for functional and structural studies of proteins. In the pentapeptide insertion mutagenesis strategy, an *in vitro* Mu transposition reaction is utilized to produce random transposon insertions into a target plasmid containing the gene of interest (Poussu et al. 2004). Custom-designed transposon ends contain NotI sites, which enables the removal of the transposon core sequence. Self-ligation of digested clones results in a 15-bp insertion in the target DNA, and when the transposon is inserted in protein-coding regions, the 15-bp insertion encodes five additional amino acids. This system can be used, for example, to screen important sites for protein-protein interactions. Based on this strategy, a commercial product from Thermo Fisher Scientific is available for the construction of linker scanning libraries (MGS, Mutation Generation System). To enable the analysis of a whole mutant library en masse, without the need to construct or isolate mutants separately, pentapeptide insertion mutagenesis can be combined with yeast two-hybrid screening and PCR-based genetic footprinting (Pajunen et al. 2009).

With the gene truncation strategy, a Mu *in vitro* transposition reaction is utilized to produce N- and C-terminal deletion variants of proteins for functional studies (Poussu et al. 2005). To produce C-terminal deletion variants, the transposon has been designed to contain translational stop codons in all three reading frames close to each transposon end. Therefore, protein translation is terminated near the transposon insertion site, and the resulting protein variants are truncated at their C-termini. To produce N-terminally truncated protein variants, the transposon is designed to contain a NotI site at each end. In this strategy the plasmid vector has to contain unique NotI restriction site 5′from the cloned target. Following digestion with NotI, the 5′-distal part of the target gene and the transposon core is eliminated, which generates N-terminally truncated versions of the target protein. Both N- and C-terminal protein variants are useful, for example, in mapping regions involved in protein-protein interactions.

Mu *in vitro* transposition can also be used to generate novel molecular diversity and construct directed evolution libraries. In a method called triplet nucleotide removal, a Mu *in vitro* transposition reaction is utilized to introduce triplet nucleotide deletions at random positions throughout a target gene (Jones 2005; Simm et al. 2007). This method makes use of a Mu transposon that contains sites for the restriction enzyme MlyI at both ends, with the recognition sites placed 1 bp away from the site of transposon insertion. MlyI is a type IIS restriction enzyme that cuts 5 bp outside its recognition sequence to generate a blunt end. Transposition results in a 5-bp target-site duplication, and digestion with MlyI results in the removal of the transposon and 4 bp of the target gene at both ends. Ligation of the ends rejoins the target gene and results in a 3-bp deletion. Triplet nucleotide removal can be followed by replacement with a random 3-bp segment (the trinucleotide exchange method, TriNEx) (Baldwin et al. 2008), or with a new protein domain (the domain insertion strategy; Edwards et al. 2008). The trinucleotide removal, trinucleotide replacement, and domain insertion

methods are efficient approaches for generating novel molecular diversity for the construction of directed evolution libraries.

1.5.1.3 Functional genetics and genomics of viruses

Mu *in vitro* transposition technology can also be used for functional genetics and genomics studies of viruses. A method for the characterization of complete bacteriophage genomes exploits direct *in vitro* integration of transposon DNA into the target viral genomic DNA, without prior cloning of the viral genome (Kiljunen et al. 2005; Krupovic et al. 2006; Vilen et al. 2003). Following electroporation into susceptible host cells, the mutant viruses are selected for their ability to form plaques. Determination of the transposon integration sites enables mapping of essential and nonessential regions and genes for virus propagation. An alternative version of the strategy, which utilizes cloned versions of virus genomes or genome segments as targets, has been used with RNA viruses to mutagenize both a genome segment (Laurent et al. 2000) and an entire RNA virus genome (Kekarainen et al. 2002).

1.5.1.4 Construction of gene-targeting vectors

Mu transposons can be used as tools for the construction of different kinds of gene-targeting vectors in order to modify the mammalian genome, including vectors that can be used for generating null, hypomorphic, or conditional alleles (Jukkola et al. 2005; Turakainen et al. 2009; Vilen et al. 2003; Zhang et al. 2005). Compared with conventional methods for the generation of targeting constructs, which often require complex DNA manipulations, transposon-based methods are fast and technically straightforward. With a transposon-based strategy, several constructions aimed at targeting different exons can be generated simultaneously, and the procedure does not require restriction enzyme sites on genomic DNA near the exon of interest. A strategy for the generation of conditional knock-out gene-targeting vectors employs two successive transposition reactions to introduce a loxP site on one side of the exon of interest and on the other side, an antibiotic-resistance cassette (selectable both in bacteria and mammalian cells) flanked by loxP and FRT sites (Turakainen et al. 2009). Targeting constructions are linearized before electroporation into mouse embryonic stem cells, and desired clones resulting from homologous integration events undergo both positive and negative selection. FRT sites enable the removal of the resistance cassette by Flp-mediated recombination, and Cre-mediated recombination between the loxP sites enables the inactivation of the targeted gene.

1.5.1.5 Insertion mutagenesis strategy for archaea

The genetic basis of many unique archaeal features has yet to be well characterized, and studies have been hampered by a lack of efficient genetic screens. A recently established Mu transposition-based strategy provides the means for generating a random genomic insertion mutant library for archaea, as has been shown with *Haloferax volcanii* (Kiljunen et al. 2014). This insertion mutagenesis strategy

combines a Mu *in vitro* DNA transposition reaction and homologous-recombination-based gene targeting in an archaeon (Kiljunen et al. 2014). The archaeal DNA is first digested and used as a target in an *in vitro* Mu transposition reaction. From the transposition products, fragments of the desired sizes are cloned into a suitable vector to yield a plasmid library. Inserts are released by digestion and transformed into archaeal cells to generate a transposon insertion mutant library. The library can then be used for the identification of non-essential genes.

1.5.1.6 Mapping single nucleotide polymorphism

The preference of Mu for targeting single-nucleotide mismatches can be exploited to map single-nucleotide polymorphism (SNP) (Yanagihara and Mizuuchi 2002), as has been demonstrated with a butterfly genome (Orsini et al. 2007). PCR amplification of a genomic region that contains at least two alleles located at the same position generates DNA duplexes that contain mismatches. With this DNA as a target, nearly 90% of the transposon insertions occur at the mismatched sites (Yanagihara and Mizuuchi 2002). Single-nucleotide polymorphisms are an important resource for providing genetic markers for the mapping of human disease genes, as well as for other studies, and Mu mismatch-targeting provides a simple method for detecting mismatches in DNA and isolating SNP markers from as-yet uncharted genomes (Orsini et al. 2007; Yanagihara and Mizuuchi 2002).

Table 2. Applications for Mu *in vitro* transposition technology

Use	Application	References
DNA sequencing	Creation of sequencing templates	Haapa et al. 1999a, Brady et al. 2011
Functional and structural studies of proteins	Pentapeptide insertion mutagenesis strategy	Poussu et al. 2004, Pajunen et al. 2009
	Gene truncation strategy	Poussu et al. 2005
	Triplet nucleotide removal method	Jones et al. 2005, Simm et al. 2007
	Trinucleotide exchange method, TriNEx	Baldwin et al. 2008
	Domain insertion strategy	Edwards et al. 2008
Functional genetics and genomics of viruses	Whole genome analysis of bacteriophages	Vilen et al. 2003, Kiljunen et al. 2005, Krupovic et al. 2006,
	Analysis of genomic regions and entire genomes of viruses cloned on specific vectors	Laurent et al. 2000, Kekarainen et al. 2002
Construction of different kinds of gene targeting vectors	Generation of null, hypomorphic, or conditional alleles	Jukkola et al. 2005, Turakainen et al. 2009, Vilen et al. 2003, Zhang et al. 2005
Mapping single nucleotide polymorphism (SNP)	Mismatch targeting	Yanagihara and Mizuuchi 2002, Orsini et al. 2007
Identification of non-essential archaeal genes	Generation of random genomic insertion mutant library for archaea	Kiljunen et al. 2014
Species non-specific gene delivery and insertional mutagenesis	Use of *in vitro* pre-assembled transpososomes for gene delivery *in vivo*	Lamberg et al. 2002, Pajunen et al. 2005, Paatero et al. 2008, Tu Quoc et al. 2007, Wu et al. 2009

2 AIMS OF THE STUDY

This work focuses on the establishment of novel Mu transposition-based molecular genetics tools and the generation of MuA transposase variants with altered transposition frequency. Transposon-based applications are widely utilized in molecular biology, and the aim of this study was to add useful tools to the biotechnology toolbox by providing both a general test to measure the specific activity of transposition complexes and a cloning strategy for replication-deficient circular DNA. An additional goal was to improve Mu *in vitro* transposition applications by generating hyperactive MuA transposase variants.

The specific aims were:

1. To generate a general-purpose assay to measure the activity measurement of DNA transposition complexes, using phage Mu transposition as a test platform.

2. To develop a method for MuA-mediated *in vitro* cloning of circular DNA, establish reaction conditions that reduce autointegration, and test the effect of MuB on reducing autointegration.

3. To generate and characterize various Mu-transposons with different origins of replication for MuA-mediated *in vitro* cloning of circular DNA, and test the effect of the simultaneous use of two transposons on reducing autointegration.

4. To generate hyperactive MuA tranposase variants, characterize their properties in the context of the Mu transpososome, and tailor specific MuA variants for different Mu-based applications.

3 MATERIALS AND METHODS

The bacterial strains, plasmids, transposons, and oligonucleotides used in this study are described in the original publications.

3.1 *In vitro* transposition reaction

Minimal Mu *in vitro* transposition reactions were conducted that contained only the MuA transposase protein, transposon DNA, and target DNA in standard reaction conditions as described in (Haapa et al. 1999b; Savilahti et al. 1995) (Studies I, II, III, and IV). As specified in each experiment, reactions contained variable amounts of transposon DNA as well as target DNA. Some reactions also contained competitor DNA or MuB (Study II).

Reactions were stopped by the addition of 1% SDS (Studies II and III), by freezing in liquid nitrogen (Study IV), or by incubation at 75˚C (Study I). Reaction products were transformed into competent cells by heat shock (Studies I and IV) or by electroporation (Studies II, III, and IV). Following transformation, cells were grown briefly and then spread onto appropriate selection plates.

3.2 Transpososome pre-assembly

Transpososomes were pre-assembled *in vitro* as described in (Lamberg et al. 2002; Pajunen et al. 2005) (Studies I and IV). Successful transpososome assembly was verified using agarose/BSA/heparin gels as previously described (Lamberg et al. 2002). Transpososomes were either used directly (Study IV) or further concentrated using polyethylene glycol (PEG 6000) precipitation as described in (Pajunen et al. 2005; Savilahti and Bamford 1993) (Study I).

The pre-assembled transpososomes were used either for genomic integrations *in vivo* (Study IV) or as donor DNA for *in vitro* transposition reactions (Study I). Differently from the *in vitro* assay, in the *in vivo* assay, pre-assembled transpososomes encounter Mg^{2+} ions and thus become active only after electroporation into the *E. coli* cells.

3.3 Generation of MuA mutant libraries

The *MuA* gene was mutated using error-prone PCR as described in (Rasila et al. 2009) with the specifications described in study IV. Taq DNA polymerase was used under three mutagenic PCR conditions (0, 1, or 2 µl of mutagenic buffer added in the standard 50-µl reaction) and Mutazyme II DNA polymerase was used in two separate PCR reactions (5 or 10 cycles of amplification). A total of five MuA mutant libraries were generated.

3.4 Papillation assay

The transpositional activity of MuA transposase variants was determined using a papillation assay as described in (Pajunen et al. 2010) (Study IV). The standard assay included incubation at 30˚C for 115 h on LB agar plates that were supplemented with Ap (100 µg/ml), Cm (20 µg/ml), lactose (0.05%), X-gal (40 µg/ml), and arabinose (1×10^{-4}%). The papillation assay is based on mini-Mu transposon mobilization *in vivo*. The mini-Mu transposon contains *cat* gene for chloramphenicol selection, and reporter gene *lacZ*, lacking the codons for amino acids 1-8. Transposition into any expressed chromosomal gene downstream of a promoter results in activation of the *lacZ* gene and thereby formation of blue microcolonies (papillae) growing on otherwise whitish *E. coli* colonies. The mean value and standard deviation (SD) of the number of papillaes were calculated for each protein variant from six colonies.

3.5 Mutant variant sequence determination

The mutated residues within the *MuA* gene of each particular plasmid were defined by sequence analysis. DNA sequencing was done at the DNA sequencing facility of the Institute of Biotechnology, University of Helsinki.

3.6 Protein expression and purification

MuA transposase proteins were expressed in BL21(DE3)pLysS using pET3d-derived plasmids, and purified with phosphocellulose and hydroxylapatite columns by employing a procedure described in (Baker et al. 1993), with modifications described in Study IV. Protein concentration was determined spectrophotometrically using the A280 value of 1.58 = 1 mg/ml (Baker et al. 1991). The purity level of the protein preparations was examined in two ways: with excessive protein loading on an SDS-PAGE gel to determine the number of protein species, and with prolonged incubation under *in vitro* reaction conditions with supercoiled plasmid DNA to examine the level of nuclease activity.

3.7 Structural analyses

The structural and functional consequences of mutations were assessed by investigating the variations within the secondary and tertiary structures of MuA proteins and Mu transpososomes (Study IV). The NMR and X-ray structures for isolated MuA protein domains, and the transpososome crystal structure, were from the Protein Data Bank (PDB) (Berman et al. 2000). The dictionary of protein secondary structure (DSSP) was used for secondary structural element identification (Shockett and Schatz 1999). UCSF Chimera (Pettersen et al. 2004) and PyMOL (The PyMOL Molecular Graphics System Version 1.3) were used for visualizations.

4 RESULTS AND DISCUSSION

4.1 Activity measurement of DNA transposition complexes (I)

Traditionally, DNA transposition activity or the quality of DNA transposition complexes has been monitored by utilizing *in vitro* transposition reactions with a plasmid target. To score transpositional activity, i.e. transposon integration events into target plasmids, the transposition reaction products are introduced into *E. coli* cells, and selected simultaneously for transposon and target plasmid resistance markers. However, this type of measurement is dependent on resistance markers and thus limited to transposons with suitable bacterial markers. This method cannot be used to measure the activity of transposons without selectable markers and those constructed solely for eukaryotic use. Furthermore, it is not ideal to use selection that depends on the marker carried by each particular transposon when comparing the activities of different transposons. Therefore, to standardize measurements of transpositional activity, we developed a general activity measurement assay for DNA transposition, and used Mu transposition as a test platform. We also characterized the properties of the assay for a set of Mu transposons, including its linear response as a function of transposon DNA concentration and its response as a function of *E. coli* competency. The established assay can be used to directly compare transpososome activities with all types of mini-Mu transposons, and it should be also directly applicable to other transposition-based systems with a functional *in vitro* reaction.

4.1.1 Activity measurement assay standardizes the measurement of transpositional activity regardless of the specific nature of transposon employed

The developed assay is based on an *in vitro* transposition reaction with a pZErO-2 target plasmid that carries a lethal *ccdB* gene (Bernard et al. 1994; Bernard and Couturier 1992; Invitrogen by Life Technologies 2012). Expression of the CcdB protein in wild-type *E.coli* is lethal to the cells, but if transposition targets *ccdB*, this gene becomes inactivated and cells that have received such plasmids survive (Principle of the assay design in Figure 1 in I). As the assay is not dependent on the marker gene carried by the transposon, it is well-suited for all types of transposons, including those that do not contain selectable marker genes or contain markers that are functional only in eukaryotic cells. We characterized the properties of the assay with three mini-Mu transposons: Cat-Mu (size 1.3 kb, encodes chloramphenicol-resistance marker), Kan/Neo-Mu (size 1.9 kb, encodes kanamycin-resistance marker) and Puro-eGFP-Mu (size 2.1 kb, no antibiotic resistance marker for bacteria) (Figure 2 in I). To enable transpososome assembly, transposons were incubated with MuA protein for different periods of time (0, 10, 60, 120, or 240 minutes). Pre-assembled transposition complexes were incubated *in vitro* with the pZErO-2 target plasmid (carries the *ccdB*

gene and a kanamycin-resistance marker) and transposition reaction products were transformed into competent *E. coli* cells and plated on kanamycin selection plates to score integration events into target plasmids. All three transposons yielded very similar results (Figure 2 in I), regardless whether the transposition events were selected toward antibiotic marker encoded by the transposon in addition to the *ccdB* gene, or solely by the *ccdB* gene. This illustrated the assay's suitability for various types of transposons and especially for transposons that entirely lack selection markers.

However, we noticed that the assembly of transposition complexes with the longer transposons Puro-eGFP-Mu (2.1 kb) and Kan/Neo-Mu (1.9 kb) was somewhat slower than that with the shorter transposon Cat-Mu (1.3 kb). Similar decrease in transposition efficiency with increased transposon length has been reported with several transposons, including Mu (Wei et al. 2010; Chandler et al. 1982; Way and Keckner 1985; Lampe et al. 1998; Izsvak et al. 2000). Since the transposons that we used in this study were all relatively short, further studies would be warranted with appropriate sets of longer transposons to draw any further conclusions on the effect of transposon length. For future studies it would also be interesting to use the assay to directly compare the transposition activity of different transposon species, for example, Tn5 and Mu.

4.1.2 Operational range of the activity measurement assay

We defined the operational range of the assay by using different quantities of Cat-Mu transposition complexes (from 40 to 520 ng of transposon DNA per *in vitro* transposition reaction) (Figure 3 in I). The number of colonies increased linearly with the amount of transposon DNA, and reliable activity measurements were obtained even with a small amount of transpososomes. These results showed that the assay was adjustable with a wide range of DNA concentrations.

4.1.3 Influence of competence status for the activity measurement assay

The competence status of recipient *E. coli* cells is known to influence the number of colonies produced following the transformation of *in vitro* transposition reaction products into competent cells. To determine the effect of variation in competence status on the activity measurement assay, we transformed *in vitro* transposition reaction products into different batches of competent DH10B *E. coli* cells (Figure 4 in I). As expected, the number of colonies correlated linearly with the competence status of the *E. coli* cells, which enables normalization and comparison of results obtained with different cell batches. We used heat-shock transformation into chemically competent cells in our experiments, but the reaction products could as well have been electrotransformed into electrocompetent cells. However, with electrotransformations, the variation between transformations could have been greater.

In summary, this study describes a very useful assay to measure DNA transposition activity regardless of the specific nature of the transposon. In particular, the assay is aimed to ease the evaluation of transposons containing only eukaryotic markers and transposons without selectable markers. We hope that the established assay will serve as a dependable quality control measure for various applications of *in vitro* DNA transposition technology, including also future transposition-based tools.

4.2 MuA-mediated *in vitro* cloning of circular DNA (II, III)

Circular DNA is abundant in nature and DNA circles are involved in many natural processes. Therefore, characterization of such molecules and studies on their functions are important and expected to provide new insights as well as novel concepts with regard the biology of living systems. However, discovery and analysis of these circles is sometimes hindered by their relatively low quantity and lack of selectable markers. In such cases cloning provides an alternative means to produce sufficient amounts of material for DNA analyses. In transposition-aided cloning, replication-deficient DNA circles are transformed into plasmids which are able to replicate in *E. coli*. So far, Tn5 and Tn552 systems have been utilized for this process (Agron et al. 2002; Jendrisak et al. 2002; Kirby et al. 2002). The method enables simultaneous selection for circular DNA and against linear fragments. Also, the circular form of the target is retained throughout the procedure, as no double-strand breaks are introduced during the transposition reaction. Here, we established a strategy to clone circular DNA based on the Mu *in vitro* transposition reaction (II), and constructed ten novel transposons in order to enable several approaches (II, III). In addition, we examined the effect of substrate stoichiometry (II), and determined the sensitivity of the methodology with decreasing target amounts (II). In addition, to study novel approaches for the reduction of intramolecular transposition, we tested the effect of MuB and BAF (II), and the simultaneous use of two transposons (III) on the transposition reaction product profile. Our results delineate the parameters that affect the reaction product profile, define the minimum amount of circular target DNA required, and provide guidelines for the efficient utilization of the strategy.

4.2.1 Overview of the MuA-mediated cloning method and characteristics of the target plasmids (II, III)

The established method utilizes an *in vitro* transposition reaction with a custom-designed transposon in order to deliver both a selectable marker and an *E. coli* plasmid replication origin into the target plasmids (Principle of the assay design in Figure 1 in II). The transposon insertion generates replication-competent target plasmids following the recovery of recombinants by transformation into *E. coli*. As target plasmids, we used the conditionally replication-deficient plasmids pALH31 (II, III) and pLOI2227 (III). The pALH31 plasmid encodes a kanamycin-resistance gene and contains the

R6Kγ replication origin, which makes it replication-deficient in standard *E. coli* strains. The pLOI2227 plasmid encodes a kanamycin-resistance gene and contains the temperature-conditional pSC101 replication origin, which makes the plasmid replication-deficient when grown at 42˚C. We used both of these target plasmids as stand-ins for a replication-deficient DNA circle.

4.2.2 Construction of novel mini-Mu transposons (II, III)

To provide a new set of tools for establishing diverse DNA circles in *E. coli*, we constructed altogether ten novel mini-Mu transposons (Figure 1 in II and Figure 2 in III). This transposon set includes five pairs, with each pair containing one of five replication origins, and the members of a pair differing only in the presence or absence of a chloramphenicol-resistance gene (*cat*). The following five replication origins were used: p15A, R6Kγ, pUC-ori, F-ori, or P1-ori. These differ greatly in their copy numbers per cell, which vary from low (only 1-2 copies per cell) to high (several hundred per cell; Table 2 in III). The optimal choice of the replication origin may depend on the study approach and on a variety of factors, such as, will there be several plasmid types in a cell at the same time and how many copies of the molecule per cell is desired.

4.2.3 MuA-mediated transposition can be used to clone circular DNA molecules (II, III)

The feasibility of the method was first evaluated by performing *in vitro* transposition reactions with the Ori(pUC)-Cat-Mu transposon and the pALH31 target plasmid. The results showed that the MuA-mediated cloning method was capable of converting the target DNA molecules into plasmids which were able to replicate in a standard *E. coli* strain (Table 2 in II). In addition, we were able to clone circular DNA from very low target amounts despite the presence of contaminating linear DNA fragments (Table 3 in II). In study III, nine novel transposons, containing five different replication origins, were shown to be functional for the cloning of circular DNA (Table 3 in III). Together these ten novel transposons represent a comprehensive set of tools for establishing diverse DNA circles in *E. coli*. We expect that all of the constructed transposons are able to facilitate the replication of plasmids at least up to 10 kb in size, and the transposons containing P1 and F origins to be able to convert even up to 300 kb sized circles into *E. coli* compatible replicons. However, as we used only small target plasmids (~3 kb) in our experiments, the cloning system is yet to be tested with larger target molecules.

4.2.4 Intramolecular transposition generates background colonies (II, III)

Intramolecular transposition (also called autointegration) is a phenomenon in which a transposon integrates into itself. In standard transposon reactions the intramolecular

transposition products will not form replication-competent circles due to the lack of the origin of replication. With transposons that contain both a replication origin and selectable marker, autointegration products can generate circles that do not contain the external target DNA but are still replication-competent in *E. coli*. Such molecules yield background colonies on selection plates which do not contain the target of interest. With our experimental set-up, we were able to determine the ratio between intramolecular and intermolecular reaction products by selecting towards transposon and target plasmid (chloramphenicol and kanamycin or kanamycin only) or towards transposon only (chloramphenicol) and enumerating bacterial colonies on selection plates. We noticed that, although we were able to score high numbers of colonies representing intermolecular transposition events into the target plasmid, we obtained even higher numbers of colonies that were only chloramphenicol resistant (Table 2 in II). This meant that the majority of colonies did not contain the target of interest. The most logical explanation for these chloramphenicol resistant colonies was intramolecular transposition, although they could also be gained via *in vivo* circularization of the transposon or illegitimate integration into the host genome. We tested this by incubating reactions without MuA and selecting against chloramphenicol (Table 2 in II). Some chloramphenicol colonies could be detected, but as most, if not all of them were also ampicillin resistant (data not shown), it indicated that they were produced by transposon carrier plasmids. Such contaminants can follow from transposon DNA preparations and are fairly common. We could therefore conclude that the intramolecular transposition was the main mechanism producing the colonies that did not contain the target plasmid.

4.2.5 Reaction product profiles with different donor/target ratios provide the means to predict the success rate in cloning (II)

The effect of the donor/target ratio on the frequency of intramolecular transposition has not been systematically studied in any transposon system. Therefore, to study the ratio of intra- versus intermolecular reaction products, we conducted *in vitro* transposition reactions with different amounts of Mu transposon and target DNA (Figure 2 and Table 3 in II). These results can be used to estimate the number of clones needed for the screening of correct clones under different donor/target ratios. Overall, the results demonstrated that when the amount of target DNA was increased relative to donor DNA, the transposon inserted more frequently into the external target. As a radical reduction in the transposon DNA is not feasible, intramolecular transposition unavoidably dominates when the amount of target DNA is limited. Although the effect of donor/target ratio has not been thoroughly studied prior to this, it was the major factor affecting the frequency of intramolecular transposition also in previous studies with Tn5- and Tn552-based cloning applications (Agron et al. 2002; Jendrisak et al. 2002; Kirby et al. 2002; Wang et al. 2010).

4.2.6 Reducing the amount of intramolecular transposition (II, III)

In previous studies of transposition-based cloning applications, intramolecular transposition was reduced by the utilization of excess target DNA and by the enrichment of intermolecular transposition products via preparative agarose gel electrophoresis (Agron et al. 2002; Jendrisak et al. 2002; Kirby et al. 2002). However, this approach requires that the amount of target DNA is already present at sufficient levels, and gel enrichment demands that the expected intramolecular and intermolecular transposition products are sufficiently different in size as to allow separation. In order to design novel approaches for the reduction of intramolecular transposition, we tested the effects of MuB and BAF (II), as well as the simultaneous use of two transposons (III) on the transposition reaction product profile.

4.2.6.1 MuB can be used to reduce intramolecular transposition (II)

MuB is a Mu-encoded regulatory protein that provides target immunity by directing transposition into distant target sites (Adzuma and Mizuuchi 1988; Adzuma and Mizuuchi 1989; Ge et al. 2010). We examined whether MuB could be used to reduce the frequency of intramolecular transposition in our system (Figure 2 in II). MuB mediated immunity has been observed both with circular supercoiled donor DNA molecules and, to a somewhat weaker extent, linearized donors (Adzuma and Mizuuchi 1988). The achivement of maximum immunity required the presence of all six natural MuA-binding sites in the transposon ends (L1-L3 and R1-R3) (Adzuma and Mizuuchi 1988). Donor molecules with transposons that contained only sites R1 and R2 at both ends exhibited weak but significant immunity (Adzuma and Mizuuchi 1988). In our study, we utilized a linear transposon with a pre-cut configuration and the R1 and R2 MuA-binding sites at each end. In order to maximize our potential effect size, the donor/target ratio was deliberately adjusted to 10:1. Two MuB concentrations were tested, 600 nM and 1200 nM, which both clearly reduced the frequency of intramolecular transposition compared to the reaction without MuB (Figure 2 in II). Our results showed that with the set-up used MuB did not completely eliminate intramolecular transposition, but it nevertheless increased the fraction of intermolecular transposition. This indicated that MuB promoted Mu integration into target DNA that did not contain a Mu end.

4.2.6.2 BAF cannot act as an effective barrier to autointegration factor in a MuA-mediated in vitro transposition reaction (II and this thesis)

A different mechanism for generating self-immunity has been characterized in retrovirus integration. It has been shown that Moloney Murine Leukemia Virus (MoMLV) exploits barrier-to-autointegration factor (BAF) to protect itself from autointegration (Lee and Craigie 1994; Lee and Craigie 1998). BAF is a 10-kDa conserved protein encoded by mammalian genomes. It binds DNA non-specifically and bridges double-stranded DNA, thereby assembling higher-order nucleoprotein

complexes and thus compacting DNA. Within retroviral preintegration complexes, viral DNA that is compacted by BAF is inaccessible as a target for integration (Bradley et al. 2005; Zheng et al. 2000). As the assay bears similarities to the Mu *in vitro* transposition reaction, we wanted to test whether BAF could also decrease autointegration mediated by MuA. Recent *in vivo* studies have shown an effect of BAF also on the DNA transposition reactions of Sleeping Beauty (SB) and piggyBac (PB) transposons (Wang et al. 2014). The autointegration frequency was increased when the BAF gene was knocked down and decreased when the gene was overexpressed, suggesting that BAF is a general autointegration-inhibiting factor in eukaryotic cells (Wang et al. 2014). As BAF compacts DNA and binds double-stranded DNA non-specifically (Bradley et al. 2005), it was feasible to think that BAF might make any transposon a less accessible target for autointegration. Results are shown below, but they are not included in the original publications.

In these experiments the donor/target ratio was 10:1. Transpososomes were preassembled without divalent cations using 1 pmol Ori(pUC)-Cat-Mu transposon DNA. Different amounts of BAF were then added to enable DNA compaction by BAF. Strand transfer was initiated by the addition of pALH31 target DNA (0.1 pmol) and $MgCl_2$, and following a brief incubation period, reaction products were electroporated into *E. coli* cells for selection on plates that contained the appropriate antibiotic(s). Overall, the decrease with BAF in the intra-/intermolecular product ratio was not statistically significant, although at 5 pmol BAF a low level of reduction was discernible. With higher amounts of BAF (25 and 50 pmol), an increase in the product ratio was actually observed (Figure 6). Thus, BAF is not an effective barrier to autointegration in a MuA-mediated *in vitro* transposition reaction.

We believe that the data reflect fundamental differences between retrovirus integration by preintegration complexes and *in vitro* transposition by Mu transpososomes. Retroviral preintegration complexes include the virus DNA, a number of virus-encoded proteins, and certain host proteins that include BAF; instead, the Mu transpososomes used in the current study contain the transposon DNA and only one protein species, the MuA transposase. It may be that the recruitment to preintegration complexes and effective function of BAF require specific protein-protein interactions. For example, it has been suggested that in HIV-1 preintegration complexes, BAF is recruited via the viral protein *gag* (Mansharamani et al. 2003). Similarly, in SB transposition, BAF is enriched in a complex that contains SB transposase and its interactor HMGXB4 (Wang et al. 2014). Thus, the non-functionality of BAF for the inhibition of autointegration in our Mu *in vitro* transposition reactions might be due to the lack of such protein-protein interactions. It could also be that, although BAF can compact the Mu transposon DNA, the remaining free DNA sites still allow the transposon to insert into itself. However, in this study, BAF not only failed to inhibit autointegration, it actually increased the amount of intramolecular transposition relative to intermolecular transposition. This could have resulted from a MuB-like action on targeting. It has been suggested that the

bending or deformation of DNA at the ends of MuB filaments could create a favored target for Mu insertion (Mizuno et al. 2013). In summary, we conclude that BAF cannot act as an effective barrier of autointegration factor in a MuA-mediated *in vitro* transposition reaction.

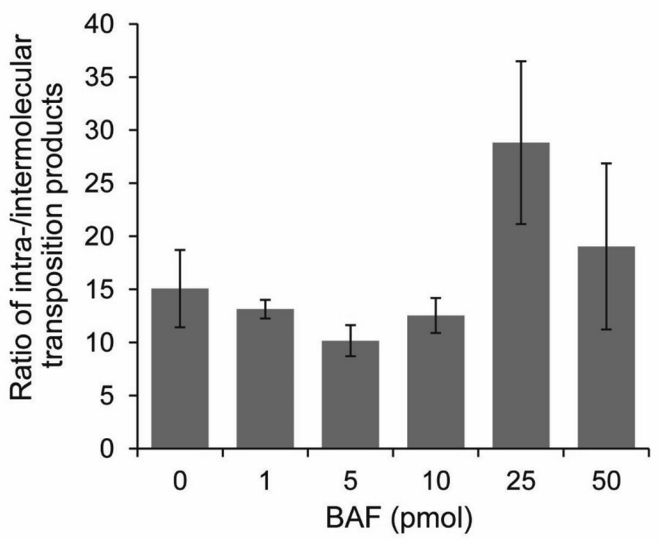

Figure 6. The effect of BAF for the ratio of intra-/intermolecular transposition reaction products. *In vitro* transposition reactions were conducted with 10:1 molar ratio of transposon Ori(pUC)-Cat-Mu to target pALH31 and different amounts of BAF. Reaction products were electroporated into DH10B *E. coli* strain and bacteria were selected with either chloramphenicol (Cm) or with double selection of chloramphenicol and kanamycin (Cm+Km). The ratio of intra/intermolecular transposition products is calculated as (Cm^R-Cm^RKm^R)/Cm^RKm^R. Results are shown as a mean of three replicates. The error bars indicate SD/2 above and below the average value of each data point.

4.2.6.3 Simultaneous use of two transposons reduces intramolecular transposition (III)

Previous experiments used transposons that contained both features that are essential for replication in *E. coli* under antibiotic selection: an origin of replication and a selectable marker. In theory, the background colonies resulting from a transposon inserting into itself could be eliminated by using two transposons, one providing the origin of replication and other the selectable marker. To test this, we incubated the ori(pUC)-Mu (contains pUC-ori) and Cat-Mu (contains the chloramphenicol-resistance marker) transposons simultaneously in one *in vitro* transposition reaction with variable amounts of the pALH31 target plasmid (Figure 5 in III). Our results demonstrated that the use of two transposons significantly reduced the frequency of, but did not eliminate intramolecular transposition. As intramolecular transposition products of neither of the two transposons can survive as stable plasmids inside the *E. coli* cells, we suggest that

the background colonies most probably contain transposon products in which one of the transposons has inserted into the other.

We have shown that the frequency of intramolecular transposition can be reduced by increasing the relative amount of target DNA, by adding MuB into the transposition reaction, or by using two transposons. All these approaches enhanced the cloning efficiency. Adjusting the target/donor ratio was the simplest means to affect the intra/inter ratio, but it required that the target DNA was available in sufficient amounts. If only a limited amount of target DNA was available, use of two transposons was the most efficient means to enhance the cloning efficiency. Use of these different means to reduce the frequency of intramolecular transposition is beneficial in cloning, as they reduce the amount of screening that is needed to discover the properly targeted circles of interest.

4.3 Generation of hyperactive MuA transposase variants (IV)

New applications for transposons in advanced molecular biology have led to the acceleration of efforts to develop hyperactive transposase variants. To improve the transposition system of Mu and facilitate the development of new applications, particularly in genomics but also in gene therapy and transgenesis, we mutagenized the MuA transposase protein and screened for hyperactive variants using an *in vivo* assay. Highly hyperactive versions of MuA were selected for further examination, and substitutions were studied both individually and in different combinations. Mapping the activity-enhancing substitutions on the Mu transpososome structure enabled us to identify the specific locations involved in both catalysis and associations between MuA subunits, as well as those between MuA and Mu DNA. Our results provide insights into the nature of hyperactivity in the context of the entire Mu transpososome, and they may help in generating hyperactive variants of other transposases.

4.3.1 Random MuA mutant libraries

To generate a variety of substitutions in MuA, we used several different error-prone PCR protocols. A total of five MuA mutant libraries were generated, which included a total of ~3×10^5 independently generated plasmid clones (Supplementary Table 4 in IV). For papillation assays, these plasmids were introduced into DH5α cells, and altogether 64,000 colonies from different mutant libraries were screened for enhanced papillation. Approximately 1-3% of the clones per mutant library clearly exhibited an enhanced papillation frequency. These results showed that the mutagenesis protocols that we used were capable of producing a large number of desired hyperactive MuA variants.

The amino-acid changes produced by any error-prone PCR mutagenesis method are most likely to be those that require only one nucleotide substitution, and therefore these

protocols do not generate all possible amino-acid changes. It has been shown that single-nucleotide mutations generate, on average, 5.7 different amino-acid substitutions for any given amino-acid residue (Miyazaki and Arnold 1999), which results in 39.5% coverage of the protein-level diversity (Wong et al. 2007). Although we were able to probe all of the important interfaces in the transpososome structure for amino-acid changes, it is possible that the most beneficial substitutions could not be identified for each amino-acid residue. An interesting topic for future studies would be to use, for example, the trinucleotide exchange method (TriNEx) to randomly replace amino-acid residues with all possible counterparts (Baldwin et al. 2008).

4.3.2 Random mutations in MuA can induce hyperactivity

To verify that the observed high-papillation-activity phenotypes were caused by mutations in the MuA gene, we selected a large set of clones for further examination. Plasmids from 222 clones were re-assayed for papillation and the MuA gene was re-cloned from 89 plasmids. Both of these re-examinations produced phenotypes identical to their respective original counterparts, showing that mutations in the MuA gene were indeed responsible for the observed phenotypes.

To reveal changes at the amino-acid level, the MuA sequence was determined from the 89 re-cloned plasmids. Of these, 71 were unique, and most contained several changes (Supplementary Table 5 in IV). To enable an accurate comparison between the transpositional activities of even the most active mutants, the plasmids were subjected to papillation analysis under stringent growth conditions (140 h incubation at 25˚C). These conditions revealed a broad spectrum of hyperactivity among the variants, and a correlation with specific amino-acid changes was apparent (Figure 3 and Supplementary Table 5 in IV). Although many changes were identified only once, certain mutations were identified frequently and a number of clear hot spots could be found. Presumably many of the changes that were identified only once were irrelevant for the hyperactive phenotype. This demonstrated that the papillation analysis was effective in screening for mutations that cause hyperactive phenotypes.

4.3.3 Single-substitution MuA variants

To study substitutions individually, we generated 47 single-substitution MuA variants. The substitutions were selected so that they represented all MuA subdomains. Other criteria were their presence in several independent clones or their presence in an individual clone of high activity (Figure 3 and Supplementary Table 5 in IV). To quantify the transpositional activity of the constructed plasmids, we subjected them to papillation analysis (115 h incubation at 30˚C) (Figure 4 in IV). At least a two-fold increase in protein activity was observed with 34 substitutions at 26 specific amino-acid residues and more than a five-fold activity enhancement was detected with 27 substitutions. The highest increase in activity was 50-fold over the wild-type level.

These results indicated that a variety of substitutions in different subdomains resulted in the hyperactivity of MuA.

4.3.4 Activity-enhancing single substitutions in different domains of MuA

To allow structure-function analysis, activity-enhancing substitutions were mapped on the recently resolved Mu transpososome structure. Each of the MuA subdomains contained activity-enhancing substitutions, although domains Iα and IIIβ each contained only one substitution, and those enhanced activity only slightly. The Iα and IIIβ domains are associated with the binding of the enhancer sequence and MuB protein, respectively, and both of these features are missing in our transposition assay. It has been previously shown that domain Iα is inhibitory in the absence of the enhancer (Mizuuchi and Mizuuchi 1989; Pajunen et al. 2010; Yang et al. 1995), and N-terminal deletion variants demonstrate significantly improved activity (Pajunen et al. 2010). As the Iα and IIIβ domains are not involved in the transpososome crystal structure, a detailed functional explanation of the effect of these residues is not currently possible.

The substitutions that most enhanced activity were located in the central subdomains, from Iβ to IIIα. When mapped to the Mu transpososome structure, many of these substitutions were found to be at the interface between the Iβ domain (residues 97 and 160) of the R1-bound subunit and the IIα subdomain (residues 478, 482, 483, and 487) of the R2-bound subunit, which are located in the same DNA segment (Figure 6 in IV). These substitutions may possibly enhance activity by improving protein-protein interactions. In addition, important activity-enhancing substitutions were located at the interface between the IIβ and IIIα subdomains of the R2-bound subunit and the Mu-end recognition and binding domain Iγ (residues 232 and 233) of the R1-bound subunit (Figure 6C in IV). The most active single-amino-acid substitution variants were E233K and E233V, and the hyperactivity caused by these substitutions is possibly a result of electrostatic changes which might affect protein-protein interactions.

The vast majority of the hyperactivity-inducing mutations were found near the catalytic subdomain IIα (residues 302, 302, 335, 340, 345, 345, 374, and 374) of the R1-bound subunits (Figure 6D in IV). Several of these substitutions change the packing and flexibility properties of the residue, and they most probably affect the packing of the catalytic core. Other hyperactivity-inducing substitutions found in subdomain IIα of the R1-bound subunits (residues 254, 258, 447, 447, 464, and 466) may affect the conformational changes required for the assembly of a catalytically competent transpososome.

Some of the hyperactivity-inducing substitutions probably improve interactions with DNA on both the R1 and R2 subunits (Figure 6E and 6F in IV). For example, substitution with a less-negative amino acid at residue E179, which is close to Mu-end DNA in both R1 and R2 subunits, could improve DNA binding. Likewise, a

substitution of the Q594 residue, which is close to target DNA on the R1-bound subunit and flanking host DNA on the R2-bound subunit, may improve DNA interactions.

In summary, the results demonstrated that the activity-inducing substitutions have an effect on different phases of the transposition pathway, including tetramer assembly, catalysis, DNA binding, and folding properties of the MuA protein.

4.3.5 Synergistic effects of substitutions

Although activity enhancement was detected with the single-substitution variants, the most-active MuA variants contained multiple substitutions. This suggested that the substitutions had synergistic effects. The most-active MuA variant (clone EP3I4) contained five amino-acid substitutions (W160R, A234V, W345R, M374V, and T543A) (Figure 7 and Supplementary Table 5 in IV). Three of these substitutions (W160R, W345R, and M374V) were found to increase activity over the wild-type level when studied individually (Figure 7 and Table 2 in IV). In order to study their synergistic effects, these three substitutions were grouped in double- and triple-substitution combinations. Of these combinations, the triple-substitution mutant (W160R, W345R, M374V) generated the highest activity (~264 papillae per colony) and produced the same degree of activity as the original variant, EP3I4 (~295 papillae per colony) (Table 2 in IV). To examine the effect of combining mutations from three different subdomains, we changed the M374V mutation to E233K. This variant (W160R, E233K, W345R) was even more active (~519 papillae per colony) than the original EP3I4 (~295 papillae per colony) (Table 2 in IV). Thus, these results showed that the substitutions had synergistic effects and emphasized the independent roles of MuA subdomains in generating hyperactive variants.

4.3.6 MuA variants for application purposes

The papillation assays we conducted quantified the transpositional activity of the MuA variants *in vivo*. To assess the transpositional activity of MuA variants in two assays, *in vitro* assay and in *in vivo* assay with pre-assembled transpososomes, we purified wild type MuA, 30 single-substitution MuA variants and a triple mutant (W160R, E233K, W345R) (Figures 8 and 9, Supplementary Figures 2 and 3 in IV). In the *in vitro* assay, plasmid DNA was used as a target for transposon integrations, and in *in vivo* assay, the target was chromosomal DNA. Differently from the *in vitro* assay, in the *in vivo* assay, pre-assembled transpososomes encounter Mg^{2+} ions and thus become active only after electroporation into the *E. coli* cells.

For the *in vitro* transposition reactions with the MuA variants, we used a pre-cleaved mini-Mu transposon as donor DNA and the pUC19 plasmid as a target (Figure 8A in IV). Reaction products were transformed into competent *E. coli* cells and plated on

appropriate antibiotic selection plates in order to score integration events into target plasmids (Figure 8B in IV). In addition, reaction products were analyzed qualitatively by agarose gel electrophoresis (Supplementary Figure 2 in IV). In the *in vivo* assay with pre-assembled transpososomes, we incubated the pre-cleaved mini-Mu transposon with MuA variants and electroporated the pre-assembled complexes into *E. coli* cells (Figure 9 in IV). Transpososome formation was assayed by native gel electrophoresis (Supplementary figure 3 in IV). In both assays, the vast majority of MuA variants demonstrated levels of activity that were at least equivalent to the wild type, and most were higher. However, the activity range detected *in vitro* was narrower than that detected with the papillation assay, and one variant (D320V) was totally inactive *in vitro*. In addition, the activities of the variants differed somewhat more from their *in vivo* papillation activities when quantified with the assay with pre-assembled transposition complexes, than with the *in vitro* transposition reactions.

Both the *in vitro* transposition reaction and *in vivo* papillation assay simulate the entire transposition pathway from assembly to catalysis; the main difference between the two is that the *in vitro* reaction utilizes pre-cut substrates, and therefore does not involve the donor cleavage step which occurs in the papillation assay. The differences between the results of the *in vitro* and *in vivo* (papillation) assays may reflect the influence of host factors. In the assay with the pre-assembled transposition complexes, the transposition pathway differs more from the *in vivo* papillation assay, as the complex assembly process is separate from integration into the target. Similar differences in activities between *in vivo* and *in vitro* assays have been detected also in other transposition systems with hyperactive transposases (Lampe et al. 1999).

4.3.7 Further studies on synergism

As the mutant proteins with three substitutions were highly active, we wanted to study whether the protein activity could be increased even further by adding more substitutions to the protein (Table 1 in IV). Initially, transpositional activities of the mutant combinations were assessed using the papillation assay, and all the mutant combination variants were extremely efficient (up to 6-fold activity enhancement over the highly active control triple-substitution mutant) (Figure 10 in IV). Next, the transpositional activity of purified mutant combination variants was assessed in two assays, *in vitro* assay and in *in vivo* assay with pre-assembled transpososomes (Figure 11 in IV), but the increase in transpositional activity observed with the purified mutant combination variants was not higher than that found with the purified single-substitution variants. However, even a small improvement in efficiency compared to the wild type level makes a significant contribution to transgenesis and gene therapy approaches and to the development of novel applications based on the Mu system. In addition, studies which combine analysis of structure and function greatly benefit future efforts to systematically design new mutants with features appropriate to a given task.

5 CONCLUSIONS AND FUTURE PERSPECTIVES

This study established novel Mu transposition-based molecular genetics tools and generated MuA transposase variants with altered transposition frequency.

The main findings and conclusions are as follows:

1. A general-purpose test to measure the specific activity of transposition complexes was established using phage Mu transposition as a test platform. This assay utilizes a common suicide target, is applicable to all types of transposons that have a functional *in vitro* reaction, and allows the direct comparison of transpositional activities within or between particular experimental systems. The assay is particularly useful with transposons that do not contain selectable marker genes or contain markers that are functional only in eukaryotic cells.

2. A Mu transposition-based cloning strategy for replication-deficient circular DNA was developed. This method utilizes a transposon that contains an origin of replication and an antibiotic-selection marker gene. The method can be used for cloning or "rescuing" circular DNA that does not contain a selectable marker (including cryptic plasmids, circular recombination intermediates, and mtDNA). In essence, these circles are transfigured into plasmids or bacterial artificial chromosomes which then are able to replicate in *E. coli*. The method enables the cloning of circular DNA from very low target amounts, even in the presence of contaminating linear DNA fragments. The donor/target ratio was found to be the major factor affecting the frequency of intramolecular transposition. The use of the MuB protein enabled us to reduce intramolecular transposition.

3. A platform with ten mini-Mu transposons for the Mu transposition-based cloning strategy for replication-deficient circular DNA was established. An assortment of replication origins enables several different approaches for the study of DNA-circles of interest. The simultaneous use of two transposons (one containing an origin of replication and other containing a selection marker) in one reaction substantially reduced the amount of intramolecular transposition.

4. Hyperactive MuA variants were generated by mutagenizing the catalytic MuA transposase protein and screened using an *in vivo* papillation assay. The papillation assay enabled us to identify various individual substitutions that increased transposition activity with respect to that of the wild-type MuA protein. Structure-function studies indicated that the activity-enhancing substitutions have an effect on different phases of the transposition pathway, including tetramer assembly, catalysis, DNA binding, and folding of the MuA protein.

The four studies I have presented extend the variety of Mu-based tools available and enable more efficient use of the Mu *in vitro* transposition reaction for different applications. Furthermore, the method developed for the measurement of the specific activity of transposition complexes should also be directly applicable to other transposition-based systems with a functional *in vitro* reaction and therefore could provide a common transposon-based tool.

Transposon-based tools are under active development, with constant efforts being made to enhance transpositional activity. The hyperactive MuA variants generated in this study will improve the utility of MuA-based tools and can promote the development of new applications also in the disciplines of gene therapy and gene delivery in mammalian cells.

Although transposon-based technologies are currently more extensively used in bacterial and invertebrate models, current progress in transposon research indicates that, in the future, they will serve as indispensable tools in the biotechnology toolkit also for vertebrate models, in which the full potential of transposons as gene-delivery vehicles can be exploited.

ACKNOWLEDGMENTS

This study was carried out at the University of Turku Department of Biology in the DNA recombination laboratory. The work was funded by the Jenny and Antti Wihuri Foundation, the Oskar Öflund Foundation, the University of Turku Graduate School, and the University of Turku Foundation, which provided grants to Elsi Pulkkinen. In addition this work was financially supported by a grant from the Academy of Finland to Professor Harri Savilahti.

First, I want to thank my supervisors, Professor Harri Savilahti and Docent Saija Haapa-Paananen. Harri, thank you for allowing me to do my PhD studies in your research group. I have always appreciated you as a researcher, and I am grateful for all the knowledge I have gained from you. Saija H-P, thank you for all the help and guidance provided. Even though you were not my supervisor for the whole period of my studies, your guidance was notable with regard to the finalization of the laboratory work and its analysis. I thank Professor Marko Virta for kindly agreeing to act as my opponent, and Professor Craig Primmer for accepting to be custos at my defense. I thank Professor Maia Kivisaar and Docent Mikko Frilander for carefully reviewing this thesis.

I wish to thank all the co-authors who have been involved in the publications of my thesis: Saija Haapa-Paananen, Saija Kiljunen, Lars Paulin, Tiina Rasila, Phoebe Rice, Hilkka Turakainen, and Mauno Vihinen.

I warmly thank all former and present members of the Savilahti group at the University of Turku: Saija Haapa-Paananen, Saija Kiljunen, Mikko Nieminen, and Maria Pajunen. I would especially like to thank Saija Kiljunen for carefully reviewing the thesis manuscript and for providing insightful comments and suggestions for improvement. I am also thankful for Saija K for always finding the time to answer my questions. In addition, I want to thank the past members of the Savilahti group at the Institute of Biotechnology: Pirjo Rahkola, Hilkka Turakainen, Anja Paatero, and Tiina Rasila. Thank you for all the advice that you provided when I was at the very beginning of my studies and visiting the lab at the Institute of Biotechnology.

I thank also my other colleagues in the Genetics department, especially our lunch group: Katja Salminen, Sanna Pausio, and Heidi Viitaniemi. We have shared many moments talking about both serious and less-serious subjects and learning how to make omelets in the microwave oven.

Many thanks also to all my friends in and outside the scientific world.

Lastly I would like to thank my family members for the lifelong support that you have provided to me. My dearest thanks go to Lauri for your love and support through these studies.

REFERENCES

Adzuma K, Mizuuchi K (1991) Steady-state kinetic analysis of ATP hydrolysis by the B protein of bacteriophage mu. Involvement of protein oligomerization in the ATPase cycle. J Biol Chem 266:6159-6167

Adzuma K, Mizuuchi K (1989) Interaction of proteins located at a distance along DNA: mechanism of target immunity in the Mu DNA strand-transfer reaction. Cell 57:41-47

Adzuma K, Mizuuchi K (1988) Target immunity of Mu transposition reflects a differential distribution of Mu B protein. Cell 53:257-266

Agron PG, Sobecky P, Andersen GL (2002) Establishment of uncharacterized plasmids in Escherichia coli by in vitro transposition. FEMS Microbiol Lett 217:249-254.

Akhverdyan VZ, Gak ER, Tokmakova IL, Stoynova NV, Yomantas YA, Mashko SV (2011) Application of the bacteriophage Mu-driven system for the integration/amplification of target genes in the chromosomes of engineered Gram-negative bacteria--mini review. Appl Microbiol Biotechnol 91:857-871

Aldaz H, Schuster E, Baker TA (1996) The interwoven architecture of the Mu transposase couples DNA synapsis to catalysis. Cell 85:257-269

Allet B (1979) Mu insertion duplicates a 5 base pair sequence at the host inserted site. Cell 16:123-129

Aziz RK, Breitbart M, Edwards RA (2010) Transposases are the most abundant, most ubiquitous genes in nature. Nucleic Acids Res 38:4207-4217

Bainton RJ, Kubo KM, Feng J-, Craig NL (1993) Tn*7* transposition: target DNA recognition is mediated by multiple Tn*7*-encoded proteins in a purified *in vitro* system. Cell 72:931-943

Baker TA, Luo L (1994) Identification of residues in the Mu transposase essential for catalysis. Proc Natl Acad Sci USA 91:6654-6658

Baker TA, Mizuuchi M, Savilahti H, Mizuuchi K (1993) Division of labor among monomers within the Mu transposase tetramer. Cell 74:723-733

Baker TA, Mizuuchi K (1992) DNA-promoted assembly of the active tetramer of the Mu transposase. Genes Dev 6:2221-2232

Baker TA, Mizuuchi M, Mizuuchi K (1991) MuB protein allosterically activates strand transfer by the transposase of phage Mu. Cell 65:1003-1013

Baldwin AJ, Busse K, Simm AM, Jones DD (2008) Expanded molecular diversity generation during directed evolution by trinucleotide exchange (TriNEx). Nucleic Acids Res 36:e77

Beauregard A, Curcio MJ, Belfort M (2008) The take and give between retrotransposable elements and their hosts. Annu Rev Genet 42:587-617

Bender J, Kuo J, Kleckner N (1991) Genetic evidence against intramolecular rejoining of the donor DNA molecule following IS10 transposition. Genetics 128:687-694

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Res 28:235-242

Bernard P, Gabant P, Bahassi EM, Couturier M (1994) Positive-selection vectors using the F plasmid ccdB killer gene. Gene 148:71-4

Bernard P, Couturier M (1992) Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. J Mol Biol 226:735-745

Bhasin A, Goryshin IY, Reznikoff WS (1999) Hairpin formation in Tn*5* transposition. J Biol Chem 274:37021-37029

Biemont C, Vieira C (2006) Genetics: junk DNA as an evolutionary force. Nature 443:521-4

Boeke JD (2002) Putting mobile DNA to work: the toolbox. In: Craig NL, Craigie R, Gellert M, Lambowitz AM (eds) Mobile DNA II. ASM Press, Washington, D.C., pp 24-37

Bradley CM, Ronning DR, Ghirlando R, Craigie R, Dyda F (2005) Structural basis for DNA bridging by barrier-to-autointegration factor. Nat Struct Mol Biol 12:935-936

Brady T, Roth SL, Malani N, Wang GP, Berry CC, Leboulch P, Hacein-Bey-Abina S, Cavazzana-Calvo M, Papapetrou EP, Sadelain M, Savilahti H, Bushman FD (2011) A method to sequence and quantify DNA integration for monitoring outcome in gene therapy. Nucleic Acids Res 39:e72

Braid MD, Silhavy JL, Kitts CL, Cano RJ, Howe MM (2004) Complete genomic sequence of bacteriophage B3, a Mu-like phage of *Pseudomonas aeruginosa*. J Bacteriol 186:6560-6574

Broach J, Volkert F (1991) 6 Circular DNA Plasmids of Yeasts. Cold Spring Harb Monogr Archive 21A:297-331

Chaconas G, Harshey RM (2002) Transposition of phage Mu DNA. In: Craig NL, Craigie R, Gellert M, Lambowitz AM (eds) Mobile DNA II. ASM Press, Washington, D.C., pp 384-402

Chalmers RM, Kleckner N (1994) Tn10/IS10 transposase purification, activation, and in vitro reaction. J Biol Chem 269:8029-8035

Chandler M, Clerget M, Galas DJ (1982) The transposition frequency of IS1-flanked transposons is a function of their size. J Mol Biol 154:229-43

Chandler M, Mahillon J (2002) Insertion sequences revisited. In: Craig NL, Craigie R, Gellert M, Lambovitz AM (eds) Mobile DNA II. ASM Press, Washington, DC, pp 305-336

Chandler M, Fayet O, Rousseau P, Ton Hoang B, Duval-Valentin G (2015) Copy-out-Paste-in Transposition of IS911: A Major Transposition Pathway. Microbiol Spectr 3:10.1128/microbiolspec.MDNA3-0031-2014

Choi W, Harshey RM (2010) DNA repair by the cryptic endonuclease activity of Mu transposase. Proc Natl Acad Sci U S A 107:10014-9

Clubb RT, Schumacher S, Mizuuchi K, Gronenborn AM, Clore GM (1997) Solution structure of the Ig subdomain of the Mu end DNA-binding domain of phage Mu transposase. J Mol Biol 273:19-25

Clubb RT, Omichinski JG, Savilahti H, Mizuuchi K, Gronenborn AM, Clore GM (1994) A novel class of winged helix-turn-helix protein: the DNA-binding domain of Mu transposase. Structure 2:1041-1048

Coros CJ, Sekino Y, Baker TA, Chaconas G (2003) Effect of mutations in the C-terminal domain of MuB on DNA binding and interactions with MuA transposase. J Biol Chem 278:31210-31217

Craig NL, Chandler M, Gellert M, Lambowitz AM, Rice PA, Sandmeyer SB (2015) Mobile DNA III. ASM Press, Washington DC

Craig NL, Craigie R, Gellert M, Lambowitz AM (2002) Mobile DNA II. ASM Press, Washington, D.C.

Craig NL (1997) Target site selection in transposition. Annu Rev Biochem 66:437-474

Craigie R, Mizuuchi K (1987) Transposition of Mu DNA: joining of Mu to target DNA can be uncoupled from cleavage at the ends of Mu. Cell 51:493-501

Craigie R, Mizuuchi K (1986) Role of DNA topology in Mu transposition: mechanism of sensing the relative orientation of two DNA segments. Cell 45:793-800

Craigie R, Mizuuchi M, Mizuuchi K (1984) Site-specific recognition of the bacteriophage Mu ends by the MuA protein. Cell 39:387-394

Curcio MJ, Derbyshire KM (2003) The outs and ins of transposition: from Mu to Kangaroo. Nat Rev Mol Cell Biol 4:865-877

Dawson A, Finnegan DJ (2003) Excision of the *Drosophila mariner* transposon *Mos1*: comparison with bacterial transposition and V(D)J recombination. Mol Cell 11:225-235

Devine SE, Boeke JD (1994) Efficient integration of artificial transposons into plasmid targets *in vitro*: a useful tool for DNA mapping, sequencing and genetic analysis. Nucleic Acids Res 22:3765-3772

Dupuy AJ, Akagi K, Largaespada DA, Copeland NG, Jenkins NA (2005) Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. Nature 436:221-6

Edwards WR, Busse K, Allemann RK, Jones DD (2008) Linking the functions of unrelated proteins using a novel directed evolution domain insertion method. Nucleic Acids Res 36:e78

Elliott TA, Gregory TR (2015) Do larger genomes contain more diverse transposable elements? BMC Evol Biol 15:69-015-0339-8

Engelman A, Mizuuchi K, Craigie R (1991) HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. Cell 67:1211-1221

Engels WR, Johnson-Schlitz DM, Eggleston WB, Sved J (1990) High-frequency P element loss in Drosophila is homolog dependent. Cell 62:515-525

Fattash I, Rooke R, Wong A, Hui C, Luu T, Bhardwaj P, Yang G (2013) Miniature inverted-repeat transposable elements: discovery, distribution, and activity. Genome 56:475-486

Feschotte C, Pritham EJ (2007) DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet 41:331-368

Feschotte C, Zhang X, Wessler SR (2002) Miniature inverted-repeat transposable elements and their relationships to established DNA transposons. In: Craig NL, Craigie R, Gellert M, Lambovitz AM (eds) Mobile DNA II. ASM Press, Washington, D.C, pp 1147-1157

Filee J, Siguier P, Chandler M (2007) Insertion sequence diversity in archaea. Microbiol Mol Biol Rev 71:121-157

Finnegan DJ (1989) Eukaryotic transposable elements and genome evolution. Trends Genet 5:103-107

Fugmann SD (2010) The origins of the Rag genes—From transposition to V(D)J recombination. Semin Immunol 22:10-16

Ge J, Lou Z, Harshey RM (2010) Immunity of replicating Mu to self-integration: a novel mechanism employing MuB protein. Mob DNA 1:8-8753-1-8

Gellert M (2002) V(D)J recombination: RAG proteins, repair factors, and regulation. Annu Rev Biochem 71:101-132

Goodier JL, Kazazian HH,Jr (2008) Retrotransposons revisited: the restraint and rehabilitation of parasites. Cell 135:23-35

Goosen N, van de Putte P (1987) Regulation of transcription. In: Symonds N, Toussaint A, van de Putte P, Howe MM (eds) Phage Mu. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 41-52

Goryshin IY, Jendrisak J, Hoffman LM, Meis R, Reznikoff WS (2000) Insertional transposon mutagenesis by electroporation of released Tn*5* transposition complexes. Nat Biotechnol 18:97-100

Goryshin IY, Reznikoff WS (1998) Tn*5 in vitro* transposition. J Biol Chem 273:7367-7374

Greene EC, Mizuuchi K (2004) Visualizing the assembly and disassembly mechanisms of the MuB transposition targeting complex. J Biol Chem 279:16736-16743

Greene EC, Mizuuchi K (2002a) Dynamics of a protein polymer: the assembly and disassembly pathways of the MuB transposition target complex. EMBO J 21:1477-1486

Greene EC, Mizuuchi K (2002b) Target immunity during Mu DNA transposition: transpososome assembly and DNA looping enhance MuA-mediated disassembly of the MuB target complex. Mol Cell 10:1367-1378

Grindley NDF (2002) The movement of Tn*3*-like elements: transposition and cointegrate resolution. In: Craig L, Craigie R, Gellert M, Lambowitz A (eds) Mobile DNA II. ASM Press, Washington DC, pp 272-302

Groenen MAM, van de Putte P (1986) Analysis of the ends of bacteriophage Mu using site-directed mutagenesis. J Mol Biol 189:597-602

Haapa S, Suomalainen S, Eerikäinen S, Airaksinen M, Paulin L, Savilahti H (1999a) An efficient DNA sequencing strategy based on the bacteriophage Mu *in vitro* DNA transposition reaction. Genome Res 9:308-315

Haapa S, Taira S, Heikkinen E, Savilahti H (1999b) An efficient and accurate integration of mini-Mu transposons *in vitro*: a general methodology for functional genetic analysis and molecular biology applications. Nucleic Acids Res 27:2777-2784

Haapa-Paananen S, Rita H, Savilahti H (2002) DNA transposition of bacteriophage Mu. A quantitative analysis of target site selection *in vitro*. J Biol Chem 277:2843-2851

Hagemann AT, Craig NL (1993) Tn7 transposition creates a hotspot for homologous recombination at the transposon donor site. Genetics 133:9-16

Han YW, Mizuuchi K (2010) Phage Mu transposition immunity: protein pattern formation along DNA by a diffusion-ratchet mechanism. Mol Cell 39:48-58

Haniford DB, Ellis MJ (2015) Transposons Tn10 and Tn5. Microbiol Spectr 3:MDNA3-0002-2014

Haniford DB (2002) Transposon Tn*10*. In: Craig NL, Craigie R, Gellert M, Lambowitz A (eds) Mobile DNA II. ASM Press, Washington D.C., pp 457-483

Harshey RM (2014) Transposable Phage Mu. Microbiol Spectr 2:10.1128/ microbiolspec.MDNA3-0007-2014

Harshey RM (2012) The Mu story: how a maverick phage moved the field forward. Mob DNA 3:21-8753-3-21

Harshey RM, Jayaram M (2006) The Mu transpososome through a topological lens. Crit Rev Biochem Mol 41:387-405

Harshey RM, Cuneo SD (1986) Carboxyl-terminal mutants of phage Mu transposase. J Genet 65:159-174

Harshey RM, Getzoff ED, Baldwin DL, Miller JL, Chaconas G (1985) Primary structure of phage Mu transposase: homology to Mu repressor. Proc Natl Acad Sci U S A 82:7676-7680

Hayes F (2003) Transposon-based strategies for microbial functional genomics and proteomics. Annu Rev Genet 37:3-29

Hickman AB, Dyda F (2015) Mechanisms of DNA Transposition. Microbiol Spectr 3:MDNA3-0034-2014

Hickman AB, Chandler M, Dyda F (2010) Integrating prokaryotes and eukaryotes: DNA transposases in light of structure. Crit Rev Biochem Mol 45:50-69

Invitrogen by Life Technologies (2012) Zero Background™ /Kan Cloning Kit., Carlsbad, California

Ivics Z, Izsvak Z (2010) The expanding universe of transposon technologies for gene and cell engineering. Mob DNA 1:25-8753-1-25

Ivics Z, Li MA, Mates L, Boeke JD, Nagy A, Bradley A, Izsvak Z (2009) Transposon-mediated genome manipulation in vertebrates. Nat Methods 6:415-22

Ivics Z, Izsvák Z (2006) Transposons for gene therapy! Curr Gene Ther 6:593-607

Izsvák Z, Ivics Z, Plasterk RH (2000) Sleeping beauty, a wide host-range transposon vector for genetic transformation in vertebrates. J Mol Biol 302:93-102

Jang S, Sandler SJ, Harshey RM (2012) Mu insertions are repaired by the double-strand break repair pathway of Escherichia coli. PLoS Genet 8:e1002642

Jendrisak J, Decker J, Meis R (2002) In Vitro Insertion of a Transposon Containing an E. coli Origin of Replication Facilitates Rapid Recovery, Propagation and Sequencing of Circular DNA Molecules from Heterologous Organisms. EPICENTRE Forum 9:14-15

Jones DD (2005) Triplet nucleotide removal at random positions in a target gene: the tolerance of TEM-1 b-lactamase to an amino acid deletion. Nucleic Acids Res 33:e80

Jukkola T, Trokovic R, Maj P, Lamberg A, Mankoo B, Pachnis V, Savilahti H, Partanen J (2005) *Meox1*^{Cre}: a mouse line expressing Cre recombinase in somitic mesoderm. Genesis 43:148-153

Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J (2005) Repbase Update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110:462-7

Kahmann R, Kamp D (1979) Nucleotide sequences of the attachment sites of bacteriophage Mu DNA. Nature 280:247-250

Kapitonov VV, Jurka J (2008) A universal classification of eukaryotic transposable elements implemented in Repbase. Nat Rev Genet 9:411-2

Kekarainen T, Savilahti H, Valkonen JP (2002) Functional genomics on *potato virus A*: virus genome-wide map of sites essential for virus propagation. Genome Res 12:584-594

Kennedy AK, Guhathakurta A, Kleckner N, Haniford DB (1998) Tn*10* transposition via a DNA hairpin intermediate. Cell 95:125-134

Kiljunen S, Vilen H, Pajunen M, Savilahti H, Skurnik M (2005) Nonessential genes of phage fYeO3-12 include genes involved in adaptation to growth on *Yersinia enterocolitica* serotype O:3. J Bacteriol 187:1405-1414

Kim K, Harshey RM (1995) Mutational analysis of the *att* DNA-binding domain of phage Mu transposase. Nucleic Acids Res 23:3937-3943

Kim K, Namgoong SY, Jayaram M, Harshey RM (1995) Step-arrest mutants of phage Mu transposase. Implications in DNA-protein assembly, Mu end cleavage, and strand transfer. Trends Biochem Sci 270:1472-1479

Kirby C, Waring A, Griffin TJ, Falkinham JO,3rd, Grindley ND, Derbyshire KM (2002) Cryptic plasmids of Mycobacterium avium: Tn552 to the rescue. Mol Microbiol 43:173-186

Kobryn K, Watson MA, Allison RG, Chaconas G (2002) The Mu three-site synapse: a strained assembly platform in which delivery of the L1 transposase binding site triggers catalytic commitment. Mol Cell 10:659-669

Krementsova E, Giffin MJ, Pincus D, Baker TA (1998) Mutational analysis of the Mu transposase. Contributions of two distinct regions of domain II to recombination. J Biol Chem 273:31358-31365

Kruklitis R, Welty DJ, Nakai H (1996) ClpX protein of *Escherichia coli* activates bacteriophage Mu transposase in the strand transfer complex for initiation of Mu DNA synthesis. EMBO J 15:935-944

Krupovic M, Makarova KS, Forterre P, Prangishvili D, Koonin EV (2014) Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. BMC Biol 12:36-7007-12-36

Krupovic M, Vilen H, Bamford JK, Kivelä HM, Aalto J-, Savilahti H, Bamford DH (2006) Genome characterization of lipid-containing marine bacteriophage PM2 by transposon insertion mutagenesis. J Virol 80:9270-9278

Kulkosky J, Jones KS, Katz RA, Mack JPG, Skalka AM (1992) Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol Cell Biol 12:2331-2338

Kuo C-, Zou AH, Jayaram M, Getzoff E, Harshey R (1991) DNA-protein complexes during attachment-site synapsis in Mu DNA transposition. EMBO J 10:1585-1591

Lamberg A, Nieminen S, Qiao M, Savilahti H (2002) Efficient insertion mutagenesis strategy for bacterial genomes involving electroporation of *in vitro*-assembled DNA transposition complexes of bacteriophage Mu. Appl Environ Microb 68:705-712

Lampe DJ, Akerley BJ, Rubin EJ, Mekalonos JJ, Robertson HM (1999) Hyperactive transposase mutants of the Himar 1 mariner transposon. Proc Natl Acad Sci USA 96:11428-33

Lampe DJ, Churchill MEA, Robertson HM (1996) A purified *mariner* transposase is sufficient to mediate transposition *in vitro*. EMBO J 15:5470-5479

Lampe DJ, Grant TE, Robertson HM (1998) Factors affecting transposition of the Himar1 mariner transposon in vitro. Genetics 149:179-87

Lamrani S, Ranquet C, Gama MJ, Nakai H, Shapiro JA, Toussaint A, Maenhaut-Michel G (1999) Starvation-induced Mucts62-mediated coding sequence fusion: a role for ClpXP, Lon, RpoS and Crp. Mol Microbiol 32:327-343

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409:860-921

Laurent LC, Olsen MN, Crowley RA, Savilahti H, Brown PO (2000) Functional characterization of the human immunodeficiency virus type 1 genome by genetic footprinting. J Virol 74:2760-2769

Lavoie BD, Shaw GS, Millner A, Chaconas G (1996) Anatomy of a flexer-DNA complex inside a higher-order transposition intermediate. Cell 85:761-771

Lavoie BD, Chan BS, Allison RG, Chaconas G (1991) 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 10:3051-3059

Leach DRF (1996) Genetic Recombination. Blackwell Science Ltd, Oxford, UK

Lee E, Iskow R, Yang L, Gokcumen O, Haseley P, Luquette LJ,3rd, Lohr JG, Harris CC, Ding L, Wilson RK, Wheeler DA, Gibbs RA, Kucherlapati R, Lee C, Kharchenko PV, Park PJ, Cancer Genome Atlas Research Network (2012) Landscape of somatic retrotransposition in human cancers. Science 337:967-971

Lee MS, Craigie R (1998) A previously unidentified host protein protects retroviral DNA from autointegration. Proc Natl Acad Sci U S A 95:1528-1533

Lee MS, Craigie R (1994) Protection of retroviral DNA from autointegration: involvement of a cellular factor. Proc Natl Acad Sci U S A 91:9823-9827

Leschziner AE, Griffin TJt, Grindley NDF (1998) Tn*552* transposase catalyzes concerted strand transfer *in vitro*. Proc Natl Acad Sci U S A 95:7345-7350

Leung PC, Harshey RM (1991) Two mutations of phage Mu transposase that affect strand transfer or interactions with B protein lie in distinct polypeptide domains. J Mol Biol 219:189-199

Leung PC, Teplow DB, Harshey RM (1989) Interaction of distinct domains in Mu transposase with Mu DNA ends and an internal transpositional enhancer. Nature 338:656-658

Levchenko I, Yamauchi M, Baker TA (1997) ClpX and MuB interact with overlapping regions of Mu transposase: implications for control of the transposition pathway. Gene Dev 11:1561-1572

Maekawa T, Yanagihara K, Ohtsubo E (1996) A cell-free system of Tn3 transposition and transposition immunity. Genes Cells 1:1007-1016

Manna D, Higgins NP (1999) Phage Mu transposition immunity reflects supercoil domain structure of the chromosome. Mol Microbiol 32:595-606

Mansharamani M, Graham DR, Monie D, Lee KK, Hildreth JE, Siliciano RF, Wilson KL (2003) Barrier-to-autointegration factor BAF binds p55 Gag and matrix and is a host component of human immunodeficiency virus type 1 virions. J Virol 77:13084-13092

Mariconda S, Namgoong S-, Yoon K-, Jiang H, Harshey RM (2000) Domain III function of Mu transposase analysed by directed placement of subunits within the transpososome. J Bioscience 25:347-360

Maxwell A, Craigie R, Mizuuchi K (1987) B protein of bacteriophage Mu is an ATPase that preferentially stimulates intermolecular DNA strand transfer. Proc Natl Acad Sci U S A 84:699-703

May EW, Craig NL (1996) Switching from cut-and-paste to replicative Tn*7* transposition. Science 272:401-404

McClintock B (1987) The discovery and characterization of transposable elements: The collected papers of Barbara McClintock. Garland, New York

Miskey C, Izsvák Z, Kawakami K, Ivics Z (2005) DNA transposons in vertebrate functional genomics. Cell Mol Life Sci 62:629-641

Mitra R, Fain-Thornton J, Craig NL (2008) *piggyBac* can bypass DNA synthesis during cut and paste transposition. EMBO J 27:1097-1109

Miyazaki K, Arnold FH (1999) Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. J Mol Evol 49:716-720

Mizuno N, Dramicanin M, Mizuuchi M, Adam J, Wang Y, Han YW, Yang W, Steven AC, Mizuuchi K, Ramon-Maiques S (2013) MuB is an AAA+ ATPase that forms helical filaments to control target selection for DNA transposition. Proc Natl Acad Sci U S A 110:E2441-50.

Mizuuchi K (1992) Transpositional recombination: mechanistic insights from studies of Mu and other elements. Annu Rev Biochem 61:1011-1051

Mizuuchi K (1983) *In vitro* transposition of bacteriophage Mu: a biochemical approach to a novel replication reaction. Cell 35:785-794

Mizuuchi M, Mizuuchi K (2001) Conformational isomerization in phage Mu transpososome assembly: effects of the transpositional enhancer and of MuB. EMBO J 20:6927-6935

Mizuuchi M, Baker TA, Mizuuchi K (1995) Assembly of phage Mu transpososomes: cooperative transitions assisted by protein and DNA scaffolds. Cell 83:375-385

Mizuuchi M, Mizuuchi K (1993) Target site selection in transposition of phage Mu. Cold Spring Harbor Sym 58:515-523

Mizuuchi M, Baker TA, Mizuuchi K (1992) Assembly of the active form of the transposase-Mu DNA complex: a critical control point in Mu transposition. Cell 70:303-311

Mizuuchi M, Baker TA, Mizuuchi K (1991) DNase protection analysis of the stable synaptic complexes involved in Mu transposition. Proc Natl Acad Sci U S A 88:9031-9035

Mizuuchi M, Mizuuchi K (1989) Efficient Mu transposition requires interaction of transposase with a DNA sequence at the Mu operator: implications for regulation. Cell 58:399-408

Montano SP, Pigli YZ, Rice PA (2012) The mu transpososome structure sheds light on DDE recombinase evolution. Nature 491:413-417

Montano SP, Rice PA (2011) Moving DNA around: DNA transposition and retroviral integration. Curr Opin Struct Biol 21:370-378

Morgan GJ, Hatfull GF, Casjens S, Hendrix RW (2002) Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in *Haemophilus*, *Neisseria* and *Deinococcus*. J Mol Biol 317:337-359

Naigamwalla DZ, Coros CJ, Wu Z, Chaconas G (1998) Mutations in domain IIIa of the Mu transposase: evidence suggesting an active site component which interacts with the Mu-host junction. J Mol Biol 282:265-274

Nakayama C, Teplow DB, Harshey RM (1987) Structural domains in phage Mu transposase: identification of the site-specific DNA-binding domain. Proc Natl Acad Sci U S A 84:1809-1813

Namgoong SY, Harshey RM (1998) The same two monomers within a MuA tetramer provide the DDE domains for the strand cleavage and strand transfer steps of transposition. EMBO J 17:3775-3785

Namgoong S-, Kim K, Saxena P, Yang JY, Jayaram M, Giedroc DP, Harshey RM (1998) Mutational analysis of domain IIb of bacteriophage Mu transposase: domains IIa and IIb belong to different catalytic complementation groups. J Mol Biol 275:221-232

Ni J, Clark KJ, Fahrenkrug SC, Ekker SC (2008) Transposon tools hopping in vertebrates. Brief Funct Genomic Proteomic 7:444-453

Nicolas E, Lambin M, Dandoy D, Galloy C, Nguyen N, Oger CA, Hallet B (2015) The Tn3-family of Replicative Transposons. Microbiol Spectr 3:10.1128/microbiolspec.MDNA3-0060-2014.

Okada N, Hamada M, Ogiwara I, Ohshima K (1997) SINEs and LINEs share common 3′ sequences: a review. Gene 205:229-243

Orsini L, Pajunen M, Hanski I, Savilahti H (2007) SNP discovery by mismatch-targeting of Mu transposition. Nucleic Acids Res 35:e44

Paatero AO, Turakainen H, Happonen LJ, Olsson C, Palomaki T, Pajunen MI, Meng X, Otonkoski T, Tuuri T, Berry C, Malani N, Frilander MJ, Bushman FD, Savilahti H (2008) Bacteriophage Mu integration in yeast and mammalian genomes. Nucleic Acids Res 36:e148

Pajunen M, Poussu E, Turakainen H, Savilahti H (2009) Application of Mu in vitro transposition for high-precision mapping of protein-protein interfaces on a yeast two-hybrid platform. Methods 49:255-62

Pajunen MI, Rasila TS, Happonen LJ, Lamberg A, Haapa-Paananen S, Kiljunen S, Savilahti H (2010) Universal platform for quantitative analysis of DNA transposition. Mob DNA 1:24

Pajunen MI, Pulliainen AT, Finne J, Savilahti H (2005) Generation of transposon insertion mutant libraries for Gram-positive bacteria by electroporation of phage Mu DNA transposition complexes. Microbiology 151:1209-1218

Peters JE (2014) Tn7. Microbiol Spectr 2:10.1128/microbiolspec.MDNA3-0010-2014.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem

Piegu B, Bire S, Arensburger P, Bigot Y (2015) 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 86:90-109

Plasterk RHA, van Luenen HGAM (2002) The Tc1/mariner family of transposable elements. In: Craig NL, Craigie R, Gellert M, Lambowitz A (eds) Mobile DNA II. ASM Press, Washington, DC, pp 519-532

Polard P, Ton-Hoang B, Haren L, Betermier M, Walczak R, Chandler M (1996) IS911-mediated transpositional recombination in vitro. J Mol Biol 264:68-81

Poulter RT, Goodwin TJ (2005) DIRS-1 and the other tyrosine recombinase retrotransposons. Cytogenet Genome Res 110:575-588

Poussu E, Jäntti J, Savilahti H (2005) A gene truncation strategy generating N- and C-terminal deletion variants of proteins for functional studies: mapping of the Sec1p binding domain in yeast Mso1p by a Mu *in vitro* transposition-based approach. Nucleic Acids Res 33:e104

Poussu E, Vihinen M, Paulin L, Savilahti H (2004) Probing the a-complementing domain of *E. coli* b-galactosidase with use of an insertional pentapeptide mutagenesis strategy based on Mu *in vitro* DNA transposition. Proteins 54:681-692

Ranquet C, Toussaint A, de Jong H, Maenhaut-Michel G, Geiselmann J (2005) Control of bacteriophage mu lysogenic repression. J Mol Biol 353:186-195

Rasila TS, Pajunen MI, Savilahti H (2009) Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, Escherichia coli mutator strain, and hydroxylamine treatment. Anal Biochem 388:71-80

Rebollo R, Romanish MT, Mager DL (2012) Transposable elements: an abundant and natural source of regulatory sequences for host genes. Annu Rev Genet 46:21-42

Rice P, Mizuuchi K (1995) Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. Cell 82:209-220

Richardson SR, Doucet AJ, Kopera HC, Moldovan JB, Garcia-Perez JL, Moran JV (2015) The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes. Microbiol Spectr 3:MDNA3-0061-2014

Robert VJ, Davis MW, Jorgensen EM, Bessereau JL (2008) Gene conversion and end-joining-repair double-strand breaks in the Caenorhabditis elegans germline. Genetics 180:673-679

Sarnovsky RJ, May EW, Craig NL (1996) The Tn7 transposase is a heteromeric complex in which DNA breakage and joining activities are distributed between different gene products. EMBO J 15:6348-6361

Savilahti H, Mizuuchi K (1996) Mu transpositional recombination: donor DNA cleavage and strand transfer in *trans* by the Mu transposase. Cell 85:271-280

Savilahti H, Rice PA, Mizuuchi K (1995) The phage Mu transpososome core: DNA requirements for assembly and function. EMBO J 14:4893-4903

Savilahti H, Bamford DH (1993) Protein-primed DNA replication: role of inverted terminal repeats in the Escherichia coli bacteriophage PRD1 life cycle. J Virol 67:4696-703

Schatz DG, Swanson PC (2011) V(D)J recombination: mechanisms of initiation. Annu Rev Genet 45:167-202

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al. (2009) The B73 maize genome: Complexity, diversity, and dynamics. Science 326:1112-5

Schumacher S, Clubb RT, Cai M, Mizuuchi K, Clore GM, Gronenborn AM (1997) Solution structure of the Mu end DNA-binding Ib subdomain of phage Mu transposase: modular DNA recognition by two tethered domains. EMBO J 16:7532-7541

Schweidenback CT, Baker TA (2008) Dissecting the roles of MuB in Mu transposition: ATP regulation of DNA binding is not essential for target delivery. Proc Natl Acad Sci U S A 105:12101-12107

Seberg O, Petersen G (2009) A unified classification system for eukaryotic transposable elements should reflect their phylogeny. Nat Rev Genet 10:276-c3

Sekine Y, Eisaki N, Ohtsubo E (1994) Translational control in production of transposase and in transposition of insertion sequence IS3. J Mol Biol 235:1406-1420

Shapiro JA (1979) Molecular model for the transposition and replication of bacteriophage Mu and other transposable elements. Proc Natl Acad Sci U S A 76:1933-1937

Shockett PE, Schatz DG (1999) DNA hairpin opening mediated by the RAG1 and RAG2 proteins. Mol Cell Biol 19:4159-4166

Siguier P, Gourbeyre E, Varani A, Ton-Hoang B, Chandler M (2015) Everyman's Guide to Bacterial Insertion Sequences. Microbiol Spectr 3:MDNA3-0030-2014

Siguier P, Gourbeyre E, Chandler M (2014) Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiol Rev 38:865-891

Simm AM, Baldwin AJ, Busse K, Jones DD (2007) Investigating protein structural plasticity by surveying the consequence of an amino acid deletion from TEM-1 b-lactamase. FEBS Lett 581:3904-3908

Staveley BE, Heslip TR, Hodgetts RB, Bell JB (1995) Protected P-element termini suggest a role for inverted-repeat-binding protein in transposase-induced gap repair in Drosophila melanogaster. Genetics 139:1321-1329

Surette MG, Harkness T, Chaconas G (1991) Stimulation of the Mu A protein-mediated strand cleavage reaction by the Mu B protein, and the requirement of DNA nicking for stable type 1 transpososome formation. *In vitro* transposition characteristics of mini-Mu plasmids carrying terminal base pair mutations. J Biol Chem 266:3118-3124

Surette MG, Lavoie BD, Chaconas G (1989) Action at a distance in Mu DNA transposition: an enhancer-like element is the site of action of supercoiling relief activity by integration host factor (IHF). EMBO J 8:3483-3489

Surette MG, Buch SJ, Chaconas G (1987) Transpososomes: stable protein-DNA complexes involved in the *in vitro* transposition of bacteriophage Mu DNA. Cell 49:253-262

Taylor AL (1963) Bacteriophage-induced mutation in *Escherichia coli*. Proc Natl Acad Sci U S A 50:1043-1051

Ton-Hoang B, Bétermier M, Polard P, Chandler M (1997) Assembly of a strong promoter following IS*911* circularization and the role of circles in transposition. EMBO J 16:3357-3371

Tosi LR, Beverley SM (2000) cis and trans factors affecting Mos1 mariner evolution and transposition in vitro, and its potential for functional genomics. Nucleic Acids Res 28:784-790

Toussaint A, Gama MJ, Laachouch J, Maenhaut-Michel G, Mhammedi-Alaoui A (1994) Regulation of bacteriophage Mu transposition. Genetica 93:27-39

Tu Quoc PH, Genevaux P, Pajunen M, Savilahti H, Georgopoulos C, Schrenzel J, Kelley WL (2007) Isolation and characterization of biofilm formation-defective mutants of Staphylococcus aureus. Infect Immun 75:1079-1088

Turakainen H, Saarimaki-Vire J, Sinjushina N, Partanen J, Savilahti H (2009) Transposition-based method for the rapid generation of gene-targeting vectors to produce Cre/Flp-modifiable conditional knock-out mice. PLoS One 4:e4341

Turlan C, Chandler M (2000) Playing second fiddle: second-strand processing and liberation of transposable elements from donor DNA. Trends Microbiol 8:268-274

VandenDriessche T, Ivics Z, Izsvak Z, Chuah MK (2009) Emerging potential of transposons for gene therapy and generation of induced pluripotent stem cells. Blood 114:1461-8

Vilen H, Aalto J-, Kassinen A, Paulin L, Savilahti H (2003) A direct transposon insertion tool for modification and functional analysis of viral genomes. J Virol 77:123-134

Vink C, Yeheskiely E, van der Marel GA, van Boom JH, Plasterk RH (1991) Site-specific hydrolysis and alcoholysis of human immunodeficiency virus DNA termini mediated by the viral integrase protein. Nucleic Acids Res 19:6691-6698

Vogel JL, Li ZJ, Howe MM, Toussaint A, Higgins NP (1991) Temperature-sensitive mutations in the bacteriophage Mu c repressor locate a 63-amino-acid DNA-binding domain. J Bacteriol 173:6568-6577

Vos JC, De Baere I, Plasterk RH (1996) Transposase is the only nematode protein required for in vitro transposition of Tc1. Genes Dev 10:755-761

Wang Y, Wang J, Devaraj A, Singh M, Jimenez Orgaz A, Chen JX, Selbach M, Ivics Z, Izsvak Z (2014) Suicidal Autointegration of Sleeping Beauty and piggyBac Transposons in Eukaryotic Cells. PLoS Genet 10:e1004103

Wang Y, Stojiljkovic N, Jehle JA (2010) Cloning of complete genomes of large dsDNA viruses by in vitro transposition of an F factor containing transposon. J Virol Methods 167:95-99

Wang Z, Namgoong S-, Zhang X, Harshey RM (1996) Kinetic and structural probing of the precleavage synaptic complex (type 0) formed during phage Mu transposition. Action of metal ions and reagents specific to single-stranded DNA. J Biol Chem 271:9619-9626

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520-562

Watson MA, Chaconas G (1996) Three-site synapsis during Mu DNA transposition: a critical intermediate preceding engagement of the active site. Cell 85:435-445

Way JC, Kleckner N (1985) Transposition of plasmid-borne Tn10 elements does not exhibit simple length-dependence. Genetics 111:705-13

Wei XX, Shi ZY, Li ZJ, Cai L, Wu Q, Chen GQ (2010) A mini-Mu transposon-based method for multiple DNA fragment integration into bacterial genomes. Appl Environ Microbiol 87:1533-41

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8:973-82

Williams TL, Jackson EL, Carritte A, Baker TA (1999) Organization and dynamics of the Mu transpososome: recombination by communication between two active sites. Gene Dev 13:2725-2737

Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hamalainen R, Cowling R, Wang W, Liu P, Gertsenstein M, Kaji K, Sung HK, Nagy A (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 458:766-770

Wong TS, Roccatano D, Schwaneberg U (2007) Steering directed protein evolution: strategies to manage combinatorial complexity of mutant libraries. Environ Microbiol 9:2645-2659

Wozniak RA, Waldor MK (2010) Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat Rev Microbiol 8:552-563

Wu Z, Xuanyuan Z, Li R, Jiang D, Li C, Xu H, Bai Y, Zhang X, Turakainen H, Saris PE, Savilahti H, Qiao M (2009) Mu transposition complex mutagenesis in *Lactococcus lactis*--identification of genes affecting nisin production. J Appl Microbiol 106:41-8

Wu Z, Chaconas G (1995) A novel DNA binding and nuclease activity in domain III of Mu transposase: evidence for a catalytic region involved in donor cleavage. EMBO J 14:3835-3843

Wu Z, Chaconas G (1994) Characterization of a region in phage Mu transposase that is involved in interaction with the Mu B protein. J Biol Chem 269:28829-28833

Yanagihara K, Mizuuchi K (2002) Mismatch-targeted transposition of Mu: a new strategy to map genetic polymorphism. Proc Natl Acad Sci U S A 99:11317-11321

Yang J-, Jayaram M, Harshey RM (1995) Enhancer-independent variants of phage Mu transposase: enhancer-specific stimulation of catalytic activity by a partner transposase. Gene Dev 9:2545-2455

Yuan JF, Beniac DR, Chaconas G, Ottensmeyer FP (2005) 3D reconstruction of the Mu transposase and the Type 1 transpososome: a structural framework for Mu DNA transposition. Gene Dev 19:840-852

Zhang C, Kitsberg D, Chy H, Zhou Q, Morrison JR (2005) Transposon-mediated generation of targeting vectors for the production of gene knockouts. Nucleic Acids Res 33:e24

Zheng R, Ghirlando R, Lee MS, Mizuuchi K, Krause M, Craigie R (2000) Barrier-to-autointegration factor (BAF) bridges DNA in a discrete, higher-order nucleoprotein complex. Proc Natl Acad Sci U S A 97:8997-9002

Zhou L, Mitra R, Atkinson PW, Burgess Hickman A, Dyda F, Craig NL (2004) Transposition of *hAT* elements links transposable elements and V(D)J recombination. Nature 432:995-1001

Zou AH, Leung PC, Harshey RM (1991) Transposase contacts with Mu DNA ends. J Biol Chem 266:20476-20482