8
Insect Transgenesis: Mechanisms, Applications, and Ecological Safety

ALFRED M. HANDLER[1]* AND PETER W. ATKINSON[2]

[1]*Center for Medical, Agricultural, and Veterinary Entomology, Agricultural Research Service, US Department of Agriculture, 1700 SW 23rd Drive, Gainesville, FL 32608, USA and* [2]*Department of Entomology and Institute for Integrative Genome Biology, University of California, Riverside, CA 92521, USA*

Introduction

Genetic transformation has been a fundamental methodology for the analysis of gene structure–function relationships in all model systems. Transformation of pneumococcus bacteria provided the first experimental evidence for DNA being the inherited genetic material and since then has been the primary method for defining gene function (Mello and Fire, 1995). Thus, the recent explosion of available genomic sequence information, which will only increase in volume for many years, makes the need for routine transformation in species of interest critical to the meaningful understanding of this sequence data. A prime example for the various analyses possible with transformation methodology has been the application of transposon-mediated germ-line transformation in *Drosophila melanogaster*. Gene identification has been approached most straightforwardly by testing the ability of a putative recombinant wild-type allele to phenotypically rescue a mutated null allele after transformation. More detailed gene structure–function relationships have been approached by assessing the phenotypic effect of systematic sequence modifications (generally nucleotide substitution or deletion) of the recombinant allele. Gene identification has also been approached by methods generally described as 'insertional mutagenesis'. These include transposon tagging and a variety of 'trap' systems, including those for enhancers, exons, and introns (Brand and Perrimon,

*To whom correspondence may be addressed (handler@nersp.nerdc.ufl.edu)

Abbreviations: CFP, cyan fluorescent protein; DTS, dominant temperature sensitive; EMSA, electrophoretic mobility shift assays; GFP, green fluorescent protein; ITR, inverted terminal repeat sequence; L, left-hand or 5'; R, right-hand or 3'; RISC, RNA-induced silencing complex; RMCE, recombinase-mediated cassette exchange; RNAi, RNA interference; rtTA, reverse *tet* transcriptional activator; SIT, sterile insect technique; TRE, *tet* response element; tTA, *tet* transcriptional activator; YFP, yellow fluorescent protein.

Biotechnology and Genetic Engineering Reviews – Vol. 23, December 2006
0264–8725/07/23/129–156 $20.00 + $0.00 © Lavoisier/Intercept, 14 rue de Provigny, F-94236 Cachan cedex, France

1993; Brand *et al.*, 1994; Lukacsovich *et al.*, 2001; Morin *et al.*, 2001). These systems generally identify coding or regulatory sequence function by genomic vector insertions resulting in a mutation phenotype or reporter gene expression, with subsequent isolation of relevant sequences by probing for the vector DNA.

The recent extension of transposon-mediated germ-line transformation to nearly 20 non-drosophilid insect species now allows the use of these methods in non-model systems, facilitating an understanding of genetic mechanisms in diverse species that would have been intractable only a few years ago (Handler, 2001). Significantly, germ-line transformation can also be used for applied purposes by genomic integration of genetic constructs, which can alter the development, behaviour, or population size of many species that negatively or positively impact agriculture or human health (Alphey *et al.*, 2002; Handler, 2002; Atkinson *et al.*, 2004). While the types of potential transgenic strains for applied use are numerous, those used in field release programmes engender special concerns related to ecological safety (Hoy, 2003). Insects and fish, unlike most other transgenic organisms, have the ability to disperse easily and so cannot be contained or retrieved once released from cages, pens, or ponds. Thus, it is critical that we understand the mechanisms that underlie the mobility properties of the transposons used as gene vectors for most transgenic insects, and evaluate new vector systems that can ensure transgenic strain integrity and transgene stability.

Vectors for insect transgenesis

Presently, four transposon vector systems are available for use in a wide variety of insects that include *Hermes* from the house fly, *Musca domestica* (Warren *et al.*, 1994), *mariner* from *Drosophila mauritiana* (Medhora *et al.*, 1988), *Minos* from *D. hydei* (Franz and Savakis, 1991), and *piggyBac* from the cabbage looper moth, *Trichoplusia ni* (Fraser *et al.*, 1983). All of these transposons are class II elements that transpose via a DNA-mediated process where an internally encoded transposase enzyme acts upon inverted terminal repeat sequences (ITRs) and adjacent sequences to effect excision of the element and subsequent insertion into another chromosomal site. This occurs as a 'cut and paste' process known as transposition (Finnegan, 1989). All of these transposable elements have been modified in a similar fashion to create a binary transformation system that consists of non-autonomous vector and helper transposase plasmids (Rubin and Spradling, 1982). The vector includes the inverted terminal repeat sequences and subterminal sequences needed for mobility that surround a selectable marker gene and other sequences of interest. The transposase gene within the vector is either deleted or made defective, and vector transposition is dependent upon a separate helper plasmid that contains the transposase gene, but not the terminal sequences necessary for integration. When transiently expressed in the germ-line, the helper transposase acts in *trans* to catalyse integration of the vector, but is lost in subsequent cell divisions, allowing the integrated vector to remain stable.

Implicit in the applied use of transposable elements in transgenic insect strains released into the field is a comprehensive understanding of how they actually move between DNA loci. This knowledge is important if the fate of these elements in insects is to be understood properly and realistic estimates of risk arising from their release are to be calculated. Two approaches to predicting transposable element

behaviour in transgenic insects can be taken. One involves using other genetic systems to inactivate, usually through removal, those sequences of the transposable element required for transposition. Recent advances in this strategy are described in the following sections. This is, in turn, somewhat, but not necessarily, dependent on the identification of these sequences, although a fairly generic approach can be undertaken since, for example, the ITRs of transposable elements are known to be essential for transposition, and can be identified through simple DNA sequence analysis. The second approach involves determining for each transposable element how, in fact, transposition occurs. Strategies for control arising from this will, most likely, be specific for each element.

As stated above, insect transformation historically has been achieved through the use of class II transposable elements, although a serine integrase and two tyrosine recombinases have been used successfully with reasonable efficiency in at least *D. melanogaster* (Siegal and Hartl, 1996; Golic *et al.*, 1997; Groth *et al.*, 2004). Integrases, recombinases, and transposases represent different families of enzymes that cut and join DNA. Tyrosine recombinases, such as FLP from *Saccharomyces cerevisiae*, Cre from the P1 bacteriophage, and the serine integrase from the ϕC31 bacteriophage, all function by bringing together the four strands of DNA involved in any single breaking and joining between two DNA ends. DNA is not broken down during the reaction. The strand breakage and rejoining is catalysed by the recombinase which, when bound to the cleaved DNA end, stores the energy resulting from DNA cleavage. This energy is then released during subsequent strand joining to produce the final recombinant product. Consequently, no high-energy co-factors such as ATP or GTP are required for catalysis by these enzymes. Serine recombinases such as the ϕC31 integrase process the exchange between cleaved DNA ends one pair of ends at a time, while tyrosine recombinases process the simultaneous exchange of all four double strand DNA breaks.

The serine integrases and tyrosine recombinases recognize specific sequences in both the donor and the target, and so represent a binary system. However, FLP and Cre both recognize 34 bp *FRT* and *loxP* recombination sites, respectively, which include duplicated 13 bp inverted terminal repeat binding sites and an intervening 8 bp spacer region. The acceptor and donor sites are identical, with the 8 bp spacer sequence providing pairing specificity and directionality that determine the type of resulting recombination product (e.g. inversion or excision from *cis*-oriented sites). Thus, recombination sites with mutated spacers can provide heterospecificity, allowing integrations by double recombination that have been used for targeted cassette exchange, as discussed later. The ϕC31 integrase mediates recombination between a simple *attB* and complex *attP* attachment site that, except for a short identical sequence where recombination occurs, are non-identical (see Groth and Calos, 2004). For both systems, no additional host factors are required, meaning, in principle, that these systems should function in any host organism provided that an enzymatically active integrase or recombinase that accurately recognizes these sequences can be synthesized. Provided a functional enzyme is produced, the rate-limiting step for these systems in insect species may be the absence of suitable target sites, which can be supplied by transformation with a transposon vector. However, this requirement makes the system dependent on the efficiency of transposable element integration into the host genome.

Class II transposons and their mechanisms of transposition

Transposases act to transfer transposable element DNA to a new DNA target site by two sequential reactions. The first is essentially conserved amongst all transposases and creates a single strand cleavage between the 3'-OH of the transposable element and the 5'-phosphoryl ends of the flanking DNA. The second step varies between types of transposable elements and transfers, through a nucleophilic attack, this 3'-OH to the 5'-phosphoryl end of the target DNA molecule. For many transposases, the supply of energy from a cofactor, as is required for DNA ligation, is not needed, indicating that this strand transfer occurs via transesterification. The fate of the transposable element and its flanking DNA depends on whether the initial cleavage involves a single or double strand. If it is the former, then transposition leads to the formation of cointegrants in which two copies of the transposable element flanking donor sequences is created at the target site. If double strand breaks occur before strand transfer, then only the transposable element is transferred to the target DNA. Cleavage of the second strand of the transposable element varies between types of element. For the bacterial element Tn*7*, second strand cleavage is performed by the TnsA protein, which is encoded by the Tn*7* element, but is not the Tn*7* transposase (Craig, 2002). For many other transposable elements, second strand cleavage occurs through a nucleophilic attack of the 3'-OH generated from the first strand cleavage on the phosphodiester bond immediately opposite it on the second DNA strand. Bacterial transposable elements such as Tn*5* and Tn*10* transpose in this manner (Hanniford, 2002; Reznikoff, 2002). If the nucleophilic attack by the 3'-OH generated by the initial cleavage occurs at the 5'-phosphoryl end of the same strand, but at the other end of the transposable element, then a circular intermediate is formed. Transposable elements such as IS*911* transpose in this manner. Indeed, the formation of a circular IS*911* element forms a strong promoter and so increases transposase gene transcription (Duval-Valentin *et al.*, 2001).

Many of the aspects of the mechanism of transposition have been worked out for several bacterial transposable elements, such as Tn*5*, Tn*7*, Tn*10*, and IS*911* (see Craig *et al.*, 2002 for reviews of each). The structure of the co-crystal of the Tn*5* transposase and the Tn*5* ITR and adjacent sequences enabled the positioning of the ends of the Tn*5* transposable element within the catalytic core of the transposase to be determined, and also illustrated how the transposase functioned as a dimer (Davies *et al.*, 2000). While several eukaryotic transposable elements are well characterized at the genetic and biochemical level, little or no information concerning the detailed structure of their transposases and how this relates to their function is available. As for many transposases, progress has been hindered by the apparent cellular toxicity of these enzymes when they are produced in even moderate quantities in either prokaryotic or eukaryotic expression systems, with a proclivity of these to form inclusion bodies. Recently, however, some progress has been made in the analysis of transposable elements used in insect transformation. Here, we focus on the *Hermes* element of *M. domestica* and the *Mos1* element of *D. mauritiana*. We do so because recent data have shed light on their mechanism of transposition and because, in the case of *Hermes*, the crystal structure of the transposase has been determined recently.

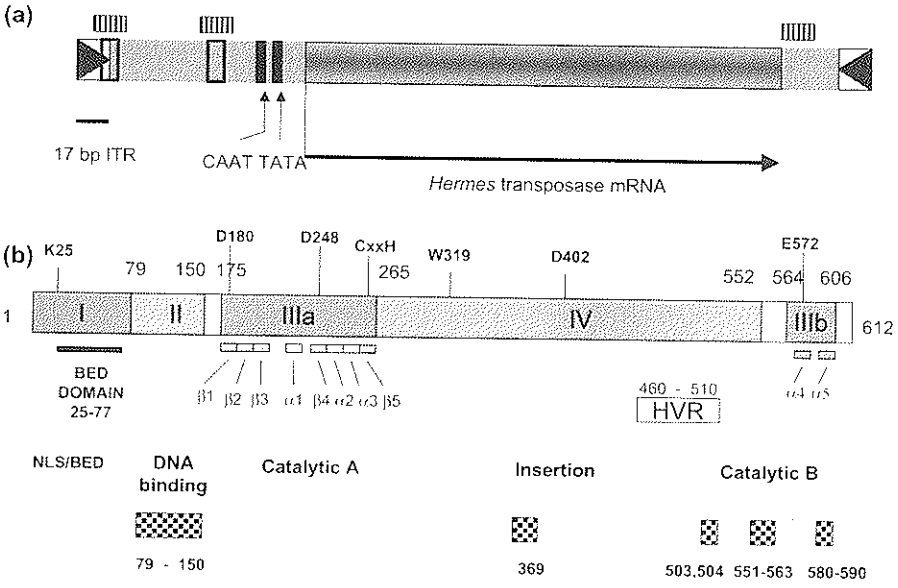

Figure 8.1. Schematic of the *Hermes* transposable element and its transposase. (a) Structure of the *Hermes* element. The transposon is 2749 bp in length with 17 bp inverted terminal repeats (ITRs) (Warren *et al.*, 1994). It encodes a transposase that comprises some 2 kb of the element. Positions of the TATA and CAAT boxes that are components of the transposase promoter are shown. The striped boxes indicate binding sites of the transposase and open squares at the left end denote a sequence that is conserved at these two locations (T.A. Laver, R.H. Hice and P.W. Atkinson, unpublished data). The shaded region between them denotes a region that increases transposase activity (T.A. Laver and P.W. Atkinson, unpublished data). (b) Structure of the *Hermes* transposase. The four domains based on the secondary structure of the full-length transposase and the crystal structure of the truncated transposase are shown (Zhou *et al.*, 2004). Amino acid coordinates indicate the approximate boundaries of these domains. Positions of the carboxylate triad of D180, D248, and E572 are shown. W319 denotes the location of a region proposed to stabilize flanking DNA adjacent to the active site. K25 denotes a residue important for nuclear localization, while the D402 residue produces an inactive transposase when mutated to N (Michel *et al.*, 2002; Michel and Atkinson, 2003). The location of the BED domain is shown, as are the locations of the five alpha helices and five beta sheets that constitute the catalytic domain of the transposase (Zhou *et al.*, 2004). The HVR block denotes a 50 amino acid region that is the most hypervariable between *hAT* transposases. The hatched blocks denote regions that are known or inferred to be important for oligomerization of the transposase to form the active hexamer (Michel *et al.*, 2003; Zhou *et al.*, 2004). Numbers refer to amino acid positions.

THE *hAT* SUPERFAMILY ELEMENT *HERMES*

The *hAT* element superfamily consists of representatives from the fungal, plant, and animal kingdoms. *hAT* elements are the most abundant class II elements found in the human genome, with this class of element comprising ~3% of the entire genome (Lander *et al.*, 2001). Plant *hAT* elements, such as *Ac* and *Tam3*, are mobile in plant species other than the original host and, similarly, the insect-based members of this superfamily, such as *Hermes*, are also mobile in insect species other than the original host (O'Brochta *et al.*, 1996; Jasinskiene *et al.*, 1998; Michel *et al.*, 2001). It is this feature of the insect *hAT* elements that has made them attractive to those who seek to introduce genes into pest insects, such as mosquitoes.

The *Hermes* element is the best-characterized insect *hAT* element (*Figure 8.1*). Its crystal structure has been determined, and biochemical and genetic data have

illustrated its principal mode of transposition (Zhou *et al.*, 2004; Hickman *et al.*, 2005). The *Hermes* element is 2749 bp in length, contains 17 bp imperfect inverted terminal repeats, and encodes a transposase that is 612 amino acids (70.1 kDa) long. Electrophoretic mobility shift assays (EMSA) reveal that the transposase binds asymmetrically to the *Hermes* ends, but at a small distance internal to the termini (T. Laver and P.W. Atkinson, unpublished data). This asymmetry is reflected in transposition assays performed in *D. melanogaster* S2 cells *in vitro*, which show a clear predisposition for deletions of the element to occur predominantly at the right end (R. Hice and P.W. Atkinson, unpublished data). The structure of a truncated but active form of the *Hermes* transposase lacking the first 78 amino acids of the protein has been determined to 2.1 Å resolution (Hickman *et al.*, 2005; Perez *et al.*, 2005). This, together with secondary structure alignments and functional studies, reveals that the full length *Hermes* transposase consists of four domains: an amino end domain that contains the nuclear localization signal and most likely a non-specific DNA binding domain (absent in the amino truncated form used to form the crystal); a DNA binding domain that binds specifically to the first 30 bp of the *Hermes* element and is also involved in dimerization of the protein; a catalytic domain; and an insertion domain (Michel and Atkinson, 2003; Zhou *et al.*, 2004; Hickman *et al.*, 2005). The insertion domain in turn contains regions involved in dimerization and also contains two α-helices, which, together with the β-sheets in the catalytic domain, form the active site of the enzyme. *Hermes* is a member of the retroviral integrase superfamily since its transposase contains an acidic catalytic triad composed of D180, D248, and E572 that form the catalytic site for coordinating two Mg^{2+} ions. These three carboxylates coordinate divalent metal cations, usually Mg^{2+}, which direct water molecules to the phosophodiester bond, which is broken and then reformed in the transesterification reaction. In the *Hermes* transposase, there is a large linear distance between the second and third members of the triad, and secondary structure alignments, together with the crystal structure, show that E572 is located on an α-helix within the insertion domain, but it is in close physical proximity to the first two aspartates of the triad located on the β-sheets, which constitute the RNase H fold of the active site (Hickman *et al.*, 2005). This finding reinforces one of the outcomes from the analysis of the crystal structure of the Tn*5* synaptic complex; that these transposases cannot be divided into simple linear domains (Davies *et al.*, 2000).

Hermes functions as a hexameric protein or, more correctly, as a trimer of dimers (Hickman *et al.*, 2005). Each monomer interacts with an adjacent monomer at three interfaces. One of these is located within the DNA binding domain in which adjacent transposase monomers interact through the three α-helices within this domain. The remaining two interfaces occupy regions of the insertion domain, with one region of one monomer interacting with a different region, still in the insertion domain, of the adjacent monomer. The total number of regions involved in oligomerization is thus five: one occupying the second domain and four occupying the fourth domain. The hexameric structure of *Hermes* transposase leads to questions concerning which of the six available catalytic sites are used for first strand cleavage, hairpin formation, second strand cleavage, and strand transfer, and whether conformation change in the protein is required for complete transpositional activity. The generation of hairpins on flanking DNA rather than on the transposon requires that, for second strand

transfer involving a nucleophilic attack of the 3'-OH of the transposon on the target site, the active site moves from the 3'-OH of the flanking DNA on the opposite strand, where first strand cleavage occurs. How this is achieved is not known; however, it has been suggested recently that this step may have been introduced to reduce the transposition frequency of this family of elements in eukaryotes (Nowotny *et al.*, 2005). *hAT* elements such as *Hermes* are structurally and functionally related to the *Rag1* recombinase that reorganizes the immunoglobulin and T cell receptor genes in vertebrates, a recombination system that is tightly regulated in the germinal cells that produce B and T lymphocytes (Zhou *et al.*, 2004). Furthermore, in both insects and humans, *hAT* transposable elements in their natural, non-engineered state are inactive since even moderate levels of transposition can result in them becoming intragenomic mutagens.

Knowledge of *Hermes* element function, together with the proposed structure of the truncated transposase, increases our ability to generate mutations of both the element and its transposase that are hyperactive, resulting in increased rates of transposition. Placing hyperactive transposase mutants under the control of a promoter that confines expression to the insect germ-line may promote driving of *Hermes* transgene vectors through an insect genome, which make the host insect incapable of vectoring disease, as discussed later. Three approaches to generating hyperactive *Hermes* element mutants can be envisaged. The first involves random mutagenesis of the transposase, followed by the selection of hyperactive mutants in expression systems that permit their identification. A yeast-based expression system that uses restoration of prototrophy on defined media enables these mutants to be identified in the *Hermes* transposase (P. Atkinson and N. Craig, unpublished data). These can then be tested for retention of hyperactivity in insect cell culture or in whole insects through the microinjection of insect embryos using established techniques. A second approach is to use site-directed mutagenesis based on the structure of the transposase and identification of the amino acids most likely to interact directly with the *Hermes* element. While accurate identification of these awaits co-crystallization of the transposase with either the donor substrate or with the target DNA, changes to amino acids in the catalytic site and changes to the DNA binding domain which may be predicted to increase or decrease binding affinity can be made, and their effects measured directly either in yeast or in insects. The final strategy targets not the *Hermes* element directly, but examines how the cell regulates transposable element activity. Transposons that are mobile can be mutagenic, and the final rate of activity is a balance between the host needing to minimize or eliminate the source of this genetic lethal load and the element needing to retain the ability to transpose in order to propagate. Cellular mechanisms identified as being important for the regulation of transposable elements act through RNA inhibition, DNA methylation, or the acetylation of histones to promote heterochromatin formation. Modification of any of these processes to permit increased *Hermes* mobility is likely to have effects beyond the *Hermes* element, and so may result in a phenotype that is undesirable and ineffective for any genetic control strategy. Knowledge of how each of these mechanisms interacts with transposable elements, especially those introduced into genomes as part of a genetic control programme, is important, however, for the efficacy of these programmes.

THE *MARINER* FAMILY ELEMENTS, *Mos1* AND *Himar1*

Members of the *mariner* family of transposable elements are widespread through insects and, in several cases, compelling phylogenetic evidence exists that they have been horizontally transferred between species (Robertson and MacLeod, 1993). Indeed, the *Mos1* element is active when introduced into *Leishmania* and chickens (Gueiros-Filho and Beverley, 1997; Sherman *et al.*, 1998). Many *mariners* are capable of reaching high copy numbers within genomes, suggesting that subsequent to initial invasion they can rapidly increase their copy number. A corollary to this is that the overwhelming majority of these elements found within genomes are inactive, suggesting that regulatory mechanisms have rapidly evolved to restrict *mariner* mobility. Indeed, to date, only one active form of *mariner* has been found in any insect, this being the *Mos1* element from *D. mauritiana* (Medhora *et al.*, 1991). The *Himar1* element was made synthetically based on the sequences of a number of inactive *mariner* elements from the hornfly, *Haemotoba irritans*, and is also active (Lampe *et al.*, 1996).

The *Mos1 mariner* element is 1286 bp in length, has 28 bp ITRs, and encodes a 41 kDa transposase. It possesses a DDD motif at the centre of its catalytic domains. This is a variant of the DDE catalytic triad within the RNAse H fold found in retroviral integrases, many DNA polymerases, *hAT* element transposases, the *Rag1* recombinase, HIV reverse transcriptase, Tn*5*, and Argonaute (Rice and Baker, 2001; Parker *et al.*, 2004; Song *et al.*, 2004). Argonaute is a protein required for the function of the RNA-induced silencing complex (RISC) critical for the degradation of RNA in RNA inhibition pathways.

The *Mos1* transposase has been crystallized but, as yet, there are no published reports of its structure (Richardson *et al.*, 2004). Purification of this transposase was facilitated by a T216A mutation, which resulted in a biologically active enzyme that was soluble. The transposase contains two domains (Zhang *et al.*, 2001; Richardson *et al.*, 2004). The N-terminal domain includes the first 118–120 amino acids that bind to the inverted terminal repeats of the transposable element. The C-terminal domain, consisting of amino acids 119 and 121–345, contains the catalytic core and so is responsible for the excision and integration of the element. It contains the RNase H fold and the catalytic triad of carboxylates. Unlike bacterial transposable elements, *Mos1* transposase initially cleaves at the 5' phosphate of the terminal nucleotide of the element yielding a 3'-OH on the flanking DNA strand (Dawson and Finnegan, 2003). How this generates second strand cleavage and then strand transfer is unknown; however, there is no hairpin formation at either end of the double strand break in *Mos1* transposition (Dawson and Finnegan, 2003). This mechanism of transposition does share some similarities with the action of *Rag1* recombinase in V(D)J recombination in that initial cleavage occurs at the 5'-P; however, the absence of hairpin formation distinguishes it from the mechanism of the *Rag1* recombinase and the *hAT* transposases.

Evidence has been obtained indicating that the *Mos1* transposase and the *Himar* transposase work as tetramers (Lipkow *et al.*, 2004; Auge-Gouillou *et al.*, 2005a). This contradicts initial suggestions based on related prokaryotic elements that the synaptic complex would contain dimers of the transposase. The ITRs of *Mos1* were found to allow adjacent binding of two *Mos1* transposase molecules, meaning that a

synaptic complex would contain four molecules (Auge-Gouillou *et al.*, 2005b). The process by which this occurs is open to speculation, although a recent model proposes that the *Mos1* transposase contains separate domains that lead to *cis*-dimerization, in which transposases adjacently bound to a single ITR interact, and *trans*-dimerization, in which transposases bound to separate *Mos1* transposons interact (Auge-Gouillou *et al.*, 2005a). A conformational change in the transposase is proposed to mediate the transition from *cis*- to *trans*-dimerization, and so lead to the formation of a synaptic complex consisting of *Mos1* transposase tetramers. EMSA studies indicate that the N-terminal region of the *Mos1* transposase is required for both forms of dimerization (Auge-Gouillou *et al.*, 2005a).

The sequences required for *Mos1* excision and transposition differ between *in vivo* and *in vitro* assays. For *in vivo* assays performed in transgenic *Drosophila*, several hundred base pairs of L (left-hand or 5') and R (right-hand or 3') ends are necessary, with three regions distinct from the 28 bp ITRs being essential for *mariner* mobilization (Lohe and Hartl, 2002). The relative spacing between these regions may also be a factor contributing to the efficiency of mobilization. These data are in stark contrast to the requirements for *Mos1* transposition *in vitro*, in which only the 28 bp ITRs and an additional 38 bp and five bp of internal *Mos1* DNA at the L and R ends, respectively, are required for transposition (Tosi and Beverley, 2000). The 28 bp ITRs contain four mismatches, with the right ITR able to out-compete the left ITR for binding to the transposase (Auge-Gouillou *et al.*, 2001; Zhang *et al.*, 2001). The ends of *Mos1* are therefore asymmetric in their ability to bind transposase, at least when examined biochemically.

The role that the ITRs and internal sequences play in *Mos1* transposition has also been investigated in *Escherichia coli* and in the mosquito, *Aedes aegypti* (Pledger *et al.*, 2004). The normal, imperfect ITRs, together with 3 bp and 7 bp of internal *Mos1* sequences from the L and R ends, respectively, are sufficient for transposition in *E. coli*. Indeed, the transposition frequency is not substantially different from that seen when several hundred bp of internal *Mos1* DNA are also present. The minimal *Mos1* end sequences of 31 bp (L) and 35 bp (R) perhaps further refine the minimal sequences found necessary from the *in vitro* assays. Interestingly, replacing the L end with the R end to create perfect ITRs increases transposition, with a 50-fold increase seen when large amounts of flanking internal sequence are present but, unexpectedly, with an almost 650-fold increase when only 3 bp and 7 bp of internal sequences were present at the L and R ends, respectively (Pledger *et al.*, 2004). Both minimal *Mos1* elements were, however, incapable of transposition in *A. aegypti* in which only *Mos1* elements with several hundred bp of flanking internal sequence could transpose, irrespective of whether the ITRs were perfect or imperfect. The increase in *Mos1* transposition seen in *E. coli* when two right ITRs were used is in agreement with previous observations where transposition of a *Mos1* element containing two R ITRs increased some 1200-fold relative to the control with the normal imperfect ITRs; however, it is not clear from these studies whether any additional *Mos1* element DNA internal to the ITRs were present (Auge-Gouillou *et al.*, 2001). Nevertheless, the general observation is maintained that the right ITR, to which the transposase binds more strongly, provides a better substrate for transposition when it is present at both ends of the transposon. Recently, it has been suggested that the major difference between the L and R ITRs is centred on the 16th base pair of the ITR

located between a palindromic sequence, which is better conserved within the R end than in the L end (Bigot *et al.*, 2005). This may account for the stronger binding of the *Mos1* transposase to the R end, but does not explain the very different requirements for transposition between bacteria and insects. One must also bear in mind the intrinsic differences in experimental design between: (1) *in vitro* experiments in which free ITRs located on oligonucleotides are supplied in large amounts; (2) experiments in which *Mos1* transposable element sequences are located on plasmids which are not affected by protein modifications that occur to chromosomal DNA; and (3) experiments in which *Mos1* elements are incumbent in eukaryotic chromosomes. All present different substrates (in different amounts) to the transposase, which itself may undergo post-translation modifications in insects. Furthermore, epigenetic modifications to the transposon are yet to be fully explored in insects. In this regard, a full systematic functional analysis not only of *Mos1* but of *Hermes* too, is yet to be achieved.

Attempts to generate hyperactive forms of both the *Mos1* and *Himar1* transposases have met with limited success. An *E. coli* papillation assay was used to identify mutants of the *Himar1* transposase that showed increased ability to mobilize a non-autonomous, genetically tagged *Himar1* element from the single copy F plasmid to the chromosome (Lampe *et al.*, 1999). The level of transpositional activity of selected mutants was then measured using a mating-out assay. The transposase was mutated by error-prone PCR to generate a library of transposase genes, each ideally containing a single mutation in the transposase. Three mutations, H276R in one mutant and Q131R and E137K in a second mutant, were found to be hyperactive in this prokaryotic-based assay, with the E137K mutant alone leading to an approximately 50-fold increase in activity (Lampe *et al.*, 1999). No published information concerning the activity of these mutant transposases when introduced into insects is available, however. More recently, two of these same mutants have been introduced into the corresponding regions of the related *Mos1* transposase (the third, R131, already being present in *Mos1*) (Pledger and Coates, 2005). The E137K and E264R mutations both reside within the C-terminal domain identified as being required for catalysis rather than DNA binding and, based on secondary structure predictions, each is located on separate α–β-sheets. E137 is located upstream of the first aspartate of the catalytic DDD triad, while E264 is located between the second and third aspartates. Two *mariner* substrates were used in these investigations, one containing the wild-type imperfect 28 bp ITRs that contain four mismatches, and perfect ITRs in which these mismatches have been corrected to generate a donor *mariner* element with, in effect, two left ITRs.

Transposition assays performed in *E. coli* using the same mating-out strategy described for *Himar1* showed that both of these *Mos1* mutants were hyperactive, and more so with the donor containing the perfect ITRs, which, by itself, displayed increased levels of transposition, even when used as a substrate for wild-type transposase (Pledger and Coates, 2005). When tested in *A. aegypti* embryos, transposition assays showed less of a difference in transposition rates between both types of donor elements in the presence of wild-type *Mos1* transposase (Pledger and Coates, 2005). When the donor element containing the perfect ITRs was used in the presence of the E137K mutant, a three-fold increase in activity was observed; in the presence of the E264R mutant, a four-fold increase in activity was seen,

indicating that some level of transposase hyperactivity was retained when these transposases were tested in insects (Pledger and Coates, 2005).

These results illustrate that hyperactive mutants of eukaryotic tranposases can be generated. For the elements *Hermes*, *Mos1*, *piggyBac*, *Minos*, and *P*, the bottleneck for selecting these mutations remains the absence of a high-throughput means for testing in whole insects. Furthermore, the existence and physical separation of germ cells from somatic cells, and the possibility that transposition frequencies may differ between these two different cell types, means that current assays overwhelmingly measure transposition frequencies in somatic nuclei and not germ line nuclei. It is in the latter that initial transposition leading to the generation of transgenic individuals occurs. Thus, insect embryo and cell culture transposition assays, while being useful indicators of the ability of an element to transpose, are not necessarily good indicators of the element's ability to genetically transform the same insect species. The severity of these bottlenecks will be compounded as excision and transposition assays developed in *E. coli* and yeast to identify hyperactive mutants are more frequently adopted. These assays are high-throughput since libraries of mutated transposases can be rapidly screened using growth on defined media, generating mutants that then need to be tested in insects. A major caveat of these assays, which is clear from the *Himar* and *Mos* experiments, is that the levels of hyperactivity seen in these model organisms is not necessarily retained in insects. This is perhaps a reflection of the power of these initial selection schemes where mutants may be obtained that are adapted for activity in these organisms, but not in the target insect. This potential of generating false positives makes the development of a high-throughput procedure for the identification of hyperactive mutants in insect germ lines even more urgent.

A question arising from several studies of *Mos1* transposition is the role of host factors in this process. One explanation for the differing donor sequence requirements for *in vitro* and *in vivo* transposition is that host factors are involved with transposition. This explanation is, however, in some conflict with the distribution of *mariner* elements throughout the phyla and with the fact that, *in vitro*, no host factors are required for transposition. These observations suggest that if host factors are involved with *mariner* transposition *in vivo*, then they are sufficiently conserved across species to allow *mariner* element transposition into genomes immediately upon invasion. The identity of such host factors is open to speculation. Without question, repair of empty excision sites and transposition sites following integration can occur either through homologous recombination machinery (if repair involves identical sequences from the homologous chromosome) or through the non-homologous, end-joining pathway since excision and integration produce double stranded DNA breaks that need to be repaired if the chromosome and the host are to survive. These proteins are, however, not specific to the transposition process. The timing of transposition during the cell cycle may also lead to interactions between the transposase and the cell cycle machinery. These could occur at the level of DNA replication (for example, by providing access to DNA that is, or is about to be replicated) or may occur with proteins that regulate the cell cycle. This area remains one of the most fertile for exploration since it bears directly on questions of genome stability generated by transposable elements.

Applied use of transgenic insect strains

The potential applications of genetically modified insects are numerous, but have been focused generally on the death or sterilization of pests, or rendering them benign (Handler, 2002; Atkinson *et al.*, 2004). Genetic modification can be used in a converse fashion in beneficial insects that are predators of pests or which produce commercial products, by improving their viability and reproduction, or specific characteristics such as immune response, odour reception, or post-translational protein modification. Currently, the major emphasis in transgenic strain development is to improve biological control programmes for agricultural pests and to eliminate the vectorial capacity of vectors of human disease.

TRANSGENIC STRAINS FOR DISEASE VECTOR REPLACEMENT

A significant thrust of recombinant DNA technology in genetic control programmes developed to combat the spread of insect vector-borne disease is to introduce genes that prevent the transmission of pathogens through the tissues of the infected vector. These modified insects would thus be incapable of spreading the disease to uninfected human hosts, and the transmission cycle of disease would be broken. The generation of a transgenic insect that is refractory to disease transmission and which can spread the genes that confer this through a field population is a fascinating problem of engineering. Given that most current strategies are targeted towards mosquitoes, we will discuss only these in our examples. Consider what is required:

(1) A gene that completely prevents the transmission of a pathogen through the female mosquito. If the pathogen is a virus or a protozoan, then its population within the infected mosquito may number in the thousands to the millions.
(2) A promoter, and other regulatory sequences, that target expression of this gene to a particular tissue at a particular developmental time.
(3) A gene drive system that spreads this gene rapidly and effectively through a field population of mosquitoes.

Furthermore:

(1) The genetic construct should not disaffect the genetic fitness of the transgenic mosquito to a level that significantly increases the time required for the construct to spread through a field population.
(2) The construct should not break down over time.
(3) The construct should not lead to the misexpression of other genes, or to the development of new expression profiles in the mosquito host.
(4) The construct should not be capable of being transferred to another species.
(5) The entire vector replacement strategy should not be transient, leading to the possible re-introduction of the disease within a short period of time.
(6) The strategy should conform to all regulatory guidelines established by all relevant countries and organizations.

All of these are significant challenges and, taken together, may leave the reader with a degree of scepticism concerning whether a genetic control strategy based on the release of genetically engineered mosquitoes (in this example) can be achieved. Yet,

when balanced against the resurgence of mosquito-borne disease and the decreasing arsenal we have against these diseases, it is clear that this strategy does represent a viable option for future control of these scourges. Indeed, if we confine ourselves to issues of science and not of regulation, then the question of gene drive remains perhaps the largest unknown since practical use of gene drive mechanisms is still to be achieved, even though such phenomena exist in nature. Here, we will briefly examine the types of target genes being considered for gene replacement strategies, and will also discuss the problems of gene drive in insects.

Target genes for gene replacement

Pathogens of humans transmitted by insects such as mosquitoes can be viruses, protozoans, and worms. These can, potentially, also exact a pathogenic toll on the female mosquito and the resulting selective pressure invariably leads to the pathogen minimizing any fitness costs borne by the infected mosquito. This balance in fitness between infectivity and naivety opens up the exciting possibility of altering the mosquito's ability to defend itself against these pathogens. Attempts at increasing the immune response of mosquitoes to infection have focused on tweaking the innate immune response to enable the mosquito to recognize and eliminate these pathogens. These efforts have proceeded in concert with a characterization of regulatory sequences needed to direct expression of these genes to those tissues that encounter the pathogen. In mosquitoes, this is primarily the epidermal lining of the gut, the haemocoel (most of the cellular and humoral components of which are produced by the fat body), and the salivary glands.

Two recent technological developments have greatly accelerated this research, at least in the African malaria mosquito, *Anopheles gambiae*. One is the completion of a fairly mature draft sequence of the entire genome with the concomitant availability of microarray approaches (Holt *et al.*, 2002). These permit the identification of suites of genes that are up or downregulated in response to infection, and thus may lead to the identification of new pathways of immunity in insects. The second is the establishment of RNA interference (RNAi) techniques in this mosquito (Blandin *et al.*, 2002), which enables knock-down or knock-out phenotypes to be determined for each new gene, resulting in more certain knowledge of the gene's true function.

An illustration of this progress is seen from the identification of new genes and pathways involved in the response to plasmodium or bacterial infection in *A. gambiae*. Large-scale functional screens of genes involved in immunity have revealed hitherto unknown lectin- and leucine-rich genes that affect plasmodium development in this mosquito (Osta *et al.*, 2004). A genome-wide screen of *Anopheles* genes identified 71 genes involved in the phagocytosis of Gram-positive and Gram-negative bacteria, of which alteration of 26 affected the ability of the *Anopheles* cellular immune response to eliminate these pathogens (Moita *et al.*, 2005). The fruits from these and similar studies will no doubt lead to the generation of transgenic mosquitoes with enhanced immune responses that show decreased transmission of these pathogens. As such, these will follow the demonstrations that increased cecropin (an anti-microbial peptide) production in the mosquito midgut decreases plasmodium abundance in *A. gambiae* (Kim *et al.*, 2004); that the expression of honey bee venom phospholipase in the midgut inhibits plasmodium

abundance in *A. stephensi* (Moreira *et al.*, 2002); and that expression of the peptide SM1 in the mosquito midgut also severely disrupts plasmodium transmission in *A. stephensi* (Ito *et al.*, 2002).

Most likely, other genes and pathways will be identified and tested in transgenic insect vectors of disease. Some will provide greater levels of transmission inhibition than others and will be chosen for further studies to assess their impact on the genetic fitness of mosquitoes that contain them in their engineered form, and on their ability to prevent the transmission of field pathogens in field cages. The existence of these genes has never been in question, since natural and selected resistance to pathogens such as plasmodium has been well documented (Collins *et al.*, 1986). The long anticipated major hurdle to the field success of genetic replacement programmes in insect vectors such as mosquitoes has been the identification of genetic drive mechanisms.

Genetic drive

Genetic drive can be defined as being any mechanism that leads to a disproportionate number of offspring containing a desired genotype in excess of what would be predicted from Mendelian inheritance. It can be achieved through different mechanisms, one being meiotic drive, another being the use of transposable elements. A transposable element with a high transposition frequency that moves by replicative transposition will, during each meiosis, move to many positions throughout the genome. If it contains the desired effector gene (which we will assume is dominant and has minimal costs on fitness), then this gene will be present at least once in every gamete. As a consequence, the element, together with the gene it carries, should quickly spread through a population. The opposing forces to this spread are the fact that transposable elements act as mutagens when they insert into new genomic locations. They can inactivate genes or change their expression, and this will lead eventually to selection against these insertions. Furthermore, this internal mutagenesis forces the host organism to quickly develop mechanisms that immobilize these elements. In *Drosophila*, for example, many elements are found within heterochromatin, in which they are inactive. An engineered transposable element that has a high transient rate of transposition may be used to spread an effector gene through a population but, if this is to be practicable, it is likely that our knowledge of how elements move and how they are regulated will need to increase before such strategies can be implemented.

Meiotic drive mechanisms are present in plants and animals that lead to an asymmetric distribution of gametes following meiosis. If the effector gene is located on a chromosome that is favoured, and that is over-represented in the final gametic pool, then it will spread through a population. Meiotic drive systems have two components: a driver located on one chromosome and a responder locus that responds to the driver located on the homologous chromosome. In *D. melanogaster*, the segregator–distorter system contains the *Sd* driver gene and the responder site, *rsp*, with other loci spread along the second chromosome known to affect the ability of *Sd* to target the *rsp* locus (Ganetzky, 1977). The phenotypic consequence of segregation distortion is sperm dysfunction caused by a failure in chromatin condensation of chromosomes targeted by the *Sd* product (Tokuyasu *et al.*, 1977). This

product is a truncated form of the *RanGP* gene, the full-length form of which is involved in nuclear transport via the Ran regulatory pathway (Merrill *et al.*, 1999). While the mechanism by which this mutated gene causes segregation distortion remains unknown, it does seem clear that it most likely does so through interference with the nuclear transport of nucleic acids and/or proteins during spermatogenesis. Sex-linked meiotic drive systems have also been identified in the mosquitoes, *A. aegypti* and *Culex pipiens*, although the molecular basis of these awaits examination (Craig *et al.*, 1960; Sweeny and Barr, 1978). Systems such as these, combined with the role that centromeric-specific binding proteins such as the histone H3 variant may play in engineering drive systems, present an interesting challenge to those seeking to exploit meiotic drive systems in insects (Palmer *et al.*, 1987).

TRANSGENIC STRAINS FOR BIOCONTROL

The initial development of transgenic strains for insects of agricultural importance have focused on the introduction of attributes that would improve the most widely used system for biological control, known as the sterile insect technique, or SIT. SIT involves the field release of many millions of sterile males that render mated females infertile (Knipling, 1955). Except for the Mediterranean fruit fly (or medfly), *Ceratitis capitata* (see Franz *et al.*, 1996), sexing is not possible, or is highly inefficient for most insects, so sterile females are released with the males. Insects sterilized by irradiation also suffer from diminished overall fitness, and released insects are marked with fluorescent powder that is unreliable and harmful to humans (Hagler and Jackson, 2001). For each component of the SIT system, recombinant-based constructs could provide a meaningful improvement to the efficacy and cost effectiveness of the programme (Robinson *et al.*, 1999).

Fluorescent protein genetic markers

Initial efforts have been on the use of fluorescent proteins for genetic marking that also serve as a means to detect newly created transgenic insects. When the green fluorescent protein, GFP, from the jellyfish, *Aequorea victoria* (Prasher *et al.*, 1992; Chalfie *et al.*, 1994), or the red fluorescent protein, DsRed, from *Discosoma striata* (Matz *et al.*, 1999), were transcribed under the regulation of a *ubiquitin* gene promoter (Lee *et al.*, 1988) in several fly species, high expression levels were observed from the thoracic flight muscle (Handler and Harrell, 1999, 2001). These proteins are highly stable and can be detected under epifluorescence optics in insects several weeks after death (A. Handler and R. Harrell, unpublished), which should allow the identification of insects that have died in field traps. Determining whether trapped female insects have mated with sterile males and if sperm precedence is occurring are important to assessing programme effectiveness, and currently a reliable method to determine this is unavailable. Linking a testes-specific promoter from the *β2-tubulin* gene (Fackenthal *et al.*, 1993; Hoyle *et al.*, 1995) to DsRed results in red fluorescing sperm bundles in *Drosophila* and *Aedes*, and in *Drosophila*, the fluorescent sperm can be unambiguously detected in the seminal receptacle of mated females (A. Handler and R. Harrell, unpublished; R. Smith and P. Atkinson, unpublished). Transfer of this method to other non-drosophilids

should be straightforward, but will require isolating the *β2-tubulin* gene, or other spermatocyte promoter, from the species of interest, which has been achieved for *Heliothis virescens* and *A. aegypti* (Davis and Miller, 1988; R. Smith and P. Atkinson, unpublished). Testes-specific promoters may also find use in strains for male sterility, which would result from the testes-specific expression of gene products that cause cellular lethality, as discussed later. The sex-specific tissue expression of a fluorescent marker has the additional benefit potentially of allowing the sexing of embryos or larvae by COPAS™ fluorescent sorting (Furlong *et al.*, 2001). Sex-specific expression can be achieved in either male or female reproductive tissue, or by transgene integration into a sex-determining chromosome (e.g. Y chromosome in male flies and W chromosome in female moths).

Conditional regulation for sterility and lethality

Genes that cause cellular lethality can result in organismal death if expressed in all tissues or those that have vital functions, or result in sterility when expression is limited to reproductive tissue. Both phenotypes can play an important role in controlling insect pest populations. Genes that are lethal effectors include those involved in programmed cell death or apoptosis, dominant lethal mutations, and toxin subunits. A caveat to their use, however, is that their expression or gene product function must be conditional so that lethality or sterility only occurs under non-permissive conditions, while viability and fertility under permissive conditions allows normal strain rearing. Conditional gene expression is possible from regulatory promoters that are specifically induced or suppressed by molecules added to diet (e.g. antibiotics), while conditional function is possible from mutated genes having a lethal product that is temperature labile. The general strategy for use of these systems is to rear populations under permissive conditions with adults (or males only) released to the field. Resulting progeny are expected to die, preferably in early development, or become sterile under environmental conditions that are non-permissive. For male sterility or female lethality required for SIT, non-permissive conditions could be provided as part of the rearing regime.

For dietary regulation of sterility or lethality, a commonly used method is derived from the tetracycline-resistance operon of the *E. coli* Tn*10* transposon. This system can either activate or inhibit gene expression in the presence of tetracycline or derivatives (Gossen and Bujard, 1992; Gossen *et al.*, 1993). The *tet* transcriptional activator (tTA) is inhibited by tetracycline from binding to the *tet* response element (TRE or *tet*O), and thus tetracycline represses expression of genes linked to the TRE (known as '*tet*-off'), which is used routinely in *Drosophila* (Bello *et al.*, 1998) and has been tested successfully in the medfly (Gong *et al.*, 2005). A mutated form of tTA, *reverse* TA (rtTA), acts in a converse fashion (or '*tet*-on') where the tetracycline derivative, doxycycline, is required for its binding and activation of the TRE (Kistner *et al.*, 1996), which has also been demonstrated in *Drosophila* (Bieschke *et al.*, 1998). Use of these systems thus allows both negative and positive conditional regulation of lethal genes linked to the TRE. Promoter/enhancer sequences linked to tTA or rtTA provide tissue and developmental specificity to the *tet*-dependent lethal gene expression. Several model systems for *tet*-repressed lethality have been tested in *Drosophila* using the programmed cell death gene, *hid* (Abrams *et al.*, 1993;

Grether *et al.*, 1995; Wing *et al.*, 1998), as the lethal effector. Two tests used the female-specific *yolk protein* gene promoter to drive *hid* expression in the female fat body, a vital organ, to specifically cause female lethality in late pupae and early adults (Heinrich and Scott, 2000; Thomas *et al.*, 2000). Another approach used the promoters, *nullo* and *serendipity*, to cause death during embryonic development (Horn and Wimmer, 2003). For both systems, most insects failed to survive only when removed from a diet containing tetracycline. An interesting permutation of the *tet*-repressible lethality system is that the tTA product itself has a lethal effect at high concentrations. When placed in a positive regulatory loop where tTA is linked to the TRE, tTA drives its own expression in the absence of tetracycline, which resulted in >98% death in medfly larvae and pupae (Gong *et al.*, 2005).

Conditional regulation of lethal gene function is also achieved by dominant-acting, temperature-sensitive lethal genes that result in death in *Drosophila* at either abnormally high (28–29°C) or low (18–20°C) temperatures, but which typically can survive at moderate temperatures of 22–27°C. For this strategy, insects may be reared at moderate temperature and released, with heat-sensitive progeny dying in tropical climates, or cold-sensitive progeny failing to overwinter in mid-latitudes and higher. For the diphtheria (cold-sensitive) (Bellen *et al.*, 1992) or ricin (heat-sensitive) (Moffat *et al.*, 1992) toxin subunit genes, linkage to appropriate promoters should result in organismal lethality or tissue-specific cell death, and have been used for tissue ablation studies (Kalb *et al.*, 1993). For the *Drosophila Notch*^{ts1} cold-sensitive mutation (Fryxell and Miller, 1994) and DTS-5 heat-sensitive mutation (Saville and Belote, 1993; Covi *et al.*, 1999), organismal death has been observed at non-permissive temperatures, and for some strains of transgenic medflies homozygous for the *Drosophila* DTS-5 mutation, larval–pupal lethality occurs at 90–95% at 30°C (R. Krasteva and A. Handler, unpublished).

All of the conditional lethality systems are currently under development, and while high rates of lethality have been obtained, if high over-flooding ratios of transgenic insects (to field populations) are released, even 98–99% lethality in progeny could result in only a marginal effect on the targeted population. If the surviving offspring are revertants, a positive effect on population size could occur, exacerbating the problem. Thus, achieving 100% lethality still should remain a goal for these strategies, and one approach would be to integrate multiple lethal constructs into host strains. Modelling and large-scale rearing tests will be essential to determine actual efficacy and population dynamics at varying levels of penetrance of the lethal effect.

Stabilization of transposon vectors

Unlike gene transfer systems used for plants and animals where genomic integrations result from the fortuitous recombination of introduced DNA, non-autonomous transposon vectors can be re-mobilized by the presence of functional transposase. Transposase can arise from the undetected presence of the vector transposon in a host species, or a related transposon that can cross-mobilize the vector. While probably a rare occurrence in nature, remobilization of defective vectors is commonly induced for insertional mutagenesis studies with *P* and *piggyBac* (Brand *et al.*, 1994; Horn *et al.*, 2003), and cross-mobilization has been experimentally

demonstrated for the *hobo* and *Hermes* elements (Sundararajan *et al.*, 1999). Autonomous transposons are normally self-mobilized in nature, which is responsible for their spread among species, and which is desired for programmes requiring the transgene to be driven into a population, as described previously for vectors of disease. For *mariner*-related (Robertson, 1993; Robertson and MacLeod, 1993) and *piggyBac* (Handler and McCombs, 2000) elements in particular, nearly identical elements in distantly related species is a strong indication for transposon mobilizations and horizontal interspecies transfer. While it is unclear whether defective transposons have been cross-mobilized or are simply mutant derivatives of once functional elements, it is certainly possible for defective elements, or vectors, to be cross-mobilized. While this may occur rarely in nature, the large population scale and pressures of mass rearing for release programmes could be favourable for the selection of such events (Robinson *et al.*, 1999).

The potential for vector re-mobilization is a serious concern for transgenic strains used in biocontrol programmes since, primarily, this would result in strain instability due to loss of the vector transgene or transposition to another chromosomal site (Handler, 2004). Vector loss would result in elimination of desired attributes conferred on the transgenic strain, while intra-chromosomal movement could result in altered transgene expression (due to regulatory position effects) and cause new mutations diminishing strain fitness. Either effect could be devastating and costly to long-term programmes by reverting a 'beneficial' strain back to its pest status. A secondary effect of vector re-mobilization, though potentially more serious, is inter-organismal transfer of the vector, similar to horizontal transmission of transposons. The transfer of transgenes and the potential creation of mutations in unintended hosts could have serious ecological effects that would be a function of the particular transgene activity and the new host (Hoy, 2003). No matter how innocuous the transgene might seem to be, it is clearly impossible to estimate how any particular transgene might behave in numerous potential hosts and the ecological consequences. Fluorescent proteins, GFP and DsRed in particular, would be the minimal amount of exogenous genetic material in any transgenic insect, and they are already used routinely in experimental settings (Horn and Wimmer, 2000). GFP has been used in many transgenic organisms as reporters for gene expression, in humans to monitor gene therapeutic agents, and transgenic fish are commercially available as novelty pets (GloFish™). Yet toxicities have been established for highly expressed GFP in cell culture (Liu *et al.*, 1999), and scientific and social concern might be expected if it was used in a transgenic organism where its stability could not be assured with certainty. These concerns would be compounded in terms of the use of transgenes resulting in a lethal cellular response.

These ecological concerns were voiced in response to the environmental assessment of a contained field release of a transgenic lepidopteran pest, the pink bollworm, *Pectinophora gossypiella*. This strain was transformed with a *piggyBac*/EGFP vector to genetically mark moths used in SIT (Peloquin *et al.*, 2000). The environmental assessment for this project was widely commented upon, with particular emphasis on issues relating to transgene stability, especially for the potential movement into infecting baculoviruses (see http://www.centerforfoodsafety.org/pubs/CommentsUSDAFieldRelease7.21.2001.pdf). This was especially significant since

piggyBac was originally discovered by virtue of it transposing from a lepidopteran cell line genome into an infecting baculovirus (Fraser *et al.*, 1983, 1985).

It was with the consideration for potential vector instability and the inability to adequately assess or control for the possibility of inter-species transfer of a transposon vector that we and others considered the critical need for new vectors that could be immobilized after transposase-mediated integration. Since transposase was already deleted from the vector, the most obvious approach to immobilization was to rearrange or delete one or both of the vector terminal sequences that bind to transposase. One approach considered is use of recombination sites, such as *FRT* or *lox*P, within vectors so that terminal sequences could be rearranged or deleted after integration in the presence of the FLP or Cre recombinase, respectively. The simplest approach is to have two recombination sites within a vector in opposite orientation that would result in an inversion (Handler, 2004). However, to displace a terminal sequence in this way, one site of at least 34 bp would have to be within the terminal sequence, and this might disrupt transposase binding and the ability to generate the initial integration. Alternatively, recombination sites could be in independent vectors that, when integrated in the same chromosome in opposite orientation, create a chromosomal inversion that results in vector chimeras. This approach was successfully tested in *Drosophila* using *FRT* recombination between *piggyBac* and *Hermes* vectors inserted into the same chromosome (I. Viktorinova and E. Wimmer, personal communication). Despite the elegance of this approach and its potential use in *Drosophila*, its wide use in nondrosophilid species probably will be limited, in the short term at least, since assigning linkage groups and vector orientation (to avoid lethal deletions) is not readily possible for most other insects.

We therefore considered an alternative approach that, *a priori*, also has limitations. To stabilize transposon vectors subsequent to genomic integration, we tested a *piggyBac* vector in *Drosophila* that allowed terminal sequence deletion by introducing an internal tandem duplication of the other terminal sequence. Independent fluorescent protein markers were placed between each set of termini to create the vector plasmid pBac{L1-PUbDsRed1-L2-3xP3-ECFP-R1} (Handler *et al.*, 2004). A duplicated 5' terminal *piggyBac* sequence (L2) was inserted internal to the flanking 5' (L1) and 3' (R1) termini, with independent markers, PUbDsRed1 and 3xP3-ECFP, placed between each set of termini. Transformation with this vector can result in two types of integration: either the shorter embedded L2-3xP3-ECFP-R1 sequence may integrate using the internal L2 5' terminus, or the complete L1-PUbDsRed1-L2-3xP3-ECFP-R1 vector may integrate using the L1 5' terminus, which is the desired integration. The theoretical limitation to this approach is that shorter vectors transpose more efficiently than longer vectors, and it was conceivable that only L2-3xP3-ECFP-R1 integrations would occur. However, while the initial transformation with this vector resulted in seven lines with only the embedded L2-3xP3-ECFP-R1 vector, one line incorporating the complete L1-PUbDsRed1-L2-3xP3-ECFP-R1 vector also occurred. In a converse fashion, upon re-mobilization of the L1-PUbDsRed1-L2-3xP3-ECFP-R1 vector by mating to a *piggyBac* transposase jumpstarter strain (having a chromosomal source of the functional transposase gene), the L2-3xP3-ECFP-R1 vector was remobilized more frequently, resulting in progeny having only the L1-PUbDsRed1 vector sequence remain (*Figure 8.2*). In the absence of the R1 3' *piggyBac* terminus, the expectation is that the remaining L1-

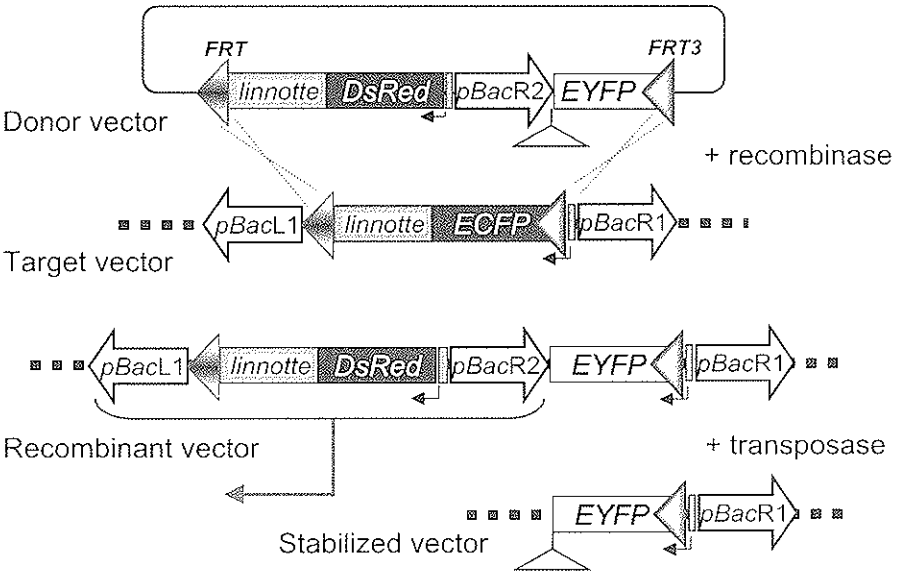

Figure 8.2. Schematic of genomic targeting by recombinase-mediated cassette exchange (RMCE) and subsequent target site vector stabilization by ITR sequence deletion. The donor vector is a pUC-derived plasmid containing heterospecific *FRT* and *FRT3* sites surrounding a 3xP3-DsRed marker gene linked to the *linnotte* homing sequence, a *piggyBac* R2 (3') terminal sequence, a cloning site (triangle), and a promoter-less EYFP gene. In the presence of FLP recombinase, double recombination cassette exchange occurs (broken lines) with the previously integrated *piggyBac* vector target site that consists of the *piggyBac* L1 (5') and R1 (3') terminal sequences surrounding heterospecific *FRT/FRT3* sites, which surround in turn the 3xP3-ECFP marker gene linked to the *linnotte* homing sequence. The *FRT3* site is positioned in between the ECFP and its promoter. The cassette exchange recombination replaces the target vector sequences internal to the *FRT/FRT3* sites with the donor vector sequences internal to its *FRT/FRT3* sites. The cassette exchange process is monitored by replacement of the EYFP phenotype with an ECFP phenotype. Stabilization of the target site vector is achieved by providing *piggyBac* transposase, resulting in mobilization of the *pBac*L1 and *pBac*R2 termini and intervening sequences (within the bracket), leaving the 3xP3-EYFP marker gene, genes of interest within the cloning site, the *FRT3* site and the *pBac*R1 terminus. In the absence of a *piggyBac* 5' terminus, the remaining target site sequence is immobilized with respect to *piggyBac* transposase.

PUbDsRed1 sequence would remain stable in the genome, irrespective of the presence of transposase. This was verified by mating the L1-PUbDsRed1 stabilized line to the jumpstarter strain, which resulted in a stable phenotype in more than 7000 progeny assayed, compared to a >5% remobilization frequency in the original L1-PUbDsRed1-L2-3xP3ECFP-R1 strain. This demonstrated stabilization of the PUbDsRed1 transgene resulting from loss of the 3' *piggyBac* vector terminus. Other genes of interest, if inserted upstream to the internal 5' sequence, would have been similarly stabilized. Stabilization vectors have since been incorporated into the medfly and caribfly, allowing a practical evaluation of this system under mass rearing regimes.

Gene vector targeting

Creation of transgenic strains for biocontrol having optimal fitness and viability

attributes is often difficult due to vector insertional mutations that disrupt vital gene functions, and suppressed or variable transgene expression due to genomic position and enhancer effects. These effects are well established by studies that show diminished viability in transgenic strains and suppressed transgene marker function (Catteruccia *et al.*, 2003; Irvin *et al.*, 2004). Several studies on transgenic lethality vectors in *Drosophila* have shown that, typically, only one of ten or more strains exhibit lethality greater than 95% (Heinrich and Scott, 2000; Horn and Wimmer, 2003), and this is attributed to position effect suppression. These disadvantages can be minimized by targeting transgene integrations to pre-defined genomic sites devoid of known coding or regulatory function and that are minimally affected by position and enhancer effects. Targeting transgene insertions has also been a major goal of those studying gene structure–function relationships that require the same genomic context for true allelic comparisons of gene expression. This motivated the development of targeting systems based upon the ϕC31 integrase (Groth *et al.*, 2004) and FLP (Horn and Handler, 2005) and Cre (Oberstein *et al.*, 2005) recombinase systems. For applied use, the targeting concept would be realized by creating several (perhaps many) target site strains by initially introducing genomic recombination sites using transposon-mediated gene transfer. These acceptor sites would be fully characterized with respect to insertion site sequence and function (i.e. coding or regulatory function) and position effect variegation, and the transgenic strains would be assessed for fitness, viability, and any other desired attributes. Optimal strains then would be used for all subsequent targeted integrations by an integrase or recombinase-mediated process.

Using a P element vector, *attP* sites have been integrated into the *Drosophila* genome, with subsequent targeted integration by *attB* plasmid in the presence of integrase (Groth *et al.*, 2004). This was the first relatively high frequency unidirectional targeting system developed for an insect achieving integrations at a rate of 55%; however, the system as described is specific for *Drosophila*, occasionally genomic 'pseudo' *attP* sites may be targeted, and single recombinations result in the entire *attB* donor vector plasmid being integrated. Thus, the integration product will typically contain extraneous DNA, including antibiotic resistance genes and plasmid origins of replication, that may be undesirable in insects used in biocontrol release programmes.

Unlike the single recombination integrase system, the FLP/*FRT* and Cre/*loxP* systems have been used in recombinase-mediated cassette exchange (RMCE) that uses a double recombination to integrate a defined cassette sequence having only genes of interest. This was developed previously using heterospecific *FRT/FRT3* sites for targeted insertions in mammalian stem cell lines to control transgene expression (Baer and Bode, 2001). Heterospecificity was conferred by nucleotide substitutions in four of the eight base pairs in the spacer sequence in *FRT* to create *FRT3*. While FLP recombinase mediates recombination between each *FRT*, cross reactivity between *FRT* and *FRT3* is eliminated.

We modified the *in vitro* RMCE system for use in organisms, and insects in particular (*Figure 8.2*), by creating genomic target sites with a *piggyBac* acceptor vector and assessing cassette exchange by interconvertible fluorescent protein markers (Horn and Handler, 2005). A homing sequence from the *Drosophila linotte* locus was incorporated to facilitate target and donor sequence pairing. The *piggyBac*

target site acceptor vector has the two heterospecific *FRT* recombination sites surrounding a cyan fluorescent protein (CFP) marker coding region regulated by the eye-specific 3xP3 promoter, and the *linotte* sequence. Once integrated into the genome, the acceptor site was targeted by a donor vector having corresponding *FRT/FRT3* sites surrounding a YFP (yellow fluorescent protein) marker-coding region that lacked a promoter (and thus could be expressed only after recombining proximal to a promoter), and *linotte* sequences. Recombination between the target and donor *FRT/FRT3* sites was mediated by FLP recombinase, with recombinants identified at a frequency of ~23% by screening for conversion of CFP to the YFP eye fluorescence marker phenotype. To stabilize targeted insertions, a new RMCE donor vector had a *piggyBac* 3'-terminus (R2) incorporated to allow post-integration deletion of the *piggyBac* 5'-terminus, as described previously.

Another RMCE system tested in *Drosophila* relies on heterospecific *loxP* recombination sites (Oberstein *et al.*, 2005). Notably, this system did not incorporate a homing sequence, but RMCE frequencies were less than half those reported for the *FRT* system. Nevertheless, this suggests that while homing sequences such as *linotte* may enhance RMCE, they are not essential. This is an important consideration for use of this system in nondrosophilids for which homing sequences have not been identified. This approach also shows that use of heterospecific *FRT* and *loxP* sites together in an acceptor site can provide numerous targets for repetitive integrations. New transgene vectors such as these, which allow genomic targeting and post-integration stabilization, should improve significantly the efficient creation of insects intended for field release and, importantly, enhance their ecological safety.

Summary

The ability to routinely genetically modify insect species holds great promise for fundamental research that explores the functional activity of genomic sequences, and the use of this information to control the viability, fitness, and behaviour of both beneficial and pest insects. Currently, almost all insect genetic modifications rely on the use of transposon vector systems, and a detailed understanding of the mechanisms that result in mobility is critical to applications that require optimal or maximum frequencies of transposition, and to applications where immobilization is necessary for vector stabilization. Great progress has been made in understanding the biophysical mechanisms and interactions between the transposase enzyme for the *Hermes* and *Mos*1 transposons and their respective ITR sequence substrates, but the relevance of this knowledge to other transposon vectors can only be speculated upon. It is clear, however, that mutations in the transposon sequence can result in their hyperactivity, and an effective means of screening for these mutations should improve our understanding and applied use of all the available vectors. Progress also has been made in testing recombinant-based constructs for their ability to diminish the vectorial capacity of mosquito disease vectors, but the ability to drive these transgenes into an endemic population is largely unknown. Genetic drive systems, such as autonomous vectors or meiotic drive, have been speculated upon, but serious testing in targeted species remains to be done. Development of transgenic strains for biocontrol has also been initiated, especially for tephritid fruit flies, and conditional lethality systems may supersede current programmes such as SIT. To do

so, nearly complete, if not complete lethality will be needed at a consistent level, and model systems have yet to achieve this. To develop such strains, repetitive introductions of transgene vectors into a host genome may be required, but a difficulty in comparing efficacy is the varying influence of different insertion sites on transgene expression and host fitness. A prospective problem for transposon-mediated vector insertions is the potential re-mobilization of the vector by an unintended source of transposase. The development of a new class of vectors that allow genomic targeting by RMCE, and transposon immobilization by ITR deletion, should have a significant impact on the efficient creation and testing of new transgenic strains, as well as minimizing the ecological risk of their release into the environment.

Acknowledgements

Grateful appreciation is extended to collaborators and current and former members of our laboratories and to the USDA-NRI-CGP (A.M.H.) and the USPHS (P.W.A.) for continual support of our research.

References

ABRAMS, J.M., AGAPITE, J., WHITE, K. AND STELLER, S. (1993). Programmed cell death during *Drosophila* embryogenesis. *Development* **117**, 29–43.

ALPHEY, L., BEARD, C.B., BILLINGSLEY, P. *ET AL.* (2002). Malaria control with genetically manipulated insect vectors. *Science* **298**, 119–121.

ATKINSON, P.W., O'BROCHTA, D.A. AND ROBINSON, A.S. (2004). Insect transformation for use in control. In: *Comprehensive insect physiology, biochemistry and pharmacology*, 2nd edition. Eds. S.S. Gill, L.I. Gilbert and K. Iatrou, pp 403–410. Amsterdam, The Netherlands: Elsevier Pergamon.

AUGE-GOUILLOU, C., HAMELIN, M.-H., DEMATTEI, M.-V., PERIQUET, M. AND BIGOT, Y. (2001). The wild-type conformation of the *Mos-1* inverted terminal repeats is suboptimal for transposition in bacteria. *Molecular Genetics and Genomics* **265**, 51–57.

AUGE-GOUILLOU, C., BRILLET, B., GERMON, S., HAMELIN, M.-H. AND BIGOT, Y. (2005a). *Mariner Mos1* transposase dimerizes prior to ITR binding. *Journal of Molecular Biology* **351**, 117–130.

AUGE-GOUILLOU, C., BRILLET, B., HAMELIN, M.-H. AND BIGOT, Y. (2005b). Assembly of the *mariner Mos1* synaptic complex. *Molecular and Cellular Biology* **25**, 2861–2870.

BAER, A. AND BODE, J. (2001). Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Current Opinion in Biotechnology* **12**, 473–480.

BELLEN, H.J., D'EVELYN, D., HARVEY, M. AND ELLEDGE, S.J. (1992). Isolation of temperature-sensitive diphtheria toxins in yeast and their effects on *Drosophila* cells. *Development* **114**, 787–796.

BELLO, B., RESENDIEZ-PEREZ, D. AND GEHRING, W.J. (1998). Spatial and temporal targeting of gene expression in *Drosophila* by means of a tetracycline-dependent transactivator system. *Development* **125**, 2193–2202.

BIESCHKE, E.T., WHEELER, J.C. AND TOWER, J. (1998). Doxycycline-induced transgene expression during *Drosophila* development and aging. *Molecular and General Genetics* **258**, 571–579.

BIGOT, Y., BRILLET, B. AND AUGE-GOUILLOU, C. (2005). Conservation of palindromic and mirror motifs within inverted terminal repeats of *mariner*-like elements. *Journal of Molecular Biology* **351**, 108–116.

BLANDIN, S., MOITA, L.F., KOCHER, T., WILM, M., KAFATOS, F.C. AND LEVASHINA, E.A.

(2002). Reverse genetics in the mosquito, *Anopheles gambiae*: targeted disruption of the defensin gene. *EMBO Reports* **3**, 852–856.

BRAND, A.H. AND PERRIMON, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415.

BRAND, A.H., MANOUKIAN, A.S. AND PERRIMON, N. (1994). Ectopic expression in *Drosophila*. *Methods in Cell Biology* **44**, 635–654.

CATTERUCCIA, F., GODFRAY, H.C. AND CRISANTI, A. (2003). Impact of genetic manipulation on the fitness of *Anopheles stephensi* mosquitoes. *Science* **299**, 1225–1227.

CHALFIE, M., TU, Y., EUSKIRCHEN, G., WARD, W. AND PRASHER, D.C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**, 802–805.

COLLINS, F.H., SAKAI, R.K., VERNICK, K.D. *ET AL.* (1986). Genetic selection of a Plasmodium-refractory strain of the malaria vector *Anopheles gambiae*. *Science* **234**, 607–610.

COVI, J.A., BELOTE, J.M. AND MYKLES, D.L. (1999). Subunit compositions and catalytic properties of proteasomes from developmental temperature-sensitive mutants of *Drosophila melanogaster*. *Archives of Biochemistry and Biophysics* **368**, 85–97.

CRAIG, G.B., HICKEY, W.A. AND VAN DE HEY, R.C. (1960). An inherited male-producing factor in *Aedes aegypti*. *Science* **23**, 1887–1889.

CRAIG, N.L. (2002). Tn*7*. In: *Mobile DNA II*. Eds. N.L. Craig, R. Cragie, M. Gellert and A.M. Lambowitz, pp 423–456. Washington, DC, USA: ASM Press.

CRAIG, N.L., CRAGIE, R., GELLERT, M. AND LAMBOWITZ, A.M. (2002). *Mobile DNA II*. Washington, DC, USA: ASM Press.

DAVIES, D.R., GORYSHIN, I.Y., REZNIKOFF, W.S. AND RAYMONT, I. (2000). Three-dimensional structure of the Tn*5* synaptic complex transposition intermediate. *Science* **289**, 77–85.

DAVIS, M.-T.B. AND MILLER, S.G. (1988). Expression of a testis beta-tubulin gene in *Heliothis virescens* during spermatogenesis. *Insect Biochemistry* **18**, 389–394.

DAWSON, A. AND FINNEGAN, D.J. (2003). Excision of the *Drosophila mariner* transposon *Mos1*: comparison with bacterial transposition and V(D)J recombination. *Molecular Cell* **11**, 225–235.

DUVAL-VALENTIN, G., NORMAND, C., KHEMICI, V., MARTY, B. AND CHANDLER, M. (2001). Transient promoter formation: a new feedback mechanism for regulation of IS911 transposition. *EMBO Journal* **20**, 5802–5811.

FACKENTHAL, J.D., TURNER, F.R. AND RAFF, E.C. (1993). Tissue-specific microtubule functions in *Drosophila* spermatogenesis require the beta 2-tubulin isotype-specific carboxy terminus. *Developmental Biology* **158**, 213–227.

FINNEGAN, D.J. (1989). Eukaryotic transposable elements and genome evolution. *Trends in Genetics* **5**, 103–107.

FRANZ, G. AND SAVAKIS, C. (1991). Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc-1-like family of transposons. *Nucleic Acids Research* **19**, 6646.

FRANZ, G.F., KERREMANS, PH., RENDON P. AND HENDRICHS, J. (1996). Development and application of genetic sexing systems for the Mediterranean fruit fly based on a temperature sensitive lethal. In: *Fruit fly pests: a world assessment of their biology and management*. Eds. B.A. McPheron and G.J. Steck, pp 185–193. Orlando, Florida, USA: St. Lucie Press.

FRASER, M.J., SMITH, G.E. AND SUMMERS, M.D. (1983). Acquisition of host-cell DNA-sequences by baculoviruses – relationship between host DNA insertions and FP mutants of *Autographa californica* and *Galleria mellonella* nuclear polyhedrosis viruses. *Journal of Virology* **47**, 287–300.

FRASER, M.J., BRUSCA, J.S., SMITH, G.E. AND SUMMERS, M.D. (1985). Transposon-mediated mutagenesis of a baculovirus. *Virology* **145**, 356–361.

FRYXELL, K.J. AND MILLER, T.A. (1994). Autocidal biological control: a general strategy for insect control based on genetic transformation with a highly conserved gene. *Journal of Economic Entomology* **85**, 1240–1245.

FURLONG, E.E., PROFFITT, D. AND SCOTT, M.P. (2001). Automated sorting of live transgenic embryos. *Nature Biotechnology* **19**, 153–156.

GANETZKY, B. (1977). On the components of segregation distortion in *Drosophila melanogaster*. *Genetics* **86**, 321–355.

GOLIC, M.M., RONG, Y.S., PETERSEN, R.B., LINDQUIST, S.L. AND GOLIC, K.G. (1997). FLP-

mediated DNA mobilization to specific target sites in *Drosophila* chromosomes. *Nucleic Acids Research* **25**, 3665–3671.

GONG, P., EPTON, M.J., FU, G. *ET AL.* (2005). A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly. *Nature Biotechnology* **23**, 453–456.

GOSSEN, M. AND BUJARD, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 5547–5551.

GOSSEN, M., BONIN, A.L. AND BUJARD, H. (1993). Control of gene activity in higher eukaryotic cells by prokaryotic regulatory elements. *Trends in Biochemical Sciences* **18**, 471–475.

GRETHER, M.E., ABRAMS, J.M., AGAPITE, J., WHITE, K. AND STELLER, H. (1995). The head involution defective gene of *Drosophila melanogaster* functions in programmed cell death. *Genes and Development* **9**, 1694–1708.

GROTH, A.C. AND CALOS, M.P. (2004). Phage integrases: biology and applications. *Journal of Molecular Biology* **335**, 667–678.

GROTH, A.C., FISH, M., NUSSE, R. AND CALOS, M.P. (2004). Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* **166**, 1775–1782.

GUEIROS-FILHO, F.J. AND BEVERLEY, S.M. (1997). Trans-kingdom transposition of the *Drosophila* element *mariner* within the protozoan *Leishmania*. *Science* **276**, 1716–1719.

HAGLER, J.R. AND JACKSON, C.G. (2001). Methods for marking insects: current techniques and future prospects. *Annual Review of Entomology* **46**, 511–543.

HANDLER, A.M. (2001). A current perspective on insect gene transformation. *Insect Biochemistry and Molecular Biology* **31**, 111–128.

HANDLER, A.M. (2002). Prospects for using genetic transformation for improved SIT and new biocontrol methods. *Genetica* **116**, 137–149.

HANDLER, A.M. (2004). Understanding and improving transgene stability and expression in insects for SIT and conditional lethal release programs. *Insect Biochemistry and Molecular Biology* **34**, 121–130.

HANDLER, A.M. AND HARRELL, R.A., II. (1999). Germline transformation of *Drosophila melanogaster* with the *piggyBac* transposon vector. *Insect Molecular Biology* **8**, 449–458.

HANDLER, A.M. AND HARRELL, R.A., II. (2001). Polyubiquitin-regulated DsRed marker for transgenic insects. *Biotechniques* **31**, 820–828.

HANDLER, A.M. AND MCCOMBS, S.D. (2000). The *piggyBac* transposon mediates germ-line transformation in the Oriental fruit fly and closely related elements exist in its genome. *Insect Molecular Biology* **9**, 605–612.

HANDLER, A.M., ZIMOWSKA, G.J. AND HORN, C. (2004). Post-integration stabilization of a transposon vector by terminal sequence deletion in *Drosophila melanogaster*. *Nature Biotechnology* **22**, 1150–1154.

HANNIFORD, D.B. (2002). Transposon Tn*10*. In: *Mobile DNA II*. Eds. N.L. Craig, R. Cragie, M. Gellert and A.M. Lambowitz, pp 457–483. Washington, DC, USA: ASM Press.

HEINRICH, J.C. AND SCOTT, M.J. (2000). A repressible female-specific lethal genetic system for making transgenic insect strains suitable for a sterile-release program. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 8229–8232.

HICKMAN, A.B., PEREZ, Z.N., ZHOU, L. *ET AL.* (2005). Molecular architecture of a eukaryotic DNA transposase. *Nature Structural and Molecular Biology* **12**, 715–721.

HOLT, R.A., SUBRAMANIAN, G.M., HALPERN, A. *ET AL.* (2002). The genome sequence of the malaria mosquito. *Anopheles gambiae*. *Science* **298**, 129–149.

HORN, C. AND HANDLER, A.M. (2005). Site-specific genomic targeting in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 12483–12488.

HORN, C. AND WIMMER, E.A. (2000). A versatile vector set for animal transgenesis. *Development, Genes and Evolution* **210**, 630–637.

HORN, C. AND WIMMER, E.A. (2003). A transgene-based, embryo-specific lethality system for insect pest management. *Nature Biotechnology* **21**, 64–70.

HORN, C., OFFEN, N., NYSTEDT, S., HACKER, U. AND WIMMER, E.A. (2003). *piggyBac*-based insertional mutagenesis and enhancer detection as a tool for functional insect genomics. *Genetics* **163**, 647–661.

HOY, M.A. (2003). Transgenic insects for pest management programs: status and prospects. *Environmental Biosafety Research* **2**, 61–64.

HOYLE, H.D., HUTCHENS, J.A., TURNER, F.R. AND RAFF, E.C. (1995). Regulation of beta-tubulin function and expression in *Drosophila* spermatogenesis. *Developmental Genetics* **16**, 148–170.

IRVIN, N., HODDLE, M.S., O'BROCHTA, D.A., CAREY, B. AND ATKINSON, P.W. (2004). Assessing fitness costs for transgenic *Aedes aegypti* expressing the GFP marker and transposase genes. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 891–896.

ITO, J., GHOSH, A., MOREIRA, L.A., WIMMER, E.A. AND JACOBS-LORENA, M. (2002). Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature* **417**, 387–388.

JASINSKIENE, N., COATES, C.J., BENEDICT, M.Q. *ET AL.* (1998). Stable transposon mediated transformation of the yellow fever mosquito, *Aedes aegypti*, using the *Hermes* element from the housefly. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 3743–3747.

KALB, J.M., DIBENEDETTO, A.J. AND WOLFNER, M.F. (1993). Probing the function of *Drosophila melanogaster* accessory glands by directed cell ablation. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8093–8097.

KIM, W., KOO, H., RICHMAN, A.M. *ET AL.* (2004). Ectopic expression of a cecropin transgene in the human malaria vector mosquito *Anopheles gambiae* (Diptera: Culicidae): effects on susceptibility to Plasmodium. *Journal of Medical Entomology* **41**, 447–455.

KISTNER, A., GOSSEN, M., ZIMMERMANN, F. *ET AL.* (1996). Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 10933–10938.

KNIPLING, E.F. (1955). Possibilities of insect control or eradication through the use of sexually sterile males. *Journal of Economic Entomology* **48**, 459–462.

LAMPE, D.J., CHURCHILL, M.E. AND ROBERTSON, H.M. (1996). A purified *mariner* transposase is sufficient to mediate transposition *in vitro*. *EMBO Journal* **15**, 5470–5479.

LAMPE, D.J., AKERLEY, B.J., RUBIN, E.J., MEKALONOS, J.J. AND ROBERTSON, H.M. (1999). Hyperactive transposase mutants of the *Himar1 mariner* transposon. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11428–11433.

LANDER, E.S., LINTON, L.M., BIRREN, B. *ET AL.* (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.

LEE, H.S., SIMON, J.A. AND LIS, J.T. (1988). Structure and expression of ubiquitin genes of *Drosophila melanogaster*. *Molecular and Cellular Biology* **8**, 4727–4735.

LIPKOW, K., BUISINE, N., LAMPE, D.J. AND CHALMERS, R. (2004). Early intermediates of *mariner* transposition: catalysis without synapsis of the transposon ends suggests a novel architecture of the synaptic complex. *Molecular and Cellular Biology* **24**, 8301–8311.

LIU, H.-S., JAN, M.-S., CHOU, C.-K., CHEN, P.-H. AND KE, N.-J. (1999). Is green fluorescent protein toxic to living cells? *Biochemical and Biophysical Research Communications* **260**, 712–717.

LOHE, A.R. AND HARTL, D.L. (2002). Efficient mobilization of *mariner in vivo* requires multiple internal sequences. *Genetics* **160**, 519–526.

LUKACSOVICH, T., ASZTALOS, Z., AWANO, W. *ET AL.* (2001). Dual-tagging gene trap of novel genes in *Drosophila melanogaster*. *Genetics* **157**, 727–742.

MATZ, M.V., FRADKOV, A.F., LABAS, Y.A. *ET AL.* (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotechnology* **17**, 969–973.

MEDHORA, M.M., MACPEEK, A.H. AND HARTL, D.L. (1988). Excision of the *Drosophila* transposable element *mariner*: identification and characterization of the *Mos* factor. *EMBO Journal* **7**, 2185–2189.

MEDHORA, M., MARUYAMA, K. AND HARTL, D.L. (1991). Molecular and functional analysis of the *mariner* mutator element *Mos1* in *Drosophila*. *Genetics* **128**, 311–318.

MELLO, C. AND FIRE, A. (1995). DNA transformation. *Methods in Cell Biology* **48**, 451–482.

MERRILL, C., BAYRAKTAROGLU, L., KUSANO, A. AND GANETZKY, B. (1999). Truncated

RanGAP encoded by the segregation distorted mutation in *Drosophila*. *Science* **283**, 1742–1745.

MICHEL, K. AND ATKINSON, P.W. (2003). Nuclear localization of the *Hermes* transposase depends on basic amino acid residues at the N-terminus of the protein. *Journal of Cellular Biochemistry* **89**, 778–790.

MICHEL, K., STAMENOVA, A., PINKERTON, A.C. *ET AL.* (2001). *Hermes*-mediated germ-line transformation of the Mediterranean fruit fly *Ceratitis capitata*. *Insect Molecular Biology* **10**, 155–162.

MICHEL, K., O'BROCHTA, D.A. AND ATKINSON, P.W. (2002). Does the proposed DSE motif form the active centre in the *Hermes* transposase? *Gene* **298**, 141–146.

MICHEL, K., O'BROCHTA, D.A. AND ATKINSON, P.W. (2003). The C-terminus of the *Hermes* transposase contains a protein multimerization domain. *Insect Biochemistry and Molecular Biology* **33**, 959–970.

MOFFAT, K.G., GOULD, J.H., SMITH, H.K. AND O'KANE, C.J. (1992). Inducible cell ablation in *Drosophila* by cold-sensitive ricin A chain. *Development* **114**, 681–687.

MOITA, L.F., WANG-SATTLER, R., MICHEL, K. *ET AL.* (2005). *In vivo* identification of novel regulators and conserved pathways of phagocytosis in *A. gambiae*. *Immunity* **23**, 65–73.

MOREIRA, L.A., ITO, J., GHOSH, A. *ET AL.* (2002). Bee venom phospholipase inhibits malaria parasite development in transgenic mosquitoes. *Journal of Biological Chemistry* **277**, 40839–40843.

MORIN, X., DANEMAN, R., ZAVORTINK, M. AND CHIA, W. (2001). A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 15050–15055.

NOWOTNY, M., GAIDAMAKOV, S.A., CROUCH, R.J. AND YANG, W. (2005). Crystal structures of RNaseH bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* **121**, 1005–1016.

OBERSTEIN, A., PARE, A., KAPLAN, L. AND SMALL, S. (2005). Site-specific transgenesis by Cre mediated recombination in *Drosophila*. *Nature Methods* **2**, 583–585.

O'BROCHTA, D.A., WARREN, W.D., SAVILLE, K.J. AND ATKINSON, P.W. (1996). *Hermes*, a functional non-drosophilid insect gene vector. *Genetics* **142**, 907–914.

OSTA, M.A., CHRISTOPHIDES, G.K. AND KAFATOS, F.C. (2004). Effects of mosquito genes on Plasmodium development. *Science* **303**, 2030–2031.

PALMER, D.K., O'DAY, K., WENER, M.H., ANDREWS, B.S. AND MARGOLIS, R.L. (1987). A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *Journal of Cell Biology* **104**, 805–815.

PARKER, J.S., ROE, S.M. AND BARFORD, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO Journal* **23**, 4727–4737.

PELOQUIN, J.J., THIBAULT, S., STATEN, R. AND MILLER, T.A. (2000). Germ-line transformation of pink bollworm (Lepidoptera: Gelechiidae) mediated by the *piggyBac* transposable element. *Insect Molecular Biology* **9**, 323–333.

PEREZ, Z.N., MUSINGARIMI, P., CRAIG, N.L., DYDA, F. AND HICKMAN, A.B. (2005). Purification, crystallization and preliminary crystallographic analysis of the *Hermes* transposase. *Acta Crystallographica* **F61**, 587–590.

PLEDGER, D.W. AND COATES, C.J. (2005). Mutant *Mos1 mariner* transposons are hyperactive in *Aedes aegypti*. *Insect Biochemistry and Molecular Biology* **35**, 1199–1207.

PLEDGER, D.W., FU, Y.Q. AND COATES, C.J. (2004). Analysis of *cis*-acting elements that affect the transposition of *Mos1 mariner* transposons *in vivo*. *Molecular Genetics and Genomics* **272**, 67–75.

PRASHER, D.C., ECKENRODE, V.K., WARD, W.W., PRENDERGAST, F.G. AND CORMIER, M.J. (1992). Primary structure of the *Aequorea victoria* green fluorescent protein. *Gene* **111**, 229–233.

REZNIKOFF, W.S. (2002). Tn5 transposition. In: *Mobile DNA II*. Eds: N.L. Craig, R. Cragie, M. Gellert and A.M. Lambowitz, pp 403–422. Washington, DC, USA: ASM Press.

RICE, P.A. AND BAKER, T.A. (2001). Comparative architecture of transposase and integrase complexes. *Nature Structural Biology* **8**, 302–307.

RICHARDSON, J.M., ZHANG, L., MARCOS, S. *ET AL.* (2004). Expression, purification and

preliminary crystallographic studies of a single-point mutant of the *Mos1 mariner* transposase. *Acta Crystallographica* **D60**, 962–964.

ROBERTSON, H.M. (1993). The *mariner* transposable element is widespread in insects. *Nature* **362**, 241–245.

ROBERTSON, H.M. AND MACLEOD, E.G. (1993). Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. *Insect Molecular Biology* **2**, 125–139.

ROBINSON, A.S., FRANZ, G. AND FISHER, K. (1999). Genetic sexing strains in the medfly, *Ceratitis capitata*: development, mass rearing and field application. *Trends in Entomology* **2**, 81–104.

RUBIN, G.M. AND SPRADLING, A.C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353.

SAVILLE, K.J. AND BELOTE, J.M. (1993). Identification of an essential gene, l(3)73Ai, with a dominant temperature-sensitive lethal allele, encoding a *Drosophila* proteasome subunit. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 8842–8846.

SHERMAN, A., DAWSON, A., MATHER, C. *ET AL.* (1998). Transposition of the *Drosophila* element *mariner* into the chicken germ line. *Nature Biotechnology* **16**, 1050–1053.

SIEGAL, M.L. AND HARTL, D.L. (1996). Transgene coplacement and high efficiency site-specific recombination with the cre/loxP system in *Drosophila*. *Genetics* **144**, 715–726.

SONG, J.J., SMITH, S.K., HANNON, G.J. AND JOSHUA-TOR, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434–1437.

SUNDARARAJAN, P., ATKINSON P.W. AND O'BROCHTA, D.A. (1999). Transposable element interactions in insects: crossmobilization of *hobo* and *Hermes*. *Insect Molecular Biology* **8**, 359–368.

SWEENY, T.L. AND BARR, A.R. (1978). Sex ratio distortion caused by meiotic drive in a mosquito, *Culex pipens* L. *Genetics* **88**, 427–446.

THOMAS, D.D., DONNELLY, C.A. WOOD, R.J. AND ALPHEY, L.S. (2000). Insect population control using a dominant, repressible, lethal genetic system. *Science* **287**, 2474–2476.

TOKUYASU, K.T., PEACOCK,W.J. AND HARDY, R.W. (1977). Dynamics of spermiogenesis in *Drosophila melanogaster*. VII. Effects of segregation distorter (SD) chromosomes. *Journal of Ultrastructure Research* **58**, 96–101.

TOSI, L.R.O. AND BEVERLEY, S.M. (2000). *cis* and *trans* factors affecting *Mos1 mariner* evolution and transposition *in vitro*, and its potential for functional genomics. *Nucleic Acids Research* **28**, 784–790.

WARREN, W.D., ATKINSON, P.W. AND O'BROCHTA, D.A. (1994). The *Hermes* transposable element from the house fly, *Musca domestica*, is a short inverted repeat-type element of the *hobo*, *Ac*, and *Tam3* (*hAT*) element family. *Genetical Research, Cambridge* **64**, 87–97.

WING, J.P., ZHOU, L., SCHWARTZ, L.M. AND NAMBU, J.R. (1998). Distinct cell killing properties of the *Drosophila* reaper, head involution defective, and grim genes. *Cell Death and Differentiation* **5**, 930–939.

ZHANG, L., DAWSON, A. AND FINNEGAN, D.J. (2001). DNA-binding activity and subunit interaction of the mariner transposase. *Nucleic Acids Research* **29**, 3566–3575.

ZHOU, L., MITRA, R., ATKINSON, P.W., HICKMAN, A.B., DYDA, F. AND CRAIG, N.L. (2004). Transposition of *hAT* elements links transposable elements and V(D)J recombination. *Nature* **432**, 995–1001.