Research article

A sensitive and rapid assay for homologous recombination in mosquito cells: impact of vector topology and implications for gene targeting

Paul Eggleston*[1] and Yuguang Zhao[2]

Address: [1]School of Life Sciences, Keele University, Huxley Building, Keele, Staffordshire, ST5 5BG, UK and [2]Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, OX3 9DS, UK

E-mail: Paul Eggleston* - p.eggleston@keele.ac.uk; Yuguang Zhao - yzhao@molbiol.ox.ac.uk

*Corresponding author

Abstract

Background: Recent progress in insect transgenesis has been dramatic but existing transposon-based approaches are constrained by position effects and potential instability. Gene targeting would bring a number of benefits, however progress requires a better understanding of the mechanisms involved. Much can be learned *in vitro* since extrachromosomal recombination occurs at high frequency, facilitating the study of multiple events and the impact of structural changes among the recombining molecules. We have investigated homologous recombination in mosquito cells through restoration of luciferase activity from deleted substrates. The implications of this work for the construction of insect gene targeting vectors are discussed.

Results: We show that linear targeting vectors are significantly more efficient than circular ones and that recombination is stimulated by introducing double-strand breaks into, or near, the region of homology. Single-strand annealing represents a very efficient pathway but may not be feasible for targeting unbroken chromosomes. Using circular plasmids to mimic chromosomal targets, one-sided invasion appears to be the predominant pathway for homologous recombination. Non-homologous end joining reactions also occur and may be utilised in gene targeting if double-strand breaks are first introduced into the target site.

Conclusions: We describe a rapid, sensitive assay for extrachromosomal homologous recombination in mosquito cells. Variations in substrate topology suggest that single-strand annealing and one-sided invasion represent the predominant pathways, although non-homologous end joining reactions also occur. One-sided invasion of circular chromosomal mimics by linear vectors might therefore be used *in vitro* to investigate the design and efficiency of gene targeting strategies.

Background

Recent progress in the development of transposon-mediated germline transformation in non-drosophilid insects has been dramatic. There are now four transposable element systems (*Mos1-mariner*, *Hermes*, *Minos* and *piggyBac*) that have been successfully deployed across a range of dipteran, lepidopteran and coleopteran insects [1]. This progress has served to focus attention onto potential ap-

plications of the technology and, as the number of transgenic strains continues to increase, consideration is being given to inherent problems of the approach and to potential alternatives. All transposon-based approaches to transformation are constrained by the quasi-random nature of the integration sites, which can give rise to insertional inactivation of essential genes and to position effects on transgene expression [2]. As an alternative approach, we have been investigating the potential for gene targeting in insect genomes through homologous recombination. In principle, this could facilitate precise modifications of a given genetic locus provided that the relevant region had previously been cloned [3,4]. This would open the way to gene replacement, knockout or repair, as well as the introduction of specific mutations at the target site. Such precision would greatly enhance the power of investigations into gene function and interaction.

Gene targeting has been pursued effectively in a range of lower eukaryotes, as well as in plants and vertebrates [1,5]. Gene targeting studies in insects are more limited but the basic machinery of homologous recombination has been demonstrated *in vitro* in *Drosophila*[6] and through the precise modification of extrachromosomal targets in mosquito cells [5,7]. The capacity of intact *Drosophila* to exploit homologous recombination has also been demonstrated through the repair of double-strand breaks mediated by excision of *P* transposable elements [8]. Precise breaks at either end of the transposon can undergo recombinational repair using information from the sister chromatid, homologous chromosome or exogenous plasmid template [8,9]. Although this is a targeted approach, it is dependent on the original location of the transposon. Gene targeting at sites that are not predetermined in this way has recently been demonstrated through two very different approaches. First, baculovirus vectors have been used to transform a green fluorescent protein (GFP) reporter gene into the silkmoth, *Bombyx mori*[10]. This approach may be limited to insects susceptible to baculovirus infection although other insect viruses may have similar potential as DNA vectors. Secondly, targeted modifications of the X-linked yellow and autosomal pugilist genes have been demonstrated in *D. melanogaster*. In these experiments, constructs carrying part of the target gene were first integrated into the host genome by means of a transposable element vector. Subsequently, the *FRT/ FLP* site-specific recombination system and a site-specific endonuclease (I-*Sce*I) were used to generate extrachromosomal DNA molecules with a double-strand break in the region of homology that were ideal substrates for gene targeting *in vivo*[11].

These experiments provide encouragement for further exploration of gene targeting strategies in insects. However, progress and optimisation of the approach will require a better understanding of the homologous recombination mechanisms involved and careful consideration of targeting vector design. Such experiments would be difficult in intact insects but much useful information can be gleaned from studies of recombination between DNA molecules introduced into cultured insect cells. Extrachromosomal recombination occurs at high frequency, making it possible to quickly obtain data from many independent events and it is relatively easy to study the effects of modifying the sequences undergoing recombination [12]. Studies in yeast and mammals, for example, have identified a number of key factors, including the benefits of using isogenic DNA sequences, longer lengths of homology and vector topologies where homologous ends are orientated inwards. One particularly useful strategy for demonstrating and quantifying homologous recombination is the regeneration of selectable marker genes from inactive substrates. In such experiments two plasmids, each carrying the same selectable marker (e.g. neomycin resistance) but deficient at unique sites, are transfected into cultured cells [13–15]. Cells that survive selection must have regenerated a functional copy of the marker gene through homologous recombination. Such experiments have been used in mammalian cells to optimise many of the parameters that affect gene targeting frequencies. One difficulty in carrying out such experiments with conventional selectable markers (e.g. neomycin resistance) is that the transfected DNA must remain in the cells for sufficient time to allow non-resistant cells to be killed and resistant cell clones to become established. Recently, variations of this technique have been developed where the restoration of luciferase reporter gene activity is used to monitor the efficiency of extrachromosomal homologous recombination. Measurements of reporter gene activity are not only rapid and convenient, but also give more precise quantitative data that help to reveal small differences in recombination efficiency.

We describe here our development of a sensitive and rapid reporter gene assay in which the restoration of luciferase activity from a pair of truncated substrates is used to study extrachromosomal homologous recombination in cultured mosquito cells. The substrates carried the firefly luciferase reporter gene driven by the *actin 5C* promoter from *Drosophila* with non-overlapping deletions of 561 bp (DL) and 371 bp (DR) respectively at the 5' and 3' ends of the luciferase gene. Thus, DL and DR, though individually defective, share a 728 bp region of homology providing the opportunity for restoration of an intact luciferase gene through homologous recombination (Fig. 1). This assay facilitated a detailed analysis of the mechanism and efficiency of homologous recombination, including the impact of topological variations in the targeting molecules. The implications of this work for the construction of insect gene targeting vectors are discussed.
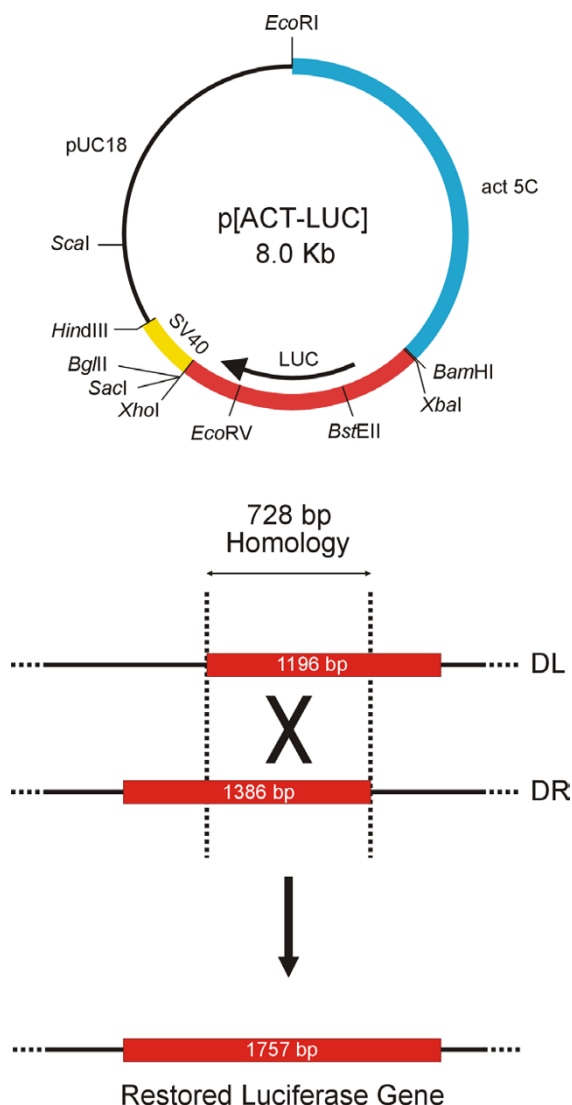
Figure 1
Design and construction of the recombination substrates. p[ACT-LUC] carries an intact transcription unit comprising the firefly (*Photinus pyralis*) luciferase coding sequence (red) driven from the *D. melanogaster actin5C* promoter (blue) with transcription terminated by the SV40 small t intron/polyA signal (yellow). The luciferase coding sequence was released from pGEM-*luc* (Promega) by digestion with *Bam*HI and *Xho*I and ligated into the same sites located between the *actin5C* promoter and SV40 termination sequence in p[ACT-SV] (unpublished data). All relevant restriction enzyme sites are indicated. The left-hand deletion substrate (DL) was generated by *Xba*I and *Bst*EII digestion to remove a 561 bp fragment at the 5' end of the luciferase coding sequence. The right-hand deletion substrate (DR) was generated by *Eco*RV and *Xho*I digestion to remove a 371 bp fragment at the 3' end of the luciferase coding sequence. The 728 bp region of homology shared by DL and DR is indicated and homologous recombination in this interval has the potential to reconstitute a functional luciferase gene.

Results

Luciferase activities derived from transfection of circular and linear substrates

Various combinations of linear and circular substrates were co-transfected into *An. gambiae* (Ag55) cells and the resulting luciferase activity recorded. Extremely high luciferase activity was seen following transfection with the positive control, a circular form of the intact luciferase gene (Fig. 2: p[ACT-LUC]^c, $6.8 \pm 1.1 \times 10^9$ cpm). Linearization of the positive control with *Hin*dIII resulted in a substantial reduction in activity compared to the circular form (Fig. 2: p[ACT-LUC]^{LH}, $3.8 \pm 1.3 \times 10^7$ cpm, $t_{(4)} = 7.53$, $0.01 > P > 0.001$), presumably due to the increased probability of degradation following the transfection of linear molecules. Although not shown in Figure 2, this reduction was even greater when the positive control was linearized with *Xho*I ($3.2 \pm 1.1 \times 10^6$ cpm), which separates the coding sequence from the polyadenylation signal and therefore interferes with normal transcript processing. Predictably, activity was essentially abolished when the positive control was linearized within the luciferase coding sequence by *Bst*EII digestion, giving an activity not significantly different from background ($4.1 \pm 1.0 \times 10^4$ cpm, $t_{(4)} = 2.21$, $P > 0.05$). In all cases, the assay background was taken to be the mean activity derived from transfections with the two deletion substrates alone ($6.3 \pm 1.4 \times 10^4$ cpm).

All transfections involving combinations of both deletion constructs gave rise to significant recoveries of luciferase activity in comparison to the assay background. Even when both deletion constructs were present as circular molecules (generally regarded as poor substrates for homologous recombination) there was a 15-fold elevation of luciferase activity compared to background (Fig. 2; DL^c + DR^c, $9.5 \pm 0.3 \times 10^5$ cpm, $t_{(4)} = 32.81$, $P < 0.001$). To investigate the impact of linearization of the substrates, double-strand breaks were introduced into different regions of DL and DR prior to their co-transfection into Ag55 cells. Linearization within the region of homology gave the best restoration of luciferase activity with an 86-fold increase over co-transfections with the circular forms (Fig. 2: DL^{LB} + DR^{LS}, $8.2 \pm 2.0 \times 10^7$ cpm, $t_{(4)} = 4.96$, $0.01 > P > 0.001$). However, luciferase activity was compromised when double-strand breaks were introduced outside the region of homology. Linearization with *Hin*dIII resulted in a 5-fold reduction in activity compared to co-transfection of circular substrates (Fig. 2: DL^{LH} + DR^{LH}, $2.1 \pm 1.1 \times 10^5$ cpm, $t_{(4)} = 7.95$, $0.01 > P > 0.001$). A similar 5-fold reduction was seen following *Sca*I linearization in the plasmid backbone (Fig. 2: DL^{LSc} + DR^{LSc}, $2.0 \pm 0.9 \times 10^5$ cpm, $t_{(4)} = 9.68$, $P < 0.001$).

In conventional gene targeting, linear vectors are directed to the chromosomal target, which is an unbroken DNA
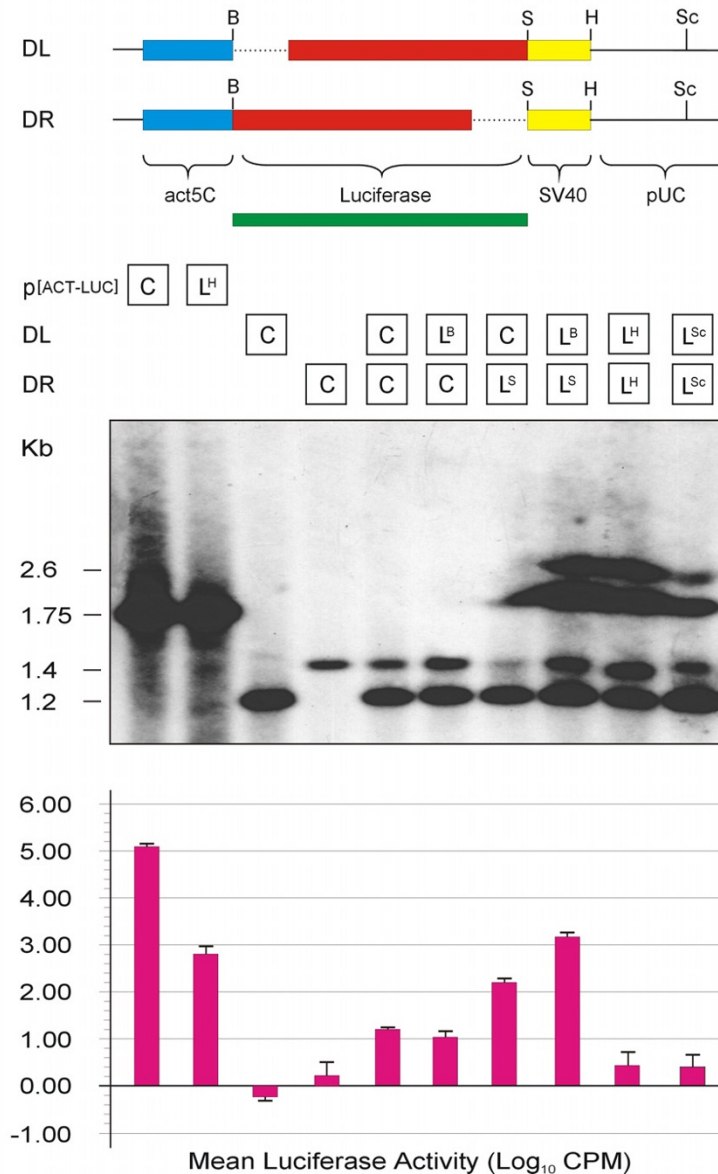
Figure 2
Functional and molecular analysis of extrachromosomal homologous recombination. The upper panel is a schematic of the left and right-hand deletion substrates (DL and DR) showing the *actin5C* promoter (blue), luciferase coding sequence (red) and SV40 termination sequence (yellow). The 5' and 3' luciferase deletions are shown as dotted lines and the pUC18 plasmid backbone as a continuous line. The intact luciferase coding sequence used as a probe for Southern analysis is shown in green. All relevant restriction sites are indicated (B, *Bam*HI; S, *Sac*I; H, *Hin*dIII and Sc, *Sca*I). The central panel shows the various combinations of plasmids transfected into mosquito cells together with a Southern analysis of the reaction products in *An. gambiae* (Ag55) cells. Transfections involved either circular (C) or linear (L) forms of the parent plasmid (p[ACT-LUC]) and deletion substrates (DL; DR), as shown, with superscripts indicating linearization by the relevant restriction enzyme (L[B], *Bam*HI; L[S], *Sac*I; L[H], *Hin*dIII; L[Sc], *Sca*I). For Southern analysis, total cellular DNA was isolated from *An. gambiae* (Ag55) cells 48 hours post-transfection, digested with both *Bam*HI and *Bgl*II, resolved on 1.5% agarose and blotted onto nitrocellulose. The membrane was probed with a ^{32}P labelled luciferase fragment, washed at high stringency (1 × SSC; 0.1 % SDS; 65°C) and exposed overnight at -70°C to X-ray film against an intensifying screen. The size of relevant signals was determined by comparison to standard markers (MBI Kilobase Ladder) and is shown in kilobase pairs (Kb). The lower panel shows the mean luciferase activities (log$_{10}$ counts per minute) recovered from the various transfections into *An. gambiae* (Ag55) cells. Values were plotted following subtraction of the assay background and the error bars represent standard deviations. A logarithmic scale was employed so that all values could be represented on the same figure.

strand. This was mimicked in our experiments by transfections with one circular and one linear substrate. Using DL as the chromosomal mimic and linearized DR, luciferase activity was 9-fold higher than in co-transfections with two circular substrates (Fig. 2: $DL^C + DR^{LS}$, $8.9 \pm 1.8 \times 10^6$ cpm, $t_{(4)} = 5.41$, $0.01 > P > 0.001$) but 9-fold lower than that when both substrates were linearized within the region of homology ($t_{(4)} = 4.46$, $0.02 > P > 0.01$). In the opposite configuration (circular DR and linearized DL), luciferase activity was not significantly different to that in co-transfections of circular substrates (Fig. 2: $DL^{LB} + DR^C$, $6.7 \pm 1.3 \times 10^5$ cpm, $t_{(4)} = 2.57$, $P > 0.05$).

Comparison of extrachromosomal homologous recombination patterns across species

Figure 3 shows that the luciferase activity patterns described for Ag55 cells are reproducible across a range of species from the genera *Anopheles, Aedes* and *Culex.* No attempt is made here at quantitative comparisons but the underlying patterns of activity are clearly very similar. In all cell lines the highest luciferase activity from co-transfections of the deletion constructs is seen following linearization within the region of homology ($DL^{LB} + DR^{LS}$). Similarly, for co-transfections of one circular and one linear substrate, linearization of the right-hand substrate ($DL^C + DR^{LS}$) always gives a higher activity than linearization of the left-hand substrate ($DL^{LB} + DR^C$). In addition, linearization outside the region of homology ($DL^{LH} + DR^{LH}$) or within the plasmid backbone ($DL^{LSc} + DR^{LSc}$) returns the poorest activity in all cell lines. Clearly, the different substrate topologies interact in very similar ways in each of the cell lines tested. Overall levels of luciferase activity do vary between lines, with some of the highest activities recorded in *Ae. aegypti* (Mos20) cells. These differences can not be a simple reflection of transfection efficiency or cellular growth and division rates since the data normalisation procedures were designed to account for such variables. A more likely interpretation might be that the cell lines differ in their promoter recognition abilities and/or suites of transcription factors that they are able to express.

Southern analysis of the molecular structure of transfected substrates

In the light of this evidence for functional restoration of luciferase activity we investigated the molecular structure of the luciferase constructs by Southern transfer. DNA was isolated from Ag55 cells that had been transfected with various combinations of luciferase substrate. The DNA was digested with both *Bam*HI and *Bgl*II, fractionated by electrophoresis, transferred to a membrane and hybridized with the intact luciferase gene. The resulting autoradiograph clearly identifies the predicted 1.75 Kb band in cells transfected with the intact luciferase construct (Fig. 2: p[ACT-LUC]^C; p[ACT-LUC]^{LH}). It also reveals the antici-
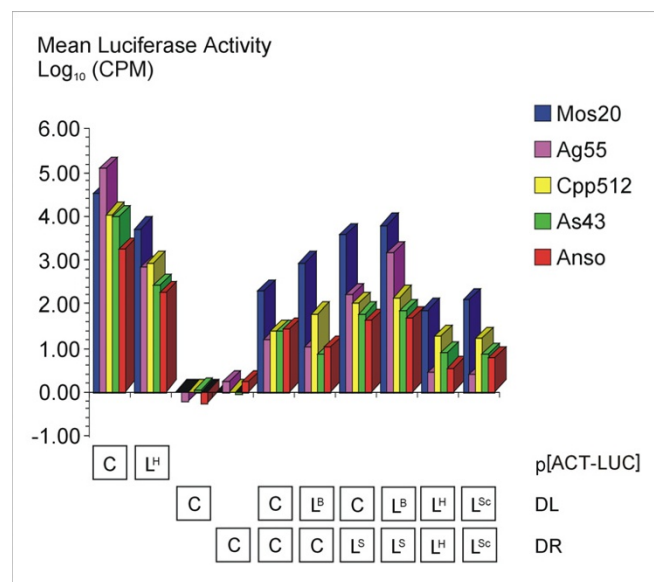
Figure 3
Comparative analysis of recovered luciferase activities across cell lines. Mean luciferase activities ($\log_{10}$ cpm) recovered from the various transfections in cell lines derived from *Ae. aegypti* (Mos20; blue), *An. gambiae* (Ag55; magenta), *Culex pipiens pallens* (Cpp512, yellow), *An. stephensi* (As43, green) and *An. sinensis* ovary (Anso, red). Standard deviations are not shown for reasons of clarity but were comparable to those shown in Fig. 2 for *An. gambiae* (Ag55) cells. Transfections involved either circular (C) or linear (L) forms of the parent plasmid (p[ACT-LUC]) and deletion substrates (DL; DR), as shown, with superscripts indicating linearization by relevant restriction enzyme (L^B, *Bam*HI; L^S, *Sac*I; L^H, *Hind*III; L^{Sc}, *Sca*I). A logarithmic scale was employed so that all values could be represented on the same figure.

pated bands at 1.2 Kb and 1.4 Kb, respectively, for the left and right-hand deletion substrates (Fig. 2) whether transfected singly or in combination. However, only some of the cells co-transfected with DL and DR showed the 1.75 Kb band that could indicate restoration of an intact luciferase gene (Fig. 2: $DL^C + DR^{LS}$; $DL^{LB} + DR^{LS}$; $DL^{LH} + DR^{LH}$; $DL^{LSc} + DR^{LSc}$). This suggests that restored luciferase sequences are present at low copy numbers, reflecting both the efficiency of homologous recombination and the relatively low numbers of cells that take up the introduced DNA. The results are consistent with the identification of the strongest signal in those cells showing the highest restoration of luciferase activity (Fig. 2: $DL^{LB} + DR^{LS}$). However, care needs to be taken in drawing any quantitative inferences from the Southern blot data since it is not possible to relate band intensities directly to functional luciferase sequences. Fragments of around 1.75 Kb may also be generated by non-homologous rearrangements between the input plasmids, particularly when introduced in linear form.

A further noticeable feature is that all co-transfections involving two linearized substrates show additional signals of approximately 2.6 Kb and of varying itensity (Fig. 2: DL^{LB} + DR^{LS}; DL^{LH} + DR^{LH}; DL^{LSc} + DR^{LSc}). Reference to the structure of the substrates (Fig. 1) and to the restriction enzymes used shows that these additional signals could be the result of end-joining reactions between linear molecules (see discussion). For example, the co-transfection (DL^{LB} + DR^{LS}) would allow an end-joining reaction to generate a partial duplication of the luciferase gene resulting in a 2.6 Kb fragment in DNA digested with *Bam*HI and *Bgl*II. These signals are not seen in co-transfections where either one or both substrates are circular since end-joining reactions are precluded in the absence of two linear substrates.

Discussion

Taken as a whole, the experiments reported here suggest that homologous recombination may be an effective mechanism in mosquito cells and encourage the further exploration of gene targeting strategies for the generation of transgenic mosquitoes. The use of different substrate topologies allowed us to investigate the process (or processes) through which homologous recombination occurs and revealed marked differences in the efficiency with which luciferase activity was restored. The data indicate that linear targeting vectors are better substrates than circular ones and that double-strand breaks can stimulate homologous recombination frequencies if they occur near, and preferably within, the region of homology. In this respect, the observed results are consistent with similar findings in mammalian cells, where double-strand breaks are thought to be effective because the broken strands can act as recipients in non-reciprocal exchanges and where their effects appear to be cumulative such that breaks in both substrates are much more efficient than breaks in only one [12,14,8,9].

The way in which extrachromosomal substrates recombine to restore luciferase activity in mosquito cells can tell us much about the cellular processes involved. Circular molecules are generally held to represent poor substrates for homologous recombination and this is clear from the results presented here. When one or both substrates are linearized, luciferase activity can be restored by one of three mutually exclusive pathways for double-strand break repair [16,17]. These are recombinational repair, single-strand annealing and non-homologous end-joining. The requirements for each pathway and, indeed, the choice of pathway, are beginning to be unravelled in mammalian and yeast systems. For example, in mouse ES cells the Ku70 and Ku80 proteins are believed to shift repair dynamics in favour of non-homologous end-joining whereas in yeast, the Rad52 and Rad51 proteins appear to shift towards recombinational repair [18]. Little is known

of the presence or action of potential homologues in insects but, from the experiments described here, it would appear that more than one recombination pathway might be involved and there is evidence that particular pathways are favoured by specific substrate topologies. There is also evidence that the choice of repair pathway may be influenced by the stage of the cell cycle at the time of repair, with a bias towards non-homologous end-joining during G_1 – early S phase and towards recombinational repair during late S phase – G_2[19].

During recombinational repair [20], initiation of homologous recombination involves a double-strand break that is enlarged to a gap by exonucleolytic degradation of both strands. The gap is then repaired by copying corresponding sequences from the homologous partner, with both sides of the gap invading the donor duplex leading to an intermediate structure with two Holliday junctions and heteroduplex DNA flanking the gap. Resolution of this structure gives equal numbers of crossover and non-crossover products and, since there is no loss of sequences, the process is described as conservative. Recent evidence has challenged the relative importance of this pathway, in particular the observation that both spontaneous and induced double-strand breaks are typically processed to long 3' single-stranded ends, rather than gaps [21]. Our data suggest that recombinational repair is not the predominant pathway in mosquito cells since, according to this model, co-transfections involving one linearized and one circular molecule (DL^C + DR^{LS} and DL^{LB} + DR^C respectively) should have restored comparable luciferase activities.

The observation that restored luciferase activity depends on the choice of substrate used as chromosomal mimic could be interpreted by a modification of recombinational repair known as one-sided invasion [22]. In this model, only one side of a double-strand break in the recipient invades the unbroken donor and primes DNA synthesis in the homologous region to generate a functional recombinant. During this process synthesis can extend beyond the region of homology and into the flanking DNA. In our experiments, linearization of DR with *Sac*I releases a free 3' end immediately downstream of the homologous region. One-sided invasion of DL, with synthesis extending towards the end of the luciferase gene, would recover an intact coding sequence and restore activity. In the opposite configuration, linearization of DL with *Bam*HI would release a free 3' end immediately upstream of the region of homology. One-sided invasion of DR with synthesis extending towards the start of the luciferase gene could also recover a functional coding sequence, although apparently at lower efficiency. This model could therefore provide an adequate explanation of the results obtained, provided only that repair efficiency is greater in one direction. In

this situation, it is possible that the strong constitutive actin 5c promoter biases the repair when initiated away from the promoter ($DL^C + DR^{LS}$).

Single-strand annealing [23] has been described as the most efficient pathway for extrachromosomal homologous recombination in mammalian cells [24] and plant cells [25–27]. It is also known to be involved in double-strand break repair in yeast, where it is dependent on direct repeats at either side of the break [28] but the requirements in higher eukaryotes are not known. Recent data suggest that single-strand annealing is the predominant pathway for double-strand break repair in mouse oocytes but that this declines by the embryonic stage [17]. The essential feature of the model is that DNA ends at a double-strand break are rendered single-stranded by a 5'-3' exonuclease or as a result of unwinding [29]. When homologous sequences are present, this process ends with complementary single-strands that are capable of re-annealing. Repair synthesis and ligation then complete the formation of the non-reciprocal homologous junction. Our data show that single-strand annealing can also be an effective pathway for double-strand break repair in mosquito cells, given appropriate substrate topology. The model is favoured by linearization of both substrates within the region of homology and this configuration ($DL^{LB} + DR^{LS}$) does provide the greatest restoration of luciferase activity. Predictably, linearization outside of the region of homology is less effective since it does not provide homologous single-strands for the annealing reaction.

Non-homologous end-joining is a pathway for the repair of double-strand breaks in which the broken DNA ends are simply re-ligated, without the need for a template molecule or region of homology. This phenomenon has been highlighted recently [30] where it was described as the predominant mechanism in zygotes and early embryos of the zebrafish, *Danio rerio* as well as in *D. melanogaster*. It is also thought that this may be a common pathway in mammalian cells, although it is error prone and may introduce small deletions at the joining site [17]. In the context of the experiments described here, end-joining would not restore a functional luciferase gene but the resulting duplication of sequence would be detectable as higher molecular weight fragments on Southern blotting.

Conclusions
The results presented here reveal the range and relative efficiencies of alternative homologous recombination pathways in mosquito cells. We show that linear targeting vectors are better substrates than circular ones and that double-strand breaks stimulate homologous recombination when introduced into, or near, the region of homology. Evidence is provided that single-strand annealing of two linear substrates represents a very efficient pathway

for homologous recombination, providing the best overall restoration of luciferase activity in these experiments. However, this utility may be diminished in the context of gene targeting, where it is not generally feasible to introduce double-strand breaks into the chromosomal target. That is to say, the target molecules are, in most cases, unbroken strands of DNA. Perhaps the best model of such a situation is to use a circular plasmid to mimic the chromosomal target site and a linear molecule as an analogue of the targeting vector [31]. Under these circumstances, we show that that one-sided invasion (a modification of the recombinational repair model) provides the best explanation for the differing levels of luciferase activity recovered from alternative chromosomal mimics. We also provide evidence for the occurrence of non-homologous end-joining reactions in mosquito cells, resulting in the creation of non-functional partial duplications of the luciferase gene that are detectable by Southern blot in all co-transfections of linear substrates. It should be noted, however, that in the context of targeted genome manipulation, end-joining reactions could only be utilised if a double-strand break was first introduced into the target genome, for example by excision of a resident transposon.

Materials and Methods
Plasmid substrates
pACT-LUC carries the firefly luciferase coding sequence under control of the *actin5C* promoter from *D. melanogaster* with transcription terminated by the SV40 small t intron and polyadenylation signal (Fig. 1). The left-hand deletion substrate (DL) was constructed by digesting pACT-LUC with *Xba*I and *Bst*EII to remove a 561 bp fragment at the 5' end of the luciferase coding sequence. The resulting 5' overhangs were filled in with the Klenow fragment of DNA Polymerase I plus dNTP's and the plasmid reconstituted by blunt-end ligation. The right-hand deletion substrate (DR) was constructed in a similar way by digesting pACT-LUC with *Eco*RV and *Xho*I to release a 371 bp fragment at the 3' end of the luciferase coding sequence, followed by blunt-end ligation (Fig. 1). All plasmids were propagated by transformation into *E. coli* XL-1 Blue (Stratagene) and plasmid DNA was purified by caesium chloride buoyant density centrifugation. To generate linear molecules for cell transfections, plasmids were digested with the appropriate restriction enzymes and digestion confirmed by gel electrophoresis. Digestion products were purified by phenol/chloroform extraction and ethanol precipitation. All plasmids used for cell transfections (linear and supercoiled) were resuspended in TE buffer and sterilized by adding a few drops of chloroform. DNA concentrations were determined spectrophotometrically and the structure and conformation of all plasmids was verified by agarose gel electrophoresis prior to use.

Cell lines and transfection protocol

Mosquito cell lines derived from *Anopheles gambiae* (Ag55, [32]), *Anopheles stephensi* (As43) and *Aedes aegypti* (Mos20, [33]) were provided originally by Dr C. Leake, London School of Hygiene and Tropical Medicine and subsequently maintained in our laboratory. Cell lines derived from *Anopheles sinensis* ovaries (Anso, [34]) and from neonate larvae of *Culex pipiens pallens* (Cpp512, [35]) were provided by Prof. M. Y. Lan, Jiangsu Institute of Parasitic Diseases, China. All cell lines were maintained following routine procedures at 27°C with 5% CO_2 in Medium 199 supplemented with foetal bovine serum (10%), lactalbumin hydrolysate (4 μg/ml), yeastolate (1 μg/ml), L-glutamine (2 mM), Penicillin (100 U/ml) and Streptomycin (100 μg/ml). Transfection was mediated by the cationic lipid DOTAP (N-[l-(2,3-Dioleoyloxy)propyl]-N,N,N-trimethylammonium methylsulphate) as described previously [36]. Briefly, 5×10^5 cells were plated in 1 ml medium in individual wells of 24 well culture plates, grown to 60% confluence and washed twice in Hanks Buffered Saline. Cells were incubated with the transfection mixture (1.25 μg DNA; 7.5 μl DOTAP in 500 μl serum-free medium) for 8 hours and then normal medium restored for 48 hours prior to luciferase assays. The primary transfection efficiency for each of the cell lines was determined by introducing a reporter plasmid expressing green fluorescent protein (pHermes A5CEGFP) as described previously [36].

Luciferase assays

Assays were conducted according to the manufacturer's guidelines (Promega). Briefly, adherent cells were rinsed twice in Hanks Buffered Saline, covered with 100 μl lysis buffer (1×) and incubated at room temperature for 10 minutes. Cell lysates were aspirated into microfuge tubes and spun briefly to pellet large debris. 10 μl cell lysate was mixed with 100 μl luciferase assay reagent (Promega) at room temperature and the reaction transferred immediately to a scintillation counter (LKB RackBeta) where light emission was measured over a period of 10 seconds. Three replicate cell samples were processed for each treatment and the average counts per minute (CPM) recorded. In addition to controlling cell confluence at the time of transfection, replicate cell lysates were assayed for protein concentration (BioRad Protein Assay Kit) to account for differences in cell density and division rates. Luciferase activities were normalised with respect to primary transfection efficiency and cellular protein concentration and conventional parametric tests (Students' t-tests) were used to determine the significance of differences between the means.

References

1. Handler AM: **A current perspective on insect gene transformation.** *Insect Biochem Mol* 2001, **31**:111-128
2. Wilson C, Bellen HJ, Gehring WJ: **Position effects on eukaryotic gene expression.** *Ann Rev Cell Biol* 1990, **6**:679-714
3. Capecchi MR: **Altering the genome by homologous recombination.** *Science* 1989, **16**:1288-1292
4. Morrow B, Kucherlapati R: **Gene targeting in mammalian cells by homologous recombination.** *Curr Opin Biotechnol* 1993, **4**:577-582
5. Eggleston P, Zhao Y: **Targeted transformation of the insect genome.** *In: Insect transgenesis: methods and applications (Edited by Handler AM, James AA) Boca Raton, Florida, CRC Press* 200029-52
6. Cherbas L, Cherbas P: **"Parahomologous" gene targeting in *Drosophila* cells: an efficient, homology-dependent pathway of illegitimate recombination near a target site.** *Genetics* 1997, **145**:349-358
7. Eggleston P, Zhao Y: **Gene targeting in mosquito cells: a demonstration of 'knockout' technology in extrachromosomal gene arrays.** *BMC Genetics* 2001, **2**:11
8. Lankenau D-H, Corces VG, Engels WR: **Comparison of targeted-gene replacement frequencies in *Drosophila melanogaster* at the *forked* and *white* loci.** *Mol Cell Biol* 1996, **16**:3535-3544
9. Kooistra R, Pastink A, Zonneveld JBM, Lohman PHM, Eeken JCJ: **The *Drosophila melanogaster* DmRAD54 gene plays a crucial role in double-strand break repair after *P*-element excision and acts synergistically with *Ku70* in the repair of X-ray damage.** *Mol Cell Biol* 1999, **19**:6269-6275
10. Yamao M, Katayama N, Nakazawa H, Yamakawa M, Hayashi Y, Hara S, Kamei K, Mori H: **Gene targeting in the silkworm by use of a baculovirus.** *Genes Dev* 1999, **13**:511-516
11. Rong YS, Golic KG: **A targeted gene knockout in *Drosophila*.** *Genetics* 2001, **157**:1307-1312
12. Bollag RJ, Waldman AS, Liskay RM: **Homologous recombination in mammalian cells.** *Ann Rev Genet* 1989, **23**:199-225
13. Small J, Scangos G: **Recombination during gene transfer into mouse cells can restore the function of deleted genes.** *Science* 1983, **219**:174-176
14. Song KY, Chekuri L, Rauth S, Ehrlich S, Kucherlapati R: **Effect of double-strand breaks on homologous recombination in mammalian cells and extracts.** *Mol Cell Biol* 1985, **5**:3331-3336
15. Baur M, Potrykus I, Paszkowski J: **Intermolecular homologous recombination in plants.** *Mol Cell Biol* 1990, **10**:492-500
16. Pastink A, Lohman PH: **Repair and consequences of double-strand breaks in DNA.** *Mut Res* 1999, **428**:141-156
17. Fiorenza MT, Bevilacqua A, Bevilacqua S, Mangia F: **Growing dictyate oocytes, but not early preimplantation embryos, of the mouse display high levels of DNA homologous recombination by single-strand annealing and lack DNA nonhomologous end joining.** *Dev Biol* 2001, **233**:214-224
18. Hiom K: **DNA repair: Rad52 – the means to an end.** *Curr Biol* 1999, **9**:R446-8
19. Takata M, Sasaki MS, Sonoda E, Morrison C: **Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells.** *EMBO J* 1998, **17**:5497-5508
20. Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW: **The double-strand-break repair model for recombination.** *Cell* 1983, **33**:25-35
21. Sugawara N, Haber JE: **Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation.** *Mol Cell Biol* 1992, **12**:563-575
22. Belmaaza A, Chartrand P: **One-sided invasion events in homologous recombination at double-strand breaks.** *Mut Res* 1994, **314**:199-208
23. Carroll D: **Homologous genetic recombination in *Xenopus*: mechanism and implications for gene manipulation.** *Prog Nucl Acid Res Mol Biol* 1996, **54**:101-125
24. Lin FL, Sperle K, Sternberg N: **Model for homologous recombination during transfer of DNA into mouse L cells: role for

DNA ends in the recombination process. *Mol Cell Biol* 1984, **4**:1020-1034

25. de Groot MJ, Offringa R, Does MP, Hooykaas PJ, van den Elzen PJ: **Mechanisms of intermolecular homologous recombination in plants as studied with single- and double-stranded DNA molecules.** *Nucl Acids Res* 1992, **20**:2785-2794

26. Bilang R, Peterhans A, Bogucki A, Paszkowski J: **Single-stranded DNA as a recombination substrate in plants as assessed by stable and transient recombination assays.** *Mol Cell Biol* 1992, **12**:329-336

27. Swoboda P, Gal S, Hohn B, Puchta H: **Intrachromosomal homologous recombination in whole plants.** *EMBO J* 1994, **13**:484-489

28. Haber JE: *In vivo* **biochemistry: physical monitoring of recombination induced by site-specific endonucleases.** *Bioessays* 1995, **17**:609-620

29. Wake CT, Vernaleone F, Wilson JH: **Topological requirements for homologous recombination among DNA molecules transfected into mammalian cells.** *Mol Cell Biol* 1985, **5**:2080-2089

30. Hagmann M, Bruggmann R, Xue L, Georgiev O, Schaffner W, Rungger D, Spaniol P, Gerster T: **Homologous recombination and DNA-end joining reactions in zygotes and early embryos of zebrafish (*Danio rerio*) and *Drosophila melanogaster.*** *Biol Chem* 1998, **379**:673-681

31. Segal DJ, Carroll D: **Endonuclease-induced, targeted homologous extrachromosomal recombination in *Xenopus* oocytes.** *Proc Natl Acad Sci USA* 1995, **92**:806-810

32. Marhoul Z, Pudney M: **A mosquito cell line (Mos 55) from *Anopheles gambiae* larva.** *Trans R Soc Trop Med Hyg* 1972, **66**:183-184

33. Varma MG, Pudney M: **The growth and serial passage of cell lines from *Aedes aegypti* (L.) larvae in different media.** *J Med Entomol* 1969, **6**:432-429

34. Lan MY, Zhao Y, Shen B, Zhou F, Fan H: **Establishment of a mosquito cell line from ovarian tissue of *Anopheles sinensis* and its characterisation.** *Chin J Parasitic Dis Cont* 1993, **6**:38-39

35. Zhao Y, Lan MG: **Establishment of a mosquito cell line of *Culex pipiens pallens* and its characteristics.** *Chin J Schist Cont* 1992, **4**:281-284

36. Zhao YG, Eggleston P: **Comparative analysis of promoters for transient gene expression in cultured mosquito cells.** *Insect Mol Biol* 1999, **8**:31-38