Xiaolin Bi
is a post doctoral research fellow at the Laboratory of Molecular Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.

Yikang S. Rong
is the principal investigator at the Laboratory of Molecular Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA.

Keywords: *homologous recombination,* Drosophila *gene targeting, FLP, I-SceI*

Genome manipulation by homologous recombination in *Drosophila*

Xiaolin Bi and Yikang S. Rong

Date received (in revised form): 9th May 2003

Abstract

By utilising a cell's recombinational machinery, researchers in many different model organisms have been able to perform gene targeting experiments in which specific sequence alterations are introduced into virtually any endogenous gene. Not only can functional knock-outs be generated by gene targeting, interesting alleles with mutations encoding specific amino acid replacements can also be made. A practical gene targeting method has only recently become available for *Drosophila*. This article reviews the *Drosophila* gene targeting method, with emphases placed on different approaches that are being used to generate different mutations.

OVERVIEW OF THE DROSOPHILA GENE TARGETING METHOD

The goal of gene targeting is to replace the chromosomal target gene, by homologous recombination, with an engineered copy on an exogenous piece of DNA molecule called 'the donor'. Besides having a modified copy of the target as sequence homology, this donor also carries a genetic marker for the identification of targeting events, as well as a double-strand break to induce recombination. In traditional gene targeting experiments in cell culture, the donor is introduced into the cell directly. For *Drosophila*, direct donor introduction by embryo injection has not led to the recovery of gene targeting events. Rong and Golic, however, successfully modified the *yellow* locus with a novel '*in vivo* donor generation' approach.[1] In this approach, the donor for gene targeting is first randomly inserted into the genome by standard P-element-mediated transformation. It is later released as a linear DNA molecule inside the cells of an intact animal. This is accomplished by the actions of two yeast enzymes expressed in *Drosophila* cells: FLP site-specific recombinase and I-SceI rare-cutting endonuclease.[2,3]

As well as carrying the common components for a donor molecule (ie targeting homology, a double-strand break and a marker, which could also serve as the marker for the initial transformation step) the *Drosophila* donor has two additional features to facilitate its release *in vivo* (Figure 1). It carries two targets — for the FLP enzyme, FLP recombination targets (FRTs) — in direct orientation, flanking the entire donor molecule. In the presence of FLP, the donor will be excised from the chromosome as a circle. The excised donor also carries a recognition site for the I-SceI enzyme. In the presence of I-SceI, the donor will be linearised at the cut site. This linear, 'free-moving', piece of DNA can participate in homologous recombination with the endogenous target locus.

The placement of the I-SceI cut site and the marker gene can be pre-selected during donor construction so that the actual targeting reaction can occur via either one of two configurations: ends-in or ends-out (Figure 1). The result of an ends-in targeting reaction is the integration of the entire donor molecule into the target, creating a tandem duplication of the target gene. By contrast, the target

Yikang S. Rong
Laboratory of Molecular Cell Biology,
National Cancer Institute,
National Institutes of Health,
Bethesda, Maryland,
20895, USA

Tel: +1 301 451 8335
Fax: +1 301 435 3697
E-mail: rongy@mail.nih.gov

locus is replaced by the donor after ends-out targeting.

In *Drosophila* gene targeting experiments, transformant lines carrying the donor P-element are first established. These are then crossed to flies with two other P-elements: a heat-inducible FLP transgene and a heat-inducible I-SceI transgene. The progeny of this cross are heat shocked early in development. Those with all three P-elements are recovered and mated to recover targeting events in the germline. Their progeny are scored for the presence of the marker gene, which indicates either a targeted event or a randomly inserted event. The two classes can be distinguished by mapping and molecular analyses.

FACTORS THAT AFFECT GENE TARGETING EFFICIENCY

After the initial success by Rong and Golic,[1] the current gene targeting method led to the modification of more than 20 loci in *Drosophila*.[1,4–8] The efficiency of gene targeting in *Drosophila* is affected by several factors that are common to traditional gene targeting practice as well. First, the presence of a double-strand break in the donor is necessary for gene targeting in *Drosophila* (unpublished results). Secondly, increasing the amount of sequence homology between the donor and the target generally improves targeting efficiency. The range of total homology that has led to successful targeting is between 2 kb and 9 kb, with >4 kb in most cases. Interestingly, unlike gene targeting in mouse embryonic stem cells, it was not essential in donor construction to use DNA which was isogenic to the targeting strain.[7] Thirdly, targeting efficiency varies from locus to locus, with no apparent trend. Lastly, targeted events generally outnumber non-targeted events, making *Drosophila* germ cells more similar to yeast cells than to mammalian cells.

The novel approach of '*in vivo* donor generation' brings about two unique factors that affect targeting efficiency in *Drosophila*. First, female germ cells are much more proficient in gene targeting than their male counterparts. The underlying mechanism for this phenomenon is not known. It is interesting (but also not understood) that different donor P-element insertions, from which the linear donor is to be liberated by FLP and I-SceI (Figure 1), can lead to vastly different targeting frequencies.[5] It is now common practice to use several independent donor transformant lines to facilitate the recovery of targeting events.

SIMPLE ENDS-OUT TARGETING VERSUS VERSATILE ENDS-IN TARGETING

Ends-in targeting was first demonstrated in *Drosophila* using the '*in vivo* donor generation' method.[1] Ends-out targeting, based on the same method, was recently accomplished.[9] The overall frequency for ends-out targeting to the *yellow* locus was similar to that for ends-in targeting to the same locus. This much welcomed invention, together with the original ends-in method, now offer *Drosophila* researchers a full range of targeting approaches such as have been available to such as their yeast and mouse colleagues for many years.

If one's goal is to generate a simple null allele of a gene, the ends-out approach is a straightforward method for doing this, since the marker gene can be inserted anywhere in the coding region to disrupt the target gene. Moreover, as the targeting homology is to be split into two pieces by the marker, one could construct a donor in which the entire target gene is replaced by the marker. By using only the flanking genomic regions as targeting homology, one can recover a complete deletion of the gene. As shown in mouse embryonic stem cells, this scheme can be used to make larger deletions in which a cluster of genes are deleted at once.[10] Another advantage of the ends-out approach is that the mutant allele will be genetically marked, which facilitates

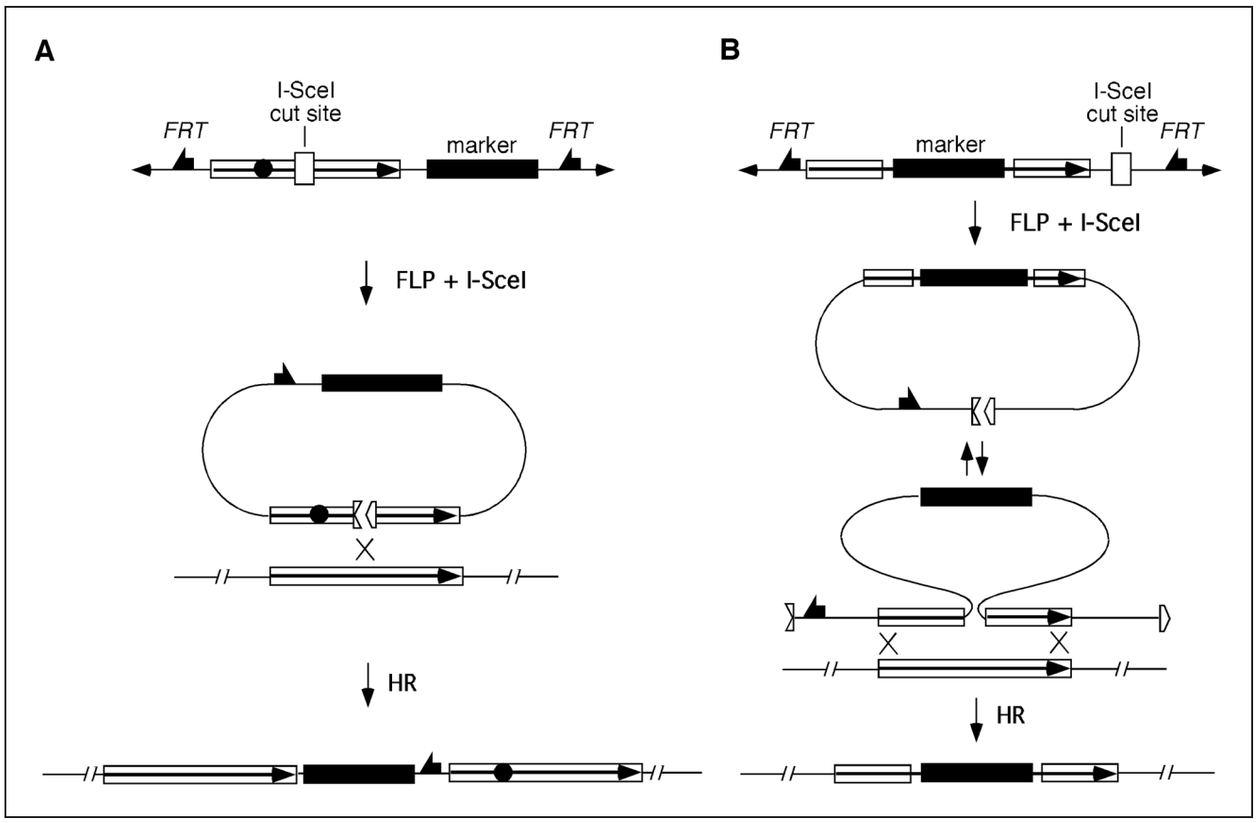

Figure 1: Two configurations for *Drosophila* gene targeting. At the top are the donor P-elements with the inverted repeats indicated as arrowheads. Open rectangular boxes represent targeting homology, with the direction of transcription indicated by an arrow inside the box. A: ends-in configuration. At the top, the targeting homology carries a point mutation indicated as a filled circle. In the middle, FLP and I-SceI generate a linear donor from the chromosomal P-element. This donor integrates into the target by a homologous recombination (HR) event with a single exchange ('X'). At the bottom is the target duplication; the right-hand copy of the target gene carries the point mutation. B: ends-out configuration. In the middle, the donor generated by FLP and I-SceI can change to the 'Ω' form to participate in ends-out targeting. A double exchange replaces the target with the donor, in which the target gene is disrupted by the marker (bottom)

future analyses. A disadvantage of this approach is that the presence of the marker could interfere with mutant analysis in some cases. This can be solved by engineering a marker in the donor construct that is flanked by direct *lox*P sites for the Cre site-specific recombinase. The marker can then be excised and eliminated in the presence of the Cre enzyme.

One-step generation of null alleles has also been accomplished by ends-in targeting in *Drosophila*.[5,8] The fact that one has to knock out both copies of the target gene imposes some limitations on this approach, however. For example, small genes would be harder to knock out. The recent development of the 'two-step allelic replacement' scheme has greatly improved the versatility of the ends-in approach.[5] The first step of this scheme is the regular ends-in targeting reaction as shown in Figure 1. After the recovery of the target duplication, the animals are crossed with ones expressing another site-specific endonuclease, I-CreI, which will inflict a double-strand break between the target gene copies at its cut site, previously imbedded in the donor construct. This break induces recombination between the target copies, leading to the reduction of the duplication to a single copy on the chromosome (Figure 2). These reduction

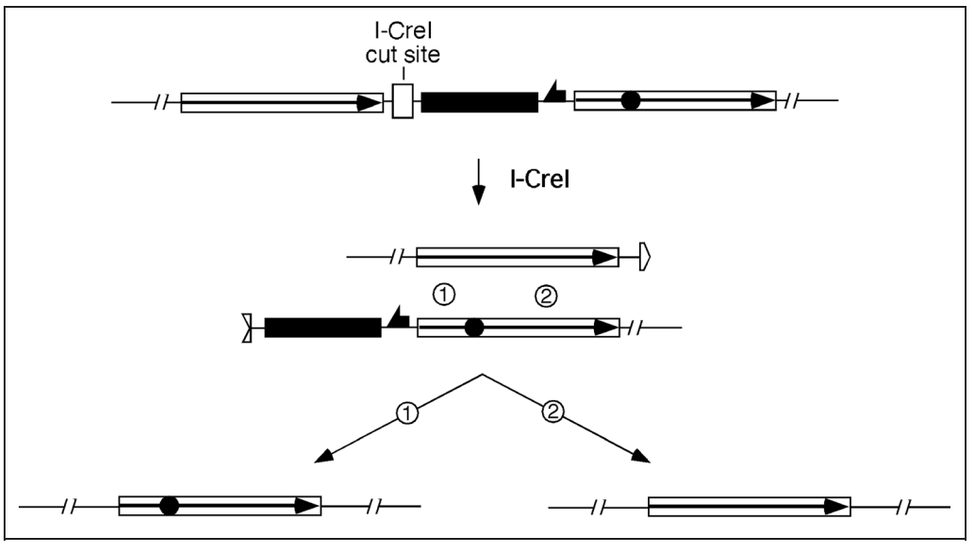

Figure 2: The reduction step of the 'two-step allelic replacement' scheme. At the top is the product of an ends-in targeting event, with a point mutation introduced to the right-hand copy of the target duplication. To induce target copy reduction, the recognition sequence for the I-CreI endonuclease is placed in the donor P-element and introduced between the duplicated copies as a result of donor integration. I-CreI generates a double-strand break at its cut site (middle). The target copies recombine to give rise to a single copy in the chromosome (bottom). If an exchange occurs in region 1 (from the start of the targeting homology to the point mutation [filled circle]), the recombination will give rise to a target gene with the point mutation. If an exchange occurs between the point mutation and the 3′ end of the targeting homology (region 2), a normal copy will be produced from the reduction

events have been recovered based on loss of the marker gene. Some of the recombination events will retain the engineered mutation in the single remaining gene. These events can be easily identified by molecular analyses.

The two-step allelic replacement method has been successful in the introduction of missense mutations that are at least 20 bp in size. This makes it suitable for situations in which small changes or point mutations are desired. One can target mutations that encode specific amino acid replacements, which would lead to interesting alleles, including conditional ones. One can mutate a specific splice site so that a particular isoform of the protein will not be produced, without interfering with the production of the other forms. Small epitope tags can be targeted so that the endogenous protein can be tagged with no loss of its function. In addition, for intron-situated genes, sequence modification by the two-step allelic replacement method would ensure that the 'other' gene is not affected. These are just a few examples in which ends-in targeting would be especially powerful. Moreover, since no exogenous sequences are left at the target locus, marker interference will not be a concern.

A bonus for performing ends-in targeting is the recovery of what were called 'class III' targeting events.[1] These were events in which deletion or insertion of sequences had been made to the target locus in the vicinity of the I-SceI cut site. The cause of class III events is not well understood. It has, however, created a series of unexpected, but useful mutations in several studies.[5–8]

CONCLUSION AND FUTURE DIRECTIONS

The implementation of two-step allelic replacement[5] and ends-out targeting[9] have provided *Drosophila* biologists with a

full range of tools for targeted mutagenesis. As more loci are targeted by different laboratories, one can expect improvement and modification of the method, making it more powerful for studies of gene functions in *Drosophila*.

Acknowledgment

We thank Dr Min Gong from our laboratory for comments on the review. Research in our laboratory is supported by the National Cancer Institute, USA. We thank researchers from various laboratories, especially the Golic laboratory at the University of Utah, for communicating unpublished results on their gene targeting experiments in *Drosophila*.

References

1. Rong, Y. S. and Golic, K. G. (2000), 'Gene targeting by homologous recombination in *Drosophila*', *Science*, Vol. 288, pp. 2013–2018.
2. Sadowski, P. D. (1995), 'The Flp recombinase of the 2-microns plasmid of *Saccharomyces cerevisiae*', *Prog. Nucleic Acid Res. Mol. Biol.*, Vol. 51, pp. 53–91.
3. Colleaux, L., d'Auriol, L., Betermier, M. *et al.* (1986), 'Universal code equivalent of a yeast mitochondrial intron reading frame is expressed into *E. coli* as a specific double strand endonuclease', *Cell*, Vol. 44, pp. 521–533.
4. Rong, Y. S. and Golic, K. G. (2001), 'A targeted gene knockout in *Drosophila*', *Genetics*, Vol. 157, pp. 1307–1312.
5. Rong, Y. S., Titen, S. W., Xie, H. B *et al.* (2002), 'Targeted mutagenesis by homologous recombination in *D. melanogaster*', *Genes Dev.*, Vol. 16, pp. 1568–1581.
6. Seum, C., Pauli, D., Delattre, M. *et al.* (2002), 'Isolation of *Su(var)3-7* mutations by homologous recombination in *Drosophila melanogaster*', *Genetics*, Vol. 161, pp. 1125–1136.
7. Egli, D., Selvaraj, A., Yepiskoposyan, H. *et al.* (2003), 'Knockout of 'metal-responsive transcription factor' MTF-1 in *Drosophila* by homologous recombination reveals its central role in heavy metal homeostasis', *EMBO J.*, Vol. 22, pp. 100–108.
8. Lankenau, S., Barnickel, T., Marhold, J. *et al.* (2003), 'Knockout targeting of the *Drosophila nap1* gene and examination of DNA repair tracts in the recombination products', *Genetics*, Vol. 163, pp. 611–623.
9. Gong, W. J. and Golic K. G. (2003), 'Ends-out, or replacement, gene targeting in *Drosophila*', *Proc. Natl. Acad. Sci. USA*, Vol. 100, pp. 2556–2561.
10. Mombaerts, P., Clarke, A. R., Hooper, M. L. and Tonegawa, S. (1991), 'Creation of a large genomic deletion at the T-cell antigen receptor beta-subunit locus in mouse embryonic stem cells by gene targeting', *Proc. Natl. Acad. Sci. USA*, Vol. 88, pp. 3084–3087.