

Making a better RNAi vector for *Drosophila*: use of intron spacers

Young Sik Lee and Richard W. Carthew*

Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, 2153 North Campus Drive, Evanston, Illinois 60208, USA

Accepted 7 February 2003

Abstract

Double-stranded RNA induces sequence-specific inhibition of gene expression at a posttranscriptional level in eukaryotes (RNAi). This natural phenomenon has been developed into a tool for studying gene function in several model organisms, including *Drosophila melanogaster*. Transgenes bearing inverted repeats are able to exert an RNAi effect in *Drosophila*, but cloning difficulties and inconsistent silencing complicate the method. We have constructed a transgene containing inverted repeats separated by a functional intron such that mRNA produced by the transgene is predicted to form loopless hairpin RNA following splicing. A single copy of the transgene effectively and uniformly silences expression of a target gene (*white*) in transgenic flies. We have developed a vector that is designed to produce intron-spliced hairpin RNA corresponding to any *Drosophila* gene. The vector is under control of the upstream activating sequence (UAS) of the yeast transcriptional activator GAL4. The UAS/GAL4 system allows hairpin RNA to conditionally silence gene expression in *Drosophila* in a tissue-specific manner. Moreover, the presence of the intron spacer greatly enhances the stability of inverted-repeat sequences in bacteria, facilitating the cloning procedure.

© 2003 Elsevier Science (USA). All rights reserved.

1. Introduction

The sequencing of the *Drosophila melanogaster* genome provides an exceptional opportunity to analyze the different functions governed by its genes [1]. Traditionally, genes are characterized by loss-of-function phenotypes caused by mutations that are induced randomly by chemical, physical, or insertional mutagenesis. The annotated sequence of the *Drosophila* genome enables reverse-genetic approaches to be used on a genome-wide scale to generate loss-of-function phenotypes. Targeted gene knockouts have recently been described in *Drosophila* [2,3], but this approach is laborious and does not allow for conditional silencing of gene expression.

Recently, RNA interference (RNAi) has been demonstrated to be an effective reverse-genetic approach to generating loss-of-function phenotypes. The presence of double-stranded RNA (dsRNA) causes the sequence-specific posttranscriptional silencing of a corresponding gene in a variety of organisms [4]. Thus, RNAi is used to

inactivate genes of interest and provides a powerful tool to study gene function. Injection of dsRNA into *Drosophila* embryos silences gene activity effectively, but its effect is transient and is not inherited in the next generation [5,6]. To overcome this problem, methods have been developed to express dsRNA stably in transgenic *Drosophila*. Most of these methods employ transgenes having an inverted-repeat (IR) configuration, which are able to produce dsRNA as extended hairpin RNA [7–10]. An alternative method has used a transgene that is symmetrically transcribed from opposing promoters [11]. A general problem with these methods is that transgenic lines often induce a variable RNAi silencing effect that exhibits incomplete penetrance and expressivity. Consequently, the copy number of silencing transgenes usually needs to be increased to observe uniform and complete gene silencing. Moreover, it is often difficult to make stable recombinant plasmids containing IRs in *Escherichia coli*. Introduction of a spacer sequence between the repeats helps stabilize some recombinant plasmids, but there are still significant reported stability problems.

In this paper, we describe an IR-based transgene designed such that the repeats are separated by a functional intron and thus are defined exons. We report that,

* Corresponding author. Fax: 1-847-467-1380.

E-mail address: r-carthew@northwestern.edu (R.W. Carthew).

in *Drosophila*, the transgene is a powerful repressor of gene activity *in vivo*, with one copy of the transgene exerting uniformly strong silencing. We further describe a modular system combining GAL4-regulated gene expression with conditional production of the silencing dsRNA to allow systematic RNAi in *Drosophila* using IR exons. With this system, inhibition of gene activity is flexibly induced in any spatial or temporal pattern, allowing for precise disruption of gene function. This technique could potentially be a powerful and economical approach to studying gene function in *Drosophila* and to manipulating gene function in specific tissues of postembryonic individuals.

2. First-generation transgenic RNAi in *Drosophila*

We initially developed a method to express dsRNA as extended hairpin-loop RNA [10]. Its design was modeled on the successful application of hairpin RNAs in generating RNAi in plants and the nematode *Caenorhabditis elegans* by expression of transgenes with IR sequences. To facilitate cloning of IRs into recombinant plasmids, we placed a 5-base nonpalindromic sequence centered at the axis of dyad symmetry that was a *Sfi*I site, GGCCATCTAGGCC (Fig. 1). This allowed us to easily ligate gene fragments together in inverted orientation, and it increased the stability of the IR DNA during plasmid replication. Sequence repeats are often deleted in *E. coli* because cruciform intermediates form during replication of plasmid DNA and are excised by the *sbcBC* gene products. Insertion of nonrepetitive se-

quence greater than 4 bp in length between IRs inhibits cruciform excision during replication [12]. Recombinant plasmids were replicated in a recombination-deficient strain. No strain is guaranteed to propagate all recombinant clones, but the SURE strain (Stratagene) is deficient in *recBC sbcBC* and eliminates all known restriction systems. Other strains we used were JM103 and JM105, which are also mutant for *sbcBC*.

To construct an IR transgene, the IR fragment is first cloned into a generic high-copy plasmid vector such as pBluescript (Stratagene) by directional *EcoRI*–*XhoI* two-way ligation. Stable recombinants are selected, and then the IR fragment is shuttled from pBluescript into the *Drosophila* transformation plasmid vector pUAST [13]. We found it more difficult to directly clone IR fragments made *in vitro* into pUAST, which we circumvented by shuttling the fragment first through pBluescript. On the 5' side of the multicloning site, pUAST contains a *Drosophila* promoter linked to GAL4-responsive upstream activating sequence (UAS) enhancer repeats and on the 3' side of the multicloning site, pUAST contains a polyadenylation signal sequence. Recombinant plasmids are then injected with helper plasmid into *Drosophila* embryos and transformant flies are generated by standard P element transformation [14]. Cloning the IR into a UAS vector allowed us to use the modular design of the GAL4/UAS system in *Drosophila* for misexpressing transgenes. Many useful lines of *Drosophila* express the yeast GAL4 protein in a variety of cells/tissues at various stages of the fly life cycle [13]. GAL4 acts as a sequence-specific transcription activator in *Drosophila*. The GAL4 line is

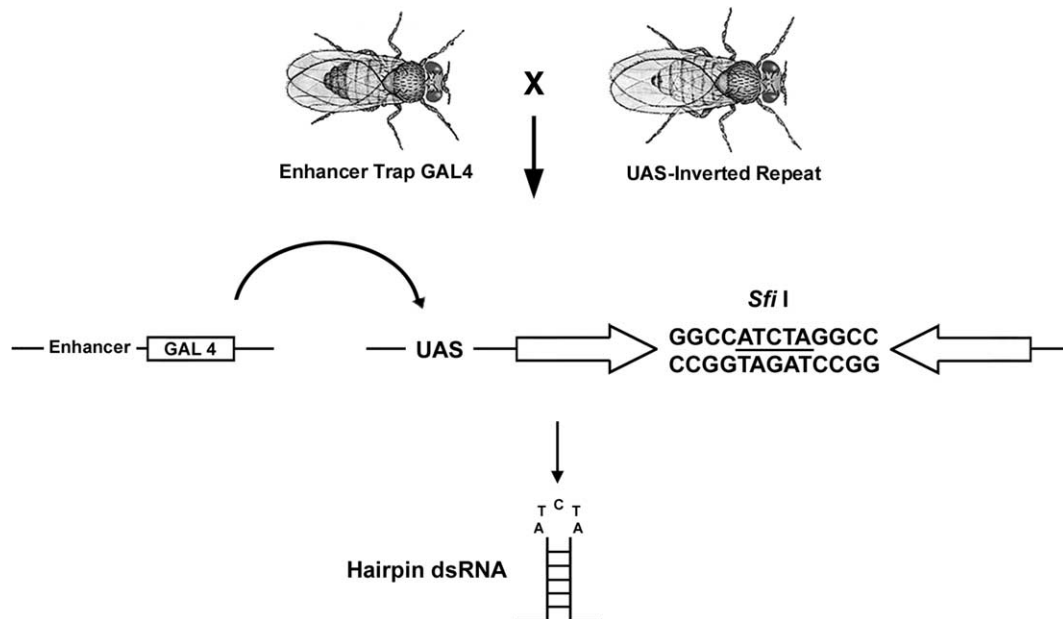


Fig. 1. First-generation transgenic RNAi. Inverted repeats are made by head–head or tail–tail ligation through a *Sfi*I site, which is introduced at one end of each repeat. The inverted repeat is placed downstream of the upstream activating sequence (UAS) promoter, and transgenic lines are made. When these UAS lines are crossed to GAL4 driver lines, the F1 heterozygotes contain both GAL4 and UAS genes [13]. Tissue-specific expression of the inverted repeats by GAL4 protein produces hairpin-loop RNA that is competent to induce RNAi in *Drosophila*.

crossed to a target UAS line carrying a single target P element inserted at a unique and random position in the genome. The target element carries a GAL4-responsive UAS enhancer, and progeny that contain both GAL4 and UAS elements express the IR sequence in cells expressing GAL4. Phenotypes due to the presence of hairpin-loop RNA in these cells can then be scored directly in flies. The RNAi system takes advantage of two very useful techniques in *Drosophila*: P element transformation and the modular GAL4/UAS system. The modular design makes analysis by RNAi flexible since hairpin-loop RNA can be produced in any spatial or temporal pattern. Moreover, RNAi is conditional, dependent on the presence of both UAS and GAL4 elements in the same individual. Thus, RNAi that might induce lethal or sterile phenotypes is conditionally generated in selected flies, and stably inherited *Drosophila* lines carrying the UAS element alone can be propagated without deleterious RNAi effect.

One feature of the target element that was critical for the success of this method was that the IR was stable in the *Drosophila* genome. This appears to be the case since UAS lines have been maintained in our lab stocks for 2 years without loss of RNAi potency when crossed with GAL4 driver lines. However, two other features of the method have proved to be more problematic. First, RNAi silencing is frequently variable, with only a fraction of treated animals exhibiting complete silencing. This partial effect is also observed at the level of target mRNA abundance in that a pooled population of treated animals might exhibit at most a fourfold reduction in mRNA levels. Thus, treated individuals have a spectrum of RNAi-induced phenotypes, which makes interpretation of gene function somewhat difficult. Moreover, there is frequently a variation in the strength of RNAi effects between different transformant lines carrying the same IR transgene. This is likely due to influence of nearby chromosomal modulation of transgene expression that depends on the point of transgene insertion. Since RNAi silencing is not complete, weak or strong IR expression translates to a corresponding weak or strong silencing effect. A second problem with this first-generation RNAi vector has been the variable success in cloning IRs from different genes. Some IRs are easily cloned even into *shcBC⁺* bacterial strains. However, some IRs have proved recalcitrant to cloning in any strain or any plasmid. To date, we have been unable to predict which sequences will produce problems when cloned as IRs. This makes transgene production a somewhat empirical trial-by-error process.

3. Transgenic RNAi with inverted exon repeats

The problems with the first-generation vector inspired us to devise a new approach to produce hairpin

RNA in vivo. It was reported that intron-spliced hairpin RNA can induce gene silencing in plants more efficiently than standard hairpin-loop RNA [15]. In a sense, the inverted repeats are structural and functional exons. The nonpalindromic intron sequence may also provide stability to the DNA construct with inverted repeats in bacteria. This led us to test an RNAi construct containing inverted repeats separated by an intron sequence, from which loopless hairpin dsRNA is predicted to be produced following splicing in *Drosophila* (Fig. 2). As a proof-of-principle demonstration of its effectiveness, we decided to test the approach on silencing the *white* gene.

The *Drosophila white* gene encodes an ABC transporter involved in localizing pigments to eye pigment granules [16]. The *white* gene was chosen because expression can be easily monitored phenotypically by changes in eye color. A *white⁺* eye is dark red in color while the eye of a null *white* mutant is completely white in color (Fig. 3). Since cloned variants of *white* are routinely used in *Drosophila* transformation vectors as the selectable marker for transformation of *white* mutant flies, we adopted an opposing transformation strategy. That is, we constructed a transformation vector with a *white* IR but no independent marker gene for selecting transformants. We then transformed *white⁺* flies with the vector and selected transformants that had a *white* loss-of-function phenotype. If an inserted *white* IR transgene successfully silenced its endogenous target gene, the transformant would be white-eyed.

The 74-nucleotide second intron of the *white* gene bears all features of a consensus *Drosophila* intron, and it was found to efficiently splice in *Drosophila* embryonic extracts in vitro [17,18]. Since *white* is normally not expressed in embryos, this result indicates that the intron can be spliced in heterologous tissues. Thus, we chose the second intron to separate inverted repeats of *white* coding sequence in our model transgene.

The 629-bp third exon of *white* was chosen to be the inverted sequence in the transgene that would mediate the RNAi effect. The third exon was amplified by PCR with unique *Pst*I and *Eco*RI sites, and it was ligated in inverted orientation upstream of a 703-bp fragment containing the *white* second intron and third exon (Fig. 2). The tail-to-tail repeat was placed into the pGMR transformation vector plasmid [19]. pGMR drives expression of transgenes specifically in the developing and adult compound eye by virtue of the eye-specific GMR promoter. This ensured that the *white* IR transgene would be expressed only in the same cells that normally express the endogenous target *white* gene. The *Drosophila* consensus sequence for a 5' splice site is AG|GTRAGT, where | designates the splice site and R indicates A or G [17]. It is noteworthy that the ligation between the two DNA fragments through the *Pst*I site does not change the consensus sequence required for

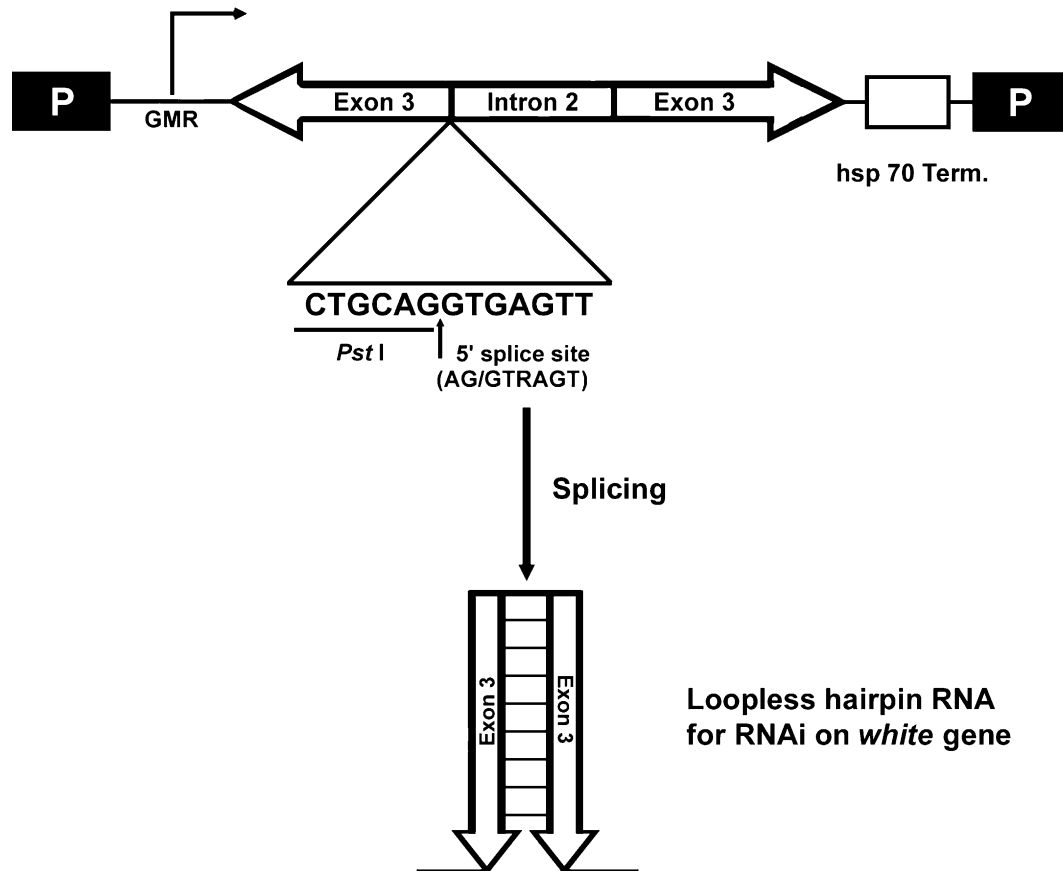


Fig. 2. Scheme for the generation of transgenic RNAi against the *Drosophila white* gene by intron-spliced hairpin RNA. Inverted repeats corresponding to the third exon of the *white* gene and separated by the second intron of the same gene were placed into the pGMR transformation vector. A fragment containing the *white* second intron and third exon was ligated to a fragment containing the inverted *white* third exon to generate a *Pst*I site at the point of ligation. Note that the *Pst*I site is coincident with the 5' splice site but does not disrupt sequences necessary for 5' splice site function. The *Drosophila* consensus sequence for a 5' splice site is shown in parentheses, in which / designates the splice site and R indicates A or G. The transgene is under transcriptional control of the GMR enhancer, which is exclusively active in the developing and adult retinal tissue that also expresses the endogenous *white* gene.

5' splice site recognition (Fig. 2). Since we intended to induce RNAi on the endogenous *Drosophila white* gene, an *Xho*I–*Nsi*I fragment corresponding to the pGMR *white* marker gene was deleted from pGMR.

Although we transformed ligation products including inverted repeats into the SURE strain (Stratagene) of *E. coli* to maximize the stability of the DNA, the repeats were also stable in a DH5 α strain, possibly as a result of the short *white* second intron sequence contributing to the stability of the inverted-repeat sequences. In contrast, attempts at cloning an IR of the *white* third exon separated by a *Sfi*I spacer into plasmids was repeatedly unsuccessful in a variety of host bacterial strains.

The pGMR-derived plasmid containing the DNA fragment for intron-spliced *white* hairpin RNA was introduced into the germ line of CantonS flies by P element transformation [14]. From approximately 1500 injected animals, eight independent transformant lines that exhibited a *white* loss-of-function phenotype were established. This transformation frequency is within an order of magnitude of the average transformation fre-

quency using a standard P element vector [14], which suggests that *white* RNAi from the IR transgene acts as a reliable marker for transformation. All eight transformant lines exhibited a yellow to pale-yellow eye color phenotype with one copy of the transgene (Fig. 3). Moreover, all individual flies from each line exhibited a uniform eye color phenotype, indicating strong penetrance and expressivity of the RNAi effect. Only females were compared to avoid any effect related to dosage compensation of the transgene. No additional or abnormal phenotypes were observed in silenced individuals, indicating that silencing was specific. The effect was stably maintained over each adult's lifetime, and silencing has been maintained over the many generations that these lines have so far been kept. Transformant adults bearing two copies of the transgene had an eye color indistinguishable from that of *white* null mutants (Fig. 3). Levels of *white* mRNA on a Northern blot were reduced in two transformant lines tested compared to wild type, to a degree consistent with their eye color phenotypes (data not shown). In conclusion, the

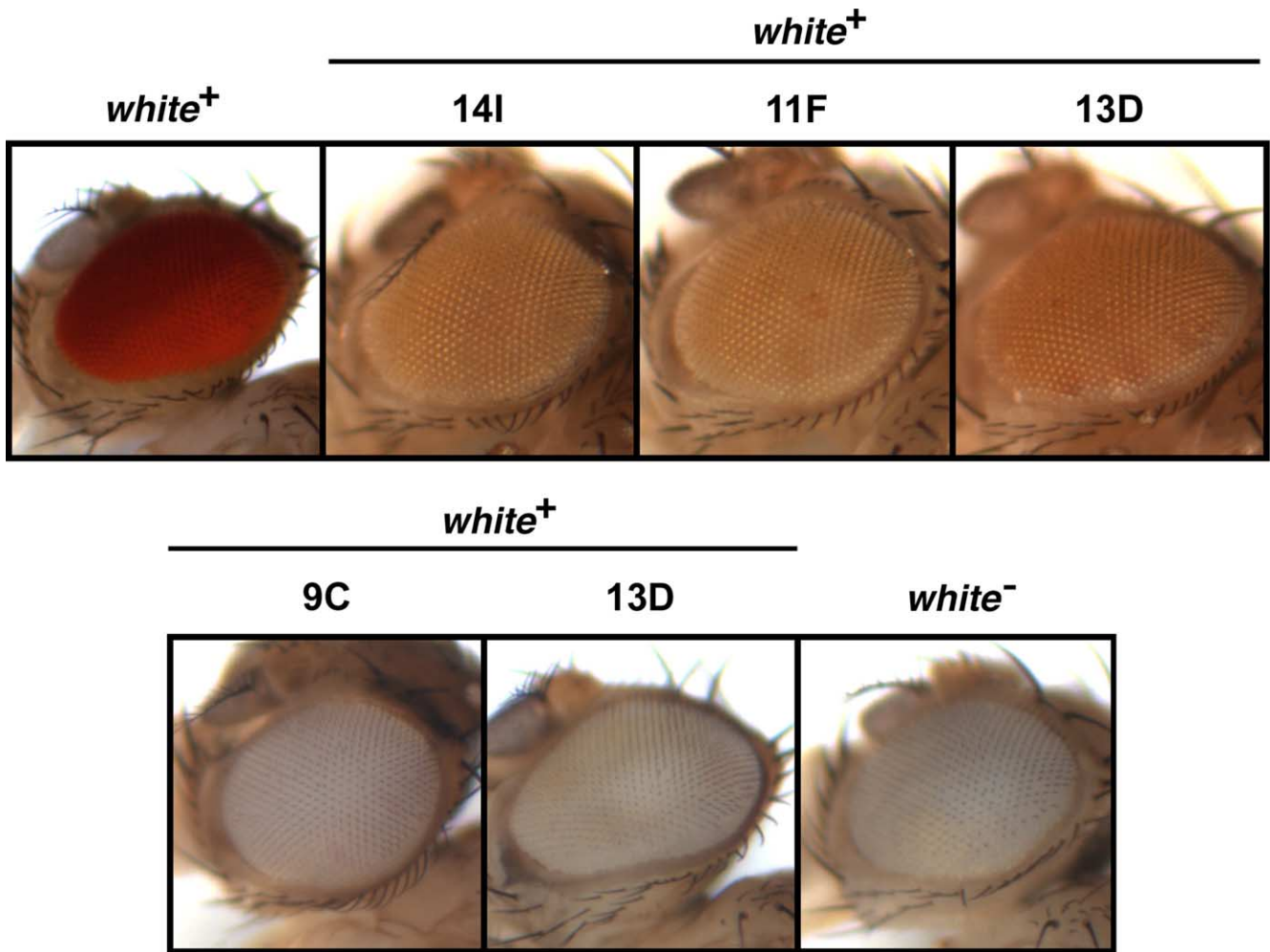


Fig. 3. Eye color phenotypes of female adults (3–5 days of age) that carry the *white* intron–hairpin GMR transgene. The top left shows a parental CantonS (wild-type *white*) fly. The lower right shows a w^{118} null mutant for the *white* gene. A P[*GMR-whiteIR*] transformant line designated 13D has the transgene on the X chromosome of the parental CantonS strain, whereas the 9C, 11F, and 14I lines have the transgene on the third chromosome of CantonS. The eye colors of transformants bearing a single copy of the transgene is shown in the top. The bottom shows eye colors of transformants bearing two copies of the transgene.

inverted exon repeat of *white* effectively silences the expression of the endogenous *white* gene in vivo.

4. A modular and multipurpose transgenic RNAi vector

Since the RNAi construct bearing an intron strongly inhibited *white* gene expression, we adapted this method to create an all-purpose RNAi vector that employs spliced hairpin RNA. The vector is derived from the pUAST transformation plasmid. This then offers the advantages of the GAL4/UAS modular expression system, as outlined earlier. We constructed the vector (pWIZ, for white intron zipper) into which gene fragments can be subcloned upstream and downstream of the 74-nucleotide *white* intron (Fig. 4). The intron is flanked by *EcoRI*, *BglII*, *NotI*, *XhoI*, *SpeI*, and *AvrII* sites on the 5' side and by *NheI*, *MluI*, and *XbaI* sites on the 3' side. The entire cassette is downstream of the UAS

enhancer–promoter and upstream of the SV40 transcription termination site. The *AvrII* and *NheI* sites in pWIZ conform to the consensus sequences for 5' and 3' splice sites, respectively. Thus, any DNA fragment inserted into the *AvrII* or *NheI* site is fully competent to be spliced as an exon. Moreover, the *SpeI*, *AvrII*, *NheI*, and *XbaI* sites are unique in pWIZ, providing convenient cloning sites for gene fragments.

To construct an IR transgene using pWIZ, a DNA fragment corresponding to the gene of interest is inserted twice into pWIZ, with inserts in opposite orientations on each side of the intron (Fig. 5). The simplest means to insert the DNA is as a PCR fragment. The system is designed so that a single PCR fragment derived from only two PCR primers can be inserted on each side of the intron. This is because *SpeI*, *AvrII*, *NheI*, and *XbaI* sites are all ligation-compatible with each other. Consequently, restriction sites compatible with *AvrII* and *NheI* sites should be placed in the PCR

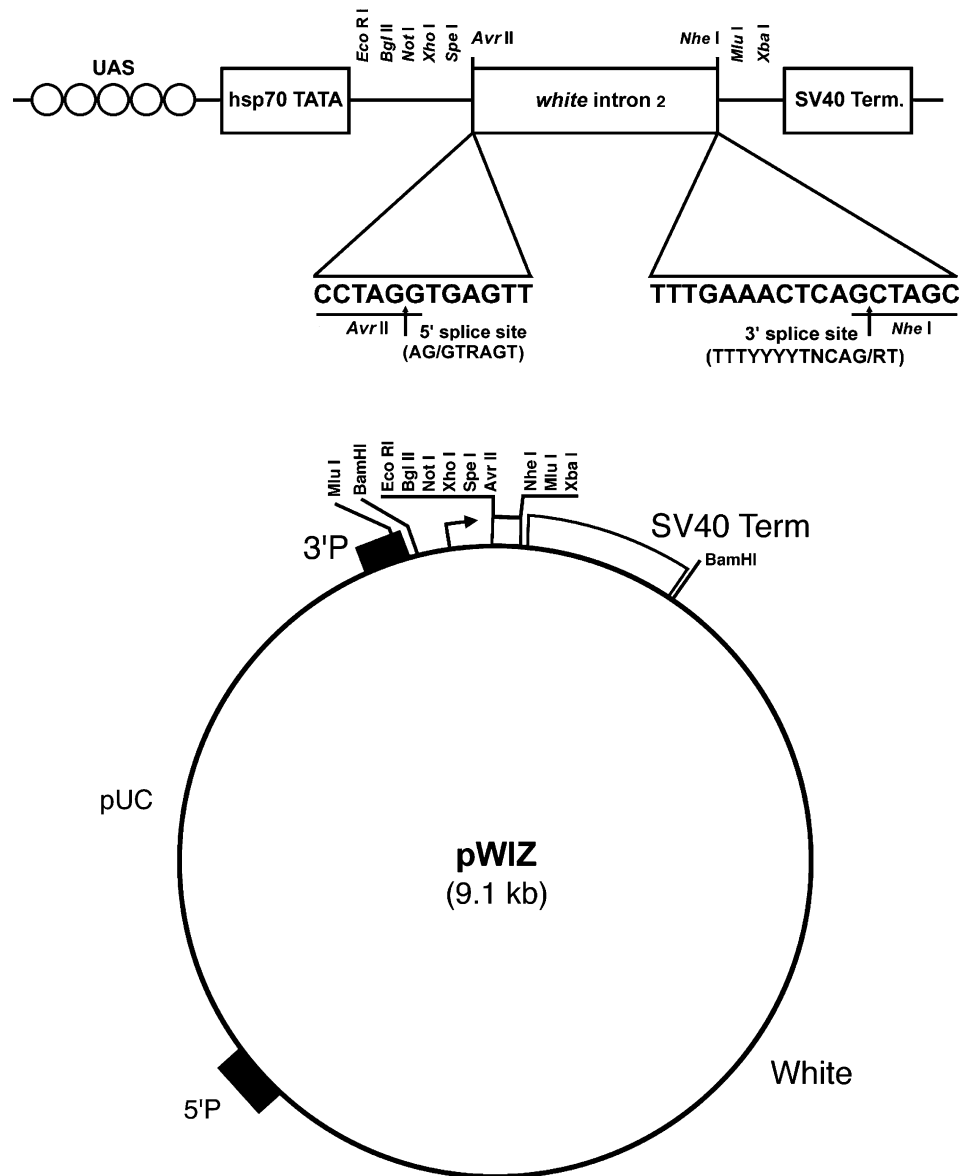


Fig. 4. Schematic representation of the pWIZ vector. The pWIZ vector was constructed by placing the 74-bp second intron of the *white* gene into the pUAST transformation vector [13]. The intron is flanked by unique *Eco*RI, *Bgl*II, *Not*I, *Xho*I, *Spe*I, and *Avr*II sites on the 5' side and *Nhe*I and *Xba*I sites on the 3' side to facilitate cloning. The sequences at the junction of the 5' and 3' splice sites in the vector are highlighted, and arrows indicate the 5' and 3' splice sites. The consensus sequences for 5' and 3' splicing are shown in parentheses: /, the splice site; R, purine; Y, pyrimidine; N, any base. Below is shown a restriction map of the pWIZ plasmid.

primers at their 5' ends. The resulting PCR product will then have *Avr*II- and *Nhe*I-compatible ends after appropriate restriction digestion. The RNAi construct is made by two sequential insertions of the same PCR product into the *Avr*II and *Nhe*I sites of pWIZ (Fig. 5). For efficient digestion, we add an extra 4 nucleotides to the 5' side of each primer restriction site. The size of amplified DNA ranges from 500 to 700 bp. Most important, we ensure that the sequence does not contain any internal restriction sites used in the PCR primers, nor should the fragment have sequences in either sense or antisense orientation that match either 5' or 3' consensus splice sites. This latter aspect is important to

prevent cryptic splicing from disrupting hairpin RNA formation. The PCR product is digested with the appropriate enzyme and ligated into the *Avr*II site of pWIZ. After a clone with the desired orientation of the insert is selected by restriction mapping, the same PCR product is ligated into the *Nhe*I site of the pWIZ derivative, and recombinants with the insert in opposite orientation to the first are screened and selected.

We have made five transgenic RNAi constructs for genes under study in our laboratory using the pWIZ vector. All of these constructs are stable as inverted repeats in *E. coli* strains such as SURE cells. Moreover, they have been introduced into the *Drosophila* genome

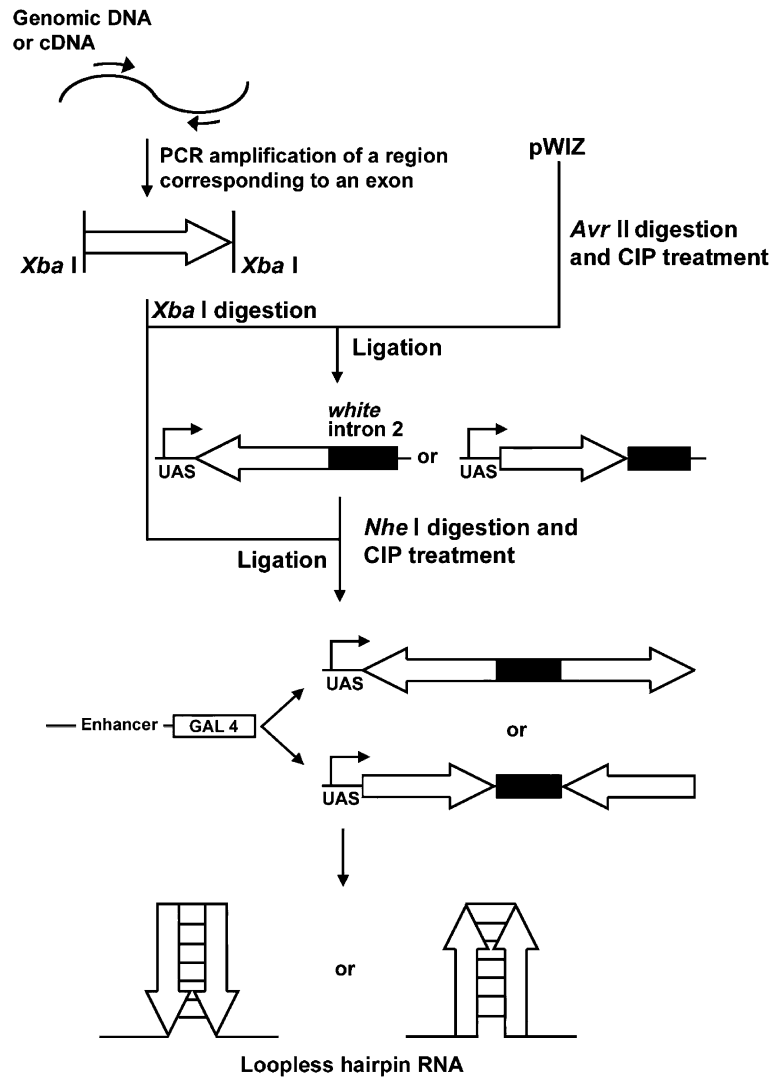


Fig. 5. A typical procedure for making a RNAi construct using the pWIZ vector. A DNA fragment corresponding to a *Drosophila* target gene is amplified by PCR. A restriction site (*SpeI*, *AvrII*, *NheI*, or *XbaI*) compatible with *AvrII* and *NheI* should be present at the 5' end of each PCR primer. The fragment can correspond to part of an exon, a complete exon, or several contiguous exons. Several criteria should be used in choosing the fragment: it should be 500 to 700 bp in length, it should not have internal restriction sites corresponding to the PCR primer sites, and it should not have sequences in either sense or antisense orientation that match a 5' or 3' consensus splice site. This latter aspect is critical to prevent cryptic splice sites from disrupting proper splicing of the hairpin RNA. As shown, an *XbaI* site is generated at each end of a PCR-amplified exon for cloning the PCR product. The PCR product is inserted twice by two ligation steps into the *AvrII* and *NheI* sites of pWIZ. CIP, calf intestinal phosphatase used to dephosphorylate the 5' ends of vector DNA prior to ligation. Recombinants are selected in the desired orientation, such that after the second ligation step, inserts are in opposite orientation on either side of the *white* intron. IRs that are head-head or tail-tail repeats might be used depending upon whether cryptic splice sites are present in the constructs. Transformation follows to generate stable transgenic lines carrying the WIZ gene. Upon mating of transgenic animals harboring the WIZ gene with animals carrying tissue- or cell-specific GAL4 drivers, the F1 progeny produce loopless hairpin RNA. This induces RNAi against target genes in tissue- and cell-specific patterns.

by P element transformation, and all constructs transformed efficiently to give stable lines.

5. Concluding remarks

In conclusion, we have developed transgenic RNAi in *Drosophila* that can be applied to many developmental and physiological processes. Hairpin RNA produced from a transgene composed of inverted repeats can spe-

cifically silence gene expression in *Drosophila*. The presence of a spacer between the inverted repeats makes for easier cloning but is offset by a weaker silencing activity in vivo [8,20]. In plants, using a functional intron as the spacer between inverted repeats strongly enhanced silencing activity of the RNAi transgene [15]. We have shown that using a functional intron as a spacer between inverted repeats produces strong and uniform RNAi silencing in *Drosophila*. A similar observation has been recently noted in *Drosophila* when inverted repeats

composed of cDNA–genomic DNA hybrids are separated by functional introns [21]. We have also described a multifunctional RNAi transformation vector (pWIZ) containing an intron spacer that makes RNAi simple to perform for the following reasons. A single PCR fragment of a gene is sufficient to construct a targeting vector; the inverted repeat sequence need not have splice sites present since they are provided by pWIZ; splice sites are preserved when the repeat fragments are inserted; the intron spacer provides stability to the inverted repeats when the plasmid is replicated in *E. coli*. Once the vector is transformed into *Drosophila*, it is conditionally quiescent until crossed with GAL4-expressing lines. Many useful GAL4-expressing lines are available, making the RNAi approach adaptable for most studies of *Drosophila*. This method is likely to be very useful for analyzing the function of the many *Drosophila* genes for which no loss-of-function mutations are available. Finally, the method provides a powerful tool to create loss-of-function phenotypes in a manner conditional for particular tissues and developmental times.

References

- [1] M.D. Adams, S.E. Celniker, R.A. Holt, C.A. Evans, J.D. Gocayne, P.G. Amanatides, S.E. Scherer, P.W. Li, R.A. Hoskins, R.F. Galle, et al., *Science* 287 (2000) 2185–2195.
- [2] Y.S. Rong, K.G. Golic, *Science* 288 (2000) 2013–2018.
- [3] Y.S. Rong, K.G. Golic, *Genetics* 157 (2001) 1307–1312.
- [4] S.M. Hammond, A.A. Caudy, G.J. Hannon, *Nat. Rev. Genet.* 2 (2001) 110–119.
- [5] J.R. Kennerdell, R.W. Carthew, *Cell* 95 (1998) 1017–1026.
- [6] L. Misquitta, B.M. Paterson, *Proc. Natl. Acad. Sci. USA* 96 (1999) 1451–1456.
- [7] G. Lam, C.S. Thummel, *Curr. Biol.* 10 (2000) 957–963.
- [8] E. Fortier, J.M. Belote, *Genesis* 26 (2000) 240–244.
- [9] S. Martinek, M.W. Young, *Genetics* 156 (2000) 1717–1725.
- [10] J.R. Kennerdell, R.W. Carthew, *Nat. Biotechnol.* 18 (2000) 896–898.
- [11] E. Giordano, R. Rendina, I. Peluso, M. Furia, *Genetics* 160 (2002) 637–648.
- [12] A. Davison, D.R. Leach, *Genetics* 137 (1994) 361–368.
- [13] A.H. Brand, N. Perrimon, *Development* 118 (1993) 401–415.
- [14] A.C. Spradling, G.M. Rubin, *Science* 218 (1982) 341–347.
- [15] N.A. Smith, S.P. Singh, M.-B. Wang, P.A. Stoutjesdijk, A.G. Green, P.M. Waterhouse, *Nature* 407 (2000) 319–320.
- [16] S.M. Mackenzie, A.J. Howells, G.B. Cox, G.D. Ewart, *Genetica* 108 (2000) 239–252.
- [17] S.M. Mount, C. Burks, G. Hertz, G.D. Stormo, O. White, C. Fields, *Nucleic Acids Res.* 20 (1992) 4255–4262.
- [18] M. Guo, P.C.H. Lo, S.M. Mount, *Mol. Cell. Biol.* 13 (1993) 1104–1118.
- [19] B.A. Hay, T. Wolff, G.M. Rubin, *Development* 120 (1994) 2121–2129.
- [20] A. Piccin, A. Salameh, C. Benna, F. Sandrelli, G. Mazzotta, M. Zordan, E. Rosato, C.P. Kyriacou, R. Costa, *Nucleic Acids Res.* 29 (2001) E55.
- [21] S. Kalidas, D.P. Smith, *Neuron* 33 (2002) 177–184.