EMERGING TECHNOLOGIES FOR GENE MANIPULATION IN *DROSOPHILA MELANOGASTER*

Koen J. T. Venken and Hugo J. Bellen

Abstract | The popularity of *Drosophila melanogaster* as a model for understanding eukaryotic biology over the past 100 years has been accompanied by the development of numerous tools for manipulating the fruitfly genome. Here we review some recent technologies that will allow *Drosophila melanogaster* to be manipulated more easily than any other multicellular organism. These developments include the ability to create molecularly designed deletions, improved genetic mapping technologies, strategies for creating targeted mutations, new transgenic approaches and the means to clone and modify large fragments of DNA.

P-ELEMENT
Transposable elements that are widely used to mutate and manipulate the genome of *Drosophila melanogaster*.

TRANSGENESIS
The process of introducing foreign DNA into a genome.

TRANSPOSON
A DNA element that can be mobilized within the genome; *Drosophila melanogaster* transposons are used for various applications, including insertional mutagenesis, gene tagging and as a carrier for transgenic DNA.

DEFICIENCY
A chromosomal aberration that is characterized by the deletion of a part of the genome.

Program in Developmental Biology, Baylor College of Medicine, Houston, Texas 77030, USA. Correspondence to H.J.B. e-mail: hbellen@bcm.tmc.edu
doi:10.1038/nrg1553

Drosophila melanogaster is an attractive experimental organism because of its rapid generation time and the ease with which it can be handled in the laboratory. As a result, analysis of mutant phenotypes can be carried out in a relatively short time. Flies also benefit from a comprehensive range of methods for carrying out molecular genetic research; these include *P*-ELEMENT-mediated TRANSGENESIS[1], a versatile gene-overexpression system that is based on the yeast Gal4–UAS (upstream activating sequence) system[2], and a tool — the Flp–*FRT* system — for carrying out site-specific recombination[3,4]. Moreover, the analysis of specific genes has been assisted by the recent publication of the complete sequence of the fly genome[5]. Despite notable progress, these molecular-genetic techniques have some limitations. Over the past 5 years, these shortfalls have been overcome, to some degree, by several ingenious technological developments that promise to make the fly an even more useful model system. The innovative technologies include methods for introducing transgenic DNA into flies, generating targeted mutations and creating designer molecular lesions. In this review, we describe several of these advances, their immediate and future applications and their potential for further development.

TRANSPOSON insertions have been exceptionally useful in *Drosophila* research, as they can be used to generate single-gene, as well as multigene deletions. Several genome-wide projects have made transposon insertions available in approximately two-thirds of all fruitfly genes. These genome-wide efforts have used (among other means) transposons, such as *P*-elements and *piggyBac*s, that have different properties and mobilization characteristics[6–8]. More recently, these rather traditional applications of transposons have been extended. For example, some of these insertions have been used to generate defined molecular deletions, therefore substantially increasing the proportion of the *D. melanogaster* genome that is covered by DEFICIENCIES[9,10]. Recently, molecularly defined transposon insertions have been shown to greatly aid the mapping of *D. melanogaster* genes that are isolated in forward-genetic screens[11]. So, one theme that we cover in this article is how transposons and the technologies that are associated with them have facilitated gene tagging, mutagenesis and mapping.

The second group of new technologies that we describe is geared towards expanding the ability to efficiently generate or identify mutations in a gene of interest. Carrying out reverse genetics in *D. melanogaster* typically means relying on one of three approaches: IMPRECISE EXCISION of an existing *P*-element, gene targeting or RNA INTERFERENCE (RNAi). Unfortunately, *P*-element insertions are not available for every gene in the genome. Furthermore, existing methods for gene targeting are time consuming, and RNAi can cause non-specific

defects or might not result in the complete knock-down of the protein-encoding mRNA. We highlight two new reverse-genetic strategies. One such method relies on 'ends-out replacement gene targeting' and allows the removal of specific sequences and the introduction of changes in a genomic locus of interest[12]. Another method, known as targeting-induced local lesions in genomes (TILLING), allows the identification of an ALLELIC SERIES of point mutations in a gene[13,14]. A third strategy is aimed at creating TEMPERATURE-SENSITIVE MUTATIONS, thereby facilitating phenotypic analyses[15].

The availability of an efficient transgenesis system, based on *P*-transposable elements[1], has allowed spectacular progress to be made in fruitfly biology (and biology in general). These transposons integrate into the genome when *P*-transposase is provided. Unfortunately, *P*-elements that contain large DNA fragments cannot integrate into the genome. We discuss two technologies that might alleviate this problem. The first is a new genome-integration method that is based on the BACTERIOPHAGE ϕC31; this bacteriophage integrates genomic constructs at defined positions in the genome[16], and might also allow the integration of large constructs. The modification of large constructs will be aided significantly by a second new method, called recombineering (recombination-mediated genetic engineering) — this cloning technology, which was first applied in the mouse, eliminates the need for restriction enzymes and DNA ligases[17,18].

The goal of this review is to summarize the technological advances made since 2002 (REF. 19) that we feel will influence fly research most significantly; we therefore do not cover all the progress that has been made over the past 3 years. The approaches and methodologies that have recently been introduced or that are currently being developed can broadly be subdivided into those that are fully dependent on the availability of the sequenced genome[5], and those that to some extent are independent of the availability of this information. It should be noted that the methods in the latter category are often based on observations and technologies that have already been developed in other systems, but that have not yet been introduced into the fly.

Transposon-associated methods

The use of transposable elements, especially *P*-elements, in *D. melanogaster* has led to many valuable applications[7]. Since the completion of the fly genome sequence in 2000 (REF. 5), the precise genomic location of transposon insertions can be determined. This has allowed gene-specific deletions to be created by imprecise excision of *P*-elements, the most commonly used method to generate single-gene deletions or mutations[19]. More recently, transposable elements have been used in various ways to generate targeted deletions that can be molecularly mapped. In this section, we describe the public and private efforts that have created new transposon insertions and molecularly mapped deletions. In addition, GENE-TRAP transposable elements that contain sequences encoding enhanced GFP (EGFP) or conditionally splicing protein introns (or INTEINS) might allow the generation of protein-expression reporters and temperature-sensitive alleles, respectively, on a genome-wide scale[15,20] (N. Perrimon, personal communication). We also briefly touch on the use of *P*-elements to map essential genes by meiotic recombination[11].

Transposon-mediated mutagenesis. The goal of the publicly financed Gene Disruption Project, initiated by the Berkeley *Drosophila* Genome Project[8,21] (BDGP), is to obtain a *P*-element insertion in every gene. There are many advantages to using *P*-elements. More than two-thirds of all *P*-element insertions occur within 400 bp of transcription start-sites[8,21]. In addition, *P*-elements hop efficiently out of the genome and generate precise and imprecise excisions at high frequency. It is predicted that more than 80% of all *D. melanogaster* genes can be tagged and mutated with *P*-elements[8]. If we combine the *P*-elements that have been generated by public efforts[8,10,22] and private efforts, such as by Thibault *et al.* at Exelixis Inc.[6] and by the pharmaceutical company Developen AG, 65% of *D. melanogaster* genes are estimated to have a *P*-element insertion (P. Hiesinger, personal communication; see also the *Drosophila* GeneTag Database in the Online links box). More than 95% of these insertion strains are already available from the Bloomington *Drosophila* Stock Center, the Szeged *Drosophila* Stock Center (see Online links box), Baylor College of Medicine, and Harvard Medical School.

One drawback of *P*-elements is their insertional preference. Approximately one-third of all *P*-element insertions are not found in hot or medium–hot-spots (genes that 'attract' numerous *P*-element insertions). Therefore, to obtain insertions in every gene, some projects are using other transposable elements, such as *piggyBac*s[23–25], which

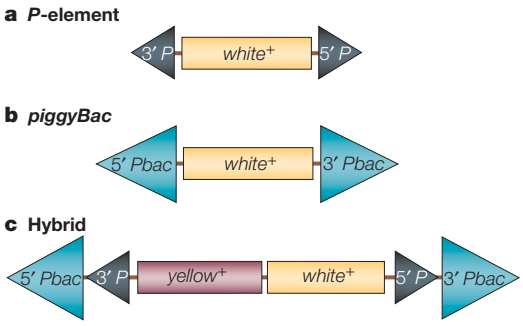


Figure 1 | **Transposon-mediated gene-disruption screens.** The gene-disruption projects used two types of transposable element as JUMP-STARTER ELEMENTS, the *P*-element (**a**) and the *piggyBac* transposon (**b**). Both transposable elements contain 5′ and 3′ ends that are necessary for efficient transposition: *P*-elements encode a 3′ *P*-transposase (3′ *P*) and a 5′ *P*-transposase (5′ *P*) site, whereas *piggyBac* elements contain a 3′ *piggyBac* transposase (3′ *Pbac*) and a 5′ *piggyBac* transposase (5′ *Pbac*) site. A hybrid element (**c**) that contains the features of both the *P*-element and *piggyBac* transposon should combine the advantages of both: precise excision and random hopping using *piggyBac* transposase, and imprecise excision using *P*-transposase. If two dominant markers are used, *yellow*$^+$ and *white*$^+$, it should be possible to screen for imprecise excisions resulting in unidirectional deletions.

IMPRECISE EXCISION
A *P*-element that is mobilized imprecisely can lead to removal of flanking genomic sequences, resulting in a local deletion.

RNA INTERFERENCE
(RNAi). A form of gene silencing in which dsRNA induces the degradation of the homologous endogenous mRNA transcripts, thereby mimicking the effect of the reduction, or loss, of gene activity.

ALLELIC SERIES
A series of alleles at the same genomic locus that can produce graded phenotypes and therefore can help to unravel different functions of the same gene.

TEMPERATURE-SENSITIVE MUTATIONS
Mutations that show a mutant phenotype at the restrictive temperature, but not at the permissive temperature. Most temperature-sensitive mutations in *Drosophila melanogaster* are heat-sensitive, but cold-sensitive mutations also occur.

BACTERIOPHAGE
A virus that infects and replicates in bacteria. Lytic bacteriophages kill the host cell, whereas so-called temperate phages can establish a stable relationship in which the bacteriophage genome is stably maintained within that of the host.

GENE TRAP
A DNA construct that contains a reporter-gene sequence downstream of a splice-acceptor site and that can integrate into random chromosomal locations. A gene trap sometimes also contains a splice donor site that is downstream of the marker. Integration of the gene trap into an intron allows the expression of a new mRNA that contains one or more upstream exons followed by the reporter gene.

INTEIN
'Protein introns' that autonomously splice themselves out of proteins (post-translationally), thereby generating a functional protein.

JUMP-STARTER ELEMENT
A transposable element that is inserted into the genome and that is used to create more insertions at other sites.

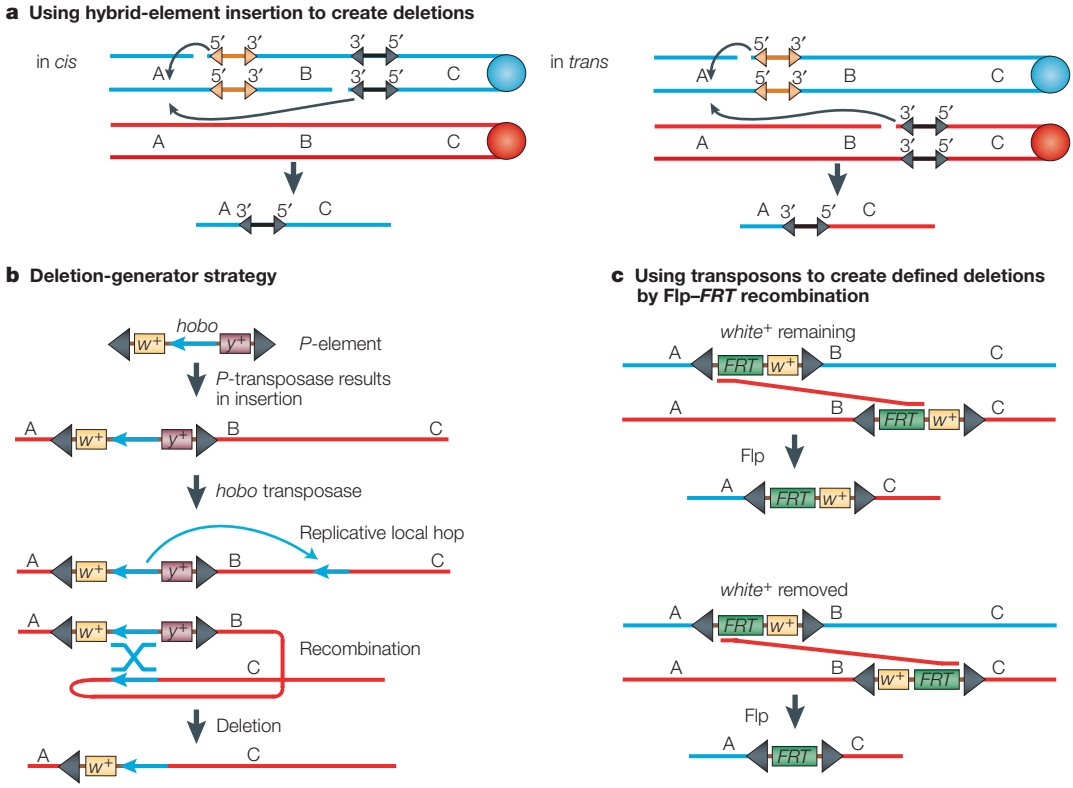


Figure 2 | **Using transposons to create chromosomal aberrations. a** | Using hybrid-element insertion to create deletions in *Drosophila melanogaster*. Two *P*-elements are located at different positions on the same chromatid (in *cis*), or on different chromatids (in *trans*) of the same chromosome. The diagram shows the chromosomes during the G2 phase of the cell cycle. The *P*-transposase cuts the 5′ end of one *P*-element and the 3′ end of the other *P*-element. The respective uncut ends are still continuous with the rest of the chromatid. Both cut ends integrate together in a nearby target location on one of the starting chromatids (curved arrows). One of the products will carry a deletion, in this case of region B (shown). The other product will carry a duplication of the same chromosome (not shown). Because *P*-transposase can result in imprecise excision of each end, a perfect deletion might not be obtained. Modified, with permission, from REF. 9 © (2004) Macmillan Publishers Ltd. **b** | The deletion-generator strategy. The hybrid transposable element consists of a *hobo* 'deleter' element that lies inside a 'carrier' *P*-element. *hobo* is flanked by two dominant markers, *white*$^+$ (w^+) and *yellow*$^+$ (y^+). The *hobo* transposase catalyses a replicative local hop of *hobo*. Intrachromosomal recombination between two *hobo* elements that are orientated in the same direction results in removal of one marker (y^+ in this case) and all genes in the intervening sequence. Intrachromosomal recombination between two inversely orientated *hobo* elements would result in an inversion (not shown). **c** | Using transposons to create defined deletions by Flp–*FRT* recombination. Two transposons, *P*-elements or *piggyBac*s, that contain *FRT* sites in the same direction on the same chromosome, are used to generate molecularly defined deletions. Both transposons are brought in *trans* in the presence of a stable Flp-expressing source. Flp catalyses the removal of the intervening sequence between both *FRT* sites. Depending on the orientation of the w^+ marker, excision results in removal of both w^+ markers or of just one. A similar strategy can be used to generate molecularly defined inversions, as well as reciprocal translocations. Modified, with permission, from REF. 9 © (2004) Macmillan Publishers Ltd.

show much less insertional specificity than *P*-elements[6,8] (FIG. 1). Unfortunately, *piggyBac*s only excise precisely, precluding their use in generating more alleles of a tagged gene. Furthermore, as for *P*-elements, many *piggyBac* insertions are not associated with any phenotype[6]. To remedy the drawbacks of both types of transposable element, we propose the use of hybrid transposons: these elements would carry a *piggyBac* backbone, to allow random insertion, and an internal *P*-element, to allow imprecise excision (FIG. 1). This approach has not yet been tested, but it might allow insertions in, and subsequent mutagenesis of genes that are refractory to *P*-element insertion.

The lack of insertional specificity of *piggyBac*s has been exploited in gene-trap screens[26,27]. The insertion of a gene-trap construct within an intron of a gene leads to a protein fusion in one sixth of cases, thereby reporting the expression, and sometimes the subcellular distribution, of the protein. The use of gene-trap transposons was pioneered in mice[28] and then adapted to *D. melanogaster* by using *P*-elements[20,29]. As *P*-elements rarely insert into introns, gene-trapping efficiency was significantly increased by using *piggyBac*s[26] (L. Cooley and A. Spradling, personal communication). Another application of gene trapping is to incorporate an intein[30,31] into the gene-trap vector. A temperature-sensitive splicing version of an intein was recently shown to be functional in *D. melanogaster*[15]. At 18°C, the intein is spliced out, producing a functional protein. However, at 29°C, splicing does not occur and the host protein might be inactive or

COMPARATIVE GENOMIC HYBRIDIZATION MICROARRAY
A high-density microarray that contains overlapping DNA clones of genomic sequences or overlapping oligonucleotides that encompass a particular genomic region. This allows chromosomal imbalances to be identified at a high resolution, and can even be used to determine exact chromosomal breakpoints.

HAPLOINSUFFICIENCY
The requirement for a diploid organism to have both functional copies of a gene or locus to produce a wild-type phenotype.

YELLOW$^+$ (y^+)
The wild-type allele of a fly gene that confers dark body colour; by contrast, *yellow*$^-$ flies have a yellowish body colour. One copy of *yellow*$^+$ is often used as a dominant genetic marker in fly trangenesis.

WHITE$^+$ (w^+)
The wild-type allele of a fly gene that confers red eyes; *white*$^-$ flies have white eyes. One copy of *white*$^+$ is often used as a dominant genetic marker in fly trangenesis.

INVERSE PCR
A method for cloning DNA that flanks a known sequence. Genomic DNA is digested and ligated into circles, and is then subjected to PCR. The PCR primers are designed to be complementary to the known sequence, but point outwards from this sequence, allowing the unknown flanking DNA to be amplified.

MEIOTIC RECOMBINATION MAPPING
Uses the recombination rate between a wild-type chromosome and a chromosome that is marked with recessive markers to obtain a recombination distance, or map unit, between two genes, mutations or markers.

MALE RECOMBINATION MAPPING
Meiotic recombination does not occur in *Drosophila melanogaster* males, so *P*-element-mediated recombination between homologous chromosomes can be used to determine the position of the mutation relative to the position of *P*-element insertions.

less active. This system provides an elegant way to generate a temperature-sensitive allele. However, because temperature sensitivity is conferred at the level of splicing, inactivation at a higher temperature is entirely dependent on protein stability. Nonetheless, gene-trap constructs containing temperature-sensitive inteins inside *piggyBac* elements could open the way to genome-wide efforts to generate temperature-sensitive alleles of many genes (N. Perrimon, personal communication).

Transposon-induced deficiencies. Chromosomal deletions represent true null alleles of each gene within the removed DNA fragment. Deficiencies are therefore extremely valuable as reagents, especially to map genes and to identify dosage-sensitive suppressors or enhancers of a certain phenotype. Deficiency stocks from the Bloomington Stock Center are the most commonly used (K. Cook and K. Matthews, personal communication). Unfortunately, many deficiency stocks were induced by X- or γ-rays, and their breakpoints have only been mapped cytologically. However, COMPARATIVE GENOMIC HYBRIDIZATION MICROARRAYS now allow us to molecularly map previously generated deficiencies, as a deletion can be detected by a localized decrease of the fluorescent signal on a DNA microarray[32]. A mapping service for available deletions that are not molecularly mapped will soon be provided by the Model System Genomics Unit of Duke University (E. Spana, personal communication; see Online links box). In addition, three new approaches have emerged in the past couple of years that allow specific deletions to be engineered, and even mapped in some cases.

The first strategy consists of creating deficiencies by deleting the genomic DNA between two *P*-element insertions by using *P*-transposase, as shown in FIG. 2a. Both *P*-elements can be located on the same chromosome in *cis*[33], or on different chromosomes in *trans*[9]. The resulting deletions often remain flanked by one of the *P*-elements, providing a starting point for molecular characterization. This strategy has been used to generate deletions that closely flank HAPLOINSUFFICIENT genes[9]. This deletion set increases the overall genomic coverage by 5–7% (REF. 9). Unfortunately, this method does not always create precise breakpoints at both *P*-element insertion sites, and excision of both *P*-elements, when it occurs, complicates the precise mapping of breakpoints.

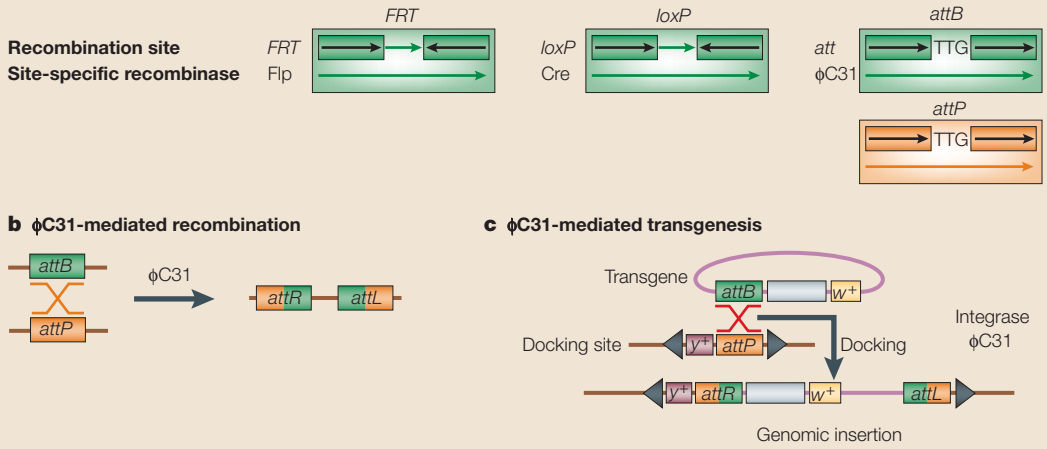
A second, more elegant technique, known as deletion-generator technology, was developed in the Gelbart laboratory[34]. This method is based on a hybrid transposable element, *P{wHy}*, that carries a *hobo* 'deleter' element that is flanked by two genetic markers, *YELLOW*$^+$ (y^+) and *WHITE*$^+$ (w^+), inside a *P*-element (FIG. 2b). *hobo* elements can undergo homologous recombination with other *hobo* elements in the presence of *hobo* transposase. So, replicative local *hobo* hopping, followed by recombination between *hobo* elements that are inserted in the same orientation, results in removal of all genes that are located between both *hobo* insertions. In addition, removal of one of the genetic markers, *yellow*$^+$ or *white*$^+$, allows the orientation of the deletion event to be determined genetically. Because one copy of *hobo* and a partial *P*-element remains, the deletion can be molecularly defined. This technology has been useful for generating nested deletions within 60 kb of the original insertion site, with endpoints that are staggered every 1–3 kb (REF. 34). Deletions that are as large as 400 kb have been reported[34,35]. The deletions that are obtained allow molecular mapping of lethal mutations in the vicinity of the original insertion site[35]. Moreover, overlapping deletions that are obtained from two neighbouring insertion sites provide an alternative method for investigating the mutant phenotype of uncharacterized genes that reside in the region between both insertions[35].

Finally, a third, powerful methodology is being used to create precise, molecularly defined deletions throughout the genome (FIG. 2c; see also BOX 1a). This effort was initiated independently in a private effort by Exelixis Inc.[9] and by DrosDel (see Online links box), a publicly financed European consortium[10]. Both have adopted the same approach that was initially developed in the Golic laboratory[4]. Two nearby transposable elements that are located in *trans* and that carry *FRT* sites in the same orientation create defined deletions after introducing Flp recombinase. Exelixis Inc. first created a library of almost 20,000 insertions, each containing an *FRT* site, and mapped the insertion sites by INVERSE PCR. A set of insertions was then selected to create 519 molecularly defined deficiencies, which average 140 kb in size and cover 56% of the genome[9]. The DrosDel project established about 3,300 molecularly defined insertions[10] and has currently generated 426 deletions with an average size of ~400 kb (S. Russell and J. Roote, personal communication). Combining both sets of molecularly defined deletions, we estimate that a genome coverage of ~80% has been achieved. These deletions will facilitate the mapping of mutations and the isolation of new genes that are identified in enhancer or suppressor screens. A current effort that was headed by Kevin Cook at the Bloomington *Drosophila* Stock Center aims to create ~1,500 more molecularly defined deletions that are on average 200 kb in size (K. Cook, personal communication). In summary, we can expect that 2,000 to 2,500 molecularly mapped deletions that will cover more than 95% of the genome will soon be one of the principal resources for fly biology.

Mapping improvements. Forward-genetic screens allow the unbiased isolation of mutants with an interesting phenotype. Genetic mapping of the corresponding loci and identification of the molecular lesions that are associated with these mutations are often difficult and labour intensive. Various methods are available for these purposes, including MEIOTIC RECOMBINATION MAPPING with marked chromosomes, MALE RECOMBINATION MAPPING and DEFICIENCY MAPPING[36,37]. Recently, two other methods have been introduced to accelerate the mapping of isolated mutations. These include SNP mapping[38–40] and mapping using defined *P*-element insertions[11].

The publication of the genome sequence of *D. melanogaster* allowed the identification of many

Box 1 | Site-specific recombinases



Site-specific recombinases (SSRs) catalyse recombination between two site-specific recombination sites[82,100]. SSRs are used to generate deletions and inversions, and can also be used to mediate the site-specific integration of a DNA sequence into the fly genome or the recombination-mediated exchange of a DNA cassette with a genomic sequence[82]. SSRs that are commonly used in different organisms are Flp, Cre (cyclic AMP-responsive element) and ϕC31 (REF. 82), which recognize *FRT*, *loxP* and attachment (or *att*) recognition sites, in corresponding order (panel a).

The importance of directionality

SSR recognition sites have a directionality, meaning that they can be left or right-orientated (long arrows in panel a). Directionality is important because it determines whether the outcome of the recombination reaction is a deletion or an inversion[82].

As shown in panel a, *FRT* and *loxP* sites each contain an inverted repeat of 13 bp (black arrows); the inverted repeat is separated by an 8-bp spacer (short green arrows), which defines the directionality of the recognition sites (long green arrows). Recombination between two *FRT* or *loxP* sites leaves behind two *FRT* or *loxP* sites. By contrast, the directionality of the *att* site (long arrows) is defined by the complete sequence of the two 84-bp *att* sites, *attB* and *attP*. The TTG consensus sequence that is found in the middle of the recognition sites of *attB* and *attP* is important for recombination; the two *att* sites otherwise do not share high levels of similarity. ϕC31 is therefore considered to be a true unidirectional integrase, because recombination between two essentially unrelated attachment sites results in two hybrid sites, *attL* and *attR* (panel b), which themselves are not a substrate for the integrase[81,100].

ϕC31-mediated integration

ϕC31-mediated integration is ideal for obtaining stable transgenic insertions[16] (panel c). A plasmid that contains an *attB* site, an exogenous sequence (grey) and a dominant marker, *white*$^+$ (w^+) integrates at defined docking sites in the genome, in this case one containing an *attP* site and another dominant marker, *yellow*$^+$(y^+). Defined docking sites are obtained after mobilization of a docking-site-containing transposon (dark grey triangles) at different locations in the genome. *white*$^+$ transgenic flies are only obtained after complete insertion of the exogenous sequence at the docking site; the *white*$^+$ selectable marker therefore allows the insertion to be traced.

DEFICIENCY MAPPING
Uses chromosomes that have different sections deleted to locate the position of a gene of interest. Without the deficiency, the normal functional gene usually masks the effect of (that is, complements) the defective or foreign copy that we wish to identify.

GENOTYPING
Comparative methodologies are used to obtain information about the sequence of a certain genomic region between different strains of the same species or between wild-type and mutagenized strains.

precisely mapped SNPs between chromosomes. They include base-pair changes, as well as single-nucleotide insertions or deletions. SNP mapping then allows a mutation to be mapped with respect to these alterations in the genome in a two-step process[38–40]. First, SNP maps are established or confirmed between two different genetic backgrounds. Second, SNPs are detected in mutant strains that have undergone meiotic recombination using high-throughput GENOTYPING. Therefore, flies that contain a mutation are crossed to a strain that has a different genetic background. Recombination products are identified by the presence or absence of specific SNPs, which allows the definition of the interval in which the mutation physically maps[38] (in this respect, the technique is similar to male recombination mapping[36]).

Visible markers such as bristle, eye and wing morphology are typically used as markers in classical genetic mapping. Because many defined *P*-element insertions contain the *white*$^+$ dominant marker, they allow the mapping of lethal mutations, as well as viable mutations with visible phenotypes[11] (FIG. 3). To map a mutation that is located on a known chromosome, two rounds of genetic mapping are required; a first round allows rough mapping to within 1 cM, which typically corresponds to a 400-kb physical interval. A second round of fine mapping usually narrows the interval to within 10–60 kb. Subsequently, overlapping primer sets are generated to amplify the DNA region and detect the associated molecular lesion. This method has proved to be cheap, efficient and very reliable in our laboratory[11,41] (H.J.B., unpublished observations).

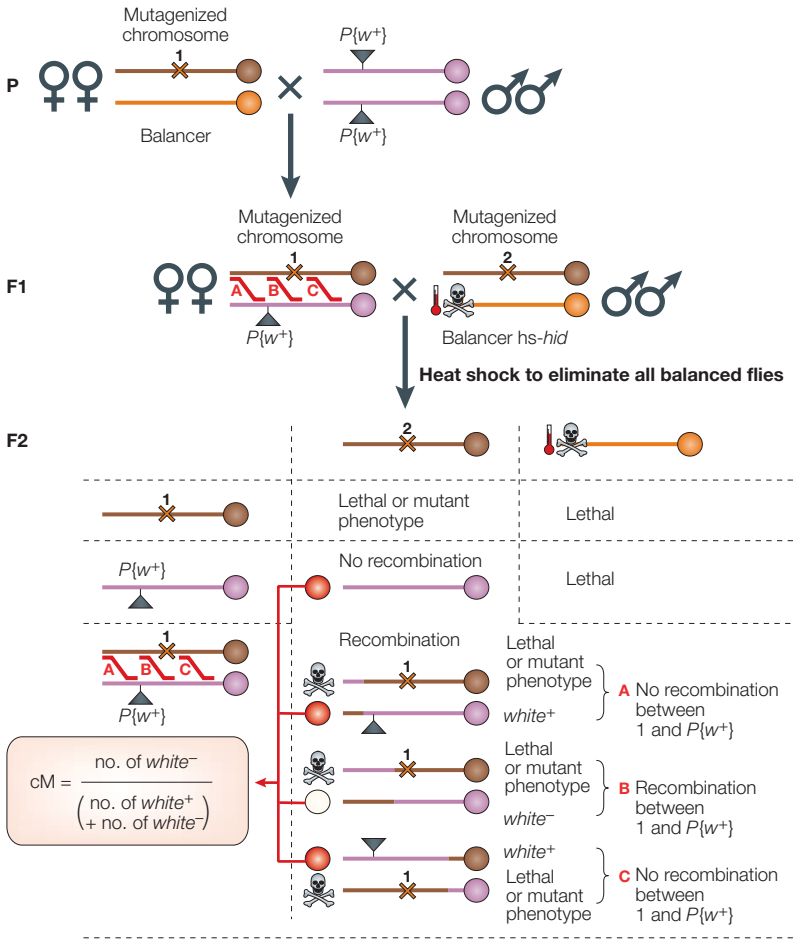


Figure 3 | **Transposon-mediated mapping.** A crossing diagram for mapping mutations in *Drosophila melanogaster* by using molecularly defined *P*-element insertions. Mutagenized chromosomes are shown in brown, chromosomes that contain *P*-element insertions in pink, and BALANCER CHROMOSOMES in orange. Balancer chromosomes contain a temperature-inducible head-involution defective (*hid*) gene, called hs-*hid*, which causes flies to die after heat-shock treatment. Two mutations, both of which are alleles of the same complementation group, are shown as brown crosses (indicated by 1 and 2) in the F1 generation. All chromosomes occur in a *white*$^-$ genetic background. In the F2 generation, transheterozygote flies and all flies that contain a hs-*hid* balancer, die or show the mutant phenotype. Only recombination between a molecularly defined *P*-element insertion and the mutation results in *white*$^-$ (white-eyed) flies (white circle). All other recombination and non-recombination events result in *white*$^+$ (red-eyed) flies (red circles). Calculating the ratio of *white*$^-$ flies with respect to the total number of flies allows the estimation of recombination distance (calculation shown in box). *P*{*w*$^+$}, *P*-elements that contain *white*$^+$.

New reverse-genetics approaches

In principle, two strategies can be used to obtain mutations — a forward- or a reverse-genetic approach. Forward-genetic approaches are driven by unbiased mutagenesis of the genome and phenotypic analysis, followed by mapping of the mutation and gene identification. Reverse genetics is driven by interest in a particular gene, and requires technology that allows selective disruption of a gene. The latter approach has gained a lot of popularity in the past 10 years, and the availability of the *D. melanogaster* genome sequence has given it an extra boost[5,19]. Three basic methods are now available for creating mutations in a gene of interest: imprecise excision of *P*-elements[19], homologous recombination[42] and TILLING[13,14]. RNAi can also be used, *in vitro* and *in vivo*, to remove gene function, and this has recently been adapted for systematic genome-wide approaches to study gene function in *D. melanogaster*[43–45]. In this section, we focus on two new strategies to create mutations in a gene: ends-out replacement gene targeting[12] and TILLING, which has only recently been applied in *D. melanogaster*[46] (S. Henikoff, B. Till and M. González-Gaitán, personal communication). The other methods have been covered extensively in other reviews[7,19,47].

Ends-out replacement gene targeting. Gene targeting is the modification of a genomic sequence by homologous recombination with an exogenous DNA fragment. This technique has been widely used in various multicellular model organisms[42]. Two forms of gene targeting have been used by the fly community: 'ends-in' or insertional gene targeting[48–50], and 'ends-out' or replacement gene targeting[12]. Ends-in and ends-out refer to the two arrangements of donor DNA that are used during gene targeting (FIG. 4). Both strategies are equally efficient, occurring in ~1 in 500 to 1 in 30,000 gametes[12]. In both methods, a transgenic donor line has to be established that contains the desired targeting construct flanked by *FRT* sites. In the first step, the Flp recombinase generates the extrachromosomal donor construct. In the second step, I-*Sce*I, a RARE-CUTTING RESTRICTION ENZYME that cleaves an 18-bp recognition site that is not present in the *D. melanogaster* genome, linearizes the targeting construct, which provides a recombinogenic template. The difference between ends-out and ends-in gene targeting is the position of the I-*Sce*I site(s) (FIG. 4). Ends-in gene targeting requires an I-*Sce*I site inside the targeting construct and results in a duplication of the targeted region, in which a *white*$^+$ dominant marker separates both parts of the duplication. Therefore, a true targeting event requires resolution of the duplication. This is mediated by I-*Cre*I, another rare-cutting restriction enzyme, which cuts a recognition site that is not present in the *D. melanogaster* genome between the duplicated fragments. Recombination between both parts removes the intervening sequence, including the *white*$^+$ dominant marker. In ends-out gene targeting, two I-*Sce*I sites are present outside each homology arm (a stretch of homologous sequence). Recombination between both homology arms results in the integration of the exogenous *white*$^+$ dominant marker and removal of the endogenous sequence.

The two strategies have different applications, advantages and disadvantages. Ends-in gene targeting allows specific mutations — deletions, insertions or point mutations — to be introduced on one or both of the homology arms (FIG. 4a). However, integration of the mutation without a duplication will only be obtained after two rounds of homologous recombination, and true removal of the gene is not obtained. By contrast, in ends-out gene targeting the gene is completely removed (FIG. 4b). Improved donor plasmids are available with a *white*$^+$ marker that is flanked by *loxP* sites (K. Golic, personal communication; see also the *Drosophila*

BALANCER CHROMOSOMES
Chromosomes that carry lethal mutations, dominant markers and multiple inversions and that are used in *trans* for a chromosome that carries a lethal or sterile mutation. They are used as genetic tools because they allow lethal mutations to be propagated indefinitely.

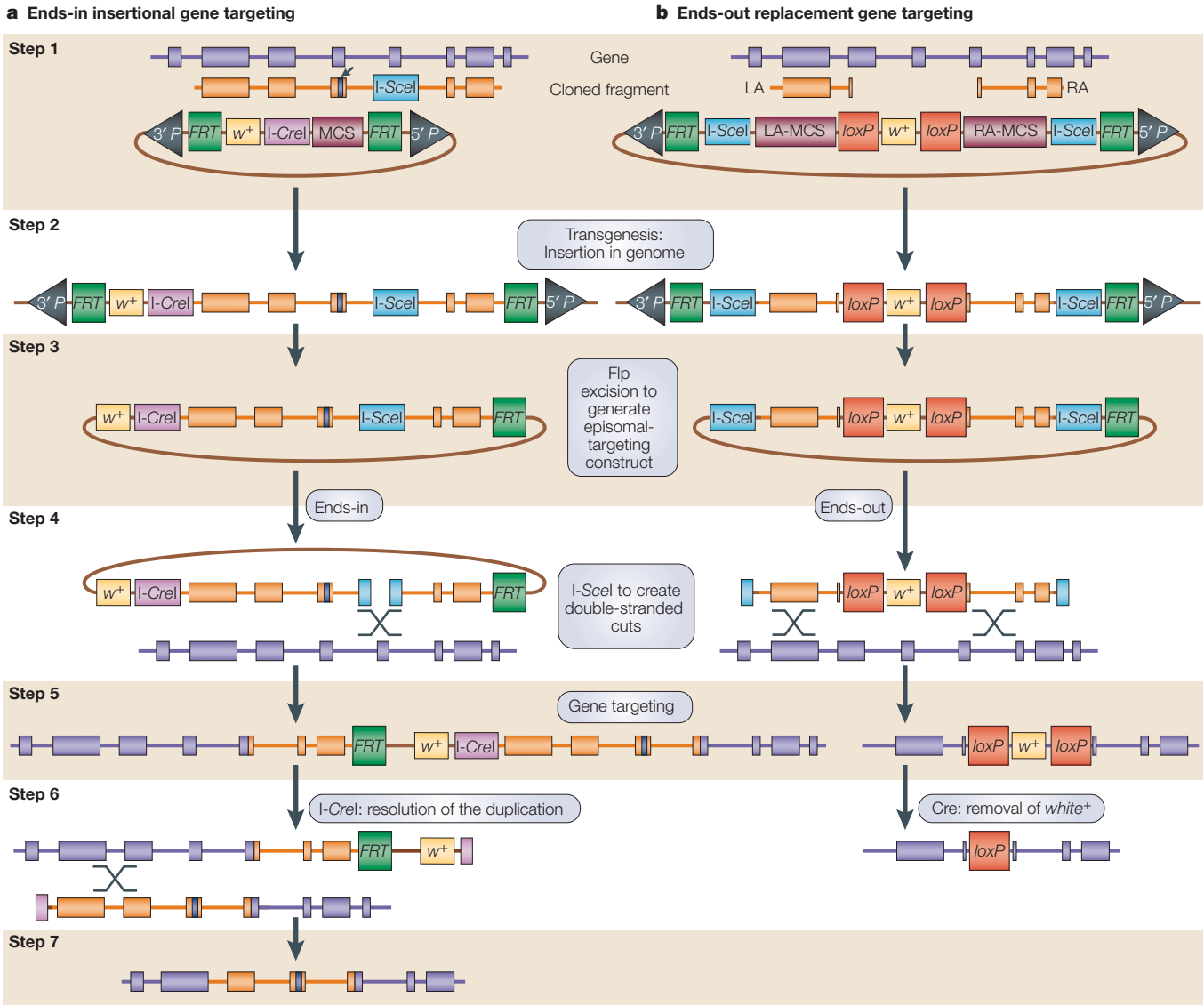


Figure 4 | **Homologous recombination. a** | Ends-in insertional gene targeting. In this strategy a genomic sequence of interest (orange) is altered to contain a desired mutation (blue band, indicated by an arrow) and an I-*Sce*I site (step 1). This fragment is then used to create a donor construct *in vitro* that contains the desired genomic region inserted in the multiple-cloning site (MCS) of a plasmid. The donor construct is also flanked on both sides by an *FRT* site, and contains an I-*Cre*I site on one side (shown in step 1). All necessary elements lie inside a *P*-element that contains the *white*$^+$ (w^+) marker. The donor construct that contains the genomic fragment is then introduced into the germline by *P*-element-mediated transformation (step 2). The targeting DNA is generated *in vivo* by Flp, which excises the donor as an extra-chromosomal circle (step 3), and I-*Sce*I, which makes a double-stranded break in the excised donor (step 4). Recombination with the endogenous target sequence generates a tandem duplication of the targeted region (step 5). One copy of the duplication carries the introduced mutation. The duplication is reduced to a single copy after a double-stranded break, which is induced by I-*Cre*I, and repair by homologous recombination (step 6). If successful, the single copy carries the introduced mutation (step 7). **b** | Ends-out replacement gene targeting. Two homology arms (left arm (LA) and right arm (RA)) of the desired genomic region (orange; step 1) are cloned *in vitro* into two multiple-cloning sites (LA-MCS and RA-MCS) of a plasmid (shown in step 1) to generate the donor construct. This construct is flanked on both sides by an I-*Sce*I site and an *FRT* site. A *white*$^+$ marker, flanked by *loxP* sites, is located between both homology arms. The donor construct is introduced into the germline by *P*-element-mediated transformation (step 2). The targeting DNA is generated *in vivo* by Flp, which excises the donor as an extra-chromosomal circle (step 3), and I-*Sce*I, which generates a linearized targeting construct (step 4). Recombination with the endogenous target sequence generates a localized insertion of *white*$^+$ in the desired genomic locus (step 5). The *white*$^+$ marker can be removed by Cre (cyclic AMP-response element) recombinase, resulting in only one *loxP* site at the replacement site (step 6).

RARE-CUTTING RESTRICTION ENZYME
Restriction enzymes that cleave a 16–18 bp recognition site. Most of these sites are not present in the host genome and can therefore be used for genome manipulations.

Genome Resource Center web page in the Online links box). The Cre (cyclic AMP-responsive element) recombinase can remove the *white*$^+$ gene, leaving the rest of the gene intact, with the exception of only one *loxP* site.

Gene targeting has recently attracted more interest in the *D. melanogaster* community, and ends-in gene targeting has been applied successfully for at least 17 loci[50–57]. The only unsuccessful attempt was for the

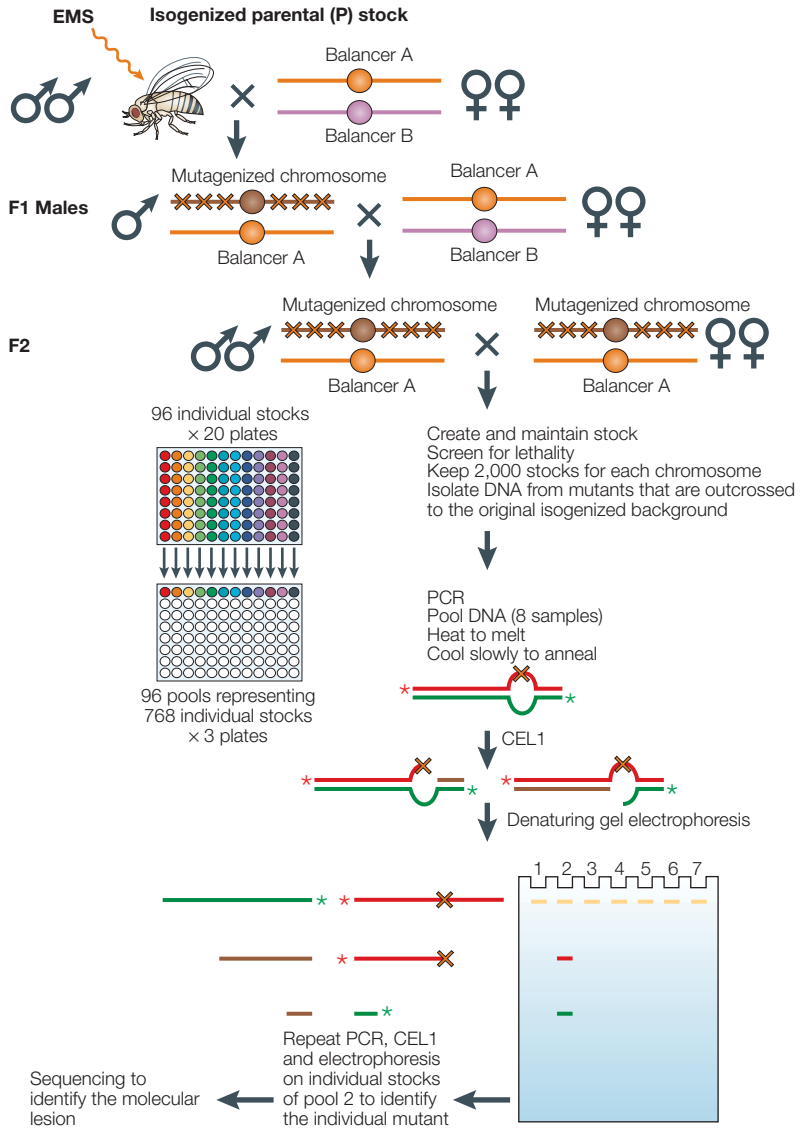


Figure 5 | **The application of TILLING in *Drosophila melanogaster*.** Isogenized male flies are mutagenized with 25–50 mM ethylmethane sulphonate (EMS) and crossed with balanced females. The F1 male progeny is then crossed with balanced females. F2 male and female progeny are sib-crossed to obtain a balanced stock. Each stock is screened for lethality to ensure the presence of at least one lethal mutation per chromosome. Approximately 2,000 homozygous lethal stocks should be obtained for each chromosome, corresponding to ~3,000 lethal mutations per chromosome. DNA is isolated and PCR is carried out on individual DNA samples using gene-specific primers. PCR products are pooled, heated and re-annealed to form DNA heteroduplexes. If a mutation is present in the amplified region, homoduplexes as well as heteroduplexes will form. CEL1, a single-strand-specific endonuclease, cuts heteroduplex DNA at mismatches. After denaturing gel electrophoresis, mutations will result in DNA fragments migrating faster compared with wild-type DNA. As PCR is carried out with two primers, each containing a differential end label (red and green), mutations are detected on complementary strands. If a mutation is present, both shorter fragments should add up to the length of the wild-type non-mutated fragment. PCR, CEL1 treatment and electrophoresis is repeated on DNA of individual stocks of a positive pool (in this case, pool 2) to identify individual mutants. Subsequent sequencing identifies the molecular lesion.

gene *synapse-associated protein of 47 kDa* (*Sap47*), indicating that some genes are unfortunately refractory to ends-in gene targeting[58]. The recently introduced, ends-out gene targeting was successfully used to target *yellow*[12], the myocardin-related transcription-factor gene, *DMRTF* (REF. 59), and the broadly expressed odorant-receptor gene, *Or83b* (REF. 60). Although both these technologies are clearly usable, the amount of work that is involved is still an important hurdle for many *D. melanogaster* biologists, especially when compared with the relative ease with which *P*-elements can be excised to create mutations in genes[7]. However, because as many as 20% of fly genes might be refractory to insertion by *P*- or *piggyBac* elements[6,8], gene targeting remains a valuable tool.

It has recently been shown that the efficiency of gene targeting in human cells and *D. melanogaster* is enhanced by zinc-finger nucleases[61,62]. These are hybrids between DNA-binding Cys_2His_2 ZINC FINGERS and DNA-cleavage domains. The zinc-finger component can be manipulated to recognize a broad range of sequences, and has therefore been used for targeted-chromosomal cleavage and mutagenesis in *D. melanogaster*[63]. The zinc fingers can recognize a unique site and induce small deletions and/or insertions precisely at the designed target by double-stranded cleavage and non-homologous end joining[63]. Homologous recombination at the double-stranded break is facilitated when a linear targeting construct, which is homologous to the flanking sequence, is provided[61,62].

Targeting-induced local lesions in genomes (TILLING). The TILLING method is based on the ability to quickly and efficiently identify single-nucleotide changes in specific genes[13,14]. TILLING was first used to identify mutations in *Arabidopsis thaliana*[64–66], but is now being adopted in other organisms such as zebrafish[67–69], rat[70], maize[71], the legume *Lotus japonicus*[72] and *D. melanogaster*[46] (S. Henikoff, B. Till and M. González-Gaitán, personal communication).

The protocol begins with the mutagenesis of an ISOGENIZED FLY STOCK with a chemical mutagen such as ethylmethane sulphonate (EMS). EMS mutations often result in G•C–A•T TRANSITIONS[73]. To obtain different mutations in *D. melanogaster*, including lethal ones, we propose the following strategy (FIG. 5). Approximately 2,000 balanced lethal stocks carrying heavily mutagenized chromosomes are created for the chromosome of interest. Each stock is outcrossed to the original isogenized chromosome. PCR is then carried out on the extracted DNA using a specific primer set in which each primer is labelled with specific fluorophores. The PCR products are pooled in groups of 8, heated and slowly cooled to allow heteroduplex formation between a mutagenized and a wild-type DNA sequence. The re-annealed DNA is treated with CEL1, a single-strand-specific endonuclease that cleaves any single base-pair mismatch[74,75]. To obtain high throughput for genomic applications, the cleaved DNA is run on modified sequencing gels to visualize mutations as differentially migrating DNA fragments[76]. Because each PCR primer is differentially labelled with a fluorophore, mutants produce two shorter bands after gel electrophoreses, each of which is labelled with its respective fluorophore. After identifying a positive pool, the cycle of PCR, CEL1 treatment

and electrophoresis is repeated on DNA of individual stocks to identify individual mutants. Subsequent sequencing identifies the molecular lesion. The number of point mutations that are typically detected for each gene is fairly high. A typical gene that encodes a 30–50-kD protein will yield an average of more than 10 point mutations if 2,000 stocks are screened (M. González-Gaitán, personal communication). However, about a third of these point mutations will be silent, and most amino-acid changes cause very subtle or no phenotypic changes.

TILLING is currently being applied on a genome-wide scale for *D. melanogaster* at the Fred Hutchinson Cancer Research Institute in Seattle (S. Henikoff and B. Till, personal communication). The flies that are being screened have recently been described[77]. In addition, a collection of 3,000 lethal mutations has been created at the Max Planck Institute in Dresden for TILLING purposes (M. González-Gaitán, personal communication).

An advantage of TILLING is that an allelic series of mutations is obtained[65]. Different mutations can affect different functions of a gene and therefore TILLING can reveal functions that are not always identifiable through forward genetics. One disadvantage of TILLING is that, as is the case for most other mutagenesis approaches, each fly stock is 'loaded' with mutations. Care should therefore be taken to carry out phenotypic analysis on 'clean' mutants. It is also recommended that two independent mutations are studied in heterozygote animals, and that the mutant phenotype is compared with that produced by molecularly defined deletions[9,10], which are true null alleles.

A variant of TILLING, known as Ecotilling, allows the identification of natural variations between different strains of the same species[14,78,79]. Ecotilling is useful for the high-throughput identification of SNPs, small insertions and deletions, and variation in small-repeat numbers between different strains. It therefore facilitates the construction of high-resolution SNP maps and aids the rapid mapping of mutations that are isolated through forward-genetic approaches (see previous section on mapping improvements).

Transgenesis

P-element-mediated transgenesis allows foreign DNA to be introduced efficiently into the *D. melanogaster* genome[1]. This has led to many important technological advances, including ENHANCER DETECTION using a *LacZ* reporter gene[80] and the PHENOTYPIC RESCUE of mutant phenotypes by genomic-rescue constructs[1]. However, there are three limitations that are associated with transposon-mediated transgenesis. First, it is difficult to transform flies with *P*-elements that are larger than 30 kb, and it is almost impossible with DNA fragments that are larger than 50 kb. Second, it is difficult to manipulate large DNA fragments, owing to the absence of unique restriction sites. Finally, integration at different sites in the genome can lead to POSITION EFFECTS, which might confound phenotypic analysis. Two methods should address these limitations: transgenesis that is mediated by the ϕC31 site-specific integrase, which integrates circular DNA at defined sites in the genome, and recombineering, which allows the cloning of large DNA fragments while minimizing the use of restriction enzymes and DNA ligase.

ϕC31-mediated transgenesis. ϕC31-mediated transgenesis is a recent addition to the *D. melanogaster* genetic arsenal[16]. As shown in BOX 1, the ϕC31 site-specific integrase mediates recombination between an attachment-B site (*attB*) that is normally present in the circular genome of *Streptomyces lividans*, and an attachment-P site (*attP*) that is normally present in the circular genome of the ϕC31 bacteriophage[81]. Recombination between the *att* sites integrates the phage genome and forms two hybrid attachment sites, *attL* and *attR*. As *attL* and *attR* are not targets for ϕC31, the integration event is irreversible unless a specific excision enzyme is expressed. This differs from Cre- or Flp-mediated recombination, which are both reversible processes[82] (BOX 1). Conveniently, the ϕC31 integrase requires no host- or cofactors and mediates both intra- and intermolecular recombination at a high frequency[81]. Recombination has been shown to take place efficiently in human and mouse tissue-culture cells[83–88] and *in vivo* in mice[89,90].

P-elements that contain wild-type *attP* sites have been integrated into the *D. melanogaster* genome and function as defined insertions sites for plasmids that contain *attB* sites[16] (BOX 1). The efficiency of ϕC31-mediated integration was high in this experiment, as 47% of the crosses resulted in transgenic progeny[16]. The efficiency of this transgenesis is approximately 5-fold higher than that mediated by *P*-transposase. In addition, because only one short attachment site in the plasmid is necessary for recombination, size limitations of the transgenic construct might be much less stringent than with *P*-element mediated transgenesis, in which the transposase has to recognize sites at either side of the *P*-element. However, this property has not been explored. Another main advantage of the ϕC31-mediated system is that plasmids will integrate into a predetermined location within the genome — that is, at an *attP* site — thereby reducing the problems that are associated with position effects. Although the ϕC31 integrase could potentially recognize pseudo *attP* sites in the *D. melanogaster* genome, integration at these sites has not been observed[16]. So, ϕC31 allows the efficient integration of genes at defined sites in the fly. This will allow a detailed comparison of various mutant constructs and might pave the way for elegant structure–function analyses. It should also allow us to carry out a better assessment of the role of regulatory elements in transcription *in vivo*, because constructs that contain different regulatory elements are integrated at an identical position in the genome.

Recombineering. One of the principal drawbacks of cloning defined DNA fragments into *P*-elements or plasmids is the limited availability of suitable restriction-enzyme recognition sites. Unfortunately, these become proportionally less abundant with increasing DNA

ZINC FINGER
A protein domain in which cysteine or cysteine–histidine residues coordinate a zinc ion. Zinc fingers are often used in DNA recognition and also in protein–protein interactions.

ISOGENIZED FLY STOCK
A starting strain of animals used for mutagenesis that contains identical pairs of chromosomes. Isogenization avoids the presence of pre-existing lethal and visible mutations.

TRANSITION
A mutation between two pyrimidines (T–C) or two purines (A–G).

ENHANCER DETECTION
A method that allows the identification of genes on the basis of their expression pattern. Engineered insertion elements carry a reporter-gene construct that is under the control of a minimal promoter that can respond to *cis*-acting regulatory elements near the insertion site.

PHENOTYPIC RESCUE
The ability of a DNA construct that contains the wild-type DNA sequence to rescue the phenotype of an identified genomic mutation.

POSITION EFFECT
The effect of the local chromosomal environment on the levels or patterns of transgene expression, possibly owing to local chromatin configuration or nearby *cis*-acting regulatory elements.

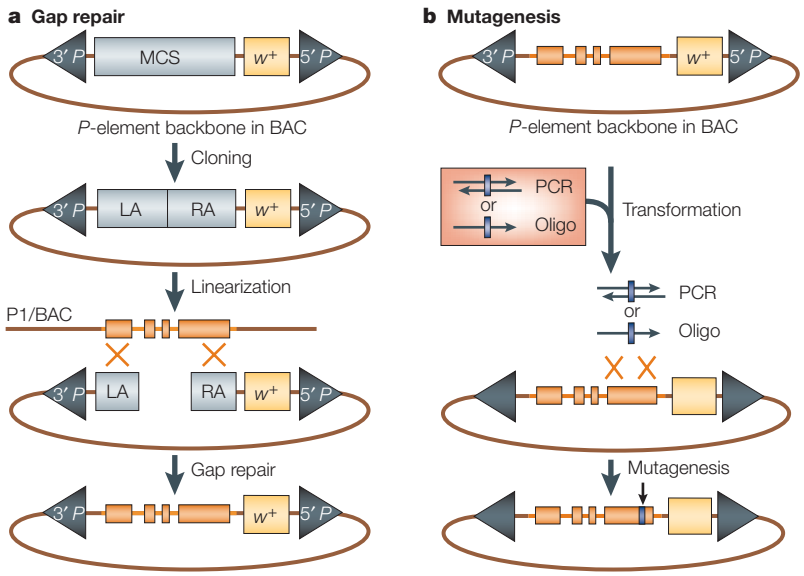


Figure 6 | **Recombineering. a** | Recombineering (recombination-mediated genetic engineering)-mediated gap repair for cloning in *Drosophila melanogaster*. To clone a desired fragment (orange boxes) from a BAC, PAC or any other target plasmid, two short homology arms, left arm (LA) and right arm (RA), at both ends of the target DNA are obtained by PCR (not shown). These are then ligated into the multiple cloning site (MCS), which resides inside a *P*-element backbone that also contains the *white*$^+$ marker (*w*$^+$) and both terminal repeats of the *P*-element (3′ *P* and 5′ *P*). This *P*-element is integrated into a conditionally amplifiable BAC. The plasmid is linearized between the LA and RA with a unique restriction enzyme, and used to transform recombination-competent bacteria. During recombination, the retrieval plasmid will obtain the desired fragment through gap repair from the target plasmid. Recombinant plasmids are selected, identified by PCR and verified by sequencing. **b** | Recombineering-mediated mutagenesis for structure–function analysis of genes. After obtaining a desired genomic fragment through gap repair inside a *P*-element backbone (according to the procedure shown in part **a**), mutations can be incorporated into the fragment by recombineering. The plasmid that contains a desired genomic DNA fragment is transformed into recombination-competent bacteria. After induction of the recombination functions, these bacteria are transformed with a PCR fragment or oligonucleotide that contain a desired mutation. Recombination will result in a mutagenized plasmid. As recombination efficiency is high, correct recombination events might be identified by PCR screening and verified by DNA sequencing. Mutations that are incorporated can be deletions, insertions or point mutations. Insertions allow the incorporation of protein tags, as well as fluorescent and colorimetric markers.

fragment size. Techniques have recently emerged that facilitate subcloning, as well as modification of large cloned inserts *in vivo*, without the need for restriction enzymes or DNA ligase[17,18]. This technology, commonly known as recombinogenic engineering or recombineering, makes use of the highly efficient, phage-based homologous recombination systems that are available in *Escherichia coli*. Many recombineering systems are available[17], and they have been used successfully to modify BACTERIAL ARTIFICIAL CHROMOSOMES (BACs) and P1 ARTIFICIAL CHROMOSOMES (PACs), as well as to subclone large DNA fragments from these vectors[17,18]. One highly efficient recombination system, which is mediated by bacteriophage λ, makes use of *E. coli* strains that contain temperature-inducible λ–recombination functions in their chromosome[91,92]. To subclone DNA, a BAC and a linearized-retrieval vector are introduced into the recombineering-competent bacteria. Gap repair of the linearized vector subclones DNA from BACs or PACs by homologous recombination between its DNA ends and the target DNA (REF. 92) (FIG. 6a). In addition, as shown in FIG. 6b, the technique allows the rapid modification of BACs through mutagenesis by using PCR products[92–96] or oligonucleotides[96–98] that contain the desired mutation as recombination templates.

Recombineering-mediated gap repair and mutagenesis have not yet been reported for *D. melanogaster* genes. However, many vectors that are currently used for fly transgenesis might be useful for recombineering. Our preliminary data show that this technology is a highly efficient mechanism for cloning DNA into modified *P*-element plasmids. Here, we briefly outline one possible application of recombineering in flies (K.J.T.V. and H.J.B., unpublished observation). The *P*-element sequences that are required for transgenesis, 5′ *P* and 3′ *P*, together with a multiple-cloning site and *white*$^+$ marker, can be integrated in a conditionally amplifiable BAC (FIG. 6a). This BAC has two origins of replication, thereby allowing the BAC to be present at single- or high-copy number, depending on the conditions and strain of bacteria used[99]. Low-copy number plasmids are useful for recombineering-mediated gap repair and mutagenesis, whereas high-copy number is necessary for cloning. As shown for gap repair (FIG. 6a), two homology arms, left arm (LA) and right arm (RA), that recognize the 5′ and 3′ ends of a desired insert, respectively, are integrated into the multiple-cloning site. Linearization between both homology arms allows the retrieval of DNA fragments from BACs or PACs by recombineering-mediated gap repair. Gap repair is carried out in a single-copy vector, allowing large inserts to be cloned and stably maintained. Because the retrieved DNA is flanked by 5′ *P* and 3′ *P* sites, these vectors can then be amplified[99] and used for *P*-element-mediated transformation.

Another important advantage of recombineering is that it allows targeted, single-nucleotide changes, as well as deletions and insertions, to be integrated easily into the cloned construct by incorporating a PCR fragment[92–96] or oligonucleotide[96–98] into recombinogenic bacteria (K.J.T.V. and H.J.B., unpublished observations) (FIG. 6b). This should allow a detailed structure–function analysis of the cloned insert *in vivo*. Finally, instead of using *P*-element-mediated transformation, adding an *attB* site in the conditionally amplifiable BAC that contains the *P*-element should allow ϕC31-integrase-mediated transgenesis. In conclusion, gap repair and subsequent mutagenesis of large DNA constructs in a conditionally amplifiable BAC, together with ϕC31-mediated transgenesis, might provide a new, powerful, general transgenesis platform for *D. melanogaster* in the near future.

Conclusions and future research

The continuing genome-wide efforts to generate a transposon insertion in every gene will create a valuable resource for the *D. melanogaster* community. This work will probably be complemented by genome-wide efforts to identify EMS-induced point mutations in the remaining genes by TILLING (FIG. 5). Both approaches should bring us a step closer to reaching SATURATION of the fly genome. Finally, homologous recombination provides a strategy to mutate those genes that are refractory to EMS and transposon-mediated mutagenesis (FIG. 4).

BACTERIAL ARTIFICIAL CHROMOSOME (BAC). A single-copy cloning vector that is derived from the F-factor of *Escherichia coli*. BACs can contain large genomic fragments. *Drosophila melanogaster* BACs carry an average insert size of 163 kb. Mapping positions of *D. melanogaster* BACs can be seen on the Flybase Genome Browser (see Online links box).

P1 ARTIFICIAL CHROMOSOME (PAC). A single-copy cloning vector that is derived from the F-factor of *Escherichia coli*. PACs can contain large genomic fragments. *Drosophila melanogaster* PACs carry an average insert size of 80 kb. The library represents a ~6-fold coverage of the genome.

SATURATION
The stage in mutagenesis at which mutations in new genes cannot be obtained by further mutagenesis. This occurs when at least one mutation in every gene has been obtained.

Our inability to study certain genes, to provide constructs for the phenotypic rescue of large genes, and to carry out systematic structure–function analyses should soon be alleviated. Recombineering allows large genomic-rescue constructs to be generated without the need for conventional cloning (FIG. 6a). In addition, it provides an alternative to mutagenesis and facilitates the structure–function analysis of large and nested genes (FIG. 6b). ϕC31-mediated transgenesis complements recombineering because mutagenized constructs can be targeted to the same chromosomal locus, circumventing the complications that are associated with position effects (BOX 1). In summary, these new technological developments will further strengthen *D. melanogaster* as a model organism and create even more opportunities to manipulate the fly genome.

Note added in proof

A recent report describes the generation of gene deletion by ends-in gene targeting[101].

1. Rubin, G. M. & Spradling, A. C. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353 (1982).
2. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
3. Xu, T. & Rubin, G. M. Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* **117**, 1223–1237 (1993).
4. Golic, K. G. & Golic, M. M. Engineering the *Drosophila* genome: chromosome rearrangements by design. *Genetics* **144**, 1693–1711 (1996).
5. Adams, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
6. Thibault, S. T. *et al.* A complementary transposon tool kit for *Drosophila melanogaster* using *P* and *piggyBac*. *Nature Genet.* **36**, 283–287 (2004).
This article describes the privately financed effort that generated *piggyBac* transposon insertions in many *D. melanogaster* genes.
7. Ryder, E. & Russell, S. Transposable elements as tools for genomics and genetics in *Drosophila*. *Brief. Funct. Genomic. Proteomic.* **2**, 57–71 (2003).
8. Bellen, H. J. *et al.* The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes. *Genetics* **167**, 761–781 (2004).
This paper describes the publicly financed effort that generated a *P*-element transposon insertion in about 40% of all *D. melanogaster* genes.
9. Parks, A. L. *et al.* Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. *Nature Genet.* **36**, 288–292 (2004).
The authors describe a privately financed effort that generated large numbers of deficiencies using *P*-elements and *piggyBac*s that contain *FRT* sites.
10. Ryder, E. *et al.* The DrosDel collection: a set of *P*-element insertions for generating custom chromosomal aberrations in *Drosophila melanogaster*. *Genetics* **167**, 797–813 (2004).
The authors discuss the generation of numerous transposon insertions that contain *FRT* sites. These insertions will be used to generate many molecularly defined deletions covering the entire *D. melanogaster* genome.
11. Zhai, R. G. *et al.* Mapping *Drosophila* mutations with molecularly defined *P* element insertions. *Proc. Natl Acad. Sci. USA* **100**, 10860–10865 (2003).
This paper describes the use of molecularly defined *P*-element insertions for the rapid and low-cost genetic mapping of mutations.
12. Gong, W. J. & Golic, K. G. Ends-out, or replacement, gene targeting in *Drosophila*. *Proc. Natl Acad. Sci. USA* **100**, 2556–2561 (2003).
The first report of ends-out replacement gene targeting in *D. melanogaster*.
13. Stemple, D. L. TILLING — a high-throughput harvest for functional genomics. *Nature Rev. Genet.* **5**, 145–150 (2004).
14. Henikoff, S., Till, B. J. & Comai, L. TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiol.* **135**, 630–636 (2004).
15. Zeidler, M. P. *et al.* Temperature-sensitive control of protein activity by conditionally splicing inteins. *Nature Biotechnol.* **22**, 871–876 (2004).
A report on the generation of a conditionally spliced intein, which can be incorporated in a host protein, making it temperature sensitive.
16. Groth, A. C., Fish, M., Nusse, R. & Calos, M. P. Construction of transgenic *Drosophila* by using the site-specific integrase from phage ϕC31. *Genetics* **166**, 1775–1782 (2004).
The first report of the use of ϕC31 integrase to obtain transgenic *D. melanogaster*.
17. Copeland, N. G., Jenkins, N. A. & Court D. L. Recombineering: a powerful new tool for mouse functional genomics. *Nature Rev. Genet.* **2**, 769–779 (2001).
18. Court, D. L., Sawitzke, J. A. & Thomason, L. C. Genetic engineering using homologous recombination. *Annu. Rev. Genet.* **36**, 361–388 (2002).
19. Adams, M. D. & Sekelsky, J. J. From sequence to phenotype: reverse genetics in *Drosophila melanogaster*. *Nature Rev. Genet.* **3**, 189–198 (2002).
20. Morin, X., Daneman, R., Zavortink, M. & Chia, W. A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in *Drosophila*. *Proc. Natl Acad. Sci. USA* **98**, 15050–15055 (2001).
21. Spradling, A. C. *et al.* The Berkeley *Drosophila* Genome Project gene disruption project: single *P*-element insertions mutating 25% of vital *Drosophila* genes. *Genetics* **153**, 135–177 (1999).
22. Beinert, N. *et al.* Systematic gene targeting on the X chromosome of *Drosophila melanogaster*. *Chromosoma* **113**, 271–275 (2004).
23. Handler, A. M. & Harrell, R. A. Germline transformation of *Drosophila melanogaster* with the *piggyBac* transposon vector. *Insect Mol. Biol.* **8**, 449–457 (1999).
24. Hacker, U., Nystedt, S., Barmchi, M. P., Horn, C. & Wimmer, E. A. *piggyBac*-based insertional mutagenesis in the presence of stably integrated *P* elements in *Drosophila*. *Proc. Natl Acad. Sci. USA* **100**, 7720–7725 (2003).
25. Horn, C., Offen, N., Nystedt, S., Hacker, U. & Wimmer, E. A. *piggyBac*-based insertional mutagenesis and enhancer detection as a tool for functional insect genomics. *Genetics* **163**, 647–661 (2003).
26. Bonin, C. P. & Mann, R. S. A piggyBac transposon gene trap for the analysis of gene expression and function in *Drosophila*. *Genetics* **167**, 1801–1811 (2004).
27. Bellen, H. J. Ten years of enhancer detection: lessons from the fly. *Plant Cell* **11**, 2271–2281 (1999).
28. Skarnes, W. C., Auerbach, B. A. & Joyner, A. L. A gene trap approach in mouse embryonic stem cells: the lacZ reported is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. *Genes Dev.* **6**, 903–918 (1992).
29. Lukacsovich, T. *et al.* Dual-tagging gene trap of novel genes in *Drosophila melanogaster*. *Genetics* **157**, 727–742 (2001).
30. Liu, X. Q. Protein-splicing intein: Genetic mobility, origin, and evolution. *Annu. Rev. Genet.* **34**, 61–76 (2000).
31. Paulus, H. Protein splicing and related forms of protein autoprocessing. *Annu. Rev. Biochem.* **69**, 447–496 (2000).
32. Yu, W. *et al.* Development of a comparative genomic hybridization microarray and demonstration of its utility with 25 well-characterized 1p36 deletions. *Hum. Mol. Genet.* **12**, 2145–2152 (2003).
33. Cooley, L., Thompson, D. & Spradling, A. C. Constructing deletions with defined endpoints in *Drosophila*. *Proc. Natl Acad. Sci. USA* **87**, 3170–3173 (1990).
34. Huet, F. *et al.* A deletion-generator compound element allows deletion saturation analysis for genomewide phenotypic annotation. *Proc. Natl Acad. Sci. USA* **99**, 9948–9953 (2002).
35. Mohr, S. E. & Gelbart, W. M. Using the *P{wHy}* hybrid transposable element to disrupt genes in region 54D–55B in *Drosophila melanogaster*. *Genetics* **162**, 165–176 (2002).
36. Chen, B., Chu, T., Harms, E., Gergen, J. P. & Strickland, S. Mapping of *Drosophila* mutations using site-specific male recombination. *Genetics* **149**, 157–163 (1998).
37. Lindsley, D. L. & Zimm, G. G. *The Genome of* Drosophila melanogaster (Academic, San Diego, 1992).
38. Berger, J. *et al.* Genetic mapping with SNP markers in *Drosophila*. *Nature Genet.* **29**, 475–481 (2001).
39. Martin, S. G., Dobi, K. C. & St Johnston, D. A rapid method to map mutations in *Drosophila*. *Genome Biol.* **2**, RESEARCH0036.1–0036.12 (2001).
40. Nairz, K., Stocker, H., Schindelholz, B. & Hafen, E. High-resolution SNP mapping by denaturing HPLC. *Proc. Natl Acad. Sci. USA* **99**, 10575–10580 (2002).
41. Verstreken, P. *et al.* Synaptojanin is recruited by endophilin to promote synaptic vesicle uncoating. *Neuron* **40**, 733–748 (2003).
42. Gong, M. & Rong, Y. S. Targeting multi-cellular organisms. *Curr. Opin. Genet. Dev.* **13**, 215–220 (2003).
43. Lum, L. *et al.* Identification of Hedgehog pathway components by RNAi in *Drosophila* cultured cells. *Science* **299**, 2039–2045 (2003).
44. Kiger, A. *et al.* A functional genomic analysis of cell morphology using RNA interference. *J. Biol.* **2**, 27 (2003).
45. Boutros, M. *et al.* Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303**, 832–835 (2004).
46. Bentley, A., MacLennan, B., Calvo, J. & Dearolf, C. R. Targeted recovery of mutations in *Drosophila*. *Genetics* **156**, 1169–1173 (2000).
The first report of targeted gene inactivation using ethylmethane sulphonate in *D. melanogaster*.
47. Carpenter, A. E. & Sabatini, D. M. Systematic genome-wide screens of gene function. *Nature Rev. Genet.* **5**, 11–22 (2004).
48. Rong, Y. S. *et al.* Targeted mutagenesis by homologous recombination in *D. melanogaster*. *Genes Dev.* **16**, 1568–1581 (2002).
49. Rong, Y. S. & Golic, K. G. A targeted gene knockout in *Drosophila*. *Genetics* **157**, 1307–1312 (2001).
50. Rong, Y. S. & Golic, K. G. Gene targeting by homologous recombination in *Drosophila*. *Science* **288**, 2013–2018 (2000).
51. Seum, C. *et al.* Isolation of *Su(var)3-7* mutations by homologous recombination in *Drosophila melanogaster*. *Genetics* **161**, 1125–1136 (2002).
52. Elmore, T., Ignell, R., Carlson, J. R. & Smith, D. P. Targeted mutation of a *Drosophila* odor receptor defines receptor requirement in a novel class of sensillum. *J. Neurosci.* **23**, 9906–9912 (2003).
53. Lankenau, S. *et al.* Knockout targeting of the *Drosophila Nap1* gene and examination of DNA repair tracts in the recombination products. *Genetics* **163**, 611–623 (2003).
54. Liu, H. & Kubli, E. Sex-peptide is the molecular basis of the sperm effect in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **100**, 9929–9933 (2003).
55. Dolezal, T., Gazi, M., Zurovec, M. & Bryant, P. J. Genetic analysis of the *ADGF* multigene family by homologous recombination and gene conversion in *Drosophila*. *Genetics* **165**, 653–666 (2003).
56. Sears, H. C., Kennedy, C. J. & Garrity, P. A. Macrophage-mediated corpse engulfment is required for normal *Drosophila* CNS morphogenesis. *Development* **130**, 3557–3565 (2003).
57. Egli, D. *et al.* Knockout of 'metal-responsive transcription factor' MTF-1 in *Drosophila* by homologous recombination reveals its central role in heavy metal homeostasis. *EMBO J.* **22**, 100–108 (2003).
58. Funk, N., Becker, S., Huber, S., Brunner, M. & Buchner, E. Targeted mutagenesis of the *Sap47* gene of *Drosophila*: flies lacking the synapse associated protein of 47 kDa are viable and fertile. *BMC Neurosci.* **5**, 16 (2004).
59. Han, Z., Li, X., Wu, J. & Olson, E. N. A myocardin-related transcription factor regulates activity of serum response factor in *Drosophila*. *Proc. Natl Acad. Sci. USA* **101**, 12567–12572 (2004).
60. Larsson, M. C. *et al. Or83b* encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron* **43**, 703–714 (2004).
61. Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764 (2003).
62. Porteus, M. H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).

63. Bibikova, M., Golic, M., Golic, K. G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* **161**, 1169–1175 (2002).
64. McCallum, C. M., Comai, L., Greene, E. A. & Henikoff, S. Targeted screening for induced mutations. *Nature Biotechnol.* **18**, 455–457 (2000).
The first article to describe the use of TILLING in *Arabidopsis thaliana*.
65. Till, B. J. *et al.* Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Res.* **13**, 524–530 (2003).
66. Greene, E. A. *et al.* Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. *Genetics* **164**, 731–740 (2003).
67. Wienholds, E., Schulte-Merker, S., Walderich, B. & Plasterk, R. H. Target-selected inactivation of the zebrafish *rag1* gene. *Science* **297**, 99–102 (2002).
68. Wienholds, E. *et al.* Efficient target-selected mutagenesis in zebrafish. *Genome Res.* **13**, 2700–2707 (2003).
69. Wienholds, E., Koudijs, M. J., van Eeden, F. J., Cuppen, E. & Plasterk, R. H. The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nature Genet.* **35**, 217–218 (2003).
70. Smits, B. M., Mudde, J., Plasterk, R. H. & Cuppen, E. Target-selected mutagenesis of the rat. *Genomics* **83**, 332–334 (2004).
71. Till, B. J. *et al.* Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol.* **4**, 12 (2004).
72. Perry, J. A. *et al.* A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* **131**, 866–871 (2003).
73. Ashburner, M. Drosophila*: a Laboratory Handbook and Manual* (Cold Spring Harbor Laboratory Press, Plainview, New York, 1989).
74. Till, B. J., Burtner, C., Comai, L. & Henikoff, S. Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Res.* **32**, 2632–2641 (2004).
75. Oleykowski, C. A., Bronson Mullins, C. R., Godwin, A. K. & Yeung, A. T. Mutation detection using a novel plant endonuclease. *Nucleic Acids Res.* **26**, 4597–4602 (1998).
76. Colbert, T. *et al.* High-throughput screening for induced point mutations. *Plant Physiol.* **126**, 480–484 (2001).
77. Koundakjian, E. J., Cowan, D. M., Hardy, R. W. & Becker, A. H. The Zuker collection: a resource for the analysis of autosomal gene function in *Drosophila melanogaster*. *Genetics* **167**, 203–206 (2004).
78. Comai, L. *et al.* Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J.* **37**, 778–786 (2004).
79. Smits, B. M., van Zutphen, B. F., Plasterk, R. H. & Cuppen, E. Genetic variation in coding regions between and within commonly used inbred rat strains. *Genome Res.* **14**, 1285–1290 (2004).
80. O'Kane, C. J. & Gehring, W. J. Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc. Natl Acad. Sci. USA* **84**, 9123–9127 (1987).
81. Thorpe, H. M. & Smith, M. C. *In vitro* site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc. Natl Acad. Sci. USA* **95**, 5505–5510 (1998).
82. Branda, C. S. & Dymecki, S. M. Talking about a revolution: the impact of site-specific recombinases on genetic analyses in mice. *Dev. Cell* **6**, 7–28 (2004).
83. Groth, A. C., Olivares, E. C., Thyagarajan, B. & Calos, M. P. A phage integrase directs efficient site-specific integration in human cells. *Proc. Natl Acad. Sci. USA* **97**, 5995–6000 (2000).
84. Belteki, G., Gertsenstein, M., Ow, D. W. & Nagy, A. Site-specific cassette exchange and germline transmission with mouse ES cells expressing ϕC31 integrase. *Nature Biotechnol.* **21**, 321–324 (2003).
85. Ortiz-Urda, S. *et al.* Stable nonviral genetic correction of inherited human skin disease. *Nature Med.* **8**, 1166–1170 (2002).
86. Ortiz-Urda, S. *et al.* ϕC31 integrase-mediated nonviral genetic correction of junctional epidermolysis bullosa. *Hum. Gene Ther.* **14**, 923–928 (2003).
87. Ortiz-Urda, S. *et al.* Injection of genetically engineered fibroblasts corrects regenerated human epidermolysis bullosa skin tissue. *J. Clin. Invest.* **111**, 251–255 (2003).
88. Thyagarajan, B., Olivares, E. C., Hollis, R. P., Ginsburg, D. S. & Calos, M. P. Site-specific genomic integration in mammalian cells mediated by phage ϕC31 integrase. *Mol. Cell Biol.* **21**, 3926–3934 (2001).
89. Olivares, E. C. *et al.* Site-specific genomic integration produces therapeutic Factor IX levels in mice. *Nature Biotechnol.* **20**, 1124–1128 (2002).
90. Hollis, R. P. *et al.* Phage integrases for the construction and manipulation of transgenic mammals. *Reprod. Biol. Endocrinol.* **1**, 79 (2003).
91. Yu, D. *et al.* An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **97**, 5978–5983 (2000).
92. Lee, E. C. *et al.* A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**, 56–65 (2001).
This is a description of a useful recombineering system that is temperature inducible, and that can be efficiently used for recombineering-mediated gap repair and mutagenesis.
93. Muyrers, J. P., Zhang, Y., Testa, G. & Stewart, A. F. Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res.* **27**, 1555–1557 (1999).
94. Muyrers, J. P. *et al.* Point mutation of bacterial artificial chromosomes by ET recombination. *EMBO Rep.* **1**, 239–243 (2000).
95. Yang, Y. & Sharan, S. K. A simple two-step, 'hit and fix' method to generate subtle mutations in BACs using short denatured PCR fragments. *Nucleic Acids Res.* **31**, e80 (2003).
96. Jamsai, D., Orford, M., Fucharoen, S., Williamson, R. & Ioannou, P. A. Insertion of modifications in the β-globin locus using GET recombination with single-stranded oligonucleotides and denatured PCR fragments. *Mol. Biotechnol.* **23**, 29–36 (2003).
97. Swaminathan, S. *et al.* Rapid engineering of bacterial artificial chromosomes using oligonucleotides. *Genesis* **29**, 14–21 (2001).
98. Zhang, Y., Muyrers, J. P., Rientjes, J. & Stewart, A. F. Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells. *BMC Mol. Biol.* **4**, 1 (2003).
99. Wild, J., Hradecna, Z. & Szybalski, W. Conditionally amplifiable BACs: switching from single-copy to high-copy vectors and genomic clones. *Genome Res.* **12**, 1434–1444 (2002).
The first article to describe a conditionally amplifiable BAC. This plasmid contains two origins of replication: one for single-copy number (useful for recombineering and clone maintenance), and another for inducible high-copy number in specific bacteria (to facilitate plasmid preparation).
100. Groth, A. C. & Calos, M. P. Phage integrases: biology and applications. *J. Mol. Biol.* **335**, 667–678 (2004).
101. Xie, H. B. & Golic, K. G. Gene deletions by ends-in targeting in *Drosophila melanogaster*. *Genetics* **168**, 1477–1489 (2004).

Acknowledgments

We thank K. Cook, L. Cooley, K. Golic, M. González-Gaitán, S. Henikoff, P. Hiesinger, K. Matthews, N. Perrimon, J. Roote, S. Russell, E. Spana, A. Spradling and B. Till for communicating unpublished results. We are particularly grateful to E. Seto and P. Verstreken for critical readings of the manuscript and K. Cook for suggestions. We apologize to colleagues whose work could not be cited owing to space limitations. H.J.B. is supported by the Howard Hughes Medical Institute, the US National Institutes of Health and NASA. We also gratefully acknowledge the constructive comments of two anonymous reviewers.

Competing interests statement

The authors declare no competing financial interests.

Online links

FURTHER INFORMATION

BACPAC Resource Center: http://bacpac.chori.org
Bloomington *Drosophila* Stock Center: http://flystocks.bio.indiana.edu
DrosDel *Drosophila* Isogenic Deficiency Kit: http://www.drosdel.org.uk
***Drosophila* GeneTag database:** http://flypush.imgen.bcm.tmc.edu/tagbase
***Drosophila* Genome Resource Center:** http://dgrc.cgb.indiana.edu
***Drosophila* TILLING Project:** http://tilling.fhcrc.org:9366/fly
Flybase Genome Browser: http://flybase.org/cgi-bin/gbrowse_fb/dmel
Flytrap — GFP Protein Trap database: http://flytrap.med.yale.edu
FlyView — expression pattern of enhancer traps: http://pbio07.uni-muenster.de
Hugo Bellen's laboratory: http://flypush.imgen.bcm.tmc.edu/lab
Model Systems Genomics: http://www.biology.duke.edu/model-system
NCI Frederick — Recombineering web site: http://recombineering.ncifcrf.gov
***P*-screen database:** http://flypush.imgen.bcm.tmc.edu/pscreen
Szeged *Drosophila* Stock Center: http://expbio.bio.u-szeged.hu/fly

Access to this interactive links box is free online.

ERRATUM

EMERGING TECHNOLOGIES FOR GENE MANIPULATION IN *DROSOPHILA MELANOGASTER*

Koen J. T. Venken and Hugo J. Bellen

Nature Reviews Genetics **6**, 167–178 (2005); doi:10.1038/nrg1553

In this article the Cre recombinase was incorrectly defined as a cyclic AMP-response element. This correction has been made to the online enhanced text and PDF version of this review.