

Chemical mutagens, transposons, and transgenes to interrogate gene function in *Drosophila melanogaster*



Koen J.T. Venken^{a,*}, Hugo J. Bellen^b

^a Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Program in Developmental Biology, Baylor College of Medicine, TX 77030, United States

^b Program in Developmental Biology, Departments of Molecular and Human Genetics, Department of Neuroscience, Howard Hughes Medical Institute, Jan and Dan Duncan Neurological Research Institute, Baylor College of Medicine, Houston, TX 77030, United States

ARTICLE INFO

Article history:

Accepted 19 February 2014

Available online 28 February 2014

Keywords:

Drosophila melanogaster

Chemical mutagens

Genetic screens

Mutation mapping

Transposons

Transposon mutagenesis

Transgenes

Genomic rescue

Enhancer bashing

Overexpression

ABSTRACT

The study of genetics, genes, and chromosomal inheritance was initiated by Thomas Morgan in 1910, when the first visible mutations were identified in fruit flies. The field expanded upon the work initiated by Herman Muller in 1926 when he used X-rays to develop the first balancer chromosomes. Today, balancers are still invaluable to maintain mutations and transgenes but the arsenal of tools has expanded vastly and numerous new methods have been developed, many relying on the availability of the genome sequence and transposable elements. Forward genetic screens based on chemical mutagenesis or transposable elements have resulted in the unbiased identification of many novel players involved in processes probed by specific phenotypic assays. Reverse genetic approaches have relied on the availability of a carefully selected set of transposon insertions spread throughout the genome to allow the manipulation of the region in the vicinity of each insertion. Lastly, the ability to transform *Drosophila* with single copy transgenes using transposons or site-specific integration using the Φ C31 integrase has allowed numerous manipulations, including the ability to create and integrate genomic rescue constructs, generate duplications, RNAi knock-out technology, binary expression systems like the GAL4/UAS system as well as other methods. Here, we will discuss the most useful methodologies to interrogate the fruit fly genome *in vivo* focusing on chemical mutagenesis, transposons and transgenes. Genome engineering approaches based on nucleases and RNAi technology are discussed in following chapters.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Drosophila melanogaster is an important model organisms long used to answer questions related to basic biology and human disease [84]. The sequencing of its entire genome [2] was the foundation for the development of a plethora of recent genetic technologies previously impossible to achieve [123,131,125,112,124]. These technologies have dramatically changed the landscape by which we approach biological questions in fruit flies. It is now possible to mutate, modify, and tag virtually every fly gene. Here we will discuss three technological approaches that have been instrumental towards scientific discovery using genome interrogation in the fruit fly. These are chemical mutagenesis, transposons and transgenes.

First, because we anticipate that there will be a renewed interest in unbiased mutagenic approaches, we will cover forward genetic screens using chemical mutagens, i.e., ethylmethane sulfonate (EMS) mutagenesis [16]. EMS screens have been used in numerous forward genetic screens based on a variety of phenotypic assays [113]. We provide outlines of typical one (F1), two (F2) and three (F3) generation screens and define recent developments that have made these screens more powerful. The section will end with strategies to map the mutant lesions and assign phenotypes to the appropriate gene.

Next, we will summarize transposon technologies. Transposons or transposable elements (TEs), historically dominated by the *P* element transposon, have been instrumental to the advancement of experimental fly biology [44]. To broaden the number of genes that can be manipulated by transposons, TEs with different insertion specificities, such as *piggyBac* [39], *Minos* [63], and *Tol2* [122], have been introduced [6]. We then discuss the different TE collections generated by individual labs and the Gene Disruption Project [7,6]. We expand on recent additions to the latter collection based on the *Minos* mediated integration cassette, better known as

* Address: Baylor College of Medicine, Houston, TX 77030, United States (H.J. Bellen).

E-mail addresses: kv134369@bcm.edu (K.J.T. Venken), hbellen@bcm.edu (H.J. Bellen).

MiMIC [130]. We outline how transposons are used in genetic screens and how they are mapped. In this section, we also explain how TEs can be used to generate large deletions with the help of the Flp recombinase [85,99,98,21], or modified by recombinase mediated cassette exchange (RMCE) [5] or site-specific integration [37,15], both catalyzed by the Φ C31 integrase.

Finally, we discuss transgenic methods based on random transposition or site-specific integration mediated by the Φ C31 integrase at defined *attP* sites distributed over the entire genome [37,15,127,67,117]. We focus on the latter system since it is highly efficient and dampens position effects [37,15]. *attP* sites have been tested and optimized for specific purposes, including genomic rescue of available mutations [28,127,126,129,117], genome-wide efforts to generate reagents for RNAi knockdown [26,77,76,78], interrogation of genomic regulatory elements [89,46,47,66], and gain-of function analysis through overexpression using different binary activation systems [14,102].

This review will not cover ends-in [94] or ends-out [35] gene targeting methods based on *in vivo* remobilization, nor will it cover custom nuclease-mediated genome editing (i.e., Zn finger, TALEN or CRISPR) [12,62,36]. Moreover, RNAi technologies based on double stranded RNA [50] or microRNA backbones (i.e., short hairpin RNAi or shRNAi) [38] will only briefly be discussed in the context of chemical and transposon mutagenesis. Gene targeting and RNAi strategies will be covered in depth in the following chapters: Genome Engineering Using Nucleases [148], and RNAi Screens [149].

2. Chemical interrogation

2.1. Ethylmethane sulfonate

Forward genetic screens based on chemically induced mutations are the least biased way to identify new players in many biological processes [113]. The most commonly used chemical mutagen is ethylmethane sulfonate (EMS) [57,113], which produces random DNA mutations through nucleotide substitution [16]. EMS is an alkylating agent that targets guanine and primarily produces point mutations. EMS is most commonly used in flies because it is effective and less toxic than most other chemical mutagens. The dose–response, germ cell sensitivity and effects of sperm storage in females have been well characterized [3]. In recent years EMS screens have reduced in popularity and this is largely due to the perception that mapping EMS-induced mutations is arduous [16]. Yet when gauged on cost, efficiency, scope, and lack of bias, EMS screens are still the most efficient of all screening technologies. Other advantages of EMS screens are their potential to be saturating and their ability to be performed in nearly any *Drosophila* lab [113].

On the contrary, genetically encoded RNAi is targeted to candidate regions of the mRNA (i.e., biased), and often results in partial knock-down and off-targeting effects as well [70]. Many transposons have inherent insertional specificity that commonly ignore significant portions of genes and genomes (i.e., they are biased), and 80% of the available TE insertions are weak loss of function alleles, possibly due to internal elements that can act as leaky promoters activating gene transcription downstream of the interruptive transposon insertion [56] (see also Section 3).

Hence, EMS is the preferred mutagen for true unbiased genotypic probing of the genome and phenotypic screening of your favorite phenotype. EMS is often used at a concentration of 25 mM [16], though lower concentrations (e.g., 10 mM) result in fewer second site hits [142]. Hexamethylphosphoramide (HMPA), an alternative chemical mutagen, introduces small deletions via cross-linking, but it is not commonly used [72]. X-rays and gamma

rays have previously been used extensively as they induce genomic rearrangements, including inversions, deletions and duplications. However irradiation produces many fewer mutations than EMS and is therefore not so popular anymore. *N*-ethyl-*N*-nitrosourea (ENU), though frequently used in other organisms, is not favored in *Drosophila*, probably because it is too toxic [3]. Other mutagens are discussed elsewhere [3]. In the next few sections, we discuss the phenotypic assays used to identify EMS-induced mutations as well as the three commonly used types of EMS screens: the single (F1), two (F2), and three (F3) generation screens.

2.2. Phenotypic assays

The phenotypic assay used in any forward unbiased EMS screen is paramount to its success, an often underappreciated fact. In most cases the goal of the screen is to isolate as many mutations in as many different genes as possible that affect the phenotype of interest, ideally identifying at least two alleles for each gene. Although we recommend using low concentrations of EMS (e.g., 10 mM) to reduce the number of second site mutations (which can hinder mapping and cloud interpretation at later stages), this will raise the number of flies to be screened although it is already often in the many thousands to tens of thousands range. Hence, primary screening assays should be quick and easy to perform, enabling one to screen hundreds or even thousands of chromosomes per week. In a well-designed screen, the primary phenotypic assay should be followed by secondary screens. These can be more tedious than the primary screen assay as typically no more than 1–2% of flies from the primary screen will be retained.

Simple phenotypic assays are the most powerful because they are easy to perform and can be completed quickly. A classical example is the assay developed to identify flies that are impaired in phototaxis [10]. One hundred mutagenized flies are loaded into a countercurrent phototaxis tool, tapped down toward the dark side of the apparatus, and then allowed to walk towards a light source. Flies that walk to the light source are separated from those that do not, and the process is repeated six times. When combined with an electroretinogram-based secondary screen, this assay previously permitted the isolation of genes that affect synaptic transmission [132]. Another powerful assay is based on flight capabilities. One hundred mutagenized flies are added to a cylinder coated with oil. Flies that were unable to fly fell to the bottom of the container and were collected for analysis of mutations affecting flight muscles [51]. Given the plethora of genetic tools that are currently available [131], more sophisticated screens can be designed that allow for equally simple phenotypic probing. They can be based on the localization of a GFP in a specific cell or cellular compartment. For example, by expressing a CD8-GFP-Shaker fusion protein at the *Drosophila* neuromuscular junction [145], third instar larva were rapidly screened through the cuticle for morphological defects of the NMJ [1].

2.3. Single (F1) generation EMS screens

The simplest and fastest EMS screens are single (F1) generation screens. These screens can be used to isolate, (1) behaviorally defective flies that carry visible X chromosome-linked mutations [10], (2) dominant suppressors and enhancers of existing mutations [107], or (3) new mutations on single chromosome arms via clonal analysis [116,75].

To isolate viable mutations on the X-chromosome, mutagenized males are crossed to attached X chromosome females [59], yielding male progeny with a paternally-derived X-chromosome. Flies with visible mutations or defective behaviors can be easily isolated [10].

In suppressor and enhancer screens, one typically starts with a known mutation that causes a subtle phenotype but which can

easily be enhanced or suppressed. A classical example was performed by using a temperature sensitive mutation of *sevenless* [107]. Normal development of the R7 photoreceptor occurred at 24 °C but not at higher temperatures. Loss of a single copy of another gene could then enhance the phenotype, observed during an F1 screen. The screen was very successful and permitted isolation of seven genes required in photoreceptor development. Similar screens can be used to identify suppressors.

Currently, most F1 screens are performed for the autosomal chromosome arms (2L, 2R, 3L, 3R) that carry an FRT site in close proximity of the respective centromeres [140] (Fig. 1). Isogenized males carrying the appropriate FRT-containing chromosome arm are mutagenized with EMS and crossed to females carrying the same FRT chromosome and a transgene driving FLP recombinase [34] in a tissue of interest. Subsequent FLP-mediated mitotic recombination results in the exchange of entire chromosome arms during cell division. This permits the generation of homozygous mutant tissue in an otherwise heterozygous animal, therefore limiting the effect of a potentially detrimental mutant phenotype at early developmental stages. The progeny of mutagenized males can be directly screened for a phenotype in the next generation. Desired mutations are isolated and immediately balanced to generate stable stocks. F1 screens are typically based on organ or cell-specific FLP expression driven by tissue-specific regulatory elements [75] or through the binary GAL4/UAS system [17], i.e., an organ- or tissue-specific GAL4 driving a UAS-FLP transgene [116]. The most widely used tissue is the eye since it is large, easily scorable, and excellent drivers like *eyeless*-FLP are available [75]. To obtain clones of sufficient size, it is important to use a driver expressed early in development. The *eyeless* regulatory elements fulfill this requirement since the gene is already expressed in embryos. The size of the mutant clone can be enlarged by using a cell lethal mutation or a Minute (i.e., a mutation often in ribosomal proteins affecting protein translation and therefore general cell health) on the homologous chromosome so that wild type clones are unable to propagate equally well [75]. The large clones in the eye allow to screen for defects during eye development [75], tissue overgrowth [48], bristle development [119], synaptic transmission [132] and neuronal wiring [68]. Other tissues can be analyzed using this technology as well, e.g., the *Ubx*-Flp transgene generates clones in all the imaginal discs among other tissues [45].

The FLP/FRT F1 screens are fast but require that heterozygous animals can be used to breed. Hence, the screen cannot be intrusive. Another disadvantage of F1 screens is that not all induced mutations breed true as DNA repair can occur in germ cells and the somatic mutations are therefore not always inherited [3]. The advantages are that F1 screens are fast and comprehensive if the primary assay is simple. A high throughput and simple assay is extremely important for powerful screens and isogenization of the chromosome of interest prior to mutagenesis is paramount [16].

2.4. Two (F2) generation EMS screens

Two (F2) generation screens are often designed to generate specific mutations and are thus more targeted than screens examining whole chromosomes or chromosome arms. The exception is the X-chromosome. Males only carry a single copy and hence F2 screens have been used to isolate viable behavioral mutations on the X-chromosome [54]. For the latter, mutagenized males are crossed to attached X females and single male progeny are backcrossed to attached X females to expand the flies carrying the mutagenized chromosome. The male progeny is then scored for behavioral phenotypes like diurnal rhythmicity [54], learning and memory [27], or defects in vision based on ERGs [83,43].

In many F2 screens, the purpose is to isolate new EMS induced mutations that fail to complement genes uncovered by a deletion [139,138] or another mutation, to isolate other alleles of a gene and to initiate structure function analyses (Fig. 2). Males are mutagenized and crossed to females that carry a second or third chromosome balancer. Single males are then crossed to a stock carrying a mutation or a deletion (and a balancer) in which the gene of interest is uncovered. Failure to complement the deletion or the mutation permits identification of specific alleles of a gene that are located within the deletion [139,138,60,24]. The inclusion of easily scorable markers (e.g., *Tubby*) can facilitate the screening process tremendously. Through complementation tests and sequencing of the gene of interest new alleles can be identified. This allows an unbiased recovery of new alleles that permit a structure function analysis [61].

2.5. Three (F3) generation EMS screens

Three (F3) generation screens are the foundation of some of the most probing experiments in developmental biology and neuroscience [80,113,9]. They remain a very powerful, if somewhat underappreciated, tool. With the advent of whole genome sequencing (WGS), they are likely to resurge in popularity since mutations

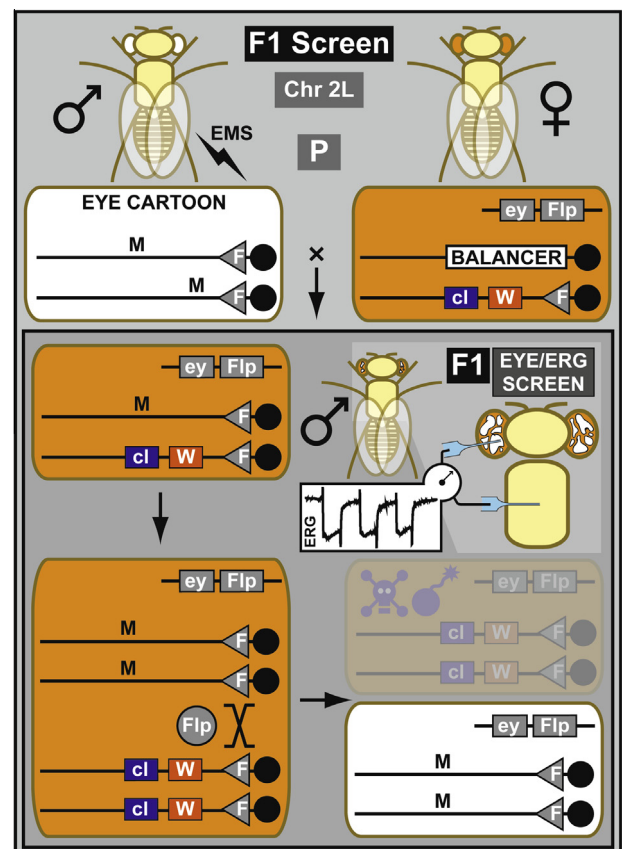


Fig. 1. Outline of a mitotic F1 screen for chromosome 2L. Mutagenized (M) males containing a chromosome 2 with an FRT site on arm 2L are crossed to females containing the same FRT carrying chromosome but with an additional dominant eye color marker, *white*⁺ (W), and recessive cell lethal (*cl*). In addition, the females carry a transgene that drives the FLP protein under the control of *eyeless* promoter (*ey*). This cross is the parental (P) generation. In the next generation (F1), the FLP generates mitotic clones exclusively in the eye. Due to the presence of a recessive cell lethal, homozygous WT clones die, and many ommatidia become homozygous mutant, facilitating phenotypic analyses based on morphology or an electrophysiological assay like the electroretinogram (ERG). Flies are recovered, and propagated for stock keeping and further analysis.

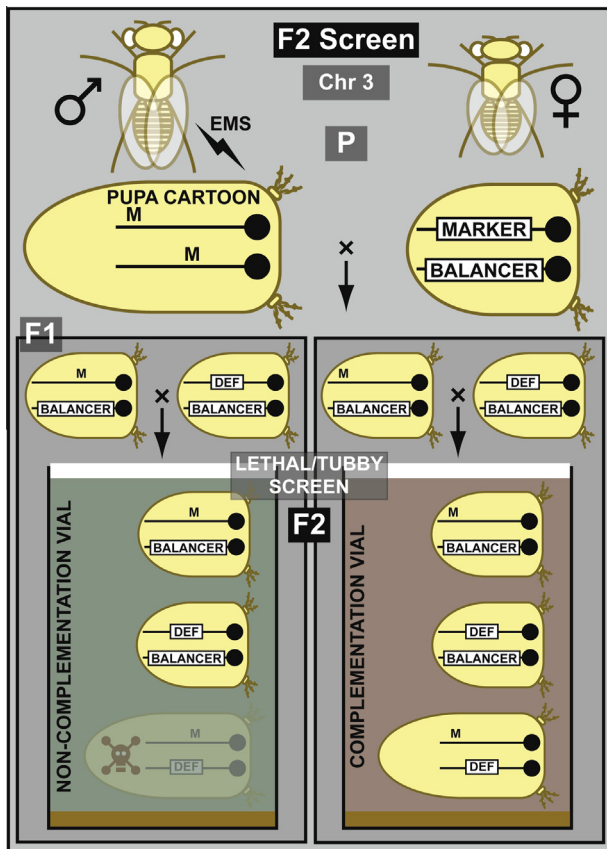


Fig. 2. Outline of a lethal phenotype F2 screen for chromosome 3. Mutagenized (M) males are crossed to balanced third chromosome *Tubby* females. In the next generation, balanced mutants are crossed to a balanced third chromosome deficiency to isolate mutations that fail to complement the deficiency. In the second generation (F2), progeny are scored for prepupal lethality. The balancer is marked with *Tubby*, hence the short *Tubby* females and pupae. The inclusion of the *Tubby* (*Tb*) marker facilitates this process since scoring can be performed immediately by looking at the pupae in the vial. The green vial carries the mutation of interest.

can be identified more easily. Recently, an extensive large-scale F3 mosaic screen was performed for the X chromosome [142] (Fig. 3). Isogenized males carrying an *y w FRT* chromosome were mutagenized with lower EMS concentrations than are typically recommended (10 mM versus 25–50 mM) [16] and crossed to balanced females. Individual mutant balanced female progeny were crossed back to balanced males to establish stocks. Chromosomes that carried homozygous lethal mutations were retained and screened in different assays by creating mutant clones in the eye or thorax. Mutations in numerous genes not previously identified or characterized on the X chromosome were discovered (Bellen, unpublished data).

2.6. Mapping of EMS mutations

Mapping EMS mutations has historically been tedious. This is no longer the case since mapping can now be done on a large scale. Mapping a few dozen genes from an EMS screen should be within reach of almost any fly lab. Depending on the chromosomal location, alleles are mapped by a variant of meiotic mapping (for the 2nd and 3rd chromosome) [143] or a variant of duplication mapping (for the X chromosome) [23,129].

Meiotic mapping is very easy to perform with a set of defined *P* element insertions [143] (Fig. 4). To map a mutation located on a known chromosome arm, two rounds of meiotic mapping are per-

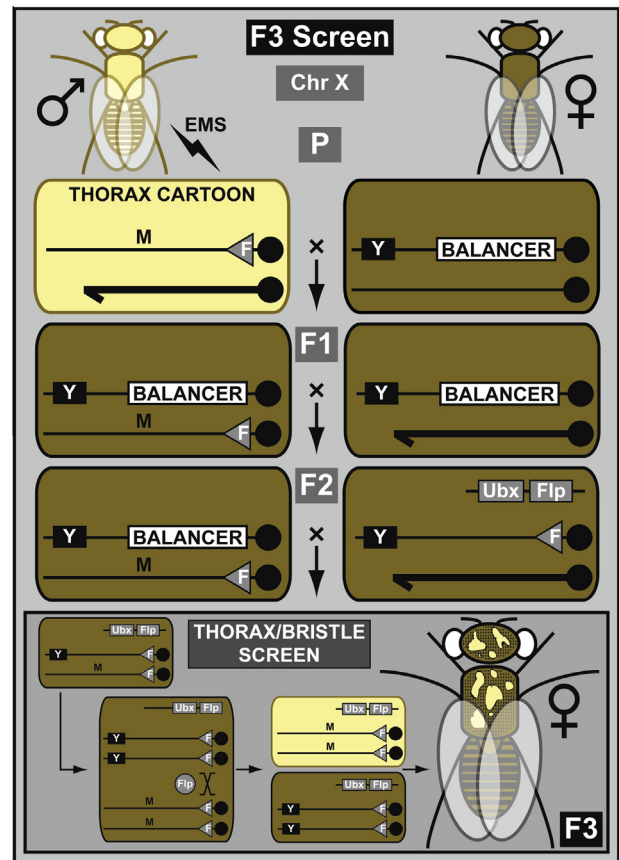


Fig. 3. Outline of a lethal phenotype F3 screen for the X chromosome. Mutagenized (M) males containing chromosome X with an *FRT* site are crossed to balanced X chromosome females. Mutant balanced females are backcrossed to balanced males to generate stocks. Stocks are used to establish X chromosome lethality. Balanced lethal mutant females are crossed to males containing the same *FRT* chromosome and a dominant body color marker, *yellow⁺* (*Y*). In addition, males have a transgene that drives the FLP protein under the control of an imaginal disc regulatory element (*Ubx*). In the next generation (F3), the FLP generates mitotic clones exclusively in the imaginal discs that can be analyzed for a morphological phenotype (e.g., bristle phenotype).

formed. In a first mapping round (i.e., rough mapping), the locus can be mapped to a 400 kb interval using 10–15 defined *P* elements whose mapping position is known. Stocks for rough mapping of mutations on the second and third chromosomes have been carefully selected so that they show strong expression of the transgenic *white⁺* marker (see also Section 3.2), resulting in a red adult eye which facilitates screening when scoring large populations. Mapping kits for each autosomal chromosome arm (i.e., 2L, 2R, 3L and 3R) are available from the BDSC (<http://flystocks.bio.indiana.edu/Browse/misc-browse/Baylor-kits.htm>). A second round of mapping permits pinpointing the mutation to a physical interval of less than 100 kb [143], but this mapping resolution is no longer required as one can use next generation sequencing approaches to analyze the data of the 400 kb interval.

Although meiotic mapping with *P* elements is feasible for the X chromosome, it becomes more complicated for lethal mutations since they cannot be recovered in males. This problem is easily circumvented by duplication mapping. Two independent duplication kits have been generated for the X chromosome [129,23]. One duplication kit was generated by sophisticated chromosomal engineering of an attached X–Y chromosome [23] (Fig. 5A), while the second duplication kit was generated by site-specific integration of an overlapping set of *P*[*acman*] clones into an *attP* site located on the third chromosome [129] (Fig. 5B) (see also Section 4.3).

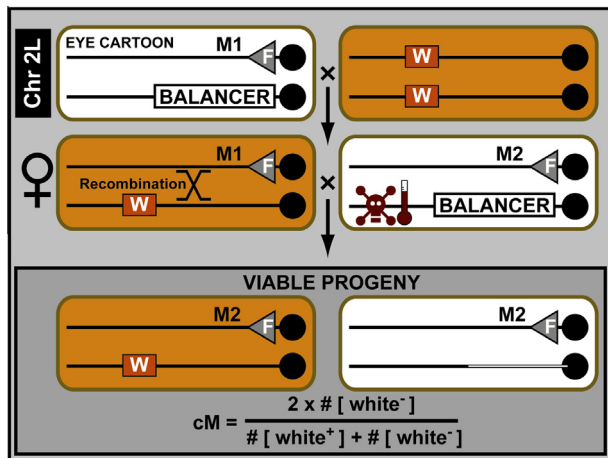


Fig. 4. Meiotic mutation mapping using *P* element insertions. A balanced unmapped lethal mutation (M1), located on chromosome arm 2L, is crossed to a mapped viable *P* element insertion marked by the dominant eye color marker *white*⁺ (W). Next, female balanced mutant progeny are crossed to males containing a second lethal mutation (M2) belonging to the same complementation group maintained over a conditionally lethal balancer chromosome (i.e., temperature induced lethality). During early development of the progeny of this cross, a heat shock removes all unwanted progeny containing the conditionally lethal balancer chromosome. The transheterozygous M1/M2 flies die as well. Since recombination between the unmapped lethal mutation (M1) and the mapped viable *P* element insertion can occur in the female germ line, the ratio of *white*⁺ flies compared to the total number of flies in the resulting progeny is a measure of the genetic distance between the unmapped mutation and the mapped *P* element insertion. Performing this crossing scheme with ~10 mapped *P* elements located on the same chromosome arm allows mapping to within 100 kb–1 Mb depending on the region of the chromosome where the mutation maps. One can then proceed with WGS for mutation identification.

The two duplication collections complement each other well. Using the large duplications (attached-XY) [23], mutations can be mapped to a 1 Mb region. Subsequent use of smaller duplications, deficiencies, and the tiled P[acman] array of duplications [129], allows one to easily narrow this interval to less than 50 kb. Sanger sequencing of overlapping PCR products ultimately reveals the gene of interest [143]. This strategy has proven to be efficient in mapping loci derived from an ongoing, large-scale X chromosome mutagenesis effort. However, the majority of mutations of this screen were mapped using WGS after rough mapping to an area of about 1–1.5 Mb, thereby greatly accelerating the gene identification process (Bellen, unpublished data).

3. Transposon interrogation

3.1. Transposons

TEs are insertional mutagens that can be easily mapped using a variety of molecular methods (see also Section 3.2). The most commonly used TE in the fly field is the *P* element, followed by *piggyBac*, and *Minos* [100]. Wild type *P* elements are composed of two 31-mer terminal inverted repeats, *P* transposase (the enzyme required to catalyze transposition), as well as other sequences essential for transposition [20] (Fig. 6A). Physical separation of the *P* transposon backbone from the *P* transposase has permitted the development of a binary transformation system [97] (Fig. 6B), allowing controlled transposition of engineered transposons into the host genome using enzyme provided *in trans*, e.g., as a helper plasmid during microinjection or as a transgene through a genetic crossing scheme [124]. Although proven to be incredibly useful, *P* elements also have drawbacks. The integration sites of *P* elements are strongly biased toward origin of replication sites [108] and the

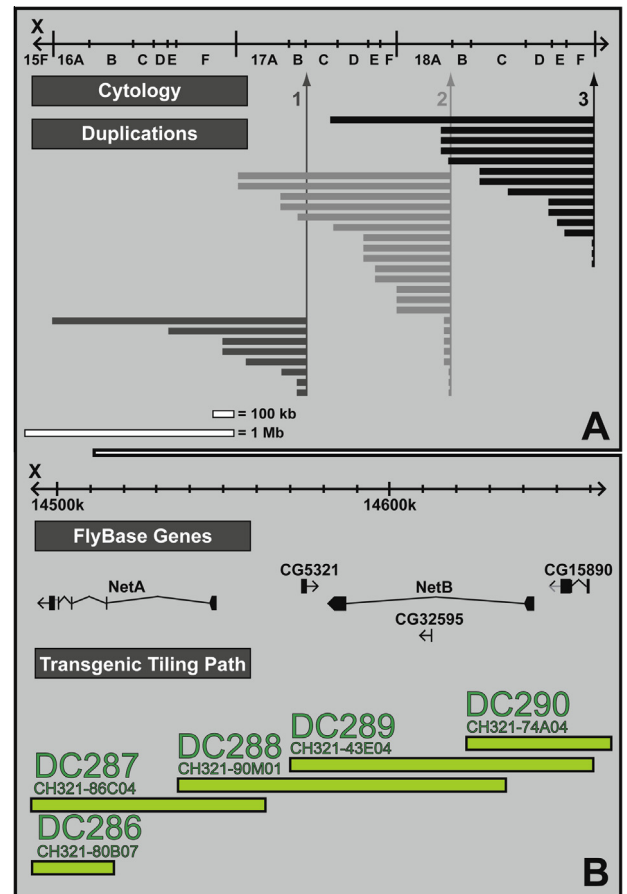


Fig. 5. Mutation mapping using duplication complementation. (A) Nested duplication mapping. A nested set of duplicated segments is generated from the same progenitor (e.g., 1). Different progenitors (i.e., 1, 2 and 3) then result in overlapping sets of nest duplications that cover majority of the X chromosome. Mutations can then be mapped by the large segments in a first round of genetic mapping and fine mapped by a nested set of duplications in a second round of genetic mapping. (B) Transgenic duplication mapping. The CHORI-321 P[acman] library contains DNA clones (CH321) that have an average insert size of 83 kb. A minimum tiling path (indicated in green) of 408 overlapping P[acman] clones for the entire 22 megabase X chromosome of *Drosophila* was selected and integrated in the same docking site on chromosome arm 3L, using the ΦC31 integrase. This transgenic duplication kit can be used to map uncharacterized X chromosome mutations. Illustrated is a small portion of the transgenic duplication kit surrounding the *Netrin-A* and *Netrin-B* genes.

5' ends of genes [6]. This insertional specificity reduces the effectiveness of mutagenesis screens and limits the number of insertions that can be used to create gene- or protein-traps [18,93,65]. Additionally, many insertions do not or only partially affect the genes near which they are inserted [79,56]. The latter is also a very important benefit in reverse genetic strategies. The lack of phenotype permits one to assess the loss of function associated with an imprecise excision of the *P* element that affects or removes the nearby gene without having to worry too much about second site mutations present on many chromosomes.

The features of *P* elements have led to the exploration of the use of other TE such as *piggyBac* [39] and *Minos* [63] (Fig. 7). Unlike *P* elements, *piggyBac* preferentially targets TTAA sequences whereas *Minos*, a Tc1/Mariner-like transposon, inserts at random [108]. The major drawback of *piggyBacs*, unlike *P* elements and *Minos*, is that they do not excise imprecisely [137], and hence do not permit the creation of other mutations in the gene of interest. Both *piggyBac* and *Minos* have been used to generate genome-wide TE collections. The *Tol2* TE has also been demonstrated to mobilize in flies [122] but has not been used in any genome wide applica-

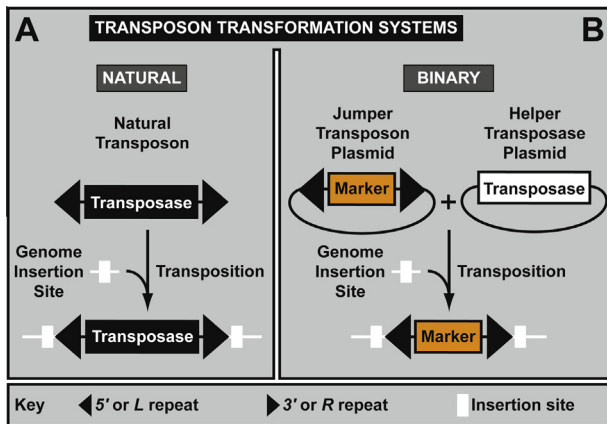


Fig. 6. Transposon interrogation. (A) Transposon transformation systems and transposons. An active natural transposon. Active natural transposons are mobile elements that consist of two inverted terminal repeats (black) that flank an open reading frame encoding a transposase. Both features are required for transposition. Transposition results in a duplication of the insertion site (white box). (B) The binary transformation system. The transposon inverted repeats and transposase can be physically separated, resulting in the binary vector/helper transposon transformation system that allows the regulated transposition of transgenes into the genome. Only when both components are provided together can productive transposition occur.

tion. To be useful for genome wide screens, a transposon has to be integrated into the genome first using its corresponding transposase, before it can be used as ammunition for subsequent remobilization mutagenesis (Fig. 8).

The transposition rates of the various transposable elements vary widely. For *P* elements, the transposition rate typically ranges from 10% to 65%, depending on the *P* element construct that is used and the level of transposase [7]. For *piggyBacs*, the transposition rate is also dependent on the nature of the element and the transposase source; it varies widely from 50% to 100% [40,118]. However, in our hands *piggyBac* transposition occurred at much lower frequencies, in the 5–10% range. For *Minos*, the transposition rate is typically 5–10% (Bellen, unpublished data), although higher rates have also been recorded previously [130,69]. High transposition rates are not necessarily beneficial, as they are associated with greater degrees of second site mutations [58]. Therefore the Gene Disruption Project has saved a very substantial number of insertions that are homozygous viable.

3.2. Dominant transposon markers

Identification of novel integration events upon the hopping or integration of a TE is dependent on the incorporation of dominant genetic markers [124] (Fig. 9). The most popular markers are those easily identified under a regular stereoscope. They include *white*⁺, a dominant eye color marker that leads to red eye pigmentation [91], and *yellow*⁺, a dominant body color marker that turns the yellowish body of a *yellow* mutant into one more tan in color [86]. Other commonly used markers are those in which fluorescent proteins have been fused to an artificial eye-specific promoter (e.g., *3xP3*) [11,41]. These markers are relatively easy to score but require a dissection microscope equipped with a UV light source and appropriate fluorescent filters.

3.3. Genome-wide transposon insertions

Details about the different TEs and TE markers, as well as the specific features of the more commonly used transposons (*P* elements, *piggyBac*, and *Minos*), are available through the Gene Disruption

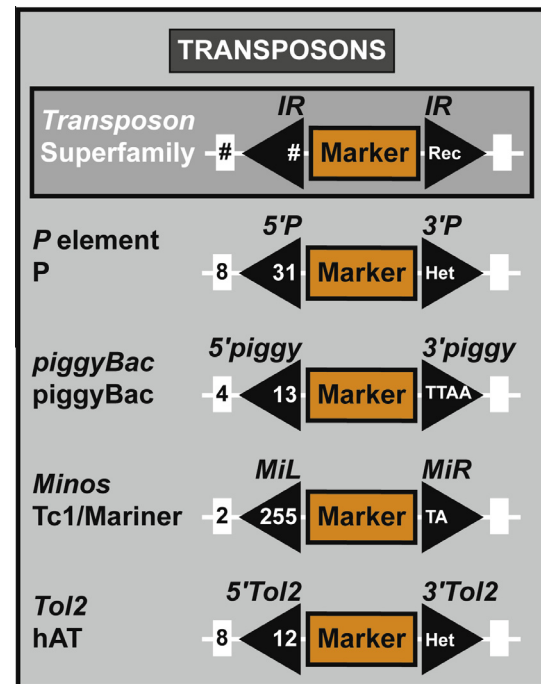


Fig. 7. Transposons. Genetic constitution of a key member of different transposon superfamilies used for genetic analysis in *Drosophila*. Transposon superfamily, size of the duplicated genome insertion site (white box, #), size of the inverted repeat (black triangle, #), and nucleotide constitution of the recognition site (black triangle, Rec) are indicated for the *P* element, *piggyBac*, *Minos*, and *Tol2*.

Project (GDP; <http://flypush.imgen.bcm.tmc.edu/pscreen/>) [6]. This project started in 1992 with the primary goal of generating at least one transposon insertion for every fly gene [6,7,110,130,111]. Initially, the project focused on *P* elements engineered by members of the community and gathered *P* element and *piggyBac* collections generated by others [7,6,110]. However, the GDP subsequently built a transposon collection using the EY *P* element which contains a UAS site, thus permitting ectopic expression of the gene near which the transposon is inserted [7]. By 2004 the diminishing returns of generating additional *P* element insertions to tag new genes became cost prohibitive. Indeed, towards the end of the project, for every 100 new insertions generated and sequenced, only 2 *P* elements could be identified in genes not already tagged. Currently, about 52% of all annotated *Drosophila* genes contain a *P* element insertion, and it is unlikely that this figure will change. All TE strains generated by the GDP are available through public stock centers, namely the Bloomington *Drosophila* Stock Center (BDSC) (<http://flystocks.bio.indiana.edu/Browse/in/GDPtop.htm>).

As an alternative to *P* elements, the GDP explored other TEs, including MiET, a *Minos* derived transposon [69]. A small pilot study suggested that MiET inserted at random in the genome [69], an aspect later confirmed with a larger screen of ~12,000 insertions, indicating that very few loci were refractory to insertion [6]. Therefore, a new version of the *Minos* transposon, the *Minos*-mediated integration cassette (MiMIC) was designed. MiMIC incorporates two inverted *attP* sites for recombinase-mediated cassette exchange (RMCE) (see also Section 3.7), allowing a broad array of versatile applications after a transposon insertion is identified [130,74]. Currently more than 12,000 MiMIC insertion strains have been generated and sequenced, half of which are available from the BDSC (<http://flypush.imgen.bcm.tmc.edu/pscreen/>). This collection consists of insertions that tag about 3200 genes so far. Other important collections based on *P* elements and *piggyBac* have also

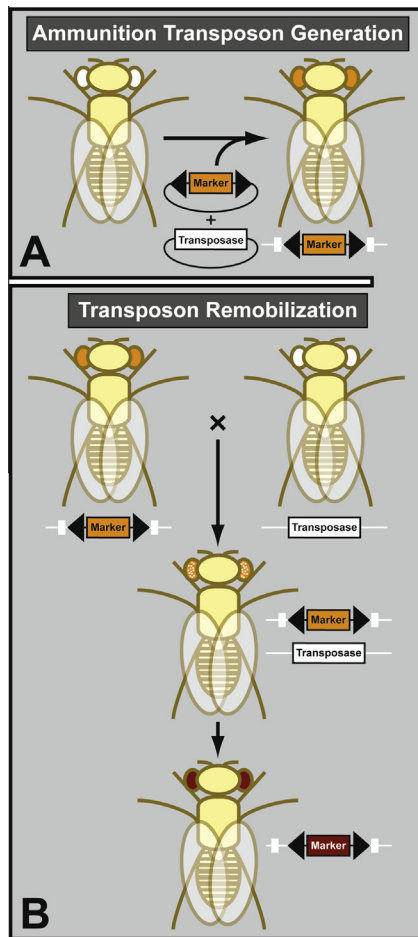


Fig. 8. Transposon mutagenesis. (A) The initial ammunition chromosome is often generated by microinjection of a plasmid containing the engineered transposon. (B) Subsequent transposon hopping is done by genetic crosses, bringing transposon and transposase together in the same fly. In the next generation, novel insertion sites are identified and the transposase segregated away.

been generated—for example, versions carrying *FRT* recombination sites [118,85,99,98]. These transposons have been used for genome-wide applications, including the development of molecularly mapped deletion and duplication kits [23]) (see also Section 3.6).

3.4. Transposon mutagenesis

Three basic TE-based approaches can be used to screen for mutations affecting specific biological processes. In the first approach a collection of TE insertions is created and subsequently screened for phenotypes, as shown for *P* elements [49] and *piggyBacs* [103]. The TE insertions are then mapped (see also Section 3.5) and molecular lesions assessed. This strategy is labor intensive and relatively inefficient compared to EMS screens. For example, while EMS can induce hundreds of mutations on each chromosome, the goal of TE hopping is often to obtain a single new insertion per chromosome. Thus, TE screens will sample less than 1% of the mutations generated in an EMS screen. Additionally, most TEs do not insert at random. *P* elements preferentially insert near the transcription start site and have hot spots, genes hit very frequently, thereby diminishing returns [6]. *piggyBacs* also have a slight preference for the transcription start site. Hence, the choice of TE may strongly bias the screen. Other weaknesses of this strategy include the fact that many TE insertions only cause a partial loss of function. This is especially true for *P* elements that insert

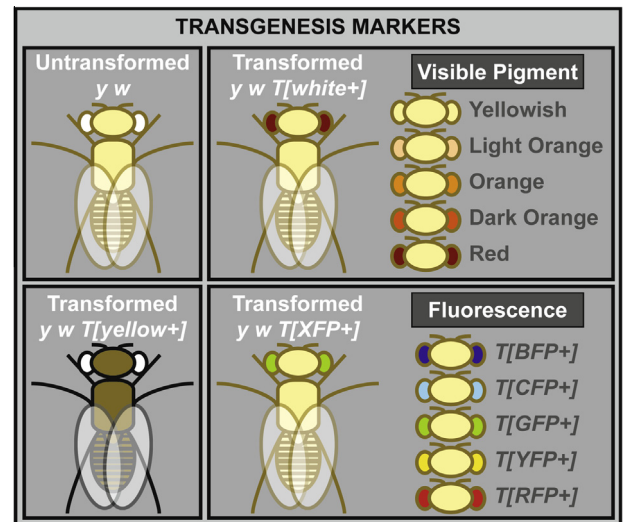


Fig. 9. Commonly used dominant genetic markers. A commonly used marker is the dominant visible eye color marker *white*⁺, which is highly under the influence of nearby genomic position effects, i.e., eye color can vary dependent on the transgene insertion site, resulting in yellowish, light orange, orange, dark orange, or WT red eye color pigmentation. Other markers often used are the dominant visible body color marker *yellow*⁺, and the dominant fluorescent markers driven by an artificial eye promoter, i.e., blue (BFP), cyan (CFP), green (GFP), yellow (YFP) or red fluorescent protein (RFP).

in regulatory regions, typically within 200 nucleotides of the transcription start site [108,6]. Finally, the main advantage of TEs—the ability to quickly map TEs using inverse PCR—does not always result in the identification of the gene of interest. Indeed, the introduction of transposases, as well as the accumulation of spurious lethals in the genetic background of some stocks, may lead to an elevated level of second site mutations not due to the TE insertion itself [58]. Consequently, rescue with a genomic fragment [129,126,117] or failure to complement a deletion [21] should be considered in the early phases of mapping project to ascertain whether the phenotype corresponds to the presumable TE insertion.

A second TE-based strategy uses a select number of TE insertion stocks already present in private collections or at the BDSC (<http://www.flystocks.bio.indiana.edu/>). Currently, TE insertions in about 11,000 genes have been created by the GDP and are available from the BDSC [7,6,130]. These stocks can be screened for specific phenotypes. Numerous screens have used this strategy for bristle number variation [79], synaptic transmission [58], hypoxia tolerance [4], and olfactory behavior [121], among others. Once interesting phenotypes have been identified, one needs to ascertain that the TE insertion causes the phenotype, preferably by doing three experiments: precise excision of the TE and reversion of the phenotype [44], failure to complement the phenotype with a deficiency [21], and rescue with a *P*[*acman*] transgene encompassing the genomic fragment [127,126,129,117]. Importantly, one need not screen all the stocks available. Rather, they can screen just a specific subtype, e.g., lethal insertions. As mentioned above, many *P* element insertions are hypomorphic and located in regulatory elements [56]. Consequently, only 18% (~2000) of all currently available TE insertion strains are homozygous lethal. The mutagenicity of different TEs, however, varies widely: the BG *P* element, which includes two gene terminating cassettes (i.e., a gene and a polyA trap) [65], only causes 8% lethality; the KG *P* element, which contains *Su*(Hw) sites engineered to interfere with gene function by buffering neighboring enhancers [96], causes 22% lethality; and the EY *P* element [7], which does not include any additional gene interruptive elements apart from the transposon itself, causes

16% lethality. MiMIC [130] has the highest mutagenic index, with 28% of the insertions associated with lethality. Thus, the number of stocks to screen varies depending if one is interested in viable or lethal mutations.

The third screening strategy uses *P* element collections in which neighboring genes can be overexpressed using UAS sites present in the TE, i.e., notably the EP [95] and EY *P* elements [7]. By integrating a UAS site in the TE, one can perform screens in which a ubiquitous or tissue-specific GAL4 driver is used to overexpress a gene near the *P* element insertion [95]. This approach has been used successfully [114,71], but relatively few novel genes have been discovered based on this strategy. Indeed it has proven difficult to associate the gain of function phenotypes with the phenotypes associated with knock-down or loss-of-function alleles needed to interpret the data [146]. Indeed, the main caveats of this type of screen are that overexpression may lead to numerous false positives and that gain-of-function phenotypes in many tissues do not correlate with gene loss.

3.5. Insertion site determination

A major advantage of TEs is that they are genetically encoded and can therefore easily be mapped to the genome using a variety of molecular methods. These methods are based on two alternative but partially overlapping procedures: inverse PCR [147] and splinkerette PCR [92] (Fig. 10). In inverse PCR genomic DNA is isolated, digested with a restriction enzyme, and self-ligated. A subsequent PCR reaction using primers that go out of the transposon backbone (hence the name “inverse PCR”) and over the unknown genomic fragment is used to determine the sequence, and ultimately the location, of the transposon insertion site. Alternatively, with splinkerette PCR common adapters are ligated to the unknown site after digestion of the genomic DNA. PCR is then conducted using one primer in the known adapter and another primer in the transposon backbone. Often a second, nested, PCR is required to amplify enough material for Sanger sequencing. Although both inverse PCR and splinkerette PCR have been successful for *Drosophila*, inverse PCR has been used to determine the genomic location of hundreds of thousands of TE insertion sites, illustrating the feasibility and ease of the technology [6,130].

3.6. Transposon induced deletions

Deletions are some of the most valuable reagents in the fly community [21,22]. They serve as a reference for null alleles and are frequently used to map genes due to their failure to complement encompassing mutations. Several strategies have been designed to generate deletions in the fly genome [123]. Currently the methodology employed most often is that based on the FLP/FRT system (Fig. 11). When two nearby mapped TEs—each containing an *FRT* site in identical orientation—are brought in *trans*, a precise deletion can be generated in the presence of the FLP recombinase. Two independently generated genome-wide transposon collections contain *FRT* sites [118]. One set was obtained by hopping a previously tested *FRT*-containing *P* element [33], generating several thousand new insertions sites. These TEs were then used to create hundreds of defined deficiencies, commonly known as the DrosDel project (<http://www.drosdel.org.uk/>) [99,98]. The other *FRT* transposon collection was generated with both *P* element and *piggyBac* [118] and resulted in hundreds of deficiencies (<http://www.flystocks.bio.indiana.edu/Browse/df/dfextract.php?num=all&symbol=exeldef>) [85]. These two deficiency sets formed the basis of a collection that was compiled at the BDSC (<http://flystocks.bio.indiana.edu/Browse/df/dftop.htm>) [21,22]. These deficiencies permit mutation mapping via complementation tests, and their coverage represents true null alleles for 98% of all fly genes.

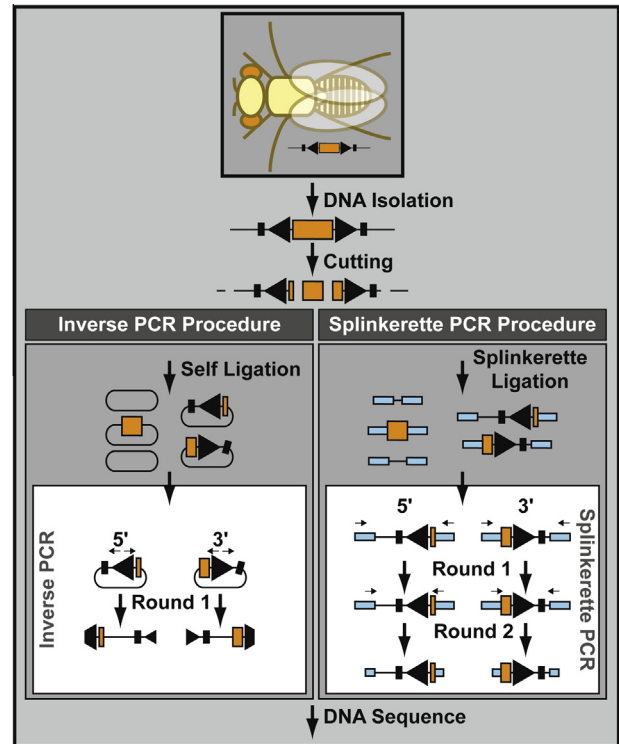


Fig. 10. Insertion site determination. Transposon insertion sites can be molecularly determined by inverse PCR or splinkerette PCR. In both cases, DNA is isolated and cut with a frequently cutting restriction enzyme. Next, using inverse PCR, self-ligated DNA is used as a PCR template. Alternatively, splinkerette adapters are ligated to the cut DNA, and ligated DNA is used as a PCR template in the splinkerette PCR procedure. In both cases, DNA sequence of the PCR fragments are determined to retrieve the transposon insertion site.

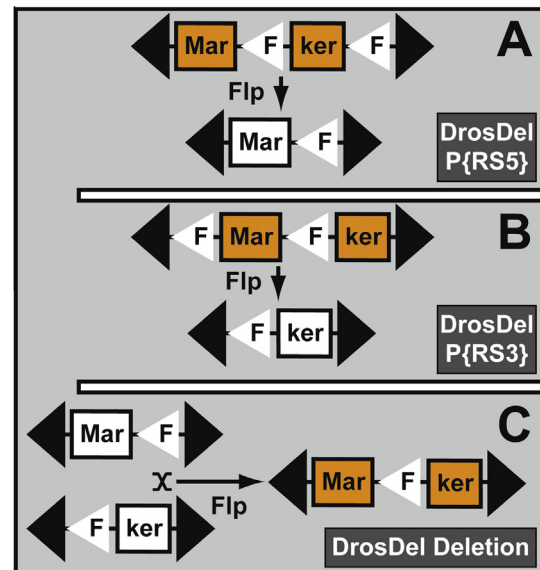


Fig. 11. Transposon induced deletion. The DrosDel deletions are generated by *white*⁺ marker reconstitution. In a first step, part of the *white*⁺ markers of two transposon insertions (P{RS5} and P{RS3}) is reduced to invisible but complementary *white*⁺ halves. In a next step, both invisible *white*⁺ halves are reconstituted into a full length phenotypically visible *white*⁺ marker, after FLP-mediated mitotic recombination.

3.7. In vivo transposon upgrading

Since transposons are genetically encoded, they can be engineered to encompass elements that allow future upgrades either

through microinjection or through genetic crosses. For example, *P* elements permit transposon conversion (i.e., *in vivo* replacement of the DNA internal to the terminal inverted repeats). This is achieved by providing a transposase source in the presence of the original acceptor *P* element and a donor DNA *P* element, each containing a different dominant marker. The transposase enzyme promotes the excision of acceptor sequences and its subsequent replacement with donor sequences, precisely swapping the content of the *P* elements [31,104]. In doing so, a *P* element containing one property (e.g., *LacZ*) can be upgraded with a novel *P* element containing a second, more sophisticated, property (e.g., the GAL4 binary transactivator). This property is so far unique to *P* elements and has not yet been reported for other transposons.

An alternate (and less labor intensive) strategy is to include two recombination (e.g., *FRT* site) (see also Section 3.6) or integration sites (e.g., *attP*) in the transposon backbone. Two transposon platforms have been engineered for limitless transposon upgrading. The first uses *MiMIC* transposon insertions and upgrades the locus through recombinase mediated cassette exchange [130,74] via the Φ C31 integrase [37,15] (Fig. 12A). The exchange replaces the existing *MiMIC* gene trap cassette located in 5'UTR or coding intron with a correction cassette that reverts lethality. Alternatively, *MiMIC* insertions located in a 5'UTR intron can be upgraded with a novel gene trap that incorporates new properties under the form of a fluorescent marker (e.g., Cherry red fluorescent protein), a binary factor (e.g., GAL4 or QF), or a recombinase (e.g., FLP) for subsequent manipulations. Most interestingly, *MiMIC* insertions located in coding intronic insertions can be upgraded with artificial exons flanked by an upstream splice acceptor site and a downstream splice donor site. The exons can encode protein tags (i.e., fluorescent markers, peptide affinity tags, or a combination of both) for microscopy visualization or protein purification.

The second upgradable TE platform uses *P* element or *piggyBac* insertions and upgrades the locus via site-specific transgenesis (Fig. 12B). These platforms are *InSITE* (Integrase Swappable *In Vivo* Targeting Element) [32] and *G-MARET* (GAL4-based mosaic-inducible and reporter-exchangeable enhancer trap) [141]. Both have been used to upgrade GAL4-containing enhancer trap insertions to novel binary factors such as LexA and QF. Interestingly, *InSITE* was engineered to allow *in vivo* exchange by simple genetic crosses avoiding microinjection experiments [32]. The FLP recombinase releases the modification construct from one location (i.e., the donor transposon), which then integrates into the acceptor transposon using the Φ C31 integrase. Finally, the Cre recombinase is used to remove the undesired portion of the acceptor transposon (e.g., GAL4), leaving the novel property behind (e.g., LexA).

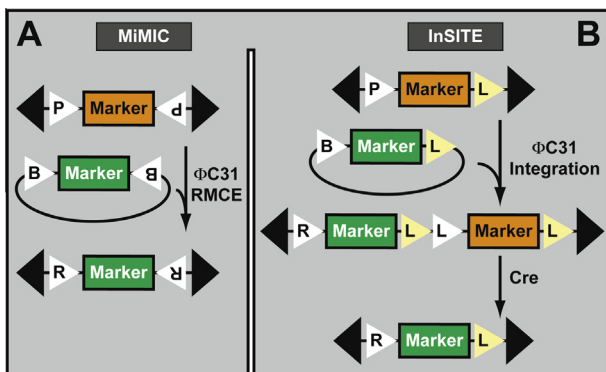


Fig. 12. *In vivo* transposon upgrading. (A) The *MiMIC* transposon insertions are upgraded by Φ C31 integrase catalyzed recombinase mediated cassette exchange. (B) *InSITE* (shown) and *G-MARET* (not shown) transposons are upgraded by Φ C31 integrase catalyzed site-specific integration followed by deletion of unnecessary genetic material.

4. Transgene libraries to assess gene function

4.1. Transgenes and transgenic libraries

In the past, transgenesis in *Drosophila* solely relied on the *P* element transposon [124]. Unfortunately, the insertion site of transposons cannot be controlled (see also Section 3.1), resulting in variable position effects caused by the surrounding genome, making it difficult to compare differently mutagenized transgenes (Fig. 13). On the contrary, the site-specific integration of transgenes via the use of a site-specific integrase eliminates this issue. More specifically, the Φ C31 integrase allows for site-specific integration of *attB* containing transgenes at defined *attP* docking sites in the genome. Since all constructs can be integrated at the same genomic location, the influence of position effects on the integrated transgene is neutralized. The Φ C31 system additionally allows for the integration of small to very large fragments. Hence, the Φ C31 system is ideally suited to compare transgenes obtained from genome-wide libraries generated for specific purposes. Hence, due to these advantages, this overview will focus on those libraries generated using vectors compatible with the Φ C31 site-specific integrase system. Currently, Φ C31-based libraries are available for several genome-wide interrogation paradigms: genomic rescue [126,28,117], RNAi [26,76,78], regulatory element interrogation [46,89,66,47], and cDNA overexpression [14,102]. These are outlined below.

4.2. Genomic rescue and recombineering

Genomic rescue refers to the use of genomic DNA fragments to revert or “rescue” phenotypes associated with a given mutation (Fig. 14). Upon evidence of phenotype recovery one can manipulate the rescue construct for various downstream applications. For example, rescue constructs can be tagged to help identify cells in which the gene of interest is expressed or determine the subcellular protein distribution. One can also create numerous mutations in the rescue construct to initiate structure function analysis, testing protein variants in a null mutant background.

The scope of genomic rescue expanded significantly with the development of the P[acman] transgenesis platform (P/ Φ C31 artificial chromosome for manipulation) [127]. P[acman] combines three powerful technologies: recombineering [105], conditionally amplifiable BACs [135], and Φ C31 site-specific integration [37,15]. First, recombineering, which allows the incorporation of virtually any genetic alteration, was used in the P[acman] system to retrieve DNA fragments up to 102 kb in size from existing BAC libraries via a single gap-repair [42] (<https://bacpac.chori.org/>). A double gap-repair procedure, wherein two pieces of overlapping BACs were joined together, allowed for the generation of a 133 kb fragment, reconstituting the Tenascin major gene [127]. Second, DNA fragments recovered from a BAC are maintained at low-copy number in a conditionally amplifiable plasmid, ensuring the stability of large constructs under normal conditions. Addition of a simple sugar solution induces plasmid replication and results in high copy number, which facilitates cloning and DNA preparation [135]. Third, the Φ C31 integrase allows the integration of P[acman] vectors into defined *attP* docking sites located in the fly genome [127,37,15,67]. While the Φ C31 integrase is able to integrate plasmids up to at least 146 kb *in vivo* [127,144], recently we were able to integrate large BACs derived from previously generated libraries [42] with an upper size limit of 212 kb. These integrations followed appropriate BAC retrofitting [53] and included the necessary elements for P[acman] transgenesis (*OriV* conditional origin of replication for conditional plasmid copy number amplification, *attB* site for site-specific integration using the Φ C31 integrase).

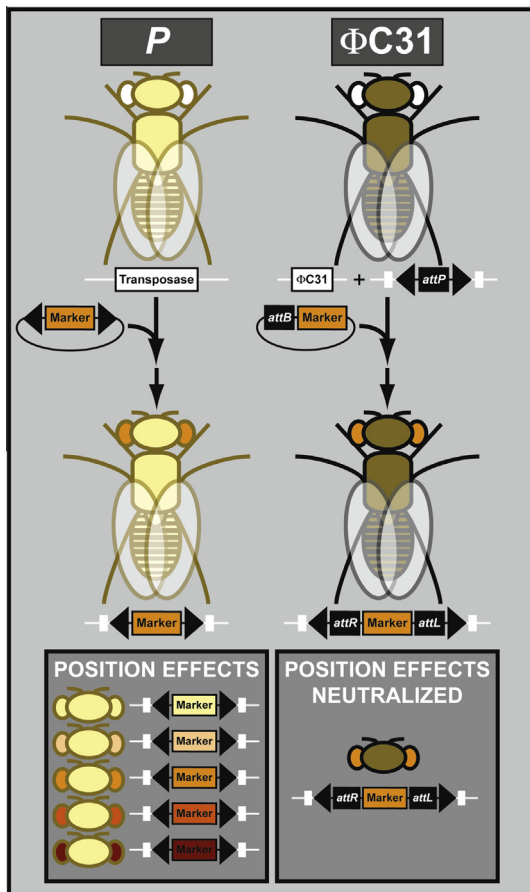


Fig. 13. Transgenesis platforms. P element transposase catalyzed transgenesis results in uncontrolled transgene integration and variable position effects, while ΦC31 integrase catalyzed transgenesis results in site-specific integration and the neutralization of position effects.

ase, and the *white*⁺ dominant eye color transgenesis marker) [127] (Venken, unpublished data).

P[acman] methodology has been used to create three genome-wide *Drosophila* P[acman] BAC DNA libraries: the CHORI-322 P[acman] library (average insert size of 21 kb), the CHORI-321 P[acman] library (average insert size of 83 kb in P[acman] (<http://www.pacmanfly.org/>) [126]), and the FlyFos library [28] (average insert size of 36 kb in a fosmid instead of the P[acman] backbone) (<http://transgenome.mpi-cbg.de/>). The three libraries have been end sequenced and annotated onto the reference fly genome. While the FlyFos library provides a 3.3-fold genome coverage, the CHORI-321 and CHORI-322 libraries represent a 12-fold coverage and allow rescue of more than 95% of annotated fly genes. Thus, the libraries largely eliminate the need for gap-repair for genomic rescue. Compared to gap-repair, which requires additional cloning steps and allows defined genomic fragments to be analyzed, clones used in the libraries were randomly generated and often contain several genes. Although this may appear to be a weakness of the libraries, small gene duplications up to 100 kb rarely cause phenotypes [129] and larger clones are more likely to include all the necessary regulatory elements of a gene. The functionality of the three libraries was tested by integrating numerous clones into *attP* docking sites and showing that they rescue phenotypes associated with known mutations in the corresponding region of interest.

The amplifiable nature of P[acman] and FlyFos clones greatly facilitates recombineering. Prior to amplification, tags and mutations are introduced at low-copy number and fixed within a bacte-

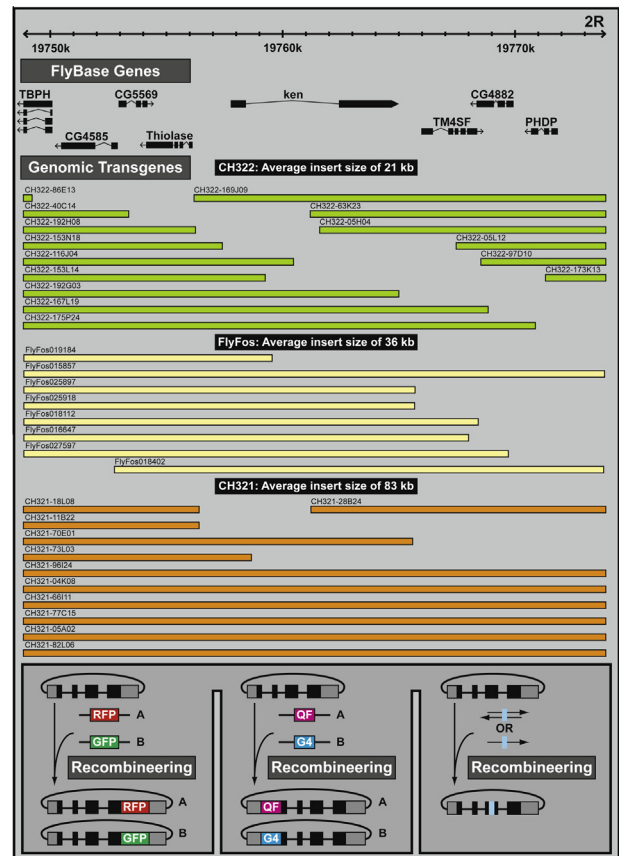


Fig. 14. Genomic rescue transgenes and recombineering. Genomic transgene fragments can be chosen from three genomic DNA libraries: CH322 clones with an average insert size of 21 kb (green), FlyFos clones with an average insert size of 36 kb (yellow), or CH321 clones with an average insert size of 83 kb (orange). Shown is a region surrounding the gene *ken* and *barbie* (*ken*). Subsequently, recombineering can be used to modify the genomic rescue fragment, to include insertions (e.g., fluorescent tags for protein visualization, and binary factors for host gene cellular visualization) or point mutations.

rial colony. These methods have already been used with P[acman] and FlyFos to (1) introduce deletions of regulatory elements for structure/function analysis [87,88,29,19], (2) incorporate protein tags for visualization of expression or acute protein inactivation [128,126,28,73,134,133], and (3) integrate binary factors (e.g., GAL4 or QF) for cellular labeling [115]. Hence, tagged clones can be used to determine the expression patterns of numerous uncharacterized genes or label specific cell populations.

4.3. Transgenic duplication kit

The X chromosome of *D. melanogaster* contains 22 Mb of euchromatic DNA, encompassing approximately 2300 protein-coding genes [2]. Mutations in essential genes or those involved in male fertility are often tedious to map unless a duplication is present on another chromosome that can be brought in through simple genetic crosses. Recently, two groups independently generated a set of duplications to expedite the mapping of X chromosome mutations (see also Section 2.6) (Fig. 5B). One group made use of the available P[acman] libraries [129]. In doing so, a minimal tiling path of 408 overlapping clones was integrated into a single *attP* docking site located on the 3rd chromosome [37,15]. Currently, 96% of the X chromosome is covered by the P[acman] duplications and contained in 382 fly strains. Approximately 100 clones were tested for rescue, and 92% rescued existing mutations

as well as entire deficiencies previously generated with the FLP/*FRT* system [85,99,98]. Interestingly, this group demonstrated that most genes are tolerated at twice the normal dosage, and the duplication collection enabled more precise mapping of two regions involved in diplo-lethality. The second group to generate an X chromosome duplication kit did so by genetic means, incorporating relatively large duplications of the X chromosome onto the Y (attached-XY). Due to the size of the duplications, this kit allows quick but rough mapping and currently covers 90% of the X chromosome [23]. Both duplication collections are available at the BDSC: <http://flystocks.bio.indiana.edu/Browse/dp/BDSC-Dps.php>.

4.4. Genome-wide RNAi

There are currently two genome-wide RNAi libraries for *Drosophila* (<http://stockcenter.vdrc.at/>)—both created in the lab of Barry Dickson. These libraries use RNAi technology to knock down genes, aiding functional analysis. The first library, the GD library, was generated using *P* element transposons [26]. However, since *P* element insertion sites cannot be controlled, a second library—KK—was created using the Φ C31 system. Other genome-wide RNAi resources are currently being generated to complement these two libraries. These include an independent collection based on long dsRNA as well as a collection based on short hairpins using a miRNA backbone [76,78]. These and other RNAi reagents are discussed in detail in a subsequent chapter [149].

4.5. Genome-wide enhancer bashing

The GAL4/UAS system has been extremely useful for interrogation of regulatory gene elements [17]. Notably, many of the driver lines used in GAL4/UAS experiments were isolated by hopping *P* element enhancer detector elements [81] carrying the binary GAL4 transactivator [17]. Since *P* elements tend to integrate into the regulatory elements of genes, the neutral promoter driving GAL4 or lacZ expression comes under the control of endogenous enhancers [8,136,13,17]. As such, many enhancer detector insertion strains display expression patterns that correspond to that of a nearby gene. However, expression can often not be limited to a small subset of cells or cell populations. This lack of cellular specificity is an issue if one desires to remove, functionally inactivate, or manipulate a few cells in a complex circuit (e.g., neurons).

To isolate the individual components of a gene's complex regulatory environment, it typically suffices to isolate smaller DNA fragments and bring them in front of a neutral promoter—the so-called enhancer bashing strategy. As position effects may influence the analysis of different regulatory regions, it is advantageous to integrate the different constructs into the same “neutral” site. Hence, a large set of Φ C31 *attP* docking sites were analyzed for position effects [67], and one, *attP2*, with virtually no leaky expression, was selected to integrate the majority of the transgenic GAL4 library [89] (Fig. 15).

The initial GAL4 collection is biased towards genes expressed in the nervous system [89], and the resulting fly strains (~7000) have been screened for expression in a variety of tissues. Expression patterns were first established in the adult brain and ventral nerve cord [46], where novel neuronal cell types were identified and brain asymmetry revealed. Interestingly, many lines expressed GAL4 in a pattern broader than anticipated, especially since regulatory elements of only 3 kb were used. In a second study, a subset of 5000 transgenic lines was used to document expression patterns in the embryonic central nervous system at two different embryonic stages (i.e., stages 9–11, and stage 16) [66]. The expression data will be valuable for determining the morphology and function of all neurons in the embryonic CNS. Finally, the lines were tested for expression patterns in third-instar imaginal discs, specifically

the eye, antenna, leg, wing, haltere, and genital discs [47]. In this study, DNA sequence motifs provided information about candidate transcription factors that may regulate enhancers. All the expression data obtained from these studies is available through Janelia Farm Research Campus (<http://flweb.janelia.org/cgi-bin/flew.cgi>). A second collection GAL4 collection, the VT collection, has also been generated using the same technology [89]. Expression data for these lines have been obtained for embryos [55] and the adult brain (<http://stockcenter.vdrc.at/control/vtlibrary>). To perform independent manipulations of different cell types, several of the fragments have been fused to an orthogonal binary activation system that consists of the transcriptional activator LexA that binds to LexO operators. Numerous LexA driven lines and reporters have been generated [141,25,90,106,32].

One of the driving goals of the GAL4 collection was to identify small subsets of labeled neurons. This was only partly achieved, and the same libraries have since been fused to the recently optimized split GAL4 system [64,90]. In the split GAL4 system, the GAL4 transcription factor is split into two hemidrivers, one containing the DNA binding domain and the other containing the activation domain. Each hemidriver is driven by separate regulatory elements. Since each hemidriver cannot activate expression of the reporter, expression is not observed. However, the location where the expression domains overlap, the two halves of GAL4 can heterodimerize, and reconstitute a functional activator that can activate the expression of a reporter gene. Thus, expression patterns can be narrowed to a small overlapping cell population shared between two much larger expression domains. Other intersectional strategies follow along a similar vein and involve combining two sets of overlapping hemi-drivers, e.g., the split LexA system [120].

4.6. Genome-wide overexpression and in vivo transgene upgrading

A final set of genome-wide reagents based on the Φ C31 integrase was recently initiated—the UAS ORFeome, a transgene collection aimed to represent at least one cDNA isoform of each *Drosophila* gene [14]. This library, which represents a large collection of genes regulating proliferation, cell cycle, and cell size,

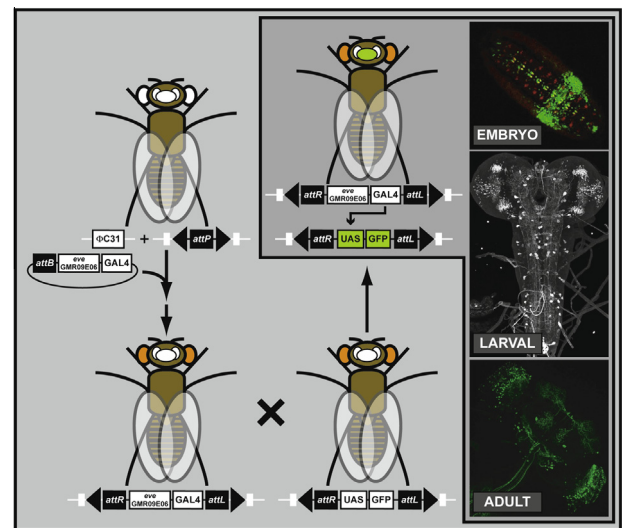


Fig. 15. Genome-wide enhancer bashing. Thousands of enhancer fragments were fused to GAL4 and assayed for gene expression patterns. The GMR09E06 enhancer fragment of the *even-skipped* (*eve*) is shown as an example. Expression patterns are shown for the embryo, larval brain and ventral nerve cord, and the adult brain. Images were obtained from the Janelia Farm Research Campus website summarizing the project (<http://www.janelia.org/gal4-gen1>).

greatly facilitates large scale overexpression studies by using any GAL4 driver of interest to drive expression of the UAS-linked open reading frames of target genes. By using misexpression of genes, problems arising from genetic redundancy are prevented and novel gene functions can be uncovered. To date, 1149 UAS-ORF fly lines have been created. As proof of principle, researchers cloned 655 growth-regulating genes in two variants—one with the open reading frame and a native stop codon, the other an open reading frame with a C-terminal 3xHA tag prior to the stop codon for affinity purification studies. Interestingly, to streamline the creation of transgenic flies, pools of plasmids were injected into embryos. Each plasmid, in addition to a single cDNA, contains a unique barcode serving as the plasmid identifier after retrieval of transgenic flies from the pooled injections.

The versatility of the ORF collection was greatly expanded with the development of an *in vivo* swapping technique (Fig. 16). Based on mitotic recombination (i.e., using the FLP recombinase), this system enables the rapid exchange of promoters and epitope tags through simple crosses. Promoters, N- or C-terminal tags, and polyadenylation signals can be exchanged at will. *In vivo* swapping was validated by changing promoters (e.g., UAS into LexO) and C-terminal protein tags (e.g., 3xHA into EGFP). To confirm the usefulness of this library for screening purposes, Wingless (Wg) pathway components were screened for a variety of phenotypic outputs [102]. Three serine/threonine kinases were identified as regulators of Wg signaling. One of these, Nek2, optimizes the Wg pathway response by direct phosphorylation of Dishevelled. While the ORF collection is likely to be a very valuable resource for studying growth regulation *in vivo*, similar caveats associated with EP and EY screens should be kept in mind (see also Section 3.4).

5. Future directions

One of the striking features of technology development is that relatively few methods are truly innovative and powerful on their own. Although exceptions such as the development of *P* element transgenesis [97,109], PCR [101] and monoclonal antibodies [52] come to mind, very often seemingly less important technical innovations can be combined to create very powerful toolsets. For example, *P* element-mediated enhancer detection [81] was shown to be a powerful methodology to identify new genes and their

expression patterns. Having a “neutral and weak” promoter controlling *LacZ* expression in a transposon permitted the detection of innumerable gene expression patterns [8,13,136]. However, it only became a major forerunner when combined with GAL4/UAS technology. This binary system from yeast had been shown to be efficient in flies [30] and mice [82], but only when the *LacZ* of the enhancer detector system was replaced with *GAL4* and a UAS-*LacZ* transgene was inserted in the genome, did both technologies (enhancer detection and GAL4/UAS binary system) develop their full potential. The GAL4/UAS methodology [17] ultimately is the most used technology in flies.

Although numerous methods are available to tackle scientific questions, often too little time is spent in weighing the quality, depth, and caveats of various approaches. Very often quick technological paradigms are available but they often have major limitations that are ignored. For example, RNAi against a given gene may be determined to cause a phenotype, but the data should be corroborated with mutations that are already publicly available. In addition, efficiency and specificity of RNAi should be tested. Similarly, when new mutations are isolated with chemical mutagenesis screens, it is important to either isolate another allele in the same complementation group or to characterize RNAis that cause the same or very similar phenotypes. Alternatively, a genomic rescue construct carrying the gene of interest should be tested. Indeed, the number of polymorphisms in some stocks is staggering, and many stocks that we tested carry mutations in two essential genes [129]. Although these second site mutations can easily be outcrossed, this often seems to be a concern solely to behavioral biologists. Given the arsenal of tools that are available [131,112], this trend should be reversed.

In summary, we believe that given the very broad tool box that is currently available in the *Drosophila* community, and in particular the innumerable combinations and permutations that can be engineered, the future of *Drosophila* to study novel and well established biological processes is extremely promising. However, we also feel that the toolsets need to be used with caution and that the users need to know the strengths and weaknesses of these resources. This chapter should provide a starting point for researchers using these tools.

Acknowledgements

We apologize to those whose work we did not cite because of our focus and space limitations. We would like to thank Pavel Tomancak for help with appropriate representation of Fig. 13, and Kristi Hoffman, Audrey Christiansen and Karen Schulze for comments on the manuscript. KJTV is supported by startup funds kindly provided by Baylor College of Medicine and the McNair Medical Institute, and Grants from the Cancer Prevention and Research Institute of Texas, and the National Institutes of Health (1R21HG006726). HJB is supported by the Howard Hughes Medical Institute.

References

- [1] H. Aberle, A.P. Haghghi, R.D. Fetter, B.D. McCabe, T.R. Magalhaes, C.S. Goodman, *Neuron* 33 (2002) 545–558.
- [2] 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, R.A. George, S.E. Lewis, S. Richards, M. Ashburner, S.N. Henderson, G.G. Sutton, J.R. Wortman, M.D. Yandell, Q. Zhang, L.X. Chen, R.C. Brandon, Y.H. Rogers, R.G. Blazej, M. Champe, B.D. Pfeiffer, K.H. Wan, C. Doyle, E.G. Baxter, G. Helt, C.R. Nelson, G.L. Gabor, J.F. Abril, A. Agbayani, H.J. An, C. Ndrews-Pfannkoch, D. Baldwin, R.M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E.M. Beasley, K.Y. Beeson, P.V. Benos, B.P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M.R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K.C. Burtis, D.A. Busam, H. Butler, E. Cadieu, A. Center, I. Chandra, J.M. Cherry, S. Cawley, C. Dahlke, L.B. Davenport, P. Davies, P.B. De, A. Delcher, Z. Deng, A.D. Mays, I. Dew, S.M. Dietz, K. Dodson, L.E. Doup, M. Downes, S. Dugan-Rocha, B.C. Dunkov, P. Dunn, K.J. Durbin, C.C.

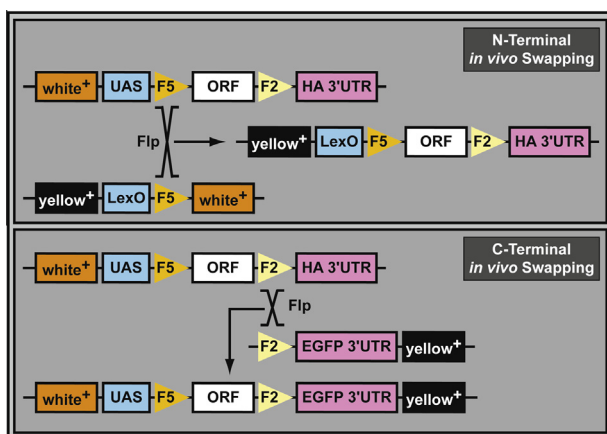


Fig. 16. *In vivo* transgene upgrading through mitotic recombination. N-terminal *in vivo* swapping, catalyzed by the FLP recombinase, results in exchange of the UAS promoter with novel promoter elements (e.g., LexO). Desired recombination events can be identified by exchange of the *white*⁺ marker with the *yellow*⁺ marker. C-terminal *in vivo* swapping, catalyzed by the FLP recombinase, results in exchange of the HA tag and the 3'UTR with the fluorescent EGFP tag and potentially a different 3'UTR. Desired recombination events can be screened by the simultaneous occurrence of the *white*⁺ and *yellow*⁺ markers.

- Evangelista, C. Ferraz, S. Ferriera, W. Fleischmann, C. Fosler, A.E. Gabrielian, N.S. Garg, W.M. Gelbart, K. Glasser, A. Glodek, F. Gong, J.H. Gorrell, Z. Gu, P. Guan, M. Harris, N.L. Harris, D. Harvey, T.J. Heiman, J.R. Hernandez, J. Houck, D. Hosten, K.A. Houston, T.J. Howland, M.H. Wei, C. Ibegwam, M. Jalali, F. Kalush, G.H. Karpen, Z. Ke, J.A. Kennison, K.A. Ketchum, B.E. Kimmel, C.D. Kodiira, C. Kraft, S. Kravitz, D. Kulp, Z. Lai, P. Lasko, Y. Lei, A.A. Levitsky, J. Li, Z. Li, Y. Liang, X. Lin, X. Liu, B. Mattei, T.C. McIntosh, M.P. McLeod, D. McPherson, G. Merkulov, N.V. Milshina, C. Mobarry, J. Morris, A. Moshrefi, S.M. Mount, M. Moy, B. Murphy, L. Murphy, D.M. Muzny, D.L. Nelson, D.R. Nelson, K.A. Nelson, K. Nixon, D.R. Nusskern, J.M. Pacleb, M. Palazzolo, G.S. Pittman, S. Pan, J. Pollard, V. Puri, M.G. Reese, K. Reinert, K. Remington, R.D. Saunders, F. Scheeler, H. Shen, B.C. Shue, I. Siden-Kiamos, M. Simpson, M.P. Skupski, T. Smith, E. Spier, A.C. Spradling, M. Stapleton, R. Strong, E. Sun, R. Svirskas, C. Tector, R. Turner, E. Venter, A.H. Wang, X. Wang, Z.Y. Wang, D.A. Wassarman, G.M. Weinstock, J. Weissenbach, S.M. Williams, T. Woodage, K.C. Worley, D. Wu, S. Yang, Q.A. Yao, J. Ye, R.F. Yeh, J.S. Zaveri, M. Zhan, G. Zhang, Q. Zhao, L. Zheng, X.H. Zheng, F.N. Zhong, W. Zhong, X. Zhou, S. Zhu, X. Zhu, H.O. Smith, R.A. Gibbs, E.W. Myers, G.M. Rubin, J.C. Venter, *Science* 287 (2000) 2185–2195.
- [3] M. Ashburner, *Drosophila: A Laboratory Handbook and Manual*, Cold Spring Harbor Laboratory Press, 1989.
- [4] P. Azad, D. Zhou, R. Zarnadt, G.G. Haddad, *G3 (Bethesda)* 2 (2012) 1169–1178.
- [5] J.R. Bateman, A.M. Lee, C.T. Wu, *Genetics* 173 (2006) 769–777.
- [6] H.J. Bellen, R.W. Levis, Y. He, J.W. Carlson, M. Evans-Holm, E. Bae, J. Kim, A. Metaxakis, C. Savakis, K.L. Schulze, R.A. Hoskins, A. Spradling, *Genetics* 188 (2011) 731–743.
- [7] H.J. Bellen, R.W. Levis, G. Liao, Y. He, J.W. Carlson, G. Tsang, M. Evans-Holm, P.R. Hiesinger, K.L. Schulze, G.M. Rubin, R.A. Hoskins, A.C. Spradling, *Genetics* 167 (2004) 761–781.
- [8] H.J. Bellen, C.J. O’Kane, C. Wilson, U. Grossniklaus, R.K. Pearson, W.J. Gehring, *Genes Dev.* 3 (1989) 1288–1300.
- [9] H.J. Bellen, C. Tong, H. Tsuda, *Nat. Rev. Neurosci.* 11 (2010) 514–522.
- [10] S. Benzer, *Proc. Natl. Acad. Sci. U.S.A.* 58 (1967) 1112–1119.
- [11] A.J. Berghammer, M. Klingler, E.A. Wimmer, *Nature* 402 (1999) 370–371.
- [12] M. Bibikova, M. Golic, K.G. Golic, D. Carroll, *Genetics* 161 (2002) 1169–1175.
- [13] E. Bier, H. Vaessin, S. Shepherd, K. Lee, K. McCall, S. Barbel, L. Ackerman, R. Carretto, T. Uemura, E. Grell, *Genes Dev.* 3 (1989) 1273–1287.
- [14] J. Bischof, M. Bjorklund, E. Furger, C. Schertel, J. Taipale, K. Basler, *Development* 140 (2013) 2434–2442.
- [15] J. Bischof, R.K. Maeda, M. Hediger, F. Karch, K. Basler, *Proc. Natl. Acad. Sci. U.S.A.* 104 (2007) 3312–3317.
- [16] C. Bokel, *Methods Mol. Biol.* 420 (2008) 119–138.
- [17] A.H. Brand, N. Perrimon, *Development* 118 (1993) 401–415.
- [18] M. Buszcza, S. Paterno, D. Lighthouse, J. Bachman, J. Planck, S. Owen, A.D. Skora, T.G. Nystul, B. Ohlstein, A. Allen, J.E. Wilhelm, T.D. Murphy, R.W. Levis, E. Matunis, N. Srivali, R.A. Hoskins, A.C. Spradling, *Genetics* 175 (2007) 1505–1531.
- [19] J.J. Cassidy, A.R. Jha, D.M. Posadas, R. Giri, K.J. Venken, J. Ji, H. Jiang, H.J. Bellen, K.P. White, R.W. Carthew, *Cell* 155 (2013) 1556–1567.
- [20] J.P. Castro, C.M. Carareto, *Genetica* 121 (2004) 107–118.
- [21] K.R. Cook, A.L. Parks, L.M. Jacobus, T.C. Kaufman, K.A. Matthews, *Fly* 4 (2010) 88–91.
- [22] R.K. Cook, S.J. Christensen, J.A. Deal, R.A. Coburn, M.E. Deal, J.M. Gresens, T.C. Kaufman, K.R. Cook, *Genome Biol.* 13 (2012) R21.
- [23] R.K. Cook, M.E. Deal, J.A. Deal, R.D. Garton, C.A. Brown, M.E. Ward, R.S. Andrade, E.P. Spana, T.C. Kaufman, K.R. Cook, *Genetics* 186 (2010) 1095–1109.
- [24] A. Diantonio, K.D. Parfitt, T.L. Schwarz, *Cell* 73 (1993) 1281–1290.
- [25] S. Diegelmann, M. Bate, M. Landgraf, *Fly. (Austin)* 2 (2008) 236–239.
- [26] G. Dietzl, D. Chen, F. Schnorrer, K.C. Su, Y. Barinova, M. Fellner, B. Gasser, K. Kinsey, S. Oppel, S. Scheiblaue, A. Couto, V. Marra, K. Keleman, B.J. Dickson, *Nature* 448 (2007) 151–156.
- [27] Y. Dudai, Y.N. Jan, D. Byers, W.G. Quinn, S. Benzer, *Proc. Natl. Acad. Sci. U.S.A.* 73 (1976) 1684–1688.
- [28] R.K. Ejsmont, M. Sarov, S. Winkler, K.A. Lipinsk, P. Tomancak, *Nat. Methods* 6 (2009) 435–437.
- [29] E.M. Enneking, S.R. Kudumala, E. Moreno, R. Stephan, J. Boerner, T.A. Godenschwege, J. Pielage, *PLoS Biol.* 11 (2013) e1001537.
- [30] J.A. Fischer, E. Giniger, T. Maniatis, M. Ptashne, *Nature* 332 (1988) 853–856.
- [31] G.B. Gloor, N.A. Nassif, D.M. Johnson-Schlitz, C.R. Preston, W.R. Engels, *Science* 253 (1991) 1110–1117.
- [32] D.M. Gohl, M.A. Silies, X.J. Gao, S. Bhalerao, F.J. Luongo, C.C. Lin, C.J. Potter, T.R. Clandinin, *Nat. Methods* 8 (2011) 231–237.
- [33] K.G. Golic, M.M. Golic, *Genetics* 144 (1996) 1693–1711.
- [34] K.G. Golic, S. Lindquist, *Cell* 59 (1989) 499–509.
- [35] W.J. Gong, K.G. Golic, *Proc. Natl. Acad. Sci. U.S.A.* 100 (2003) 2556–2561.
- [36] S.J. Gratz, A.M. Cummings, J.N. Nguyen, D.C. Hamm, L.K. Donohue, M.M. Harrison, J. Wildonger, K.M. O’Connor-Giles, *Genetics* 194 (2013) 1029–1035.
- [37] A.C. Groth, M. Fish, R. Nusse, M.P. Calos, *Genetics* 166 (2004) 1775–1782.
- [38] B. Haley, D. Hendrix, V. Trang, M. Levine, *Dev. Biol.* 321 (2008) 482–490.
- [39] A.M. Handler, R.A. Harrell, *Insect Mol. Biol.* 8 (1999) 449–457.
- [40] C. Horn, N. Offen, S. Nystedt, U. Hacker, E.A. Wimmer, *Genetics* 163 (2003) 647–661.
- [41] C. Horn, E.A. Wimmer, *Dev. Genes Evol.* 210 (2000) 630–637.
- [42] R.A. Hoskins, C.R. Nelson, B.P. Berman, T.R. Laverty, R.A. George, L. Ciesiolka, M. Naemuddin, A.D. Arenson, J. Durbin, R.G. David, P.E. Tabor, M.R. Bailey, D.R. DeShazo, J. Catanese, A. Mammoser, K. Osoegawa, P.J. de Jong, S.E. Celniker, R.A. Gibbs, G.M. Rubin, S.E. Scherer, *Science* 287 (2000) 2271–2274.
- [43] Y. Hotta, S. Benzer, *Proc. Natl. Acad. Sci. U.S.A.* 67 (1970) 1156–1163.
- [44] T. Hummel, C. Klambt, *Methods Mol. Biol.* 420 (2008) 97–117.
- [45] H. Jafar-Nejad, H.K. Andrews, M. Acar, V. Bayat, F. Wirtz-Peitz, S.Q. Mehta, J.A. Knoblich, H.J. Bellen, *Dev. Cell* 9 (2005) 351–363.
- [46] A. Jenett, G.M. Rubin, T.T. Ngo, D. Shepherd, C. Murphy, H. Dionne, B.D. Pfeiffer, A. Cavallaro, D. Hall, J. Jeter, N. Iyer, D. Fetter, J.H. Hausenfluck, H. Peng, E.T. Trautman, R.R. Svirskas, E.W. Myers, Z.R. Iwinski, Y. Aso, G.M. DePasquale, A. Enos, P. Hulamm, S.C. Lam, H.H. Li, T.R. Laverty, F. Long, L. Qu, S.D. Murphy, K. Rokicki, T. Safford, K. Shaw, J.H. Simpson, A. Sowell, S. Tae, Y. Yu, C.T. Zugates, *Cell Rep.* 2 (2012) 991–1001.
- [47] A. Jory, C. Estella, M.W. Giorgianni, M. Slattery, T.R. Laverty, G.M. Rubin, R.S. Mann, *Cell Rep.* 2 (2012) 1014–1024.
- [48] M. Kango-Singh, R. Nolo, C. Tao, P. Verstreken, P.R. Hiesinger, H.J. Bellen, G. Halder, *Development* 129 (2002) 5719–5730.
- [49] A. Kania, A. Salzberg, M. Bhat, D. D’Evelyn, Y. He, I. Kiss, H.J. Bellen, *Genetics* 139 (1995) 1663–1678.
- [50] J.R. Kennerdell, R.W. Carthew, *Nat. Biotechnol.* 18 (2000) 896–898.
- [51] T. Koana, Y. Hotta, *J. Embryol. Exp. Morphol.* 45 (1978) 123–143.
- [52] G. Kohler, C. Milstein, *Nature* 256 (1975) 495–497.
- [53] S. Kondo, M. Booker, N. Perrimon, *Genetics* 183 (2009) 1165–1173.
- [54] R.J. Konopka, S. Benzer, *Proc. Natl. Acad. Sci. U.S.A.* 68 (1971) 2112–2116.
- [55] E.Z. Kvon, G. Stampfel, J.O. Yanez-Cuna, B.J. Dickson, A. Stark, *Genes Dev.* 26 (2012) 908–913.
- [56] M.C. Lafave, J. Sekelsky, *Genetics* 188 (2011) 749–752.
- [57] E.B. Lewis, F. Bacher, *Drosophila Info. Serv.* 43 (1968).
- [58] F.L. Liebl, K.M. Werner, Q. Sheng, J.E. Karr, B.D. McCabe, D.E. Featherstone, J. Neurobiol. 66 (2006) 332–347.
- [59] D.L. Lindsley, G.G. Zimm, *The Genome of Drosophila melanogaster*, Academic, San Diego, 1992.
- [60] J.T. Littleton, H.J. Bellen, *Genetics* 138 (1994) 111–123.
- [61] J.T. Littleton, M. Stern, M. Perin, H.J. Bellen, *Proc. Natl. Acad. Sci. U.S.A.* 91 (1994) 10888–10892.
- [62] J. Liu, C. Li, Z. Yu, P. Huang, H. Wu, C. Wei, N. Zhu, Y. Shen, Y. Chen, B. Zhang, W.M. Deng, R. Jiao, *J. Genet. Genomics* 39 (2012) 209–215.
- [63] T.G. Loukeris, B. Arca, I. Livadaras, G. Dialektaki, C. Savakis, *Proc. Natl. Acad. Sci. U.S.A.* 92 (1995) 9485–9489.
- [64] H. Luan, N.C. Peabody, C.R. Vinson, B.H. White, *Neuron* 52 (2006) 425–436.
- [65] T. Lukacsovich, Z. Asztalos, W. Awano, K. Baba, S. Kondo, S. Niwa, D. Yamamoto, *Genetics* 157 (2001) 727–742.
- [66] L. Manning, E.S. Heckscher, M.D. Purice, J. Roberts, A.L. Bennett, J.R. Kroll, J.L. Pollard, M.E. Strader, J.R. Lupton, A.V. Dyukareva, P.N. Doan, D.M. Bauer, A.N. Wilbur, S. Tanner, J.J. Kelly, S.L. Lai, K.D. Tran, M. Kohwi, T.R. Laverty, J.C. Pearson, S.T. Crews, G.M. Rubin, C.Q. Doe, *Cell Rep.* 2 (2012) 1002–1013.
- [67] M. Markstein, C. Pitsouli, C. Villalta, S.E. Celniker, N. Perrimon, *Nat. Genet.* 40 (2008) 476–483.
- [68] S.Q. Mehta, P.R. Hiesinger, S. Beronja, R.G. Zhai, K.L. Schulze, P. Verstreken, Y. Cao, Y. Zhou, U. Tepass, M.C. Crair, H.J. Bellen, *Neuron* 46 (2005) 219–232.
- [69] A. Metaxakis, S. Oehler, A. Klinakis, C. Savakis, *Genetics* 171 (2005) 571–581.
- [70] S.E. Mohr, N. Perrimon, *Wiley Interdiscip. Rev. RNA* 3 (2012) 145–158.
- [71] C. Molnar, A. Lopez-Varea, R. Hernandez, J.F. de Celis, *Genetics* 174 (2006) 1635–1659.
- [72] K. Nairz, P. Zipperlen, C. Dearolf, K. Basler, E. Hafen, *Genome Biol.* 5 (2004) R83.
- [73] N. Negre, C.D. Brown, L. Ma, C.A. Bristow, S.W. Miller, U. Wagner, P. Kheradpour, M.L. Eaton, P. Loriaux, R. Sealfon, Z. Li, H. Ishii, R.F. Spokony, J. Chen, L. Hwang, C. Cheng, R.P. Auburn, M.B. Davis, M. Domanus, P.K. Shah, C.A. Morrison, J. Zieba, S. Suchy, L. Senderowicz, A. Victorsen, N.A. Bild, A.J. Grundstad, D. Hanley, D.M. MacAlpine, M. Mannervik, K. Venken, H. Bellen, R. White, M. Gerstein, S. Russell, R.L. Grossman, B. Ren, J.W. Posakony, M. Kellis, K.P. White, *Nature* 471 (2011) 527–531.
- [74] R.A. Neumuller, F. Wirtz-Peitz, S. Lee, Y. Kwon, M. Buckner, R.A. Hoskins, K.J. Venken, H.J. Bellen, S.E. Mohr, N. Perrimon, *Genetics* 190 (2012) 931–940.
- [75] T.P. Newsome, B. Asling, B.J. Dickson, *Development* 127 (2000) 851–860.
- [76] J.Q. Ni, L.P. Liu, R. Binari, R. Hardy, H.S. Shim, A. Cavallaro, M. Booker, B.D. Pfeiffer, M. Markstein, H. Wang, C. Villalta, T.R. Laverty, L.A. Perkins, N. Perrimon, *Genetics* 182 (2009) 1089–1100.
- [77] J.Q. Ni, M. Markstein, R. Binari, B. Pfeiffer, L.P. Liu, C. Villalta, M. Booker, L. Perkins, N. Perrimon, *Nat. Methods* 5 (2008) 49–51.
- [78] J.Q. Ni, R. Zhou, B. Czech, L.P. Liu, L. Holderbaum, D. Yang-Zhou, H.S. Shim, R. Tao, D. Handler, P. Karpowicz, R. Binari, M. Booker, J. Brennecke, L.A. Perkins, G.J. Hannon, N. Perrimon, *Nat. Methods* 8 (2011) 405–407.
- [79] K.K. Norga, M.C. Gurganus, C.L. Dilda, A. Yamamoto, R.F. Lyman, P.H. Patel, G.M. Rubin, R.A. Hoskins, T.F. Mackay, H.J. Bellen, *Curr. Biol.* 13 (2003) 1388–1396.
- [80] C. Nusslein-Volhard, E. Wieschaus, *Nature* 287 (1980) 795–801.
- [81] C.J. O’Kane, W.J. Gehring, *Proc. Natl. Acad. Sci. U.S.A.* 84 (1987) 9123–9127.
- [82] D.M. Ornitz, R.W. Moreadith, P. Leder, *Proc. Natl. Acad. Sci. U.S.A.* 88 (1991) 698–702.
- [83] W.L. Pak, J. Grossfield, K.S. Arnold, *Nature* 227 (1970) 518–520.
- [84] U.B. Pandey, C.D. Nichols, *Pharmacol. Rev.* 63 (2011) 411–436.
- [85] A.L. Parks, K.R. Cook, M. Belvin, N.A. Dompe, R. Fawcett, K. Huppert, L.R. Tan, C.G. Winker, K.P. Bogart, J.E. Deal, M.E. Al-Herr, D. Grant, M. Marcinko, W.Y. Miyazaki, S. Robertson, K.J. Shaw, M. Tabios, V. Vysotskaia, L. Zhao, R.S. Andrade, K.A. Edgar, E. Howie, K. Killpack, B. Milash, A. Norton, D. Thao, K.

- Whittaker, M.A. Winner, L. Friedman, J. Margolis, M.A. Singer, C. Kocpczynski, D. Curtis, T.C. Kaufman, G.D. Plowman, G. Duyk, H.L. Francis-Lang, *Nat. Genet.* 36 (2004) 288–292.
- [86] J.S. Patton, X.V. Gomes, P.K. Geyer, *Nucleic Acids Res.* 20 (1992) 5859–5860.
- [87] K.L. Pepple, M. Atkins, K. Venken, K. Wellnitz, M. Harding, B. Frankfort, G. Mardon, *Development* 135 (2008) 4071–4079.
- [88] M.W. Perry, A.N. Boettiger, J.P. Bothma, M. Levine, *Curr. Biol.* 20 (2010) 1562–1567.
- [89] B.D. Pfeiffer, A. Jenett, A.S. Hammonds, T.T. Ngo, S. Misra, C. Murphy, A. Scully, J.W. Carlson, K.H. Wan, T.R. Laverty, C. Mungall, R. Svirskas, J.T. Kadonaga, C.Q. Doe, M.B. Eisen, S.E. Celniker, G.M. Rubin, *Proc. Natl. Acad. Sci. U.S.A.* 105 (2008) 9715–9720.
- [90] B.D. Pfeiffer, T.T. Ngo, K.L. Hibbard, C. Murphy, A. Jenett, J.W. Truman, G.M. Rubin, *Genetics* 186 (2010) 735–755.
- [91] V. Pirrotta, *Biotechnology* 10 (1988) 437–456.
- [92] C.J. Potter, L. Luo, *PLoS One* 5 (2010) e10168.
- [93] A.T. Quinones-Coello, L.N. Petrella, K. Ayers, A. Melillo, S. Mazzalupo, A.M. Hudson, S. Wang, C. Castiblanco, M. Buszczak, R.A. Hoskins, L. Cooley, *Genetics* 175 (2007) 1089–1104.
- [94] Y.S. Rong, K.G. Golic, *Science* 288 (2000) 2013–2018.
- [95] P. Rorth, *Proc. Natl. Acad. Sci. U.S.A.* 93 (1996) 12418–12422.
- [96] R.R. Roseman, J.M. Swan, P.K. Geyer, *Development* 121 (1995) 3573–3582.
- [97] G.M. Rubin, A.C. Spradling, *Science* 218 (1982) 348–353.
- [98] E. Ryder, M. Ashburner, R. Bautista-Llacer, J. Drummond, J. Webster, G. Johnson, T. Morley, Y.S. Chan, F. Blows, D. Coulson, G. Reuter, H. Baisch, C. Apelt, A. Kauk, T. Rudolph, M. Kube, M. Klimm, C. Nickel, J. Szidonya, P. Maroy, M. Pal, A. Rasmuson-Lestander, K. Ekstrom, H. Stocker, C. Hugentobler, E. Hafen, D. Gubb, G. Pflugfelder, C. Dorner, B. Mechler, H. Schenkel, J. Marhold, F. Serras, M. Corominas, A. Punset, J. Roote, S. Russell, *Genetics* 177 (2007) 615–629.
- [99] E. Ryder, F. Blows, M. Ashburner, R. Bautista-Llacer, D. Coulson, J. Drummond, J. Webster, D. Gubb, N. Gunton, G. Johnson, C.J. O’Kane, D. Huen, P. Sharma, Z. Asztalos, H. Baisch, J. Schulze, M. Kube, K. Kittlaus, G. Reuter, P. Maroy, J. Szidonya, A. Rasmuson-Lestander, K. Ekstrom, B. Dickson, C. Hugentobler, H. Stocker, E. Hafen, J.A. Lepesant, G. Pflugfelder, M. Heisenberg, B. Mechler, F. Serras, M. Corominas, S. Schneuwly, T. Preat, J. Roote, S. Russell, *Genetics* 167 (2004) 797–813.
- [100] E. Ryder, S. Russell, *Brief Funct. Genomic. Proteomic.* 2 (2003) 57–71.
- [101] R.K. Saiki, D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, H.A. Erlich, *Science* 239 (1988) 487–491.
- [102] C. Schertel, D. Huang, M. Bjorklund, J. Bischof, D. Yin, R. Li, Y. Wu, R. Zeng, J. Wu, J. Taipale, H. Song, K. Basler, *Dev. Cell* 25 (2013) 207–219.
- [103] O. Schuldiner, D. Berdnik, J.M. Levy, J.S. Wu, D. Luginbuhl, A.C. Gontang, L. Luo, *Dev. Cell* 14 (2008) 227–238.
- [104] K.J. Sepp, V.J. Auld, *Genetics* 151 (1999) 1093–1101.
- [105] S.K. Sharan, L.C. Thomason, S.G. Kuznetsov, D.L. Court, *Nat. Protoc.* 4 (2009) 206–223.
- [106] H.K. Shearin, A.R. Dvarishkis, C.D. Kozeluh, R.S. Stowers, *PLoS One* 8 (2013) e77724.
- [107] M.A. Simon, D.D. Bowtell, G.S. Dodson, T.R. Laverty, G.M. Rubin, *Cell* 67 (1991) 701–716.
- [108] A.C. Spradling, H.J. Bellen, R.A. Hoskins, *Proc. Natl. Acad. Sci. U.S.A.* 108 (2011) 15948–15953.
- [109] A.C. Spradling, G.M. Rubin, *Science* 218 (1982) 341–347.
- [110] A.C. Spradling, D. Stern, A. Beaton, E.J. Rhem, T. Laverty, N. Mozden, S. Misra, G.M. Rubin, *Genetics* 153 (1999) 135–177.
- [111] A.C. Spradling, D.M. Stern, I. Kiss, J. Roote, T. Laverty, G.M. Rubin, *Proc. Natl. Acad. Sci. U.S.A.* 92 (1995) 10824–10830.
- [112] D. St Johnston, *Wiley Interdiscip. Rev. Dev. Biol.* 2 (2013) 587–613.
- [113] D. St Johnston, *Nat. Rev. Genet.* 3 (2002) 176–188.
- [114] N. Staudt, A. Molitor, K. Somogyi, J. Mata, S. Curado, K. Eulenber, M. Meise, T. Siegmund, T. Hader, A. Hilfiker, G. Bronner, A. Ephrussi, P. Rorth, S.M. Cohen, S. Fellert, H.R. Chung, O. Piepenburg, U. Schafer, H. Jackle, G. Vorbruggen, *PLoS Genet.* 1 (2005) e55.
- [115] R.S. Stowers, *Fly (Austin)* 5 (2011).
- [116] R.S. Stowers, T.L. Schwarz, *Genetics* 152 (1999) 1631–1639.
- [117] J. Szabad, H.J. Bellen, K.J. Venken, *G3 (Bethesda)* 2 (2012) 1095–1102.
- [118] S.T. Thibault, M.A. Singer, W.Y. Miyazaki, B. Milash, N.A. Dompe, C.M. Singh, R. Buchholz, M. Demsky, R. Fawcett, H.L. Francis-Lang, L. Ryner, L.M. Cheung, A. Chong, C. Erickson, W.W. Fisher, K. Greer, S.R. Hartouni, E. Howie, L. Jakkula, D. Joo, K. Killpack, A. Laufer, J. Mazzotta, R.D. Smith, L.M. Stevens, C. Stuber, L.R. Tan, R. Ventura, A. Woo, I. Zakrajsek, L. Zhao, F. Chen, C. Swimmer, C. Kocpczynski, G. Duyk, M.L. Winberg, J. Margolis, *Nat. Genet.* 36 (2004) 283–287.
- [119] A.C. Tien, A. Rajan, K.L. Schulze, H.D. Ryoo, M. Acar, H. Steller, H.J. Bellen, *J. Cell Biol.* 182 (2008) 1113–1125.
- [120] C.Y. Ting, S. Gu, S. Guttikonda, T.Y. Lin, B.H. White, C.H. Lee, *Genetics* 188 (2011) 229–233.
- [121] N.E. Tunstall, A. Herr, B.M. de, C.G. Warr, *PLoS One* 7 (2012) e35641.
- [122] A. Urasaki, T. Mito, S. Noji, R. Ueda, K. Kawakami, *Gene* 425 (2008) 64–68.
- [123] K.J. Venken, H.J. Bellen, *Nat. Rev. Genet.* 6 (2005) 167–178.
- [124] K.J. Venken, H.J. Bellen, *Development* 134 (2007) 3571–3584.
- [125] K.J. Venken, H.J. Bellen, *Methods Mol. Biol.* 859 (2012) 203–228.
- [126] K.J. Venken, J.W. Carlson, K.L. Schulze, H. Pan, Y. He, R. Spokony, K.H. Wan, M. Koriabine, P.J. de Jong, K.P. White, H.J. Bellen, R.A. Hoskins, *Nat. Methods* 6 (2009) 431–434.
- [127] K.J. Venken, Y. He, R.A. Hoskins, H.J. Bellen, *Science* 314 (2006) 1747–1751.
- [128] K.J. Venken, J. Kaspruwicz, S. Kuenen, J. Yan, B.A. Hassan, P. Verstreken, *Nucleic Acids Res.* 36 (2008) e114.
- [129] K.J. Venken, E. Popodi, S.L. Holtzman, K.L. Schulze, S. Park, J.W. Carlson, R.A. Hoskins, H.J. Bellen, T.C. Kaufman, *Genetics* 186 (2010) 1111–1125.
- [130] K.J. Venken, K.L. Schulze, N.A. Haelterman, H. Pan, Y. He, M. Evans-Holm, J.W. Carlson, R.W. Levis, A.C. Spradling, R.A. Hoskins, H.J. Bellen, *Nat. Methods* 8 (2011) 737–743.
- [131] K.J. Venken, J.H. Simpson, H.J. Bellen, *Neuron* 72 (2011) 202–230.
- [132] P. Verstreken, T.W. Koh, K.L. Schulze, R.G. Zhai, P.R. Hiesinger, Y. Zhou, S.Q. Mehta, Y. Cao, J. Roos, H.J. Bellen, *Neuron* 40 (2003) 733–748.
- [133] P. Verstreken, C.V. Ly, K.J. Venken, T.W. Koh, Y. Zhou, H.J. Bellen, *Neuron* 47 (2005) 365–378.
- [134] P. Verstreken, T. Ohyama, C. Haueter, R.L. Habets, Y.Q. Lin, L.E. Swan, C.V. Ly, K.J. Venken, C.P. De, H.J. Bellen, *Neuron* 63 (2009) 203–215.
- [135] J. Wild, Z. Hradecna, W. Szybalski, *Genome Res.* 12 (2002) 1434–1444.
- [136] C. Wilson, R.K. Pearson, H.J. Bellen, C.J. O’Kane, U. Grossniklaus, W.J. Gehring, *Genes Dev.* 3 (1989) 1301–1313.
- [137] A. Witsell, D.P. Kane, S. Rubin, M. McVey, *Genetics* 183 (2009) 1187–1193.
- [138] R.C. Woodruff, M. Ashburner, *Genetics* 92 (1979) 117–132.
- [139] R.C. Woodruff, M. Ashburner, *Genetics* 92 (1979) 133–149.
- [140] T. Xu, G.M. Rubin, *Development* 117 (1993) 1223–1237.
- [141] R. Yagi, F. Mayer, K. Basler, *Proc. Natl. Acad. Sci. U.S.A.* 107 (2010) 16166–16171.
- [142] S. Yamamoto, W.L. Charng, N.A. Rana, S. Kakuda, M. Jaiswal, V. Bayat, B. Xiong, K. Zhang, H. Sandoval, G. David, H. Wang, R.S. Haltiwanger, H.J. Bellen, *Science* 338 (2012) 1229–1232.
- [143] R.G. Zhai, P.R. Hiesinger, T.W. Koh, P. Verstreken, K.L. Schulze, Y. Cao, H. Jafar-Nejad, K.K. Norga, H. Pan, V. Bayat, M.P. Greenbaum, H.J. Bellen, *Proc. Natl. Acad. Sci. U.S.A.* 100 (2003) 10860–10865.
- [144] L. Zheng, Y. Michelson, V. Freger, Z. Avraham, K.J. Venken, H.J. Bellen, M.J. Justice, R. Wides, *PLoS One* 6 (2011) e22956.
- [145] K. Zito, D. Parnas, R.D. Fetter, E.Y. Isacoff, C.S. Goodman, *Neuron* 22 (1999) 719–729.
- [146] J.F. de Celis, C. Molnar, A cautionary tale on genetic screens based on a gain-of-expression approach: The case of *LanB1*, *Fly (Austin)* 4 (2010) 24–29.
- [147] H. Ochman, A.S. Gerber, D.L. Hartl, Genetic applications of an inverse polymerase chain reaction, *Genetics* 120 (1988) 621–623.
- [148] K.J. Beumer, D. Carroll, Targeted genome engineering techniques in *Drosophila*, *Methods* 68 (1) (2014) 29–37.
- [149] S.E. Mohr, RNAi screening in *Drosophila* cells and in vivo, *Methods* 68 (1) (2014) 82–88.