

Flying in the face of total disruption

P Robin Hiesinger & Hugo J Bellen

More than 50% of the 13,666 *Drosophila melanogaster* genes are now reported to contain a *P*-element or *piggyBac* insertion. Some of these insertions have been used to create molecularly defined deletions spanning more than 50% of the genome.

For geneticists, the information gained by disrupting a gene of interest is often valuable, but targeted mutagenesis is labor-intensive in all multicellular eukaryotic model systems. In *Drosophila melanogaster*, two approaches are most often used to target mutations of a known gene: imprecise excision of transposable *P* elements¹, and homologous recombination². *P*-element insertions in the vicinity of or within a gene are the most common means of disrupting genes. Hence, genome-wide screens to obtain insertions in every *D. melanogaster* gene are a lofty but valuable goal. Three large projects, two publicly sponsored (the Berkeley *Drosophila* Gene Disruption project (GDP)^{3,4} and the European DrosDel project⁵) and one privately sponsored (Exelixis), were initiated a few years ago. Though not yet complete, these three collections are now made publicly available through team efforts with the Bloomington *Drosophila* Stock Center (GDP and Exelixis) and the Szeged Stock Center (DrosDel). On pages 283 and 288, the data of the privately sponsored efforts are reported. In the first paper, Thibault *et al.*⁶ describe many thousand new molecularly defined gene 'tags' created by *P*-element and *piggyBac* insertions. The second paper, by Parks *et al.*⁷, reports the generation of more than 500 molecularly defined genomic deletions (deficiencies) derived from the transposable elements described by Thibault *et al.*⁶.

P. Robin Hiesinger is at the Howard Hughes Medical Institute and Department of Molecular and Human Genetics and Hugo J. Bellen is at the Howard Hughes Medical Institute, the Department of Molecular and Human Genetics and the Program in Developmental Biology, Baylor College of Medicine, Houston, Texas, 77030, USA. e-mail: hbellen@bcm.tmc.edu

In these genome-wide gene disruption projects, the strategy is based on the availability of the finished genome sequence⁸, now in its third version (Release 3.1). Each randomly introduced transposable element insertion is sequenced and compared with all previously isolated events. Insertions that target single genes that were not previously targeted are normally saved. A comparison of the new collection with the stocks already publicly available should soon allow us to estimate the fraction of the genes that is now disrupted or amenable to targeted disruption (Fig. 1).

Tag, you're it!

P elements can easily be mobilized to insert in the genome and, through imprecise exci-

sion, can produce deletions of a specific locus. But *P* elements have two disadvantages: they have non-random insertion preferences, and they cause second-site lesions, also known as 'hit-and-run' events, which may occur in as many as 30% of strains⁹. This prompted Thibault *et al.* to switch to another transposable element that was recently added to the *D. melanogaster* toolbox: the species-nonspecific *piggyBac* element^{10,11}. *piggyBacs* seem to insert more randomly in the genome than *P* elements, although *piggyBacs* have been reported to insert preferentially into introns and cause a higher frequency of lethal insertions (9%) than would be expected for truly random insertions¹¹. Thibault *et al.* now report the

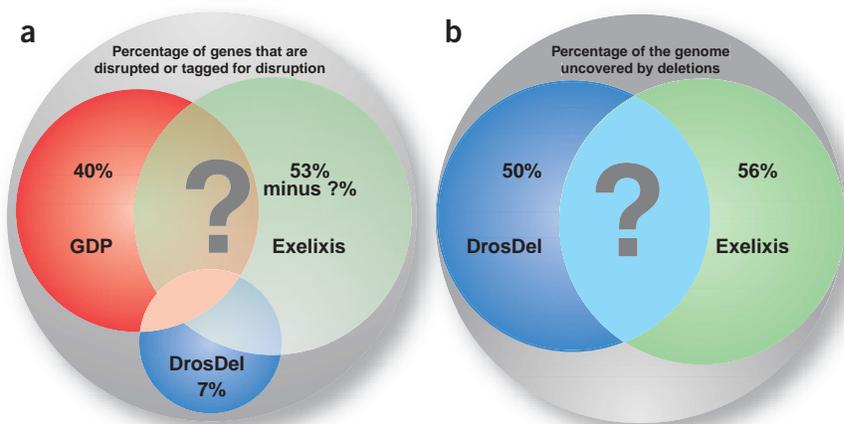


Figure 1 Genome coverage of the publicly available fly strains with tagged genes (a) and molecularly defined deletions (b). The new Exelixis collections^{6,7} (green) are compared with the publicly funded GDP^{3,4} (red) and DrosDel (E. Ryder, personal communication; blue) projects. The overlaps of these collections are not yet known. The new Exelixis transposon insertion collection (a) is depicted as an incomplete circle, as the actual number of disrupted genes or genes amenable to disruption has not yet been determined (53% minus ?%).

generation of ~20,000 *piggyBac* and 10,000 *P*-element insertions. They molecularly mapped 89% of the 29,682 insertions and found that 22% of the *piggyBacs* caused recessive lethality. This frequency is more than twice that previously observed with either *piggyBac*- or *P* element-based screens¹². The authors argue that these lethal insertions are not due to secondary lesions and that the lethality could be reverted by precise excision of the *piggyBacs* in six test cases. They also found that *piggyBacs* had insertional specificity, as they identified 26 insertional hot spots. Unfortunately, the authors did not specify how many insertions fell within these hot spots and whether they contributed to the high frequency of isolated lethal insertions.

The question of how many *piggyBac* insertions represent independently mutated genes is important, as *piggyBacs* cannot be excised imprecisely. Consequently, insertions causing no disruption or no obvious alteration in phenotype, though formally considered alleles, cannot easily be used to remove the gene. Hence, the 53% gene saturation stated by Thibault *et al.* means that the authors tagged 53% of the 13,666 *D. melanogaster* genes with transposon insertions. Although this is a substantial figure by any measure (the GDP has currently reached 40%; ref. 4), a high percentage of these genes are not disrupted nor can they be disrupted by imprecise excision (in contrast to the GDP collection). A comparison of the new collection with the stocks already available to the public should allow us to estimate the fraction of genes that is now disrupted or amenable to targeted disruption (Fig. 1). Fortunately, as reported by Parks *et al.*⁷, the collection of *piggyBac* insertions described by Thibault *et al.* provides a tool for the generation of molecularly defined deletions. We

anticipate that this deletion set will become the most used part of this collection and consider it a true leap forward in the fly field.

Defining deficient

Deficiencies in *D. melanogaster* have been used for decades to map genes, to identify suppressor and enhancers through dominant modifier screens and to define null phenotypes. Yet the deficiency kits distributed by the Bloomington *Drosophila* Stock Center consist mostly of irradiation- or chemical mutagen-induced aberrations in genetically complex strains from various sources, and their molecular breakpoints are largely unknown. Kevin Cook and his colleagues therefore attempted to fill the gaps and collect more defined deficiencies by creating deletions of genomic DNA between *P*-element insertions. Parks *et al.* report the generation of 45 new deficiencies, most created by excision of *P* elements in *trans*. This strategy was originally based on the observation that *P*-element insertions on the same chromosome (in *cis*) can be excised in tandem to create deletions¹³. This strategy was adapted (J. Roote, D. Gubb & M. Ashburner, personal communication) to create deletions between *P* elements in *trans*, thereby avoiding the tedious preliminary recombination step. Unfortunately, this methodology does not always generate precise breakpoints at the *P*-element insertion sites.

In contrast, the method used by the Exelixis group⁷ (and the DrosDel consortium⁵) to generate deletions expands on the use of FRT sites in the transposable elements. Introduction of FLP recombinase will induce defined deletions through homologous recombination on recognition of two nearby transposable elements in *trans* carrying FRT sites in the same orientation¹⁴.

This method allowed the Exelixis team to isolate 519 molecularly defined deficiencies, averaging 140 kb and covering 56% of the genome in an isogenic background (Fig. 1). In an ongoing effort, the DrosDel project has so far generated 3,243 molecularly defined insertions and is aiming to generate 600 tiled deletions with an average size of 600 kb (ref. 5). So far, their *P*-element insertions are associated with 977 distinct genes, and their deletion collection covers 50% of the genome, with deletions averaging 463 kb (E. Ryder, personal communication; Fig. 1). The overlap of the new collections is not yet known. In addition, Cook and his colleagues have selected 3,800 lines from the Exelixis collection to generate a total of 2,500 tiled deletions averaging 220 kb. This will essentially create a kit of molecularly defined deletions that should have a breakpoint every seven genes on average (K. Cook, personal communication). These new deficiency kits, together with the possibility to custom-design smaller deletions^{5,7,15}, will greatly strengthen the genetic foundations on which the fly community relies.

1. Robertson, H.M. *et al. Genetics* **118**, 461–470 (1988).
2. Rong, Y.S. *et al. Genes Dev.* **16**, 1568–1581 (2002).
3. Spradling, A.C. *et al. Genetics* **153**, 135–177 (1999).
4. Bellen, H.J. *et al. Genetics* (in the press).
5. Ryder, E. *et al. Genetics* (in the press).
6. Thibault, S.T. *et al. Nat. Genet.* **36**, 283–287 (2004).
7. Parks, A.L. *et al. Nat. Genet.* **36**, 288–292 (2004).
8. Adams, M.D. *et al. Science* **287**, 2185–2195 (2000).
9. Salzberg, A. *et al. Genetics* **147**, 1723–1741 (1997).
10. Handler, A.M. & Harrell, R.A. 2nd. *Insect Mol. Biol.* **8**, 449–457 (1999).
11. Hacker, U., Nystedt, S., Barmchi, M.P., Horn, C. & Wimmer, E.A. *Proc. Natl. Acad. Sci. USA* **100**, 7720–7725 (2003).
12. Bellen, H.J. *Plant Cell* **11**, 2271–2281 (1999).
13. Cooley, L., Thompson, D. & Spradling, A.C. *Proc. Natl. Acad. Sci. USA* **87**, 3170–3173 (1990).
14. Golic, K.G. & Golic, M.M. *Genetics* **144**, 1693–1711 (1996).
15. Huet, F. *et al. Proc. Natl. Acad. Sci. USA* **99**, 9948–9953 (2002).