

Mapping and identification of essential gene functions on the X chromosome of *Drosophila*

Annette Peter, Petra Schöttler, Meike Werner, Nicole Beinert, Gordon Dowe, Peter Burkert, Foteini Mourkioti, Lore Dentzer, Yuchun He¹, Peter Deak^{2,3}, Panayiotis V. Benos^{4,+}, Melanie K. Gatt^{2,3}, Lee Murphy⁵, David Harris⁵, Bart Barrell⁵, Concepcion Ferraz⁶, Sophie Vidal⁶, Christine Brun⁶, Jacques Demaille⁶, Edouard Cadieu⁷, Stephane Dreano⁷, Stéphanie Gloux⁷, Valerie Lelaure⁷, Stephanie Mottier⁷, Francis Galibert⁷, Dana Borkova⁸, Belen Miñana⁸, Fotis C. Kafatos⁸, Slava Bolshakov^{8,9}, Inga Sidén-Kiamos⁹, George Papagiannakis⁹, Lefteris Spanos⁹, Christos Louis^{9,10}, Encarnación Madueño¹¹, Beatriz de Pablos¹¹, Juan Modolell¹¹, Alain Bucheton⁶, Debbie Callister³, Lorna Campbell³, Nadine S. Henderson³, Paul J. McMillan³, Cathy Salles³, Evelyn Tait³, Phillippe Valenti³, Robert D.C. Saunders^{3,12}, Alain Billaud¹³, Lior Pachter¹⁴, Robert Klapper¹⁵, Wilfried Janning¹⁵, David M. Glover^{2,3}, Michael Ashburner^{2,4}, Hugo J. Bellen¹, Herbert Jäckle[‡] & Ulrich Schäfer[‡]

Max-Planck-Institut für Biophysikalische Chemie, Abt. Molekulare Entwicklungsbiologie, Am Fassberg, 37077 Göttingen, ⁸European Molecular Biology Laboratory, 69012 Heidelberg, ¹⁵Institut für Allgemeine Zoologie und Genetik, Westfälische Wilhelms-Universität, Schlossplatz 5, 48149 Münster, Germany, ¹Howard Hughes Medical Institute, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, ¹⁴Department of Mathematics, University of California at Berkeley, Berkeley, CA, USA, ²Department of Genetics, University of Cambridge, Cambridge CB2 3EH, ⁴EMBL Outstation—The European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, ⁵Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, ³Department of Anatomy and Physiology, CRC Cell Cycle Genetics Group, University of Dundee, Dundee DD1 4HN, ¹²Department of Biological Sciences, The Open University, Milton Keynes MK7 6AA, UK, ⁶Montpellier University Medical School, Institut de Génétique Humaine, CNRS, 114 rue de la Cardonille, 34396 Montpellier Cedex 5, ⁷UPR 41, CNRS, Recombinations Génétiques, Faculté de Médecine, 2 Avenue du Pr Leon Bernard, 35043 Rennes Cedex, ¹³Fondation Jean Dausset—CEPH (Centre d'Etude du Polymorphisme Humain), 27 rue Juliette Dodu, 75010 Paris, France, ⁹Institute of Molecular Biology and Biotechnology, FORTH, Heraklion, ¹⁰Department of Biology, University of Crete, Heraklion, Greece and ¹¹Centro de Biología Molecular Severo Ochoa, CSIC and Universidad Autónoma de Madrid, 28049 Madrid, Spain

Received October 11, 2001; revised and accepted November 22, 2001

The *Drosophila melanogaster* genome consists of four chromosomes that contain 165 Mb of DNA, 120 Mb of which are euchromatic. The two *Drosophila* Genome Projects, in collaboration with Celera Genomics Systems, have sequenced the genome, complementing the previously established physical and genetic maps. In addition, the Berkeley *Drosophila* Genome Project has undertaken large-scale functional analysis based on mutagenesis by transposable P element insertions into autosomes. Here, we present a

large-scale P element insertion screen for vital gene functions and a BAC tiling map for the X chromosome. A collection of 501 X-chromosomal P element insertion lines was used to map essential genes cytogenetically and to establish short sequence tags (STSs) linking the insertion sites to the genome. The distribution of the P element integration sites, the identified genes and transcription units as well as the expression patterns of the P-element-tagged enhancers is described and discussed.

⁺Present address: Department of Genetics, School of Medicine, Washington University, 4566 Scott Avenue, St Louis, MO 63110, USA

[‡]Corresponding authors. Tel: +49 551 201 1482; Fax: +49 551 201 1755; E-mail: uschaef@gwdg.de or hjaeckl@gwdg.de

INTRODUCTION

Drosophila melanogaster has been a model organism for genome research for almost 90 years, starting with Sturtevant's demonstration that recombination frequencies could be used to map genes in a linear order along the chromosome (Sturtevant, 1913). Physical maps that relate genes to physical sites on the chromosomes (Bridges, 1935), together with the precise mapping of cloned DNA fragments by *in situ* hybridization to polytene chromosomes, enabled the first chromosomal maps of cloned genomic DNA segments (Wensink *et al.*, 1974). These achievements, the development of procedures for molecular screening (Grunstein and Hogness, 1975) and the assembly of large chromosomal contigs (Bender *et al.*, 1983) have facilitated systematic attempts to construct P1-based physical maps of the *Drosophila* genome. These maps cover ~90% of the euchromatic genome and are punctuated by short sequence tag (STS) sites that are, on average, 50 kb apart (Kimmerly *et al.*, 1996). More recently, a whole genome shotgun sequencing project, a collaborative effort between Celera Genomics and the Berkeley *Drosophila* Genome Project (BDGP), has provided a 95% coverage of the *Drosophila* genome (Adams *et al.*, 2000; Rubin *et al.*, 2000b). In addition, expressed sequence tag (EST) projects have generated sequence information from >100 000 transcripts (Rubin *et al.*, 2000a).

To assign function to the DNA sequences, the BDGP has undertaken large-scale functional analysis of the genome through gene-disruption projects that are based on mutagenesis by transposable P element insertion. These projects have focused on the isolation of essential genes on autosomes. The current collection of strains disrupt at least 1200 different genes, representing ~30% of the estimated 3600 genes on the autosomes that are mutable to easily scorable phenotypes, mainly sterility or lethality (Spradling *et al.*, 1995, 1999). Here, we report a large-scale P element insertion screen to identify essential genes on the X chromosome, as well as the construction of an X-chromosomal BAC tiling map, which complements other ongoing studies of the *Drosophila* genome.

RESULTS AND DISCUSSION

Tiling map of the X chromosome

A physical map of the X chromosome was constructed by hybridizing 521 probes to high-density filters bearing both the EDGP and BDGP BAC libraries. 5771 BAC clones were identified by hybridization. Eleven contigs were assembled using a combination of manual analysis of hybridization data and *in situ* hybridization. These span the X chromosome from the subtelomeric region of 1A to 20C near the base of the chromosome arm (Figure 1; supplementary information available on the website <http://www2.open.ac.uk/biology/research/molecular-genetics/EDGPmap.html>). The longest contig, contig 1, spans divisions 1–6 and comprises both cosmid clones (Siden-Kiamos *et al.*, 1990) and BAC clones. Divisions 1–3 were mapped to yield template clones for sequencing (Benos *et al.*, 2000, 2001), and in this region the BAC libraries were screened principally to link cosmid contigs.

Probe sequences were selected from several sources. The EDGP cosmid mapping project determined many sequence tags

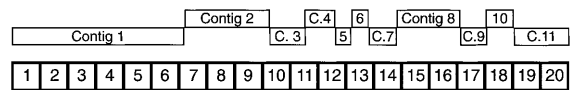


Fig. 1. Diagram showing the contig distribution along the X chromosome. Contigs are indicated by the boxes above the schematic representation of the X chromosome. Contig 1 is composed of cosmids and BACs.

from the termini of cosmid clone inserts. These sequence tags correspond to cosmid clones of known cytological location, with apparently single-copy inserts that are members of contigs (Madueno *et al.*, 1995). Some additional tags were determined from isolated cosmid clones of known cytological location. Additional probes were derived from sequences reported in databases, generally corresponding to genes of known cytological map position and from sequences of P element insertions, again of known cytological position. Probes used were selected partly on the basis of cytological position, in order to enhance the distribution of probe sequences along the chromosome.

Example for inter-contig gaps

Two probes (ES00262, mapping to 7B1-2, and ES00263, mapping to polytene chromosome bands 7B1-6) fail to hybridize to BAC clones from either library. Both probes are derived from mapped P1 clones. No BAC clones hybridizing to probes on either side of ES00262 and ES00263 were observed. The explanation may be that the sequences from which these probes were derived are incorrect or that there may be polymorphism between the *Drosophila* strains used in the construction of the BAC and P1 libraries. This gap represents a separation between contig 1 (1A to 7A) and contig 2 (7B to 10A). Other gaps generally correspond to intensely staining polytene chromosome bands (e.g. those in 10B, 11CD, 13A and 17A) and presumably represent long regions with few probes. In a number of cases, clones repeatedly failed to give signals with probes to which, on the basis of cytological location and other hybridization data, they might be expected to hybridize. Possible explanations include clones being absent on certain filters, cloned DNA segments bearing segmental deletion, and other failures in hybridization.

Generation of P element insertions

The euchromatin of the *Drosophila* X chromosome comprises ~22 Mb of DNA, containing 2182 predicted protein coding genes (Adams *et al.*, 2000). As an estimated two-thirds of all *Drosophila* genes show no obvious loss-of-function phenotype (Miklos and Rubin, 1996), ~800 of the X-chromosomal genes are mutable to a scorable phenotype such as sterility or lethality. In order to assess such genes in a systematic manner, we generated a collection of lethal P element insertions on the X chromosome.

The crossing scheme for generating X-chromosomal *P{lacW}* insertion lines is depicted in Figure 2. Females with a reinsertion of the *P{lacW}* element, as shown by the absence of the dominant second chromosome marker *Curly* and the *white*⁺ eye colour from *P{lacW}*, were individually mated to *FM7c* males.

A. Peter et al.

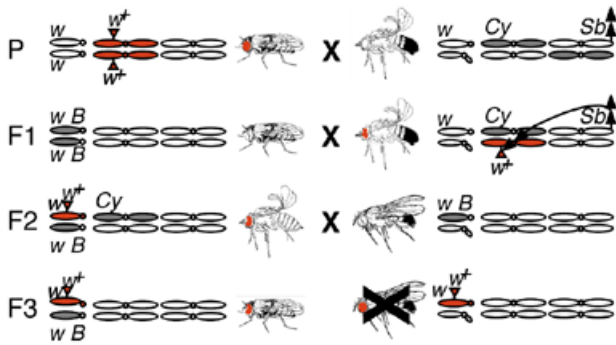


Fig. 2. Crossing scheme for isolating lethal X-chromosomal insertion lines. The sex chromosomes (left pair in each genotype) and the two large autosomes are shown schematically with females on the left. Relevant mutations are labelled accordingly. The $P\{lacW\}$ insertion is indicated by the triangle and its $white^+$ selectable marker is represented by the darker chromosome. The transposase source is indicated by the double-headed arrow. Balancer chromosomes are hatched. Absence of the crossed-out males in the F_3 generation indicates a potential lethal $P\{lacW\}$ insertion. For further details, see Methods.

57 105 such crosses were initially set and 39 900 (69.9%) produced offspring. From these crosses, 501 strains (1.3%) were derived that were either hemizygotously lethal or semilethal (defined as <20% viability compared to the sibling balancer males). Taking into account that the X chromosome represents roughly one-seventh of the genomic target size, 1.3% corresponds to a frequency of ~9% lethal P element insertions, a rate similar to those reported for autosomal insertions with comparable P element vector (Cooley et al., 1988; Bellen, 1999). In addition, 73 lines were isolated that carried a visible mutation or that were sterile in one sex.

Distribution of P element insertions

The insertion sites of 401 strains were determined by *in situ* hybridization on polytene chromosome squashes. In addition, DNA from the insertion sites was isolated either by plasmid rescue or by inverse PCR to generate an STS for all the insertion lines. Figure 3 summarizes the distribution of the P integrations along the X chromosome, showing a higher frequency of $P\{lacW\}$ insertions in regions closer to the telomere. An observation that remains valid even when considering only non-redundant insertions. This skewed distribution is not mirrored by the relatively even distribution of transcription units along the X chromosome. Hence, it is likely that an inherent insertion site preference of P elements (Liao et al., 2000) is responsible for the observed distribution.

The quality of a P element collection depends largely on the proportion of strains that harbour single $P\{lacW\}$ insertions. *In situ* hybridizations to polytene chromosomes were carried out for 401 strains. 89.5% (359 lines) contain single insertions, 9.7% (39 lines) contain two insertions and 0.7% (three lines) contain three $P\{lacW\}$ insertions. However, for those 14 strains that showed a second hybridization signal in polytene region 3A, we were unable to confirm the second insertion site by molecular means. Conversely, molecular analysis identified eight strains

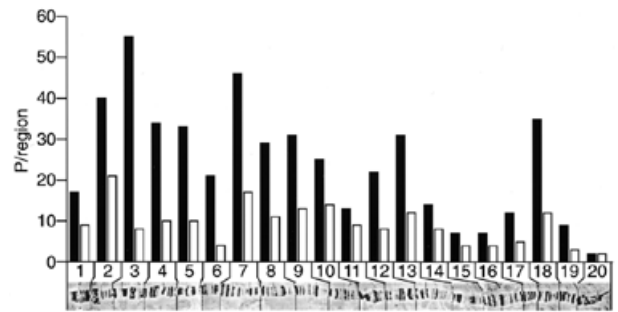


Fig. 3. $P\{lacW\}$ distribution along the X chromosome. The localization is based on the molecularly defined integration site. Open bars represent all insertions in a given polytene region whereas shaded bars indicate the number of genes affected.

that contain a second integration site and two containing three integrations where only one *in situ* hybridization signal was observed. In eight such cases, the insertion sites are <50 kb apart, and thus likely to be below the limit of resolution of the *in situ* hybridization technique. In four of these closely linked strains, the same transcription unit was hit by two P elements, indicative of local hopping during the remobilization of the starting $P\{lacW\}$. Collectively, these results indicate that >90% of the lines contain single P element insertions. Hence, this collection of P element insertion strains is comparable with the best P element collections generated to date (Spradling et al., 1999).

$P\{lacW\}$ insertion and an associated lethal phenotype can be causally linked if excision of the $P\{lacW\}$ restores wild type function. We tested a total of 206 lines by combining a transposase source $P\{ry^+ \Delta 2-3\}(99B)$ with the the $P\{lacW\}$ element (Bellen et al., 1989). 134 strains (65.0%) produced fully viable male progeny that lost the $white^+$ gene, suggesting that the P element excised precisely. The 65% value should be regarded as a minimum, since the reversion tests were performed with only 10 females bearing both the $P\{lacW\}$ insertion and the transposase source. Some P elements are difficult to excise and hence exhibit low mobilization frequencies and will almost certainly have been missed in this type of protocol.

Characterization of the lines

As the β -galactosidase gene in $P\{lacW\}$ can report the expression pattern of adjacent genes, we immunohistochemically stained embryos with anti- β -gal antibody. 398 of the 501 strains (79.4%) expressed the enzyme. This is considerably higher than the 64% reported after random autosomal insertion of the same P element (Bier et al., 1989) and may result from the fact that the X-chromosomal lines are selected for insertions affecting essential genes. Representative patterns for each line were deposited in the FlyView database, <http://flyview.uni-muenster.de/> (Janning, 1997), where they can be searched by various criteria.

The lethal phase was determined in 497 of the 501 lines. The most frequent stage where development stops is the larval stage, with 29.2% of the analyzed lines (see Supplementary data). This is not unexpected, since homozygous mutants can often survive

up to this stage by relying on maternal contribution of gene products deposited by the heterozygous mother. In addition, the majority of the lines probably do not carry a null allele, but rather hypomorphic mutations.

We have generated an STS flanking the P element insertion site for 496 of the 501 lines by either plasmid rescue or by inverse PCR (the sequences were submitted to the EMBL nucleotide database; see Supplementary data). The integration sites were determined by BLAST search to the published *Drosophila* genomic sequence (Adams *et al.*, 2000). In cases where the gene annotation is supported by EST data, we can unambiguously identify the affected gene. For predicted genes whose open reading frame is only annotated we assumed that the insertion is integrated in the vicinity of a putative transcription start site upstream of the nearest translational start codon, since there is a strong preference for P integrations in the 5'-end of genes (Spradling *et al.*, 1995).

Since the value of this collection is, in part, determined by the number of new genes that are affected, we will briefly discuss our mapping data. A total of 513 STS sequences were generated from 496 strains. In five strains, repeated attempts to isolate flanking sequences failed, and in 16 the identification of the integration site was uninformative, since the P element is inserted in repetitive DNA. In 11 of these lines, the insertion is within a *yoyo* retrotransposon (Whalen and Grigliatti, 1998). The other 497 STSs were generated from the remaining 480 lines and are derived from unique X-chromosomal sequences. Five insertions were >7 kb away from any annotated gene and it is unclear whether the P integration is responsible for the observed lethal phenotype. Three additional P insertions occurred in the vicinity of known EP insertions but were some distance from the next annotated gene, suggesting that the P elements have inserted into a putative promoter region of an unknown gene. Fourteen P elements, representing five gene pairs, may affect two genes. In each of these cases, one gene is localized within a large intron of another gene.

183 *P{lacW}* insertions affect 52 different genes that were previously characterized at both the molecular and genetic level, e.g. *Notch*, *pebbled* and *short gastrulation*. Phenotypically well characterized mutations of previously unidentified genes form a second group, consisting of two loci. These are *trol*, formerly known as *zw1*, with 12 insertions and *stardust* with one insertion. The largest group of insertions disrupts genes for which no mutation has previously been reported. In most cases, some molecular information is available or the genes have been predicted by computer algorithms (Adams *et al.*, 2000). In summary, 301 insertions affect 130 genes (including the five nested gene pairs) that remain to be characterized. Among this class are transcription units that code for proteins putatively involved in signaling cascades, e.g. *Rala*, protein degradation, or *Rpt4*, intracellular transport, or *Ntf-2*, as well as many other cellular and developmental processes.

To determine the frequency by which genes were hit, we relied on molecular information. Indeed, allelism cannot easily be determined by genetic criteria, as complementation tests between X chromosome lethals are not possible without first introducing a duplication that rescues the lethality. Genes on the X chromosome that are hit only once represent 96 out of 490 insertions (19.6%). This compares to a rate of 25.8 and 28.8% in the second and third chromosome collection, respectively

(Spradling *et al.*, 1999). The 60 so-called warm spots (2–5 insertions per gene; Spradling *et al.*, 1999) correspond to 169 X-chromosomal insertions (34.5%), similar to what was observed for the second and third chromosomes (35.8 and 34.2%, respectively). There are 23 hot spots on the X chromosome that are hit at least six times, resulting in a total of 225 insertions (45.9%). The frequencies for the second and third chromosome hot spots are 39.1 and 38.1%, respectively. The six genes that were hit most frequently are: *Notch*, 22 insertions; *Trf2*, 17; *inx2*, 15; *ras*, 14; and *act5C*, *ctp* and *trol* 12 each. In total, these six hot spots contained 104 insertions, more than all the transcription units that were hit once. In summary, in this X-chromosomal collection, single hits are under-represented and hot spots are over-represented, when compared to the autosomal P insertion collections. Taken together, of the X-chromosomal genes mutated in our screen each is hit, on average, 2.7 times, whereas the figures for the autosomal collection are 2.3 and 2.2, respectively.

CONCLUSIONS

Saturation mutagenesis studies give a ratio of lethal complementation groups to polytene chromosomes bands of 0.81 (Ashburner *et al.*, 1999). This estimate predicts ~820 essential genes for the X chromosome (1012 bands). Hence, our collection represents mutations in slightly <25% of the X-chromosomal essential genes and is similar to the data obtained for the autosomal collection of lethal P insertions (Spradling *et al.*, 1999). The P element screen for X-chromosomal genes required for adult viability and the BAC tiling map reported here represent a valuable tool for the *Drosophila* research community. The gene disruption collection provides the first opportunity to link ~130 X-chromosomal genes with a phenotype. The majority of the lines are available from the Bloomington stock center, together with their *in situ* hybridization data, and the enhancer trap expression patterns of these strains have been deposited in the FlyView database. Using these resources, the *Drosophila* research community can utilize the collection for the functional and molecular characterization of affected genes. The BAC tiling map can, on the other hand, serve as a reference point for mapping sequences and, more importantly, it will be essential as a source for a particular genomic fragment, e.g. for germ line transformation. The combined efforts of the community will not only increase our knowledge of the model organism *Drosophila* but will also, by virtue of the evolutionary conservation of many genes and processes, shed light on human gene function.

METHODS

***Drosophila* strains and embryos.** The following fly strains (Lindsley and Zimm, 1992; FlyBase, 1999) were used in the experiments. The starting element, *P{lacW}* (Bier *et al.*, 1989) originated from two male sterile insertion strains, 1040B and 1260A, which each contain a single transposon on chromosome 2. The strain used to supply transposase activity was *w; CyO/wg^{Sp}; TM6/Sb P{ry⁺ Δ2-3}(99B)* (Robertson *et al.*, 1988). The female fertile *FM6* and the female sterile *FM7c* chromosomes were employed as X chromosome balancers. In addition to being fertile in both sexes, *FM6* has the advantage of carrying a *white* null allele, thereby facilitating the detection of low expression of the *white⁺* marker gene.

A. Peter et al.

Crossing scheme (see also Figure 2)

P: *w/w; P{lacW}/P{lacW} × w/Y; wg^{SP}/CyO; ry Sb P{ry⁺ Δ2-3}(99B)/TM6B*

F₁: *FM6/FM6 × w/Y; P{lacW}/CyO; ry Sb P{ry⁺ Δ2-3}(99B)/+*

F₂: *w/FM6; CyO/+ (with P{lacW} somewhere) × FM7c/Y*

The absence of non-balancer bearing sons in the F₃ generation indicates a potential lethal insertion in the X chromosome.

Immunohistochemistry. Enhancer trap expression pattern was analyzed by immunohistochemistry. Embryos were collected on apple juice agar plates, fixed and stained with anti-β-galactosidase antibody (Cappel). The secondary peroxidase-coupled antibody was a biotinylated anti-rabbit Ig (Organon Teknika).

Molecular analysis. DNA was isolated from adult flies with QIAGEN DNeasy according to the instructions of the manufacturer. For isolation of flanking sequences by plasmid rescue (Pirrota, 1987), *EcoRI*-digested DNA was used, resulting in DNA from the 3'-end of *P{lacW}*. Inverse PCR (Silver, 1991) was performed essentially as described in the BDGP webpage (<http://www.fruitfly.org/>). The construction of the BAC tiling map and their application is described as supplementary information on the website ftp://ftp.ebi.ac.uk/pub/databases/edgp/200111/bac_probes.txt; the P insertion data are stored in the same directory under <ftp://ftp.ebi.ac.uk/pub/databases/edgp/200111/PX-lines.txt>.

Supplementary data. Supplementary data are available at *EMBO reports* Online.

ACKNOWLEDGEMENTS

We thank many colleagues for their help. This work was supported by the German Human Genome Project (grant 01 KW 9632/9) and a contract from the European Commission.

REFERENCES

- Adams, M.D. et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science*, **287**, 2185–2195.
- Ashburner, M. et al. (1999) An exploration of the sequence of a 2.9-Mb region of the genome of *Drosophila melanogaster*: the *Adh* region. *Genetics*, **153**, 179–219.
- Bellen, H.J. (1999) Ten years of enhancer detection: lessons from the fly. *Plant Cell*, **11**, 2271–2281.
- Bellen, H.J., O’Kane, C.J., Wilson, C., Grossniklaus, U., Pearson, R.K. and Gehring, W.J. (1989) P element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes Dev.*, **3**, 1288–1300.
- Bender, W., Spierer, P. and Hogness, D.S. (1983) Chromosomal walking and jumping to isolate DNA from the *Ace* and *rosy* loci and the *bithorax* complex in *Drosophila melanogaster*. *J. Mol. Biol.*, **168**, 17–33.
- Benos, P.V. et al. (2000) From sequence to chromosome: the tip of the X chromosome of *D. melanogaster*. *Science*, **287**, 2220–2222.
- Benos, P.V. et al. (2001) From first base: the sequence of the tip of the X chromosome of *Drosophila melanogaster*, a comparison of two sequencing strategies. *Genome Res.*, **11**, 710–730.
- Bier, E. et al. (1989) Searching for pattern and mutation in the *Drosophila* genome with a *P-lacZ* vector. *Genes Dev.*, **3**, 1273–1287.
- Bridges, C.B. (1935) Salivary chromosome maps with a key to the banding of the chromosome of *Drosophila melanogaster*. *J. Hered.*, **26**, 60–64.
- Cooley, L., Kelley, R. and Spradling, A. (1988) Insertional mutagenesis of the *Drosophila* genome with single P elements. *Science*, **239**, 1121–1128.
- Flybase (1999) The FlyBase database of the *Drosophila* genome projects and community literature. *Nucleic Acids Res.*, **27**, 85–88.
- Grunstein, M. and Hogness, D.S. (1975) Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl Acad. Sci. USA*, **72**, 3961–3965.
- Janning, W. (1997) FlyView, a *Drosophila* image database, and other *Drosophila* databases. *Semin. Cell. Dev. Biol.*, **8**, 469–475.
- Kimmerly, W. et al. (1996) A P1-based physical map of the *Drosophila* euchromatic genome. *Genome Res.*, **6**, 414–430.
- Liao, G.C., Rehm, E.J. and Rubin, G.M. (2000) Insertion site preferences of the P transposable element in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA*, **97**, 3347–3351.
- Lindsley, D.L. and Zimm, G.G. (1992) *The Genome of Drosophila melanogaster*. Academic Press, San Diego, CA.
- Madueno, E. et al. (1995) A physical map of the X chromosome of *Drosophila melanogaster*: cosmid contigs and sequence tagged sites. *Genetics*, **139**, 1631–1647.
- Miklos, G.L. and Rubin, G.M. (1996) The role of the genome project in determining gene function: insights from model organisms. *Cell*, **86**, 521–529.
- Pirrota, V. (1987) Vectors for P-mediated transformation in *Drosophila*. In Rodriguez, R.L. and Denhardt, D.T. (eds) *Vectors: A Survey of Molecular Cloning Vectors and their Uses*. Butterworths, Boston, MA, pp. 437–456.
- Robertson, H.M., Preston, C.R., Phillis, R.W., Johnson-Schlitz, D.M., Benz, W.K. and Engels, W.R. (1988) A stable genomic source of P element transposase in *Drosophila melanogaster*. *Genetics*, **118**, 461–470.
- Rubin, G.M., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M. and Harvey, D.A. (2000a) A *Drosophila* complementary DNA resource. *Science*, **287**, 2222–2224.
- Rubin, G.M. et al. (2000b) Comparative genomics of the eukaryotes. *Science*, **287**, 2204–2215.
- Siden-Kiamos, I. et al. (1990) Towards a physical map of the *Drosophila melanogaster* genome: mapping of cosmid clones within defined genomic divisions. *Nucleic Acids Res.*, **18**, 6261–6270.
- Silver, J. (1991) Inverse polymerase chain reaction. In McPherson, M.J., Quirke, P. and Taylor, G.R. (eds) *PCR: A Practical Approach*. IRL Press, Oxford, UK, pp. 137–146.
- Spradling, A.C., Stern, D.M., Kiss, I., Roote, J., Lavery, T. and Rubin, G.M. (1995) Gene disruptions using P transposable elements: an integral component of the *Drosophila* genome project. *Proc. Natl Acad. Sci. USA*, **92**, 10824–10830.
- Spradling, A.C., Stern, D., Beaton, A., Rhem, E.J., Lavery, T., Mozden, N., Misra, S. and Rubin, G.M. (1999) The Berkeley *Drosophila* Genome Project gene disruption project: single P element insertions mutating 25% of vital *Drosophila* genes. *Genetics*, **153**, 135–177.
- Sturtevant, A.H. (1913) The linear arrangement of six X-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool.*, **14**, 43–59.
- Wensink, P.C., Finnegan, D.J., Donelson, J.E. and Hogness, D.S. (1974) A system for mapping DNA sequences in the chromosomes of *Drosophila melanogaster*. *Cell*, **3**, 315–325.
- Whalen, J.H. and Grigliatti, T.A. (1998) Molecular characterization of a retrotransposon in *Drosophila melanogaster*, nomad, and its relationship to other retrovirus-like mobile elements. *Mol. Gen. Genet.*, **260**, 401–409.

DOI: 10.1093/embo-reports/kvf012