RESEARCH RESOURCES FOR *DROSOPHILA*: THE EXPANDING UNIVERSE

Kathleen A. Matthews, Thomas C. Kaufman* and William M. Gelbart‡*

Abstract | *Drosophila melanogaster* has been the subject of research into central questions about biological mechanisms for almost a century. The experimental tools and resources that are available or under development for *D. melanogaster* and its related species, particularly those for genomic analysis, are truly outstanding. Here we review three types of resource that have been developed for *D. melanogaster* research: databases and other sources of information, biological materials and experimental services. These resources are there to be exploited and we hope that this guide will encourage new uses for *D. melanogaster* information, materials and services, both by those new to flies and by experienced *D. melanogaster* researchers.

WHOLE-GENOME SHOTGUN SEQUENCING
A method of sequencing large genomes in which an effectively random sampling of sequencing reads is collected from a target genome. The sequence of the original intact DNA is inferred on the basis of overlap among the fragments.

FINISHED SEQUENCE
Complete sequence of a clone or genome, with a defined level of accuracy and contiguity.

**Department of Biology, Indiana University, Bloomington, Indiana 47405-3700, USA.*
‡Biological Laboratories, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA.
Correspondence to K.A.M.
e-mail: matthewk@indiana.edu

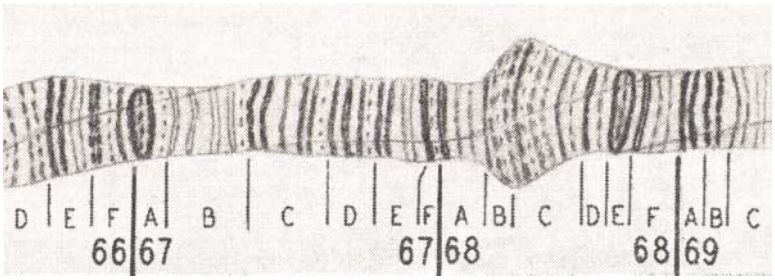
The position of the fruitfly *Drosophila melanogaster* as one of the most important genetic models in modern biology is unassailable. Few systems have the staying power of *D. melanogaster* at the leading edge of research into the mechanisms of inheritance, the construction of the animal body plan, the formation of the complex biological pattern, the function of the nervous system and the forces acting on genetic variation in natural populations. *Drosophila* is even more central to understanding the biology of higher animals today than it was when Thomas Hunt Morgan's group led the way in demonstrating the validity of the chromosome theory of inheritance[1]. *Drosophila melanogaster* provides an experimental system with a genome that is, on average, 5% of the size of mammalian genomes, and that shares most gene families and pathways with mammals, as well as some of the same tissues and organ systems[2,3]. The truly remarkable extent of similarity between invertebrates and vertebrates that has been revealed in recent years, not only between proteins but between entire cellular, developmental and behavioural pathways, ensures that *Drosophila* biology will continue to inform human biology and can be expected to have a direct impact on our understanding of human health[4].

A further factor that has contributed to the success of the fly as a biological model is the culture of free communication, large-scale production of research materials and the open sharing of resources that was established by the first *Drosophila* geneticists[5,6]. Almost 100 years ago, the fly group that was founded by Morgan established the principle of openness as an engine of progress, and the fruit of this tradition in the *Drosophila* community is still evident in the wealth of publicly available material and information for *Drosophila* research (see BOX 1 for an abbreviated history of resource development and sharing by members of the fly community). The financial support of national governments and charitable organizations has allowed access to these resources to be either free of charge (as is the case for all online information that is cited in this article) or available at a modest cost (in the case of most stocks, molecular materials and services).

The genomic era has markedly accelerated the development of research resources. Two exceptional milestones in the history of *Drosophila* as a research model were the public release of the assembled WHOLE-GENOME SHOTGUN SEQUENCE of *D. melanogaster* in 1999 by Celera Genomics and the Berkeley *Drosophila* Genome Project (BDGP) and the release of the annotated genome[7] that followed shortly after. Over the following 2 years, a FINISHED SEQUENCE[8] was produced by BDGP, the Human Genome Sequencing Center at Baylor College of

Box 1 | Seventy years of *Drosophila* resources

A comprehensive history of research milestones and resource development is outside the scope of this review; however, the sampling of notable contributions listed below provides a sense of the forces of community-mindedness that is at work in the field of *Drosophila* research.



Early reference catalogues

Information relating to research on *Drosophila melanogaster* was compiled for the first time by Calvin Bridges and Milislav Demerec in 1934 with the publication of the *Drosophila Information Service*[32] (DIS). Volume 1 of this collection consisted of contributions by most of the *Drosophila* researchers of the day and included the first published *Drosophila* stock list (572 stocks from the fly group at the California Institute of Technology[61]). Bridges' genetic linkage maps and POLYTENE CHROMOSOME cytological maps were first published the following year[62]. The figure shows a segment of the 3L polytene chromosome of *D. melanogaster* that was drawn by Bridges. In 1944, Bridges and Katherine Brehme[63] published the first document that outlined the genetic nomenclature that should be used for *D. melanogaster*, and a complete catalogue of mutant alleles, with map locations, phenotypic descriptions and in some cases illustrations (largely drawn from information that was previously contributed to the DIS). This work was later carried on with heroic effort by Dan Lindsley and colleagues[64,65], before passing the torch to FlyBase[16] in 1992. Irwin Herskowitz compiled and published *Drosophila* bibliographies for more than 35 years, beginning in 1947 (REF. 66).

Genetic tools

From 1948, a concerted effort by the California Institute of Technology *Drosophila* Stock Center (the predecessor of the Bloomington Stock Center) produced a set of greatly improved BALANCER CHROMOSOMES[61]. Thomas Alderson's demonstration of the value of ethylmethane sulphonate (EMS) as a mutagenic agent in *D. melanogaster*[67] in 1965, followed in 1968 by Edward Lewis and Fran Bacher's simple protocol for mutagenesis[68], led to the publication of almost 15,000 new mutant alleles[16]. In the early 1970s, Lindsley, Larry Sandler and colleagues[69] created a ground-breaking collection of nearly 400 Y–autosome translocations that allowed them, and the rest of the fly community, to create DEFICIENCIES and complementary duplications for 68% of the *D. melanogaster* genome (85% of chromosomes 2 and 3). A decade later the efforts of Christiane Nüsslein-Volhard, Eric Weischaus and colleagues to systematically identify genes that are involved in embryonic development[70–73] produced a plethora of information and materials that fuelled the study of development in *D. melanogaster* for the next two decades.

Germline transformation to whole-genome sequence

The development of a method for *D. melanogaster* transgenesis in the early 1980s by Allan Spradling and Gerald Rubin[74,75], and the general release of the requisite materials by these researchers, opened the *D. melanogaster* genome to a markedly improved level of manipulation. The next two decades saw huge growth in the quantity of both information and materials, a trend that was only accelerated with the completion of the *D. melanogaster* sequence in 2000 (REF. 7).

Image from REF. 62 © (1935) the American Genetic Association.

POLYTENE CHROMOSOME
A giant chromosome that is formed by many rounds of replication of the DNA. The replicated DNA molecules tightly align side-by-side in parallel register, which creates a non-mitotic interphase chromosome that is visible by light microscopy.

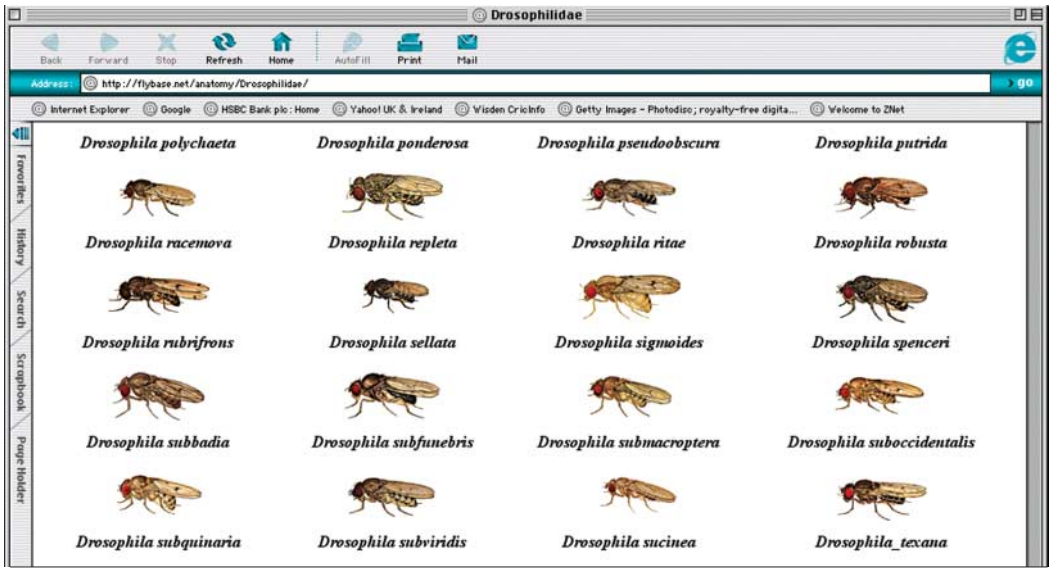
BALANCER CHROMOSOMES
Structurally rearranged chromosomes that prevent meiotic recombination between homologues. Because typical balancers are recessive lethal or sterile and carry visible mutations, they are used to create and maintain stable heterozygous stocks of lethal and sterile mutants.

DEFICIENCY
(Also known as a deletion.) A chromosome aberration in which the DNA encoding a portion of two or more genes is missing from the chromosome. Single gene or intragenic deletions are treated as alleles in the *Drosophila melanogaster* nomenclature.

Medicine and Celera Genomics, and was completely reannotated by FlyBase[9] (the principal online repository and access point for *Drosophila* information). The availability of a high-quality annotated genome sequence (which continues to improve in quality and richness) has transformed fly genetic analysis from a primarily gene-by-gene endeavour to one that operates on a genome-wide scale.

The momentum generated by the fly community during the past century makes this an extremely exciting time for *Drosophila* research. Powerful new tools for analysing gene function are available or are on the horizon[10]. A relatively complete sequence of the heterochromatic component of the genome will provide new insights into genome organization, modes of gene regulation and methods for characterizing highly repetitive DNA. With the addition of the complete annotated genome sequence of *Drosophila pseudoobscura*[11] in 2004, and with the genomes of another ten *Drosophila* species (see the *Drosophila* Species Genomes (DroSpeGe) web site) due to be released later in 2005, the fly is providing an unrivalled testing tool for comparative genomics. New computational tools for genome analysis will be developed that can then be applied to vertebrate genomes. Comparative analysis of these *Drosophila* genomes will refine our characterizations of coding sequences and markedly improve our understanding of regulatory sequences. Moreover, the resequencing of 7.3 Mb of DNA from 50 *D. melanogaster* genomes (see the *Drosophila* Population Genomic Project web site) will provide a unique resource for population genomics and contribute to an understanding of complex traits. The availability of these inter- and intraspecific genomes and their integration into the well-studied biology of flies will drive research into the mechanisms of speciation, the genetics of adaptation, genotype–phenotype

Box 2 | **'Working with *Drosophila*' bibliography**



Although a lot of information on *Drosophila* species is available on the World Wide Web, various printed compendia are also valuable. We list below some of the publications that are most likely to be useful to those starting to work with flies, and should certainly be at hand for most, if not all, laboratories that work on *Drosophila* species. Admittedly, some of the individual contributions in the edited volumes are a bit out of date; however, those covering anatomy and development, for example, are timeless. Moreover, the volumes on protocols can be extremely valuable to the neophyte.

Fly husbandry

Drosophila*: A Practical Approach*[77] | *Fly Pushing: the Theory and Practice of* Drosophila *Genetics*[78]

Fly genetics, anatomy and development

Biology of Drosophila[79]

The Genetics and Biology of Drosophila[80,81] | *The Genetics and Biology of* Drosophila[82] | *The Embryonic Development of* Drosophila melanogaster[83] | *The Development of* Drosophila melanogaster[84] | *Atlas of* Drosophila *Development*[85] | Drosophila *Eye Development*[86]

Laboratory protocols

Drosophila*: A Laboratory Handbook and Manual*[87] | Drosophila*: A Laboratory Handbook*[88] | Drosophila melanogaster: *Practical Uses in Cell and Molecular Biology*[89] | Drosophila *Protocols*[90] | Drosophila *Cells in Culture*[91]

Chromosomes

Chromosome Maps of Drosophila[92] | Drosophila *Cytogenetic Protocols*[93]

Evolution

Evolution in the Genus Drosophila[94] | *Progress and Prospects in Evolutionary Biology: the* Drosophila *Model*[95]

The figure is taken from the anatomy and images section of FlyBase (see Drawings of Drosophilidae in the Online links box), and shows a selection of Drosophilidae from southwestern USA and from Mexico. Image courtesy of FlyBase © (1993–2005) The Genetics Society of America.

mapping and chromatin organization, and they will enhance studies on developmental and biochemical mechanisms.

Here we review the information, material and service resources that are available for *Drosophila* research. Our primary goal is to provide a first-stop reference for identifying the best sources of information and materials for *Drosophila* research and to recommend some strategies for the effective navigation of those resources. We cover the principal online research databases that house information on the biology and genetics of *Drosophila*, plus a few sources of more specialized information such as neuroanatomy. Sources of wild-type and genetically defined fly stocks, cell lines, as well as molecular constructs, microarrays and antibodies are described; for each online source we highlight the distinguishing features of materials that are available. Options for outsourcing experimental work such as RNA interference (RNAi) screening are described, and should be relevant to both the neophyte and to the experienced *Drosophila* researcher. We attempt to describe the domain of each resource in sufficient detail to allow the reader to readily identify a resource to meet a given need. We include a few examples of how these resources might be used on a day-to-day basis in the laboratory, with an emphasis on how those new to flies might exploit these resources. We

Box 3 | Online resources on *Drosophila melanogaster* and related species

Below is a list of online sources of data and other information.

Genome

- Trace archive access: Assembly/Alignment/Annotation of 12 *Drosophila* Genomes (AAA) | Berkeley *Drosophila* Genome Project (BDGP) | *Drosophila* Species Genomes (DroSpeGe) | Ensembl | Entrez
- Sequence archive access: AAA | BDGP | *Drosophila* Polymorphism Database (DPDB) | DroSpeGe | Ensembl | Entrez | FlyBase
- Sequence searches: DPDB | DroSpeGe | Ensembl | Entrez | FlyBase | University of California, Santa Cruz Genome Browser (UCSC GB)
- Gene structure: *Drosophila melanogaster* Exon Database (DEDB) | Ensembl | Entrez | FlyBase | UCSC GB
- Annotated genes: DroSpeGe | Ensembl | Entrez | euGenes | FlyBase | FlyMine | UCSC GB
- SNPs: BDGP | Entrez | FlySNP
- Homologies: DroSpeGe | Ensembl | euGenes | FlyMine | Homophila | UCSC GB
- Maps: *Drosophila* Heterochromatin Genome Project (DHGP) | DroSpeGe | Ensembl | Entrez | euGenes | FlyBase | UCSC GB

Proteome

- Sequence archive: Entrez | Integr8 | UniProt
- Sequence searches: Entrez | Integr8 | UCSC GB | UniProt
- Annotated proteins: Entrez | Integr8 | PANTHER | UCSC GB | UniProt
- Protein interactions: Biomolecular Interaction Network Database (BIND) | CuraGen | Fly General Repository for Interaction Datasets (Fly GRID) | FlyMine | PIMRider

Other

- Sequence analysis tools: AAA | BDGP | DPDB | Ensembl | Entrez | Integr8 | UCSC GB | UniProt
- Gene expression: ArrayExpress | BDGP | FlyEx | FlyMine | FlyPNS | FlyTrap | flytrap | Gene Expression Omnibus (GEO) | Neuroblast lineages | PANTHER | Yale *Drosophila* Development Gene Expression Timecourse (Yale Dev Ex)
- Genetic/molecular interactions: BIND | CuraGen | FlyBase | Fly GRID | FlyMine | PANTHER | PIMRider
- Gene function/phenotypes: euGenes | FlyBase | FlyMine | FlyTrap | GenomeRNAi | Homophila | Integr8 | Interactive Fly | tracheal database (tracheal db) | UCSC GB | UniProt
- Pathways: BIND | PANTHER
- Human disease gene homology: Homophila | Yale Dev Ex
- Anatomy: BDGP | FlyBase | Flybrain | FlyMove | FlyPNS | flytrap | Neuroblast lineages | tracheal db
- Development: BDGP | FlyBase | Flybrain | FlyMove | FlyPNS | Interactive Fly | Neuroblast lineages | tracheal db | Yale Dev Ex
- Taxonomy: DroSpeGe | Entrez | Integr8 | TaxoDros
- Literature: Entrez | FlyBase | Flybrain | FlyMine | FlyPNS | Homophila | Integr8 | Neuroblast lineages | TaxoDros | tracheal db | UniProt
- *Drosophila* laboratory web sites: WWW Virtual Library: *Drosophila*

Below is a list of genetic stocks, cell lines, clones, libraries and arrays (for *D. melanogaster* unless noted otherwise).

Mutant alleles

- Point mutations: Bloomington *Drosophila* Stock Center (Bloom) | *Drosophila* Genetic Resource Center (Kyoto) | Tucson *Drosophila* Species Stock Center (Tucson)* | Zuker Collection
- Transposable-element insertion alleles: Bloom | Exelixis *Drosophila* Stock Collection at Harvard Medical School (Exelixis-HMS) | Gene Disruption Project *P*-Screen Database (GDP) | Kyoto | Szeged *Drosophila* Stock Center (Szeged)
- Upstream activating system alleles: Bloom | GDP | Kyoto
- GFP-tagged proteins: FlyTrap | Kyoto

Other stocks

- Deficiency kits: Bloom | Kyoto | Szeged
- Duplication kits: Bloom
- GAL4 drivers: Bloom | Exelixis–HMS | Kyoto | Szeged
- *FRT* insertions: Bloom | Kyoto | Szeged
- Heterochromatic insertions: DHGP
- Gene-expression reporters: Bloom | FlyTrap | FlyView | Kyoto
- Deficiency generators: Bloom
- Other aberrations: Bloom | Kyoto | Tucson*
- Wild-type lines: Bloom | Ehime *Drosophila* Species Center of Japan (Ehime)‡ | Kyoto‡ | Tucson*

Cell lines and antibodies

- Cell lines: American Type Culture Collection (ATCC) | *Drosophila* Genomics Resource Center (DGRC)‡
- Hybridomas/antibodies: ATCC | Developmental Studies Hybridoma Bank (DSHB) | East Asian Distribution Center (EADC)

Clones, libraries and microarrays

- cDNA molecules/ESTs: DGRC | geneservice*
- BACs: Arizona Genomics Institute (AGI)* | BACPAC Resources Center (BPRC)‡ | geneservice*
- Cosmids: geneservice*
- Fosmids: BPRC* | DGRC*
- Vectors: DGRC
- RNAi libraries: geneservice* | Open Biosystems*
- Microarrays: Affymetrix | DGRC | NimbleGen

Below is a list of services.

Genetic screening

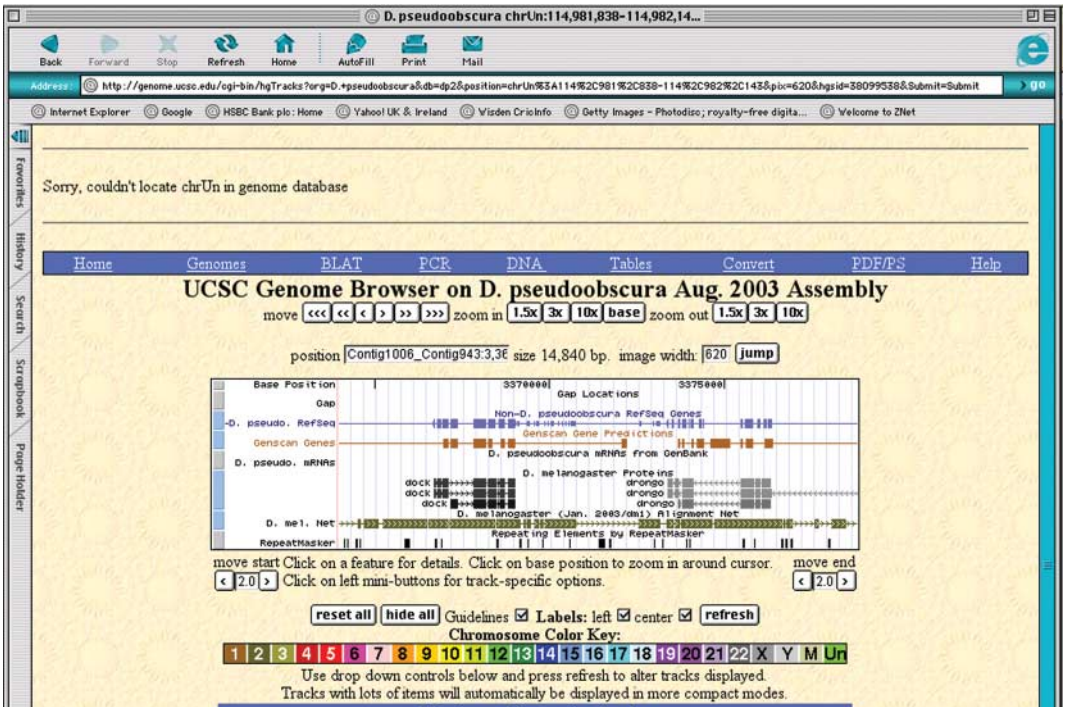
- EMS: *Drosophila* TILLING Project (Fly-TILL)
- RNAi: *Drosophila* RNAi Screening Center
- Custom: Genetic Services

Other

- Transgene construction: Genetic Services
- Germline transformation: Duke University Model System Genomics (Duke U. MSG) | EMBL | Genetic Services
- dsRNA injection: Duke U. MSG | Genetic Services
- *P*-element mapping: Duke U. MSG | geneservice
- Other sequencing: geneservice
- Immunohistochemistry: Genetic Services
- Embryo sorting: Duke U. MSG
- Stock construction: Genetic Services
- Microarray analysis: Duke U. MSG | geneservice
- Bioinformatics: Duke U. MSG

*Resources available for *Drosophila* species other than *D. melanogaster*.
‡Resources available for *D. melanogaster* and related species.

Box 4 | **Other genome-focused databases and web-based resources**



The web sites listed below provide integrated data across genera, allowing *Drosophila* data to be viewed in the context of genomic data from other organisms.

- UCSC Genome Bioinformatics[96] (see UCSC Genome Browser in BOX 3) contains the reference sequence and working draft assemblies for a large collection of genome sequences, including those for *Drosophila* species, and provides tools to view, extract and analyse sequence, annotation, expression and homology data. The image is of the UCSC Genome Browser (dp2 release, August 2003), showing the *Drosophila pseudoobscura* genome with selected annotations.
- Ensembl[97] presents up-to-date, automatically annotated sequence data for metazoan genomes, including *Drosophila* species.
- euGenes[98] provides a uniformly formatted summary of gene and genomic information from eukaryotic organism databases, including FlyBase.
- FlyMine provides genomic, expression and proteomics data for *D. melanogaster*, mosquito and other insects. The current release includes genome annotation data for *D. melanogaster* and *Anopheles gambiae*, yeast two-hybrid data for *D. melanogaster* and *Caenorhabditis elegans* (providing access to the *C. elegans* RNAi phenotypes), and orthologue/PARALOGUE data for these species.

The following resources provide data and links specific to continuing *Drosophila* genome projects.

- The Assembly/Alignment/Annotation of 12 *Drosophila* Genomes web page provides links to the genome sequencing centres and other sites that are involved in *Drosophila* sequencing and annotation, and other information relevant to comparative analysis of *Drosophila* genomes.
- The Berkeley *Drosophila* Genome Project provides information on libraries, cDNAs and other materials used in the sequencing and annotation of the *D. melanogaster* genome, an archive of sequence and annotation data files, information from the Gene Disruption Project, SNP maps and embryonic gene-expression data.

Image courtesy of UCSC Genome Bioinformatics © (2003) University of California, Santa Cruz.

also provide a bibliography of publications that focus on working with *Drosophila*; these self-help references range from primers on fly husbandry to exhaustive tomes on the nitty-gritty and the esoterica of *Drosophila* as a research organism (BOX 2).

An overview of all of the resources covered in this article and their uses are presented in BOX 3, and every resource is hyperlinked to the relevant web site in the online version of this article (the URL for each resource mentioned is provided in the Online links box).

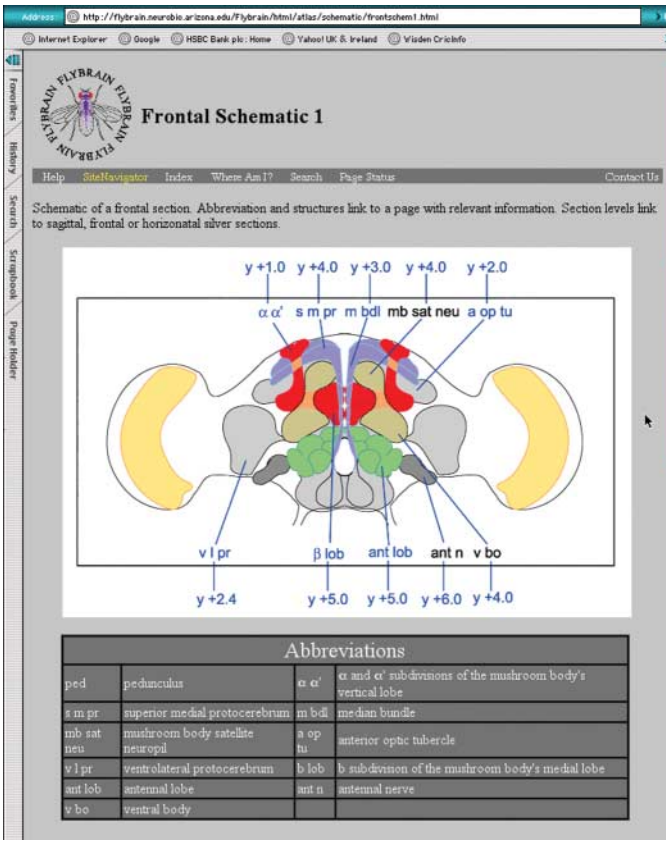
Information resources

The category of 'information resources' includes various archives and databases for the entire *Drosophila* genus, such as those containing nucleic-acid and protein data, plus various *Drosophila* species-specific resources. The genome annotation sets that are included in FlyBase are currently available for *D. melanogaster* and *D. pseudoobscura*, and have been adopted as the reference sequence by the species-wide genome databases. It must be noted that the annotated

PARALOGUE
One of a set of homologous genes in the same species that have evolved from a gene duplication and can be associated with a subsequent divergence of function.

Box 5 | Specialty *Drosophila*-based databases and web-based resources

The depth and breadth of work on *Drosophila* has spawned various specialized, web-accessible, *Drosophila* databases. The following list highlights some of these databases, but it is not exhaustive and new resources appear regularly. See the FlyBase External *Drosophila* Resources page for further resources. Unless indicated otherwise in the text below, data at these sites are from *Drosophila melanogaster*.



- The *Drosophila* Heterochromatin Genome Project (DHGP) is working to assemble, map and annotate high-quality, 'finished' sequences for a large portion of the heterochromatin (~30% of the genome). The site provides sequence and annotation data for heterochromatin before its incorporation into FlyBase, maps of the heterochromatin, and information on heterochromatic transgenic insertions that are not available elsewhere.
- *Drosophila melanogaster* Exon Database presents information on exons in a graphical format that incorporates all of the alternative splice variants, allowing users to easily determine the types and effects of alternative splicing.
- *Drosophila* Polymorphism Database provides a collection of all existing polymorphic sequences in the genus *Drosophila*.
- Homophila[99] provides information on the *D. melanogaster* cognates of human disease genes.
- InterActive Fly[100] focuses on the function, expression and interaction of developmental genes. It provides an overview of the genes, the biochemical processes they direct and the resulting morphological changes (this is particularly useful for those who are new to development).
- FlyMove[101] describes developmental processes with pictures, animation and movies.
- Flybrain[102] contains neuroanatomical descriptions of the CNS and PNS, and information on enhancer-trap insertions that are expressed in the brain. The image shows a schematic frontal section of the adult brain of *D. melanogaster*.
- flytrap documents the expression patterns of GAL4 enhancer-trap insertions in the adult brain through the use of images, movies and text descriptions.
- Neuroblast Lineages and Maps provide data on neuroblast development in the embryonic CNS, including a cell-by-cell description of its development, and expression patterns of neuroblast lineages genes.
- FlyPNS describes the embryonic and larval PNS, addressing discrepancies in the literature and providing the current view of the development and morphology of PNS-sensory neurons. PNS-specific antibodies and enhancer-trap lines are also listed.
- tracheal database offers an expert's guide to trachaeal development.
- FlyTrap[51] contains micrographs and text descriptions of GFP-tagged protein localization in ovaries. The insertions are useful for many cell biology applications.
- GenomeRNAi is a database of RNAi phenotypes that were obtained in cells in culture.
- Fly General Repository for Interaction Datasets provides data on gene and protein interactions.
- FlySNP[44] provides a high-density genome-wide map of SNPs and describes the assays that are available for SNP genotyping.
- TaxoDros covers the taxonomy of the Drosophilidae, including descriptions, biodiversity data and references.
- The WWW Virtual Library: *Drosophila* includes links to researchers' web sites and other *Drosophila* resources. Those who are new to flies and do not have a *Drosophila* colleague at their institution can use this site to identify fly workers on the basis of their institution and research interest.

Image courtesy of Flybrain[102] © (1995–2000) Flybrain.

sequence is a work-in-progress and is therefore revised regularly; these changes are propagated to the principal sequence databanks and every attempt is made to prevent data divergence. So, although the resources that are listed here include *Drosophila* genome data in different contexts, the underlying genome sequence and annotation data are the same across the sites. BOX 3 provides an overview of the online public resources that contain research information on *Drosophila* species. Many of these web sites are described in more detail below; others are briefly described in BOXES 4,5. Related software applications that must be downloaded and run locally, such as the Apollo genome-annotation viewer and editor[12], are not covered here, but can be found on some of the included sites.

The resources that are used and how they are used depends on what starting information is available and the nature of the query. In many cases, particularly for those not already working with *Drosophila*, that information will be a DNA or protein sequence and the goal will generally be to find known sequences — in any organism — that are similar to one's own. In this situation, the starting point will be one of the cross-species sequence databanks, which are described in the next section.

The International Nucleotide Database. A consortium consisting of the European Bioinformatics Institute (EBI, UK), the National Center for Biotechnology Information (NCBI, USA) and the National Institute of Genetics (NIG, Japan) has produced the International Nucleotide Database, which is an annotated collection of all publicly available nucleic-acid sequences. The complete nucleotide sequences of *Drosophila* species genomes reside in this database, along with full-length cDNA sequences, EST sequences, sequences that flank transposable-element insertion sites and other sequences that are relevant to *Drosophila* research. The data are archived by each group and can be accessed through their web sites — GenBank[13], EMBL Nucleotide Database[14] or DNA Data Bank of Japan[15] — or through specialized sites such as FlyBase[16] (discussed in a later section). At GenBank, several distinct sequence sets are available for similarity searches using BLAST[17] (see also Online links box). The best choice will depend on the query sequence and the goals of the search. The sequence sets include one ('nr') that allows you to search most of the sequences that have been contributed by the research community, finished genome sequences and RefSeq sequences (a comprehensive set of non-redundant sequences from the most-studied research organisms, provided by NCBI). Other sets allow narrower searches: one ('chromosome') is limited to RefSeq sequences that are derived from complete genomes, chromosomes or CONTIGS, four sets ('est', 'est_human', 'est_mouse' and 'est_other') are specific for 'single pass' cDNA sequences (ESTs), one ('gss') is for transgene insertion sites, and another set ('htsg') is restricted to unfinished, high-throughput genomic sequences. Searches of the Trace Archives provide the only way to access unassembled sequences from genome projects; this category of data will be important if there are sequences that are similar to your sequence of interest in regions of the genome that are difficult to assemble, such as heterochromatin.

Entrez Genome at the NCBI provides a map-based data-entry point for those species with complete genome sequences. Map-based displays allow gene data to be viewed in the context of genome organization. For *D. melanogaster*, Entrez MapView cross-references include links to FlyBase gene reports, to lists of similar-sequence genes across species and to Model Maker. Model Maker displays the evidence on which a predicted gene structure — a gene model — is based, and allows alternative models to be created from that evidence. At the EBI, standardized views of richly annotated and cross-referenced data for those species with complete genome sequences are provided by the Genome Reviews database, although this does not yet include any *Drosophila* sequences.

Protein resources. There are many specialized databases that are dedicated to some aspect of protein information. A general review of protein resources is beyond the scope of this article, but we recommend the 2004 *Nucleic Acids Research* database issue[18] as a good starting point for more extensive information on protein resources than is included here.

The primary resource for information on the sequence and function of proteins, including those of *D. melanogaster*, is the Universal Protein Resource, UniProt[19]. UniProt Knowledgebase contains extensive curated data on protein classification and function, and is cross-referenced to other databases. Protein sequences and other information can be browsed, searched (by text searches and BLAST sequence-similarity searches) and downloaded from the UniProt site. Links are also provided to UniProt partner sites that offer more specialized tools for protein identification and analysis. EBI's Integr8 browser provides an integrated view of data from Genome Reviews and UniProt proteome sets for organisms for which whole-genome sequences are available. Proteome data and analysis (but not Genome Reviews data) for *D. melanogaster* are currently available through Integr8.

NCBI also maintains a protein-sequence database and related datasets. Entrez Protein provides access to protein records, which are compiled from various external databases and translations from the annotated coding regions that are present in GenBank[13]. The Clusters of Orthologous Groups (COG) database[20] clusters proteins from complete eukaryotic genomes into groups that contain sequences that are mutual best hits in sequence similarity searches between species. Protein data can also be accessed through the UCSC Proteome Browser, a component of the UCSC Genome Bioinformatics web site, and the PANTHER database[21] provides data and tools for high-throughput analysis of protein sequences.

FlyBase. If an interrogation of the nucleic acid or protein databases identifies similarities between a given sequence and those in *D. melanogaster* or another

BLAST
(Basic Local Alignment Search Tool.) A sequence comparison algorithm, optimized for speed, that is used to search sequence databases for optimal local alignments to a query.

CONTIG
Overlapping series of clones or sequence reads (for a clone contig or sequence contig, respectively) that correspond to a contiguous segment of the source genome.

Drosophila species, FlyBase[16] will probably be the next stop. FlyBase is the primary source of genetic and genomic information on *D. melanogaster* and related species of the family Drosophilidae. The FlyBase web site provides an integrated view of data captured from the scientific literature, the sequence databanks, large-scale data producers and providers of *Drosophila* material resources, such as mutant stocks or cell lines. Curated FlyBase datasets include information on genes, gene sequences, genome annotation, chromosome aberrations, natural transposons and their insertion sites in the reference genome, transgene constructs and their insertions in experimental genomes, terms for anatomical structures and developmental stages, images illustrating many of those terms, and references to the literature. A user-maintained section of contact information for *Drosophila* workers is also available. Previous releases of FlyBase data (dating back to August 2003) are accessible through the FlyBase Archives.

Genome annotation data that is currently available from the FlyBase web site includes predicted gene and pseudogene sequences, exon–intron structures, transcripts and proteins, regulatory regions, DNA sites that affect chromosome architecture, point mutations, transgene-insertion sites, aberration junctions, and regions of ORTHOLOGY and conserved SYNTENY between *Drosophila* species. In addition, FlyBase gene data includes synonyms, map location, protein domains, developmental stage and location of gene expression, gene function (using GENE ONTOLOGY[22] terms), mutant alleles, stocks, genetic interactions, transgene constructs, insertions, references and other data.

The Anatomy and Images section of FlyBase provides various tools for learning about *Drosophila* anatomy and development, and accessing genetic data that are based on associations with developmental stages and anatomical features. This section includes options to search or browse anatomy and development terms, search gene-expression data that are based on these terms, browse a glossary, or browse annotated images that illustrate *Drosophila* anatomy and development.

The DroSpeGe web page, mentioned in a previous section, is the product of a collaboration between FlyBase and the US National Human Genome Research Institute-financed Genome Sequencing Centers. This service, which includes prepublication, as well as published data, is intended to simplify access to both the existing and rapidly developing new datasets for the genomes of the 12 *Drosophila* species that are currently under analysis. From this page you can access the raw data, search the available sequences with BLAST[17] and view the available annotations in GBrowse[23] (BOX 6). Multiple sequence alignments and annotations of new genomes will be available on this site as the collaborative analysis proceeds.

The breadth and depth of information in FlyBase is both its strength and a challenge for users. Pointers to some of the best entry points to the data in FlyBase are provided in BOX 6. Access to the full complexity of FlyBase data is gained only with some investment of effort by users. Our advice would be to read the information on the query pages, familiarize yourself with the alternative report formats, explore the available links, and if all else fails, look at page-specific help links or the FlyBase-wide documentation (FlyBase Reference Manual).

Although FlyBase is the most comprehensive source of information for *Drosophila* work, there are many other options for accessing *Drosophila* genome data, analysis tools and other resources. Most of the sites with information on gene expression and molecular interactions are described below. More resources are listed in BOXES 4,5.

Gene-expression and molecular-interaction data. As noted in the introduction, the functional analysis of *Drosophila* genes has moved from the single-gene approach, which was successful but slow, to whole-genome approaches that produce large volumes of data in a single experiment. We have some distance to go before reaching the goal of understanding each component of every biological pathway, but the data are accumulating at a high rate. Several sites provide access to large datasets that focus on gene expression or molecular interactions. A word of warning is needed, however. The technologies used to produce some of these data are new and under development, and in some cases it is difficult to correlate results from one experimental platform to another. Additionally, it is wise to confirm the interactions that are stated on these sites by using alternative approaches.

The results of *Drosophila*-specific global expression studies that have been carried out either by using TRANSCRIPTOME ARRAYS or high-throughput analyses of *in situ* gene expression are available from two sources. The Yale *Drosophila* Development Gene Expression Timecourse (Yale Dev Ex)[24] provides access to array-based gene-expression patterns for ~33% of the *D. melanogaster* genome during a complete time course of development. Although gene-by-gene access to the Yale expression data is available through FlyBase, at the Yale site, the data can be downloaded or queried and viewed on the basis of other criteria, such as expression clustering or human disease gene homology. The second source, BDGP's Patterns of Gene Expression in *Drosophila* Embryogenesis database, includes microarray data, and images and annotations of *in situ* hybridizations to whole embryos using ESTs from the BDGP *Drosophila* Gene Collection[25] (*D. melanogaster* genes). Links to these data are also available through FlyBase gene reports.

Quantitative data for embryonic expression of *D. melanogaster* SEGMENTATION GENES is available from the FlyEx database[26]. Data are presented for individual embryos and can be viewed as a table, a graph or as a reconstructed image. The antibodies used to generate these data are available from the East Asian Distribution Center (see below).

Drosophila data are also included in the two principal sources for pan-organismal gene-expression data. NCBI's Gene Expression Omnibus (GEO)[27] is a data repository and access point for high-throughput gene-expression data that is generated with various platforms. EBI's ArrayExpress[28] is a public repository for annotated microarray data.

ORTHOLOGY
Describes genes in different species that derive from a common ancestor.

SYNTENY
The state of genes being located on the same chromosome. In current usage, this typically refers to the order of genes within a chromosomal segment. Syntenic relationships among genes can be conserved over large evolutionary distances.

GENE ONTOLOGY
Three structured, controlled vocabularies (ontologies) that describe gene products in terms of their associated biological processes, cellular components and molecular functions in a species-independent manner.

TRANSCRIPTOME ARRAYS
Microarrays that are designed to include probes to the entire mRNA complement that is expressed by an organism.

SEGMENTATION GENE
One of a group of genes that specify the segmental pattern within the anterior–posterior body axis of *Drosophila* and other arthropods.

Box 6 | How to access the data in FlyBase

Data on *Drosophila* species can be accessed through FlyBase in various ways. The information you already have – a sequence, a gene, a molecular function, a physiological function – is the most important factor in determining the best way to find relevant information, but your preferences for display types and level of expertise with flies will also affect how you access the data.

Map-based entry points

Map-based search and browsing options are good choices when the starting information is physically tied to the genome.

- GBrowse is a graphical map interface that supports queries that are based on sequence coordinates, sequence accession numbers, cytological coordinates and gene symbols, among other input options. GBrowse maps in FlyBase (see figure), display genes and many other sequence-based genome features. GBrowse also provides cytology-based access to genes, transgene insertions and aberrations that have not been mapped to the sequence. Datasets are available in GBrowse for *Drosophila melanogaster* and *Drosophila pseudoobscura*, and further annotated genomes will be added in 2005. The screen shot shows a FlyBase GBrowse display of a region of *D. melanogaster* chromosome 3L.
- CytoSearch is a query tool that produces customizable text-based maps of genes, transgene insertions, duplicated and deleted segments, and aberration breakpoints that are based on cytological location, sequence region or the length of a gene-transcription unit. Note that great care should be taken when comparing map locations of sequence-mapped objects with cytology-mapped objects. Cytology is, by its nature, imprecise relative to sequence coordinates, and the quality of primary cytological mapping data is highly variable. Both experimentally observed cytology and FlyBase Computed Cytology should be taken as approximations.
- BLAST *Drosophila* sequence datasets — organized as chromosome arms, SCAFFOLDS, predicted proteins, EST sequences or transgene-insertion sites, for example — can be queried using BLAST sequence-similarity searches. When 'chromosome arms' is the chosen feature type, resulting matches include links to a GBrowse map that displays your query sequence on the reference sequence.

Text-based entry points

Text-based searches are available for all FlyBase data classes. The search tool on the homepage supports a search of all fields of all data sections; with this search it is not necessary to know which data section holds the information you want, but it is more likely to produce more irrelevant results than data-class-specific searches. Each data class can be searched from the homepage, with two options: all fields or valid symbols, synonyms and full names.

Advanced query options. Each data section includes a query form with more extensive field-specific options; for example, the 'genes' query supports searches by species, class of gene, map location, Gene Ontology terms, mutant-allele characteristics and more. Although more challenging to use, these data-class-specific forms support more focused questions than the quick search found on the home page.

Bulk-data query options. The FlyBase Batch Download by ID query tool accepts lists or files of any valid FlyBase identifier as the query input. Several output options for content and format are available, including selected fields and downloadable files in spreadsheet or database formats. This tool is particularly useful for converting one type of FlyBase identifier to another; for example, CG SYMBOLS to valid gene symbols or database identifiers (FBgns).

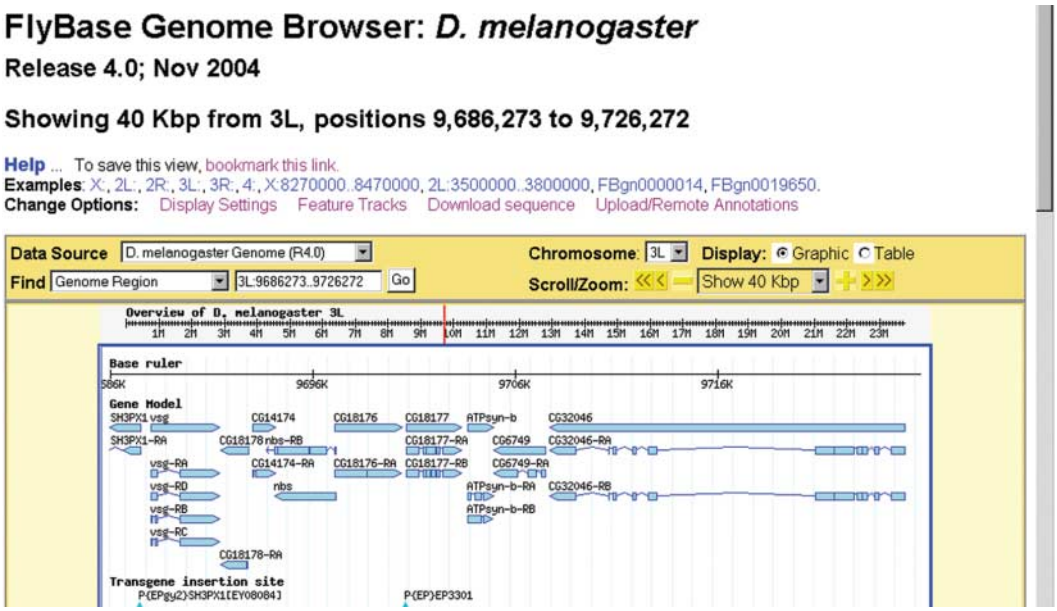


Image courtesy of FlyBase[16] © (1993–2005) The Genetics Society of America.

SCAFFOLDS
(Also known as supercontigs.) Genomic units that are composed of one or more contigs that have been ordered and orientated using end-read information.

CG SYMBOL
A unique symbol composed of an integer prefixed with CG that is assigned to each *D. melanogaster* gene identified through annotation of the whole genome shotgun sequence. A CG symbol is the valid gene symbol in the absence of gene symbols that are based on mutant phenotype, molecular feature or determined function.

Two sites provide comprehensive results of potential and verified protein–protein interaction studies in *D. melanogaster* that are based on the yeast two-hybrid system[29]. The *Drosophila* Interaction Database, by CuraGen Corporation and collaborators, provides direct access to the protein-interaction map for *D. melanogaster* that was produced by this company[30]. The *D. melanogaster* protein-interaction data by Hybrigenics S.A. and collaborators are available from the PIMRider web site. Links to both sets of data are also available through FlyBase gene reports.

The Biomolecular Interaction Database[31] collects information on molecular interactions from both high-throughput sources and the literature. The database classifies molecular interactions in three ways: direct interactions, complexes that are made up of one or more interactions and pathways that are defined by a specific sequence of two or more interactions.

Material resources

Material resources cover the gamut from genetic mutations that were created early in the twentieth century to the latest transgenic constructs for sophisticated manipulation of the genome. For the most part, materials will interest those who are actively working with *Drosophila*; consequently, the information below tends to be specialized. Some materials, however, such as *D. melanogaster* cell lines, or the full-length cDNAs of the 'Gold' version of BDGP's *Drosophila* Gene Collection[25], will have a broader audience among non-*Drosophila* researchers. For those new to *Drosophila* genetics, the document Genetic Nomenclature of *D. melanogaster*, which is maintained by FlyBase, explains the symbolic notation that is used to describe *Drosophila* genotypes and other aspects of *Drosophila* nomenclature. BOX 3 provides an overview of public resources and the available materials.

Genetic stocks. The current toolkit for *Drosophila* began with the collection of mutants[32] that was generated by Morgan's group (BOX 1). Genetically defined stocks of mutant alleles, chromosomal aberrations and now also a wide variety of transgenic lines, remain one of the fundamental resources for analysing gene function in the fly. Wild-type strains are crucial for evolutionary and ecological studies. There are currently five international stock centres, three focusing on *D. melanogaster* and two focusing on other *Drosophila* species. (Note that all live *Drosophila* entering the United States must be accompanied by an import permit from the US Department of Agriculture; see the Import Permit Information page on the Bloomington *Drosophila* Stock Center web site for information and forms.)

Currently the largest (~20,000 stocks) and most diverse of the *D. melanogaster* collections, the Bloomington *Drosophila* Stock Center (at Indiana University, USA) maintains transposable-element insertion alleles that are derived from the Gene Disruption Project[33], Exelixis[34] and others (~5,400 genes currently represented) and non-insertion alleles that together provide mutations in nearly 7,000 genes. Other kinds of stocks that are available from Bloomington include: deletions (~2,000 total), including the Bloomington deficiency kit, as well as the Exelixis[35] and DrosDel[36] sequence-mapped deficiencies; deletion-generating insertions[37] (~600); duplications (~600); defined UPSTREAM ACTIVATING SEQUENCE (UAS) transgenic strains (~300) and GAL4 DRIVER LINES[38] (~60), *FLP* RECOMBINASE TARGETS (*FRT*) and inducible-*FLP*[39,40] insertion stocks, and MOSAIC ANALYSIS WITH A REPRESSIBLE CELL MARKER (MARCM)[41,42] transgenic lines that allow sophisticated MOSAIC ANALYSIS[43] (~60), SNP-mapping[44,45] stocks (~25), a *P*-element point mutation mapping kit[46] (~40), and various other stocks that are useful for genetic analysis. The Bloomington Stock Center web site provides a browsing page of these and other specialized subsets of stocks as an aide for users wanting to identify potentially useful stocks.

The Szeged *Drosophila* Stock Centre (at Szeged University, Hungary) houses a collection of about 8,000 *D. melanogaster* stocks consisting of transposable-element insertion lines (which are, for the most part, unique to this collection) and DrosDel[36] deficiency strains (~500). The insertions that are available include the full set of Rørth's[47] UAS-activated endogenous gene-misexpression lines (~2,300), the lethal lacZ ENHANCER-TRAP INSERTIONS of Deak *et al.*[48] (~1,300 unique insertions, some in combination with basal *FRT* insertions to support mosaic analysis[43]) and Oh *et al.*[49] (~700 lines), and the full set of DrosDel insertions[36] (~1,800) that can be used to construct deficiencies with defined end-points[10].

The collection of the *Drosophila* Genetic Resource Center at the Kyoto Institute of Technology, Japan, consists of approximately 13,000 primarily *D. melanogaster* stocks, about 10,000 of which are unique to Kyoto. The collection includes a wide range of alleles and aberrations, including the Bloomington deficiency kit and a large number of transposable-element insertion strains that are unique to this centre. Insertion sets include that of Torok *et al.*[50], which contains lethal lacZ enhancer-trap lines that are not part of the Gene Disruption Project set[33] (~900), protein-trap insertion lines that produce GFP-tagged proteins[51,52] (~140), the Toba *et al.*[53] UAS-activated endogenous gene-misexpression lines (~5,000), and a subset of the GAL4 enhancer-trap insertion lines that were produced by the NP Consortium (~2,000). A few *Drosophila simulans* stocks are also available from Kyoto.

The Tucson *Drosophila* Species Stock Center collection at the University of Arizona, USA, currently includes approximately 1,300 stocks that represent 35 species groups from 5 genera. Wild-type stocks make up ~77% of the collection; location and date information for the collection are available for most stocks. Mutant alleles and a few aberrations are available for 26 species (~300 stocks). The strains that are used for non-*D. melanogaster* whole-genome shotgun sequencing (11 species) or BAC library production (20 species) are now available from the Tucson collection or are expected to be available sometime in 2005. The Tuscon Center hosts an annual *Drosophila* Species Identification Workshop that focuses on anatomy of various *Drosophila* species and the use of TAXONOMIC KEYS to identify them.

UPSTREAM ACTIVATING SEQUENCE
A yeast regulatory region that binds the transcriptional activator GAL4 and, when bound, activates the expression of an associated gene. Used in *Drosophila* in conjunction with GAL4 to control the induction of gene expression.

GAL4 DRIVER LINE
A line that expresses the yeast transcriptional activator GAL4 under the control of an inducible or a tissue-specific promoter. When crossed to an effector line that carries the upstream activator sequence (UAS) fused to a gene of interest, progeny with both the GAL4 and UAS components express the gene of interest in an activator-specific manner.

FLP RECOMBINASE TARGET (*FRT*). Components of a recombination system that is adapted from the *Saccharomyces cerevisiae* 2μ plasmid. *FLP* encodes a site-specific recombinase, and *FRT* is the target site for the Flp recombinase. Expression of *FLP* mediates excision of any sequence that is flanked by *FRT* sites.

MOSAIC ANALYSIS WITH A REPRESSIBLE CELL MARKER
A system for mosaic analysis that allows the visualization of mutant cells but not heterozygous parent cells or homozygous wild-type siblings.

MOSAIC ANALYSIS
The process of creating and assessing the fate of clonally derived groups of cells that are genetically distinct from cells of the surrounding tissues.

ENHANCER-TRAP INSERTIONS
Genomic insertions of a transgenic construct used to identify genes that are expressed in specific tissues. When the construct inserts near a tissue-specific enhancer, the weak promoter on the construct comes under the control of the enhancer, resulting in tissue-specific expression of the reporter gene.

The Ehime *Drosophila* Species Center at Ehime University, Japan, holds a collection of about 400 wild-type stocks that represent 50 species. Most of the strains were collected in Japan and the location of the collection sites and dates are available for most stocks.

Special laboratory collections. Most *Drosophila* laboratories make strains available to others on an informal, small-scale basis. A few laboratory lists are searchable through the FlyBase Stocks Search. Those needing published materials that are not in any of the public stock lists should request these from the corresponding authors of relevant publications.

A few groups host specialized large-scale collections of stocks. Those with public web sites are listed in the Stocks section of FlyBase. The Gene Disruption Project (GDP) web site at Baylor College of Medicine provides access to their database and to insertions that have not yet been transferred to the Bloomington Stock Center. The Exelixis *Drosophila* Stock Collection, which is maintained by the Artavanis-Tsakonas laboratory at Harvard Medical School, consists of ~16,500 stocks from the Exelixis insertion collection[34]. The Zuker Collection[54] at the University of California, San Diego, was developed to yield mutations in genes that are not required for viability. It consists of ~12,000 lines in which one of the two main autosomes, having been heavily mutagenized with EMS and initially selected for viability, carries an average of 5 to 6 mutant loci. The *Drosophila* Heterochromatin Genome Project has ~600 insertions into heterochromatin that have been mapped to cytological bands on mitotic chromosomes.

Cell lines. Tissue culture has become an increasingly important tool over the past few years for the functional analysis of *Drosophila* genes[55,56]. The number of available lines has grown markedly, and knowledge of the tissue types that are represented has improved. The primary source of cultured *D. melanogaster* and related species cell lines is the *Drosophila* Genomics Resource Center at Indiana University (DGRC). Almost 100 lines are currently available that are derived from embryos, IMAGINAL DISCS, the CNS, and HAEMOCYTES, including lines from 9 closely related species. Two embryonic epithelial lines are available from the American Type Culture Collection (ATCC). Both resources provide protocols and advice on using these materials.

Antibodies. Antibodies that are useful for *Drosophila* work, although not typically raised against *Drosophila* antigens, but that do cross-react to fly proteins, are available from various commercial sources. These sources sometimes explicitly state that an antibody reacts to *Drosophila* proteins but often provide no further information. It is therefore wise to determine, by either contacting other potential users or the supplier directly, whether any specific reagent is known to bind before making a purchase. A large collection of lines (~100) that produce MONOCLONAL ANTIBODIES that are raised against *Drosophila* proteins are maintained by the Developmental Studies Hybridoma Bank (DSHB; see HYBRIDOMA in Glossary) at the University of Iowa. A specialized collection of antibodies against *Drosophila* segmentation gene proteins is available from the East Asian Distribution Center at the National Institute of Genetics, Mishima, Japan.

Clones and libraries. Extensive *Drosophila* DNA resources are in the public domain, including materials for *Drosophila* species other than *D. melanogaster* (BOX 3). Genomic BAC libraries for *D. melanogaster* and 20 other *Drosophila* species, and FOSMID libraries for some of these species, are (or will be) available from either the BACPAC Resources Center (BPRC) at the Children's Hospital Oakland Research Institute, the Arizona Genomics Institute (AGI), or from the DGRC. An independently generated *D. melanogaster* embryo genomic BAC library is available from MRC geneservice at the Wellcome Trust Genome Campus.

Drosophila melanogaster cDNA collections that resulted from large-scale EST projects[25,57] are available from several sources. The BDGP *Drosophila* Gene Collection[25] (DGC), is available from the BPRC, Open Biosystems, MRC geneservice and the DGRC. DGRC's *Drosophila* Gene Collection sets include the Gold subset, which consists of ~6,000 clones with verified full-length open reading frames that are identical in amino-acid sequence to the reference genome. Testis cDNAs[58] from the B. Oliver laboratory at the US National Institutes of Health are available from the MRC geneservice.

Clones from the CuraGen Corporation collection of yeast two-hybrid constructs (~21,000) that were used by Giot *et al.*[30] to produce a protein-interaction map for *D. melanogaster* are available individually from the DGRC. The PCR primer pairs (~14,000) that were used to generate this set of clones are available from the DGRC as a complete set.

A broad collection of vectors for *Drosophila* work is also available from the DGRC. These include vectors for germline and cell-line transformation, for epitope tagging, for RNAi, for construction of GAL4–UAS and Flp–*FRT* transgenes, and for various miscellaneous purposes.

Microarrays. There are at least three sources for *Drosophila* microarrays that are available internationally. The Affymetrix GeneChip *Drosophila* Genome 2.0 array is an *in situ* synthesized oligonucleotide array that targets 18,500 transcripts that are predicted in release 3.1 of the annotated *D. melanogaster* genome sequence. The *D. melanogaster* microarray that is currently available from the DGRC also targets the transcriptome, but is an amplicon-based array spotted on a glass slide. The Version 1 array consists of 14,151 DNA fragments that are amplified from genomic DNA using gene-specific oligonucleotide primer pairs that were designed against Release 1 of the annotated sequence, targeting approximately 75% of the genes that were predicted in Release 3.1 of the annotated sequence. Customized, high-density *D. melanogaster* 24mer arrays are available from NimbleGen Systems. These arrays are manufactured on demand and are based on the most current

TAXONOMIC KEY
A device that is used to identify unfamiliar organisms and is usually based on a series of paired statements that describe contrasting morphological characteristics. Also known as a dichotomous key.

IMAGINAL DISCS
In holometabolous insect larvae, these are tissues that give rise to the external adult structures, such as the wing, eye and antennae.

HAEMOCYTES
Blood cells of invertebrates.

MONOCLONAL ANTIBODIES
Purified antibodies that are derived from single clones that recognize single antigens.

HYBRIDOMA
A hybrid cell line that is created by fusing a mortal antibody-producing B-lymphocyte with an immortalized myeloma line. The hybridoma line is immortal and produces a continuous supply of a particular monoclonal antibody.

FOSMID
A low-copy-number cosmid vector that is based on the *Escherichia coli* F-factor replicon. Cloned sequences are more stable in fosmids than in multi-copy vectors.

data in the nucleotide-sequence databanks at the time that the order is placed. Multiple genomes can be incorporated into NimbleGen array designs. FlyChip provides array services to researchers in the United Kingdom.

Experimental services

There are several providers of experimental services for fly work. Outsourcing routine work such as transgene construction and embryo injections could save time for any project, but will be of particular value if you are new to *Drosophila* and do not have local access to the necessary equipment and expertise. Some services are highly specialized and would require significant investment from even large fly groups to replicate locally. Most of these providers are institutional service projects that have public funding. BOX 3 provides an overview of service providers for fly work, which, although not extensive, covers a wide range of services.

Genetic screening. If mutations are not already available in a gene of interest or you would like to assay for genes that affect a given process, help might be available on a contract basis. The *Drosophila* TILLING Project (Fly-Till)[59] at the Fred Hutchinson Cancer Research Center might be able to provide an allelic series of EMS mutations in a specific gene. The *Drosophila* RNAi Screening Center (DRSC)[60] at Harvard Medical School accepts applications for genome-wide RNAi screens for functional defects in tissue-culture cells[55]. Both projects are publicly financed and impose certain requirements on users of their services. Genetic Services is a private company that offers customized genetic screening in *Drosophila* and other model organisms.

Transgene construction and injection services. Customized transgene construction is offered by Genetic Services. Genetic Services, Duke University Model System Genomics (MSG) and EMBL *Drosophila* Injection Service provide embryo injection services for germline transformation; the first two facilities will also inject dsRNA for customers. More providers that limit their services to local laboratories are listed on the FlyBase External *Drosophila* Resources web page.

Other services. As indicated in BOX 3, further services are available from MRC geneservice, Duke University MSG and Genetic Services. The services that are available include sequencing and sequence-based mapping of transgene insertions, stock construction, immunohistochemistry, fluorescence-based embryo sorting, microarray analysis, and customized bioinformatics, such as microarray analysis. More information can be found on each provider's web site.

Examples of the uses of *Drosophila* resources

It is not possible to anticipate and illustrate all of the ways in which the resources referred to above might be used, given the wide range of biological questions that can be asked. Here we provide a few examples to illustrate how the available resources can be used to accomplish a few specific tasks. Although *Drosophila* material resources are usually unique, because information is readily acquired from one public data source and redistributed in a different context or format, in almost every case there are several paths to the same information. Here we focus on one path to each goal that is likely to serve most users well. Individual interests might well fall into the domain of one of the alternative sites that are noted in BOXES 4,5; in which case, one of these alternatives might be a better starting place.

Finding a 'knock-out' for a* Drosophila *gene. If your starting point is a sequence, you could use a BLAST search (from any of the public sites that are described in this article) that identifies a *Drosophila* gene to link to a FlyBase 'gene report' (use BLASTp or BLASTx to search for similarity to protein rather than DNA sequence to identify related genes across large evolutionary distances). If your starting point is a gene name, symbol, synonym or identifier, use any of the FlyBase text searches (BOX 6) to retrieve the relevant gene report. You should then select the 'alleles' subsection from the list of 'available reports'; ideally, one or more alleles of the class 'amorph' or 'loss of function' will be listed there, but for many alleles this information is not reported in the literature and therefore not available from FlyBase (note that, when it is reported, this classification is often not rigorous). More often, you will need to read about the known mutant alleles and determine which ones are likely to be the strongest available alleles. Weak *P*-element transposon insertion alleles can be used to create null alleles by imprecise excision of the insertion. If a candidate is identified, check the allele report or the 'stocks' subsection of the gene report for its availability from one of the public stock centres. If a stock is not listed there, search for the allele in FlyBase Stocks because newly added stocks data will appear there before they are fully incorporated into other FlyBase reports. Because many insertions are not yet in FlyBase, it is important that you also check the *P*-Screen Database of the Gene Disruption Project, the Harvard Medical School Exelixis *Drosophila* Stock Collection web site and the DrosDel Project web site (view 'show all deletions' for the chromosome arm that carries your gene to see the available insertion for that arm; note that 'available deletions' refers to deletions that could be generated from available insertions rather than those that have been generated); if your gene maps to heterochromatin, then use the *Drosophila* Heterochromatin Genome Project web site. If no useful alleles or insertions are available for your gene you will need to investigate the existence of deficiencies, as described in the next situation.

Finding deficiencies for a* Drosophila *gene. Deficiencies are extremely important tools in *Drosophila* research and the available resources are extensive. The 'best' deficiency will be the one that deletes a gene entirely and the fewest other genes. The smallest deficiencies with sequence-mapped end points are those of the Exelixis

set, which are available from the Bloomington Stock Center. Use the Exelixis Deficiencies browsing page on the Bloomington web site to identify candidate deletions for your gene (always confirm that your gene is indeed deleted before investing significant work in any deficiency; for example, by carrying out an appropriate PCR assay). DrosDel deletions are also a good choice and are second here only because they are larger, on average, than the Exelixis deficiencies. Use the DrosDel home page to view all available deletions. Most DrosDel deletions are available from both the Szeged and Bloomington Stock Centers; search for the aberration symbol (for example, *Df(2L)ED695*) on the FlyBase Stocks web site to retrieve relevant stock reports.

If you are unsuccessful in identifying one of these smaller, sequence-defined deletions, try the Bloomington Stock Center's deficiency kit. These are typically large, cytologically defined deficiencies. Browse the list from the Bloomington Stock Center web site. If you identify a candidate, look at the complete list of Bloomington deficiencies for that chromosome to see if smaller deletions are also available for that interval. More deficiencies might be identified in FlyBase (search in 'aberrations') but if they delete regions that are not represented in the Bloomington deficiency kit they are probably no longer available in stock. If you cannot find an existing deficiency, see the Review by Venken and Bellen[10] in this issue for the newest ways to target specific regions of the *D. melanogaster* genome for deletion.

Finding functional information on a* Drosophila *gene. Many kinds of data yield functional information about genes. If you are interested in a particular gene, a good place to start is the FlyBase gene report. If the product of your gene has been associated with a molecular function, biological process or cellular component — for example, transcription factor, ageing or centrosome — that information will appear in the FlyBase Gene Report. You will find the terms in the body of the synopsis report; open the 'Gene Ontology' subsection report for the basis of the association and references. Open the 'expression and phenotypes' subsection to find data on the wild-type expression patterns of transcripts and proteins, the cellular components, developmental stages, anatomical structures that are affected by mutant alleles and free text descriptions of mutant phenotypes of alleles. Most of the resources described here that generate gene-function data make links to their data available to FlyBase (search these sites directly for new data since the latest FlyBase update, or new since revised links were last provided to FlyBase). These links are located in the 'external database links' section of the gene report. Links to BDGP, Yale Dev Ex and NCBI GEO provide *in situ* and/or array expression data for the gene of interest. DRSC and the Heidelberg Genome RNAi *Drosophila* Resources links, if present, provide information on phenotypes that are associated with RNAi in cultured cells. Fly General Repository for Interaction Datasets (Fly GRID) and Hybrigenics PIMrider links provide genetic and/or physical interaction data. PANTHER links display a PANTER gene-information page that links to gene-expression data, if available, that is generated through quantitative PCR assays.

If you are starting with a function, process, or cellular component and want to see all the genes that have been associated with a particular term, use the Gene Ontology (GO) page in the FlyBase Genes section to search or browse for relevant genes. For example, if you are interested in receptor-mediated endocytosis, selecting '-R-' from the 'genes associated with biological process' list displays a page that includes a list of all *D. melanogaster* genes that have been annotated with the term receptor-mediated endocytosis. Alternatively, you can select 'Gene Ontology' from the 'browsing' section of the FlyBase GO page and browse through the GO-term hierarchy; a link to a list of all fly genes that are annotated with that term is presented for every term.

To identify *Drosophila* genes that are expressed at a given stage of development or in a given anatomical structure, use the gene expression page in the genes section of FlyBase. Here you can search for genes that meet your criteria or browse the hierarchy of developmental and anatomical terms, which includes links to lists of genes that are annotated with that term (not all genes in the *D. melanogaster* genome have been analysed to the point that their function or anatomical expression patterns are known and therefore the results of the last two searches will not necessarily be exhaustive). Also search the BDGP Patterns of Gene Expression in *Drosophila* Embryogenesis database for developmental stage or anatomical features. Although links to these data are included in FlyBase gene records, BDGP expression data are not available to FlyBase search and browse functions.

Conclusions

The challenges of traversing the complex array of informational resources that are available over the internet continually increase in importance and in difficulty. As we invent new and more sophisticated ways to interrogate genomes and their expressed products, the public database resources — as the main portals to the results of these interrogations — take on a role of increasing importance to the global biomedical research enterprise. For this reason, support for these resources and for standards that increase interoperability and linkages among the principal databases is vital to ensure that the community has access to the vast amount of information that is available.

It is often not appreciated how much twenty-first century biomedical research is being driven by the production and dissemination of large amounts of material and informational resources. Although these resources are not inexpensive to maintain, they have become an essential part of the high-throughput landscape that is the hallmark of the genomic era.

1. Bridges, C. B. Direct proof through non-disjunction that the sex-linked genes of *Drosophila* are borne by the X-chromosome. *Science* **40**, 107–109 (1914).
The author reports experiments with *D. melanogaster* that provide the first direct proof of the chromosome theory of inheritance.
2. Kida, Y., Maeda, Y., Shiraishi, T., Suzuki, T. & Ogura, T. Chick Dach1 interacts with the Smad complex and Sin3a to control AER formation and limb development along the proximodistal axis. *Development* **131**, 4179–4187 (2004).
3. De Velasco, B., Shen, J., Go, S. & Hartenstein, V. Embryonic development of the *Drosophila* corpus cardiacum, a neuroendocrine gland with similarity to the vertebrate pituitary, is controlled by *sine oculis* and *glass*. *Dev. Biol.* **274**, 280–294 (2004).
4. Bier, E. *Drosophila*, the golden bug, emerges as a tool for human genetics. *Nature Rev. Genet*. **6**, 9–23 (2005).
5. Oliver, C. P. in *The Genetics and Biology of* Drosophila Vol. 1a (eds Ashburner, M. &Novitski, E.) 3–23 (Academic, London, 1976).
Reviews the main contributions of *Drosophila* research to genetics and biology during the first half of the twentieth century.
6. Kohler, R. E. *Lords of the fly:* Drosophila *genetics and the experimental life* (Univ. Chicago Press, Chicago, 1994).
7. Adams, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
The authors report the first whole-genome shotgun sequence of a large, complex genome and thereby provide an important new tool for *D. melanogaster* research.
8. Celniker, S. E. *et al.* Finishing a whole genome shotgun: release 3 of the *Drosophila melanogaster* euchromatin genome sequence. *Genome Biol.* **3**, research0079.1–0079.14 (2002).
9. Misra, S. *et al.* Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. *Genome Biol.* **3**, research0083.1–0083.22 (2002).
This reports the first release of human-curated annotations of the *D. melanogaster* genome. Computational results were combined with experimental data, including extensive cDNA evidence, to generate high-quality genome annotations.
10. Venken, K. J. T. & Bellen, H. J. Emerging technologies for gene manipulation in *Drosophila. Nature Rev. Genet*. **6**, 167–178 (2005).
11. Richards, S. *et al.* Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and *cis*-element evolution. *Genome Res*. **15**, 1–18 (2005).
The authors describe the whole-genome sequence of a second *Drosophila* species, *D. pseudoobscura*, and compare its genome organization and sequence features with *D. melanogaster.*
12. Lewis, S. E. *et al.* Apollo: a sequence annotation editor. *Genome Biol.* **3**, research0082.1–0082.14 (2002).
13. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. GenBank: update. *Nucleic Acids Res*. **32**, D23–D26 (2004).
14. Kulikova, T. *et al.* The EMBL Nucleotide Sequence Database. *Nucleic Acids Res*. **32**, D27–D30 (2004).
15. Miyazaki, S., Sugawara, H., Ikeo, K., Gojobori, T. & Tateno, Y. DDBJ in the stream of various biological data. *Nucleic Acids Res*. **32**, D31–D34 (2004).
16. Drysdale, R. A. *et al.* FlyBase: genes and gene models. *Nucleic Acids Res*. **33**, D390–D395 (2005).
17. McGinnis, S., & Madden, T. L. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res*. **32**, W20–W25 (2004).
18. Bateman, A. (ed.) Database issue. *Nucleic Acids Res.* **32**, D1–D599 (2004).
This special issue of *Nucleic Acids Research*, which is part of a series on databases, provides extensive information on protein resources including those for *Drosophila*.
19. Apweiler, R. *et al.* UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **32**, D115–D119 (2004).
20. Tatusov, R. L. *et al.* The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**, 41 (2003).
21. Paul, D. *et al.* PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res*. **13**, 2129–2141 (2003).
22. Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res*. **32**, D258–D261 (2004).
23. Stein, L. D. *et al.* The generic genome browser: a building block for a model organism system database. *Genome Res.* **12**, 1599–1610 (2002).
24. Arbeitman, M. N. *et al.* Gene expression during the life cycle of *Drosophila melanogaster*. *Science* **297**, 2270–2275 (2002).
25. Stapleton, M. *et al.* The *Drosophila* gene collection: identification of putative full-length cDNAs for 70% of *D. melanogaster* genes. *Genome Res*. **12**, 1294–1300 (2002).
This paper describes the generation of the *Drosophila* Gene Collection, a non-redundant set of full-length cDNA clones that is an important resource for functional genomics studies.
26. Kosman, D., Reinitz, J. & Sharp, D. H. in *Proceedings of the 1998 Pacific Symposium on Biocomputing* (eds Altman, R., Dunker, K., Hunter, L. and Klein, T.) 6–17 (World Scientific, Singapore, 1999).
27. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. **30**, 207–210 (2002).
28. Alvis Brazma, A. *et al.* ArrayExpress — a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res*. **31**, 68–71 (2003).
29. Fields, S. & Song, O. A novel genetic system to detect protein–protein interactions. *Nature* **340**, 245–246 (1989).
30. Giot, L. *et al.* A protein interaction map of *Drosophila*. *Science* **302**, 1727–1736 (2003).
31. Bader, G. D., Betel, D. & Hogue, C. W. BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res*. **31**, 248–250 (2003).
32. Bridges, C. B. & Demerec, M. Editorial and other contributions. *Dros. Inf. Serv.* Vols 1–10 (1934–1938).
33. Bellen, H. J. *et al.* The BDGP Gene Disruption Project: single transposon insertions associated with 40% of *Drosophila* genes. *Genetics* **167**, 761–781 (2004).
This paper describes efforts to create a transposon-insertion allele of every gene in the *D. melanogaster* genome. This collection is an important resource for functional genomics.
34. Thibault, S. T. *et al.* A complementary transposon tool kit for *Drosophila melanogaster* using *P* and *piggyBac*. *Nature Genet.* **36**, 283–287 (2004).
35. Parks, A. L. *et al.* Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. *Nature Genet.* **36**, 288–292 (2004).
36. Ryder, E. *et al.* The DrosDel Collection: a set of *P*-element insertions for generating custom chromosomal aberrations in *Drosophila melanogaster*. *Genetics* **167**, 797–813 (2004).
37. Huet, F. *et al.* A deletion-generator compound element allows deletion saturation analysis for genomewide phenotypic annotation. *Proc. Natl Acad. Sci. USA* **99**, 9948–9953 (2002).
38. Duffy, J. B. GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. *Genesis* **34**, 1–15 (2002).
39. Chou, T. B. & Perrimon, N. Use of a yeast site-specific recombinase to produce female germline chimeras in *Drosophila*. *Genetics* **131**, 643–653 (1992).
40. Xu, T. & Rubin, G. M. Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* **117**, 1223–1237 (1993).
41. Luo, L., Lee, T., Nardine, T., Null, B. & Reuter, J. Using the MARCM system to positively mark mosaic clones in *Drosophila*. *Dros. Inf. Serv*. **82**, 102–105 (1999).
42. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker (MARCM) for *Drosophila* neural development. *Trends Neurosci*. **24**, 251–254 (2001).
43. Blair, S. S. Genetic mosaic techniques for studying *Drosophila* development. *Development* **130**, 5065–5072 (2003).
44. Berger, J. *et al.* Genetic mapping with SNP markers in *Drosophila*. *Nature Genet*. **29**, 475–481 (2001).
45. Hoskins, R. A. *et al.* Single nucleotide polymorphism markers for genetic mapping in *Drosophila melanogaster*. *Genome Res.* **11**, 1100–1113 (2001).
46. Zhai, R. G. *et al.* Mapping *Drosophila* mutations with molecularly defined *P* element insertions. *Proc. Natl Acad. Sci. USA* **100**, 10860–10865 (2003).
47. Rørth, P. A modular misexpression screen in *Drosophila* detecting tissue-specific phenotypes. *Proc. Natl Acad. Sci. USA* **93**, 12418–12422 (1996).
48. Deak, P. *et al. P*-element insertion alleles of essential genes on the third chromosome of *Drosophila melanogaster*: correlation of physical and cytogenetic maps in chromosomal region 86E–87F. *Genetics* **147**, 1697–1722 (1997).
49. Oh, S. W. *et al.* A *P*-element insertion screen identified mutations in 455 novel essential genes in *Drosophila*. *Genetics* **163**, 195–201 (2003).
50. Torok, T., Tick, G., Alvarado, M. & Kiss, I. *P*-lacW insertional mutagenesis on the second chromosome of *Drosophila melanogaster*: isolation of lethals with different overgrowth phenotypes. *Genetics* **135**, 71–80 (1993).
51. Kelso, R. J. *et al.* Flytrap, a database documenting a GFP protein-trap insertion screen in *Drosophila melanogaster*. *Nucleic Acids Res*. **32**, D418–D420 (2004).
52. Morin, X., Daneman, R., Zavortink, M. & Chia, W. A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in *Drosophila*. *Proc. Natl Acad. Sci. USA* **98**, 15050–15055 (2001).
53. Toba, G. *et al.* The gene search system. A method for efficient detection and rapid molecular identification of genes in *Drosophila melanogaster*. *Genetics* **151**, 725–737 (1999).
54. Koundakjian, E. J., Cowan, D. M., Hardy, R. W. & Becker, A. H. The Zuker collection: a resource for the analysis of autosomal gene function in *Drosophila melanogaster*. *Genetics* **167**, 203–206 (2004).
55. Boutros, M. *et al.* Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303**, 832–835 (2004).
56. Cherbas, L., & Cherbas, P. in Drosophila*: a Laboratory Manual* 2nd edn (eds Sullivan, W., Ashburner, M. & Hawley, R. S.) (Cold Spring Harbor Laboratory Press, New York, 2000).
57. Rubin, G. M. *et al.* A *Drosophila* complementary DNA resource. *Science* **287**, 2222–2224 (2000).
58. Andrews, J. *et al.* Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Res*. **10**, 2030–2043 (2000).
59. Henikoff, S., Till, B. J. & Comai, L. TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiol.* **135**, 630–636 (2004).
60. Kiger, A. *et al.* A functional genomic analysis of cell morphology using RNA interference. *J. Biol*. **2**, 27 (2003).
61. Lewis, E. B. Grant application submitted to the National Science Foundation, funded as BIR-9403760 (1980).
62. Bridges, C. B. Salivary chromosome maps with a key to the banding of the chromosomes of *Drosophila melanogaster*. *J. Hered*. **26**, 60–64 (1935).
These beautiful representations of the cytological features of polytene chromosomes, which were isolated from the salivary gland, were the first physical map of the *D. melanogaster* genome.
63. Bridges, C. B. & Brehme, K. S. *The mutants of* Drosophila melanogaster (Carnegie Institution, Washington DC, 1944).
64. Lindsley, D. L. & Grell, E. H. *Genetic variations of* Drosophila melanogaster (Carnegie Institution, Washington DC, 1968).
65. Lindsley, D. L. & Zimm, G. G. *The Genome of* Drosophila melanogaster (Academic, San Diego, California, 1992).
66. Herskowitz, I. H. *Drosophila* bibliography. *Dros. Inf. Serv*. **21**, 95–98 (1947).
67. Alderson, T. Chemically induced delayed germinal mutation in *Drosophila*. *Nature* **207**, 164–167 (1965).
68. Lewis, E. B. & Bacher, F. Methods of feeding ethyl methane sulfonate (EMS) to *Drosophila* males. *Dros. Inf. Serv.* **43**, 193 (1968).
69. Lindsley, D. L. *et al.* Segmental aneuploidy and the genetic gross structure of the *Drosophila* genome. *Genetics* **71**, 157–184 (1972).
Reports the first whole-genome approach to genetic manipulation and analysis in *D. melanogaster*.
70. Nüsslein-Volhard, C., Wieschaus, E. Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795–801 (1980).
71. Nüsslein-Volhard, C., Wieschaus, E., Kluding, H. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. I. Zygotic loci on the second chromosome. *Rouxs Arch. Dev. Biol.* **193**, 267–282 (1984).
72. Jurgens, G., Kluding, H., Nüsslein-Volhard, C., & Wieschaus, E. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. II. Zygotic loci on the third chromosome. *Rouxs Arch. Dev. Biol*. **193**, 283–295 (1984).
73. Wieschaus, E., Nüsslein-Volhard, C., Jurgens, G. Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*. III. Zygotic loci on the X-chromosome and fourth chromosome. *Rouxs Arch. Dev. Biol.* **193**, 296–307 (1984).
74. Spradling, A. C. & Rubin, G. M. Transposition of cloned *P*-elements into *Drosophila* germ line chromosomes. *Science* **218**, 341–347 (1982).
The first report of *Drosophila* germline transformation using engineered transposable elements, a crucial technique that opened the genome to easy manipulation.
75. Rubin, G. M. & Spradling, A. C. Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353 (1982).
76. Lefevre, G. in *The Genetics and Biology of* Drosophila Vol. 1a (eds Ashburner, M. & Novitski, E.) (Academic, London; New York, 1976).
77. Roberts, D. B. Drosophila*. A practical approach* (Oxford Univ. Press, Oxford, 1986).
78. Greenspan, R. J. *Fly pushing: the theory and practice of* Drosophila *genetics* 2nd edn (Cold Spring Harbor Laboratory Press, New York, 2004).
79. Demerec, M. *Biology of* Drosophila (Cold Spring Harbor Laboratory Press, New York, 1994).

80. Ashburner, M. & Novitski, E. *The Genetics and Biology of* Drosophila Vols 1a–c (Academic, London; New York, 1976).
81. Ashburner, M. & Wright, T. *The Genetics and Biology of* Drosophila Vols 2a–d (Academic, London; New York, 1978–1980).
82. Ashburner, M., Carson, H. & Thompson, J. *The Genetics and Biology of* Drosophila Vols 3a–d (Academic, London; New York, 1981–1986).
83. Campos-Ortega, J. A. & Hartenstein, V. *The Embryonic Development of* Drosophila melanogaster 2nd edn (Springer, Berlin; Heidelberg, 1997).
84. Bate, M. & Martinez Arias, A. *The Development of* Drosophila melanogaster Vols 1,2 (Cold Spring Harbor Laboratory Press, New York, 1993).
85. Hartenstein, V. *Atlas of* Drosophila *development* (Cold Spring Harbor Laboratory Press, New York, 1993).
86. Moses, K. Drosophila *Eye Development* (Springer, New York, 2002).
87. Ashburner, M. Drosophila*: a Laboratory Handbook and Manual* (Cold Spring Harbor Laboratory Press, New York, 1989).
88. Ashburner, M., Golic, K. G. & Hawley, R. S. Drosophila*: a Laboratory Handbook* 2nd edn (Cold Spring Harbor Laboratory Press, New York, 2004).
89. Goldstein, L. S. B. & Fyrberg, E. A. (eds) Drosophila melanogaster*: Practical Uses in Cell and Molecular Biology* Vol. 44 (eds) (Academic, London; New York, 1994).
90. Sullivan, W., Ashburner, M. & Hawley, R. S. Drosophila *Protocols* (Cold Spring Harbor Laboratory Press, New York, 2000).
91. Echalier, G. Drosophila *cells in culture* (Academic, London; New York, 1997).
92. Sorsa, V. *Chromosome maps of* Drosophila Parts 1,2 (CRC Press, Boca Raton, Florida, 1988).
93. Henderson, D. S. Drosophila *Cytogenetic Protocols* (Humana, Totowa, New Jersey, 2003).
94. Patterson, J. T. & Stone, W. *Evolution in the Genus* Drosophila (Macmillan, New York, 1952).
95. Powell, J. R. *Progress and Prospects in Evolutionary Biology: the* Drosophila *Model* (Oxford Univ. Press, New York, 1997).
96. Karolchik, D. *et al.* The UCSC Genome Browser database. *Nucleic Acids Res.* **31**, 51–54 (2003).
97. Birney, E. *et al.* An overview of Ensembl. *Genome Res.* **14**, 925–928 (2004).
98. Gilbert, D. G. euGenes: a eukaryote genome information system. *Nucleic Acids Res*. **30**, 145–148 (2002).
99. Chien, S., Reiter, L. T., Bier, E. & Gribskov, M. Homophila: human disease gene cognates in *Drosophila*. *Nucleic Acids Res*. **30**, 149–151 (2002).
100. Brody, T. The Interactive Fly: gene networks, development and the Internet. *Trends Genet.* **15**, 333–334 (1999).
101. Weigmann, K. *et al.* FlyMove — a new way to look at development of *Drosophila*. *Trends Genet*. **19**, 310–311 (2003).
102. Armstrong, J. D. *et al.* Flybrain, an on-line atlas and database of the *Drosophila* nervous system. *Neuron* **15**, 17–20 (1995).

Acknowledgements

This work was supported by grants from the US National Science Foundation's Living Stock Collections Program and the following US National Institutes of Health bodies: National Human Genome Research Institute (NHGRI), National Center for Research Resources (NCRR), National Institute of General Medical Sciences (NIGMS) and National Institute of Child Health and Human Development (NICHHD).

Competing interests statement

The authors declare no competing financial interests.

Online links

INFORMATION RESOURCES

ArrayExpress: http://www.ebi.ac.uk/arrayexpress
Assembly/Alignment/Annotation of 12 *Drosophila* Genomes: http://rana.lbl.gov/drosophila/multipleflies.html
Berkeley *Drosophila* Genome Project: http://www.fruitfly.org
cDNA and EST projects | Patterns of Gene Expression in *Drosophila* Embryogenesis | SNPs
Biomolecular Interaction Network Database: http://bind.ca
Clusters of Orthologous Groups: http://www.ncbi.nlm.nih.gov/COG/new
DNA Data Bank of Japan: http://www.ddbj.nig.ac.jp
***Drosophila* Heterochromatin Genome Project:** http://www.dhgp.org/index.html
***Drosophila* Interaction Database (CuraGen):** http://portal.curagen.com/cgi-bin/interaction/flyHome.pl
***Drosophila melanogaster* Exon Database:** http://proline.bic.nus.edu.sg/dedb
***Drosophila* Polymorphism Database:** http://bioinformatica.uab.es/dpdb/dpdb.asp
***Drosophila* Population Genome Project:** http://www.dpgp.org
***Drosophila* Species Genomes:** http://flybase.net/species
EMBL Nucleotide Sequence Database: http://www.ebi.ac.uk/embl
Ensembl Genome Browser: http://www.ebi.ac.uk/ensembl
Entrez: http://www.ncbi.nlm.nih.gov/Entrez
Genome | Protein | MapView
euGenes: http://eugenes.org
FlyBase: http://flybase.org
Anatomy and Images | Archives | BLAST | Batch Download by ID | Computed Cytology | CytoSearch | Drawings of Drosophilidae | External *Drosophila* Resources | GBrowse | Genes Query | Genetic Nomenclature | Reference Manual | Stocks Search | Stocks section
Flybrain: http://flybrain.neurobio.arizona.edu
FlyEx: http://urchin.spbcas.ru/flyex
Fly General Repository for Interaction Datasets: http://biodata.mshri.on.ca/fly_grid/servlet/SearchPage
FlyMine: http://www.flymine.org
FlyMove: http://flymove.uni-muenster.de
FlyPNS: http://www.princeton.edu/~vorgogoz/FlyPNS/page1.html
FlySNP: http://flysnp.imp.univie.ac.at
flytrap: http://www.fly-trap.org
FlyTrap: http://flytrap.med.yale.edu
Gene Expression Omnibus: http://www.ncbi.nlm.nih.gov/geo
Genome Reviews: http://www.ebi.ac.uk/GenomeReviews
Heidelberg GenomeRNAi *Drosophila* Resources: http://www.dkfz-heidelberg.de/signaling/ernai/ernai.html
Homophila: http://superfly.ucsd.edu/homophila
Integr8: http://www.ebi.ac.uk/integr8/EBI-Integr8-HomePage.do
Interactive Fly: http://flybase.org/allied-data/lk/interactive-fly/aimain/1aahome.htm
NCBI GenBank: http://www.ncbi.nlm.nih.gov/About/tools/restable_nuc.html
Neuroblast lineages: http://www.neuro.uoregon.edu/doelab//lineages.htm
PANTHER Classification System: https://panther.appliedbiosystems.com/navigation.jsp
PIMRider *Drosophila* Protein Interaction Map: http://pim.hybrigenics.com/pimriderext/droso/index.html
TaxoDros: http://taxodros.unizh.ch
tracheal db: http://www.biozentrum.unibas.ch/affolter/trachea
UCSC Genome Bioinformatics: http://genome.ucsc.edu/index.html
Genome Browser | Proteome Browser
UniProt: http://www.pir.uniprot.org
Knowledgebase
WWW Virtual Library — *Drosophila*: http://www.ceolas.org/fly
Yale *Drosophila* Developmental Gene Expression Timecourse: http://genome.med.yale.edu/Lifecycle

MATERIAL RESOURCES

Affymetrix GeneChip Arrays: http://www.affymetrix.com/products/arrays/index.affx?Drosophila
American Type Culture Collection: http://www.atcc.org
Arizona Genomics Institute: http://genome.arizona.edu
BACPAC Resources Center: http://bacpac.chori.org/dromel98.htm
DGC
Bloomington *Drosophila* Stock Center: http://flystocks.bio.indiana.edu
Bloomington Browsing Page | Bloomington Deficiency Kit | Exelixis Deficiencies | Import Permit Information
Developmental Studies Hybridoma Bank: http://www.uiowa.edu/~dshbwww
DrosDel *Drosophila* Isogenic Deficiency Kit: http://www.drosdel.org.uk
***Drosophila* Genetic Resource Center (Kyoto):** http://www.dgrc.kit.ac.jp/en/index.html
***Drosophila* Genomics Resource Center:** http://dgrc.cgb.indiana.edu
Microarrays | DGC
East Asian Distribution Center: http://www.nig.ac.jp/labs/DevGen/segmentation
Ehime *Drosophila* Species Center of Japan: http://kyotofly.kit.jp/ehime
Exelixis *Drosophila* Stock Collection at Harvard Medical School: http://*Drosophila*.med.harvard.edu
FlyChip: http://www.flychip.org.uk
Gene Disruption Project P-Screen Database: http://flypush.imgen.bcm.tmc.edu/pscreen
MRC geneservice: http://www.hgmp.mrc.ac.uk/geneservice/index.shtml
BACs | DGC | Testis cDNAs
NimbleGen Systems: http://www.nimblegen.com/products/fly.html
NP Consortium: http://flymap.lab.nig.ac.jp/~dclust/getdb.html
Open Biosystems: http://www.openbiosystems.com/drosophila_rnai_collection.php
DGC
Szeged *Drosophila* Stock Centre: http://expbio.bio.u-szeged.hu/fly/index.php
Tucson *Drosophila* Species Stock Center: http://stockcenter.arl.arizona.edu
Zuker Laboratory Collection: contact Charles Zuker; czuker@ucsd.edu

RESEARCH SERVICES

***Drosophila* RNAi Screening Center:** http://flyrnai.org/RNAi_index.html
***Drosophila* TILLING Project:** http://tilling.fhcrc.org:9366/fly
Duke University Model System Genomics: http://www.biology.duke.edu/model-system/services.htm
EMBL *Drosophila* Injection Service: http://www.embl-heidelberg.de/~voie
Genetic Services: http://www.geneticservices.com

FURTHER INFORMATION

Celera Genomics: http://www.celera.com
Human Genome Sequencing Center at Baylor College of Medicine: http://www.hgsc.bcm.tmc.edu/projects
***Nucleic Acids Research* Database Issue:** http://nar.oupjournals.org/content/vol32/suppl_1
Access to this interactive links box is free online.