

TACACAATCAGTTAGTTTCCACCGACAGTCCGCAGAAACCATTTCGACGGC  
GTCGGCAATCCGTAAGATAGCCAAATATATTATTGTTTCAGATACTCACT  
AGCAGACAACATGCAGATCCGCTTCGAGTGTTCGAAATCAGTGAAATTC  
TAAAGTTCCAGCGATTCGAAAGGAAATCAGTGAAATTCGAAATTC  
ATCAGTACGAGTAAAGAAATTCGAAATTCGAAATTCGAAATTC  
ATTTCGCGCAAGCGGACTTTTGGGAATGAATGAAATAAAAAAA  
AATAATAAAAAACAACAACAGTGCAACAACAGCCGGGCATCTTCATAGAT  
AACTTCTGCCTGCACCTGGTATATGTACTTATCACATAGACATATATATA



# FlyBase

## FlyBase 2023:

New features and tools to accelerate your research

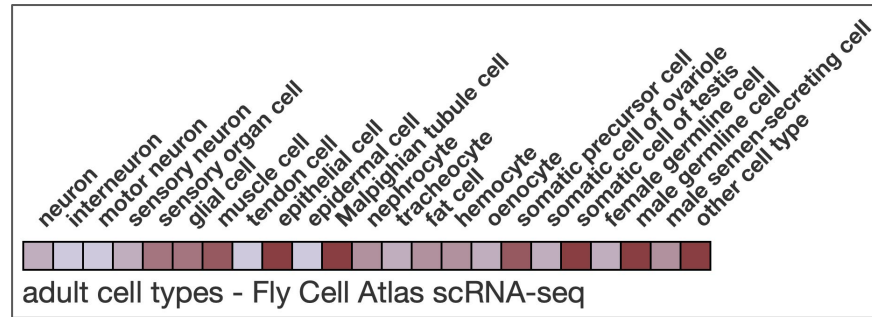
Steven Marygold

[sjm41@cam.ac.uk](mailto:sjm41@cam.ac.uk)

# Curation overview

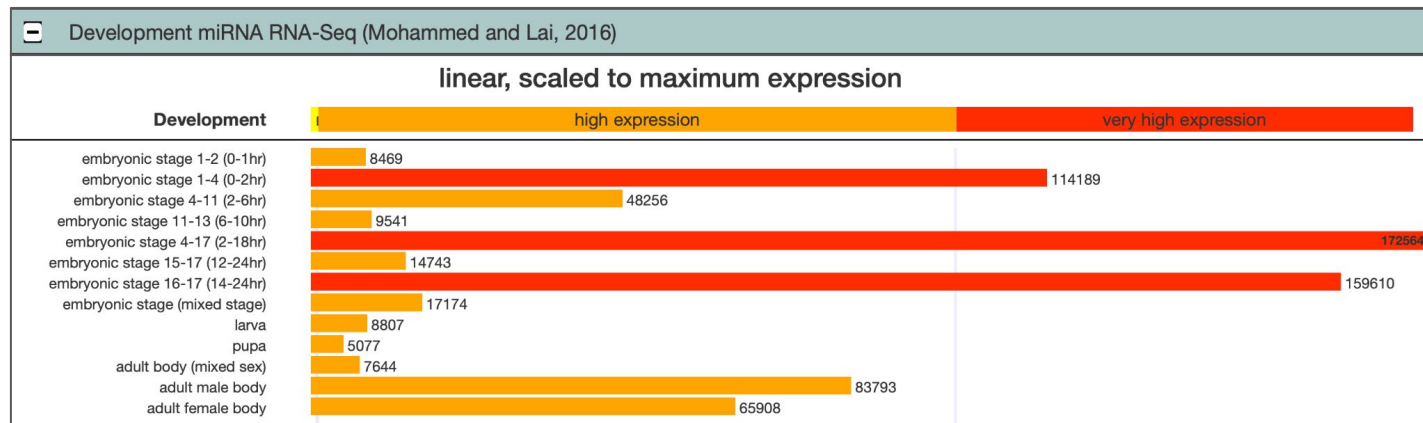
<b>Bibliography</b> ~2,870 references	<b>Ref-to-gene curation</b> ~1,330 papers (curators) ~1,150 papers (authors)	<b>Genetic reagents</b> ~23,220 alleles ~1,460 papers	<b>Phenotypes</b> ~7,660 annotations ~295 papers
<b>Functional data (GO)</b> ~3,020 annotations from ~715 papers	<b>Gene expression</b> ~1,570 annotations from ~1100 papers	<b>Gene Groups &amp; Signaling Pathways</b> ~70 new gene groups 1 new pathway	<b>scRNA-seq datasets</b> ~10 datasets ~370 single cell clusters
<b>Experimental Tools</b> ~80 tools	<b>Genome features</b> ~640 features from ~210 papers	<b>Disease models</b> ~1,510 allele assoc. ~65 new disease rep.	<b>Physical interactions</b> ~845 interactions from ~180 papers
<b>Gene models</b> ~10 new models	<b>Cell lines</b> ~320 associations ~10 new reports	<b>Chemicals</b> ~5,370 associations from ~1,490 papers	<b>Ontologies</b> ~482 fly anatomy terms ~46 other terms

# Expression: Fly Cell Atlas scRNA-seq data

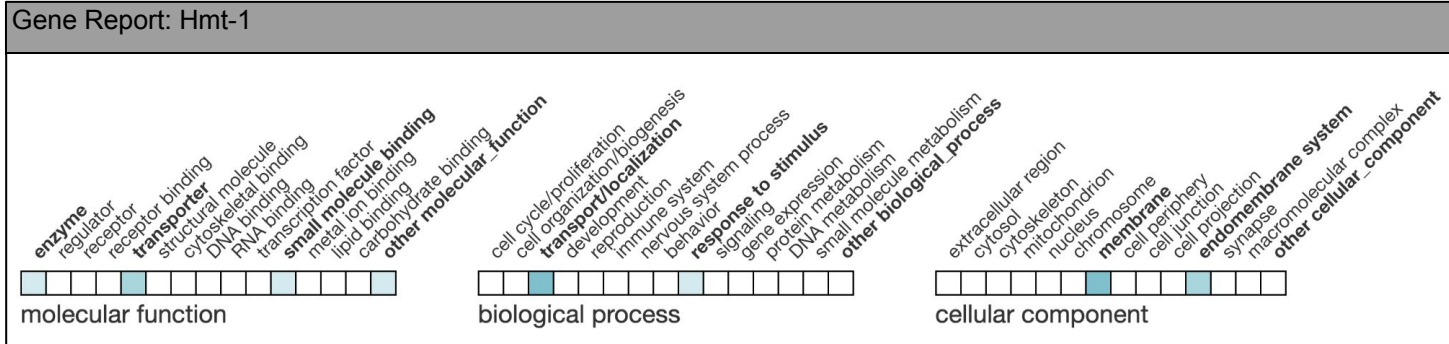


Proportion (%) of cells with detectable expression	cell type	Average level in cells with detectable expression
16	neuron	610
10	interneuron	870
19	motor neuron	429
15	sensory neuron	570
17	sensory organ cell	649
26	glial cell	744
31	muscle cell	983
13	tendon cell	992
25	epithelial cell	664
14	epidermal cell	848
66	Malpighian tubule cell	334
35	nephrocyte	538
23	tracheocyte	732
36	fat cell	531
18	hemocyte	617
25	oenocyte	460
34	somatic precursor cell	595
16	somatic cell of ovariole	482
58	somatic cell of testis	405
24	female germline cell	346
60	male germline cell	701
14	male semen-secreting cell	426

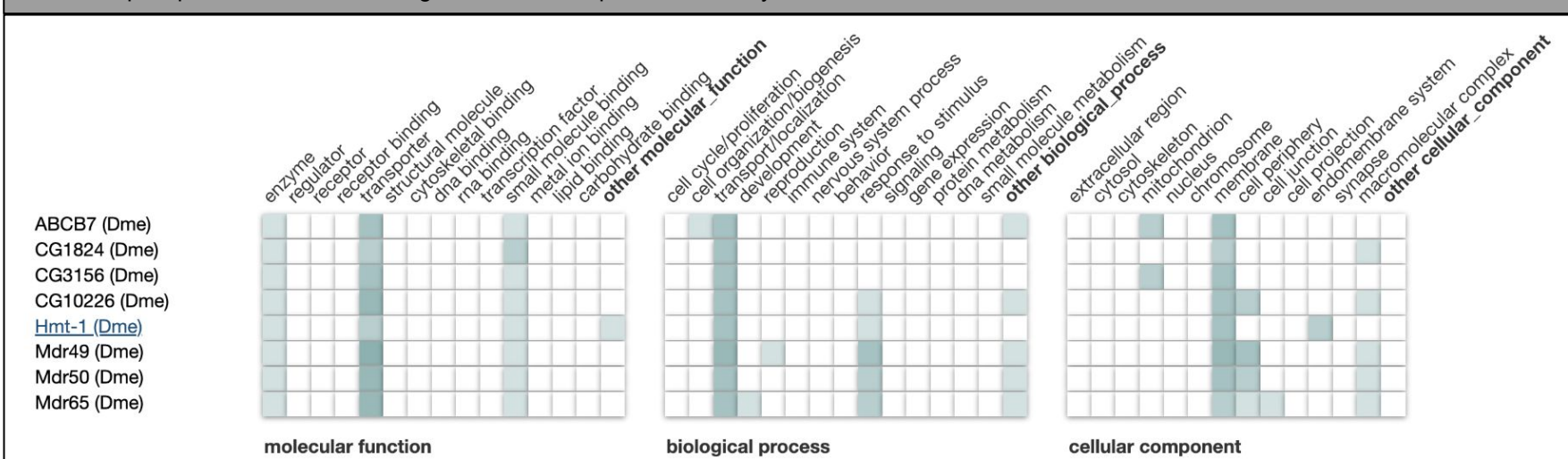
# Expression: miRNA histograms



# Function: GO ribbons and stacks



## Gene Group Report: ABCB ATP-Binding Cassette Transporter Subfamily



# Function: GO ribbon stack from HitList

Convert ▾ Export ▾ Analyze ▾

selected items to a FlyBase tool:

32 selected

Sequence Downloader

**GO Ribbon Stack Viewer**

Batch Download

QueryBuilder

FeatureMapper

Symbol

ion ID

selected items as a file:

rdhB ID list (download)

CAP FlyBase records crossreferencing table

Scer/Ga selected items to an external tool:

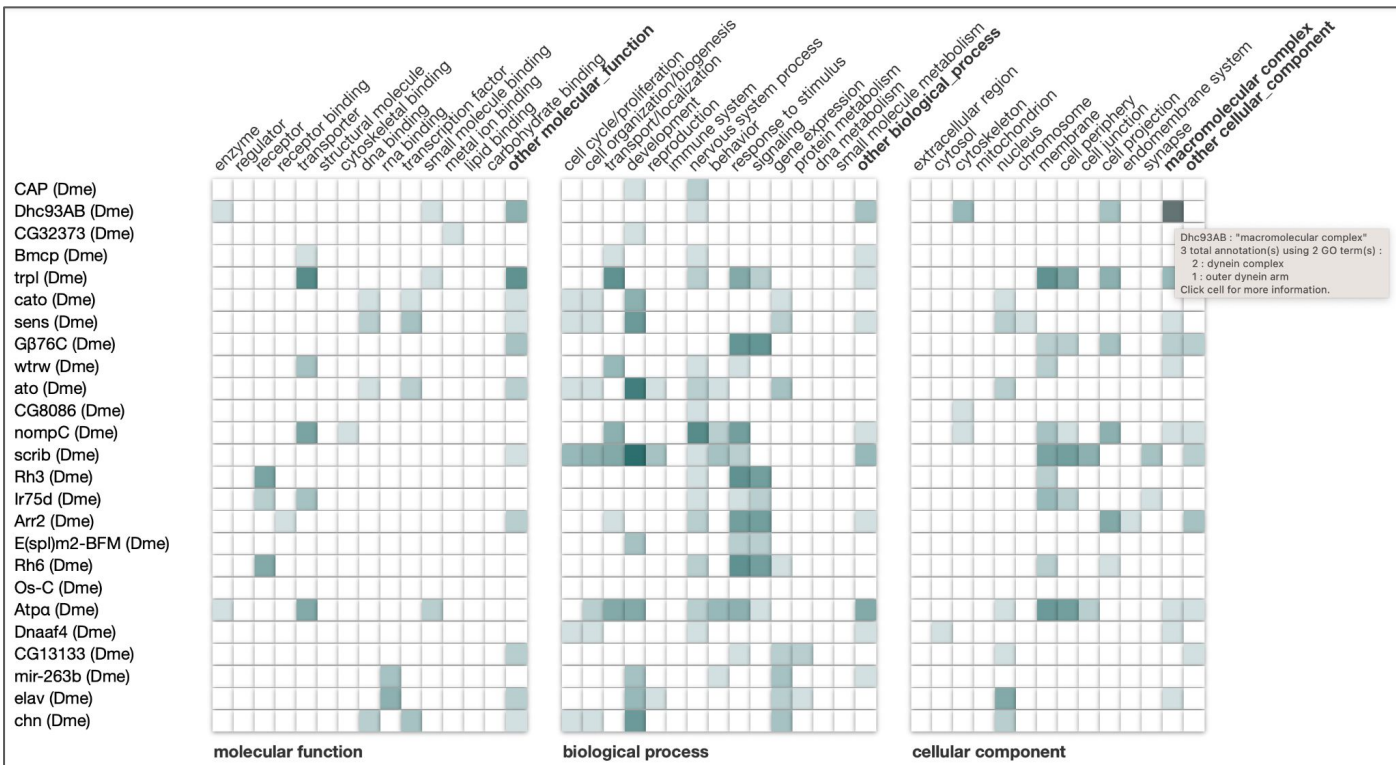
DCX-EMAP PANGEA Enrichment Tool (DRSC)

pinta prolonged depolarization afterpotential (PDA) is not apparent

santa-maria scavenger receptor acting in neural tissue and majority of rhodopsin is absent

tlIB touch insensitive larva B

twy twitchy



# Function: Export to PANGEA (GSEA tool) from HitList

Convert Export Analyze

selected items to a FlyBase tool:

- 32 selected
- Sequence Downloader
- GO Ribbon Stack Viewer
- Batch Download
- QueryBuilder
- FeatureMapper

selected items as a file:

- rdhB ID list (download)
- CAP FlyBase records crossreferencing table
- Scer/G... selected items to an external tool:
- PANGEA Enrichment Tool (DRSC)**

Symbol	Description	Gene ID
DCX-EMAP	Doublecortin-domain-containing echinoderm-microtubule-associated protein	CG42247
pinta	prolonged depolarization afterpotential (PDA) is not apparent	CG13848
santa-maria	scavenger receptor acting in neural tissue and majority of rhodopsin is absent	CG12789
tilB	touch insensitive larva B	CG14620
twy	twitchy	CG5964



**PANGEA**  
(Pathway, Network and Gene-set Enrichment Analysis)

Fly Worm Zebrafish Human Mouse Rat

### Search Single Gene List

Enrichment analysis for a single list of genes. If you wish to analyze multiple lists at once, please use the [Multiple Search](#) page.

- Select Gene Identifier: Gene Id Mapping**  
Species Specific Gene ID
- Enter Genes:** [Use Example Genes](#)  
FBgn0033504  
FBgn0013812  
FBgn0052373  
FBgn0036199  
FBgn0005514  
FBgn0024249  
FBgn002573  
FBgn0004823
- Enter your own background genes (optional) (?)**  
[Use Example Background](#)
- Choose Gene Sets:** [Use Defaults](#)  
[see Gene Set Descriptions](#)

### Gene Ontology sets

Using GO hierarchy

- GO Biological Processes
- GO Cellular Component
- GO Molecular Function

**Direct GO term only**

- Direct GO Biological Processes
- Direct GO Cellular Component
- Direct GO Molecular Function

### Gene Ontology Subsets

Generic GO consortium Subsets (GO slim)

- SLIM1 GO BP
- SLIM1 GO CC
- SLIM1 GO MF

### Other Gene sets

Gene group collections

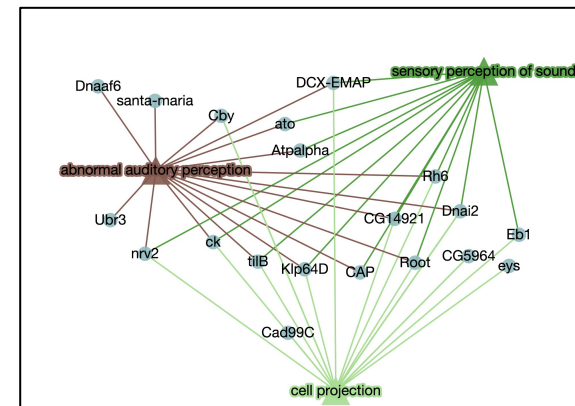
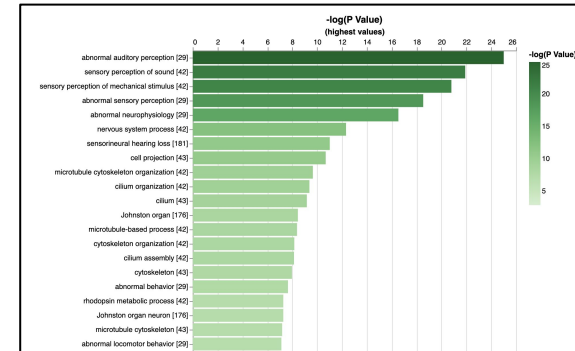
- DRSC GLAD Gene Group
- FlyBase Gene Group

Pathway Resources

- DRSC PathON signaling pathway core components
- DRSC PathON signaling pathway target genes
- FlyBase signaling pathway (experimental evidence)
- KEGG Pathway D.mel
- PANTHER pathway D.mel
- REACTOME pathway

Protein annotation

- DRSC COMPLEAT protein complex (literature)
- DRSC COMPLEAT protein interaction cluster (computed)



# Function: enzyme reaction graphics

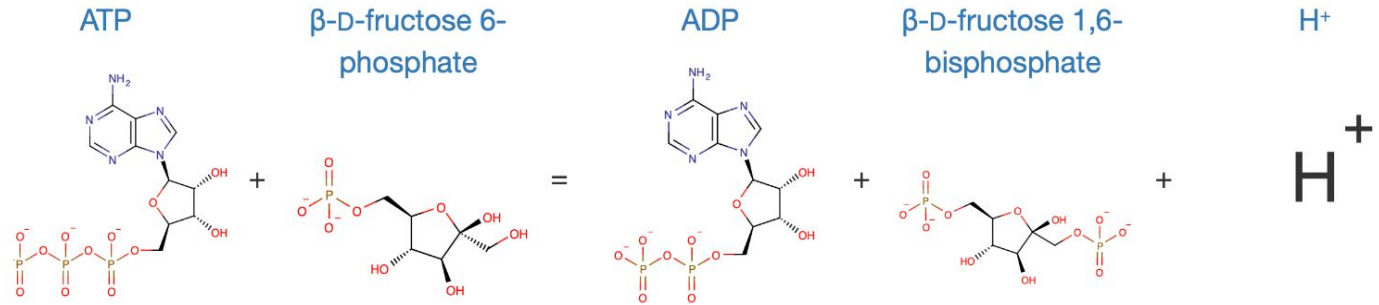
## Function

### Catalytic Activity (EC/Rhea)

#### 6-phosphofructokinase activity

ATP + beta-D-fructose 6-phosphate = ADP + beta-D-fructose 1,6- bisphosphate + H(+) (2.7.1.11)

RHEA 16109:



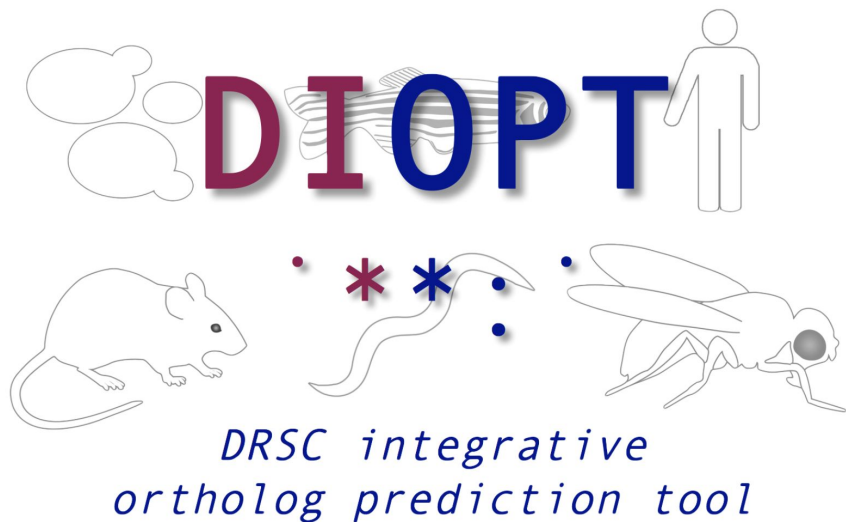




# Variant molecular consequences

Variant	Affected Genes	Related Alleles	Affected Transcripts
deletion C > C <a href="#">3R:11,963,124..11,964,369</a> NT_033777.3:g.11963124_11964369del	<a href="#">aurA</a> <b>consequence:</b> splice_acceptor_variant, splice_donor_variant, start_lost, 5_prime_UTR_variant, intron_variant <b>impact:</b> HIGH	<a href="#">aurA<sup>ST</sup></a>	<a href="#">aurA-RA</a> <b>consequence:</b> splice_acceptor_variant, splice_donor_variant, start_lost, 5_prime_UTR_variant, intron_variant <b>impact:</b> HIGH
insertion C > C. <a href="#">3R:11,963,170..11,963,171</a> NT_033777.3:g.11963170_11963171ins	<a href="#">aurA</a> <b>consequence:</b> 5_prime_UTR_variant <b>impact:</b> MODIFIER	<a href="#">aurA<sup>EY03490</sup></a> assoc. with <a href="#">P{EPgy2}aurA[EY03490]</a>	<a href="#">aurA-RA</a> <b>consequence:</b> 5_prime_UTR_variant <b>impact:</b> MODIFIER
point_mutation C > T <a href="#">3R:11,963,298</a> NT_033777.3:g.11963298C>T	<a href="#">aurA</a> <b>consequence:</b> missense_variant <b>impact:</b> MODERATE	<a href="#">aurA<sup>3</sup></a>	<a href="#">aurA-RA</a> <b>consequence:</b> missense_variant <b>impact:</b> MODERATE
point_mutation A > C <a href="#">3R:11,963,623</a> NT_033777.3:g.11963623A>C	<a href="#">aurA</a> <b>consequence:</b> missense_variant <b>impact:</b> MODERATE	<a href="#">aurA<sup>3</sup></a>	<a href="#">aurA-RA</a> <b>consequence:</b> missense_variant <b>impact:</b> MODERATE

# Orthology updates



# OrthoDB

- Orthologs in humans and model organisms
- Updated to v9.1
  - 14 different algorithms
  - *A. gambiae* and *E. coli* added
- Orthologs in other insects & arthropods
- Now retrieved via live API call
  - Always up-to-date (v11)

# Outreach: 'New to Flies' and FAQ update

## FlyBase:New to Flies

**FlyBase** is a database of genetic and molecular data for *D. melanogaster* and other *Drosophila* species. The resources listed here may be especially useful to a broad audience of research professionals. The resources listed here may be especially useful to you.

For a quick look at answers to common issues, check out the **Frequently Asked Questions**.

### Contents [hide]

- 1 Learning how to use *Drosophila*
- 2 Learning *Drosophila* genetics
- 3 External Resources on the FlyBase Wiki
- 4 How to use FlyBase
- 5 Social media and online groups

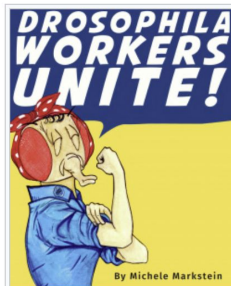
## Learning how to use *Drosophila*

A basic introduction to *Drosophila* can be found at these websites:

- An introduction to *Drosophila melanogaster*
- *Drosophila Workers Unite!* A laboratory manual for working with *Drosophila*
- *Fly Basics*, from the *Fly Research Portal*
- A quick and simple introduction to *Drosophila melanogaster*
- *FlyMove*: Images and movies describing *Drosophila* development

### Online books:

- *Experiments with *Drosophila* for Biology Courses*, eds. SC Lakhota and HA Ranganath: features 85 protocols commonly used in *Drosophila* labs, for use by both undergraduate labs and researchers
- *Drosophila melanogaster - Model for Recent Advances in Genetics and Therapeutics*, ed. FK Perveen: a disease model centered introduction to *Drosophila*
- *The Encyclopedia of North American Drosophilids Volume 1: Drosophilids of the Midwest and Northeast*, eds. T Werner, T Steenwinkel, and J Jaenike



*Drosophila Workers Unite!* A laboratory manual for working with *Drosophila*.

## FlyBase:FAQ

### FlyBase FAQ

1. Bulk data retrieval [\[Expand\]](#)
2. *D. melanogaster* as a model organism [\[Expand\]](#)
3. Fast-Track Your Paper [\[Expand\]](#)
4. FlyBase Community Advisory Group [\[Collapse\]](#)
  - 4.1. How can I join the FlyBase Community Advisory Group (FCAG)?  
Please fill in the FCAG registration form.
  - 4.2. I am a current member of the FCAG. I have moved to a new institution, so could you please update my details?  
Please fill in the FCAG registration form, so that we can update your information in our database.
5. FlyBase fee [\[Collapse\]](#)
  - 5.1. How can I pay the FlyBase fee? Can I pay the fee for my entire lab/institution/company?  
You can pay the FlyBase fee at [this link](#). We also have a dedicated [FlyBase Fees FAQ](#). Please contact us to discuss institutional/departmental/corporate fees.
6. FlyBase people database [\[Collapse\]](#)
  - 6.1. Is FlyBase people database still active?  
No, the FlyBase people database was retired.
7. Genome browser (JBrowse and GBrowse) [\[Collapse\]](#)
  - 7.1. Is GBrowse being discontinued?  
Yes, GBrowse is no longer being maintained, and GBrowse access will be discontinued in FlyBase release FB2202\_06. All GBrowse tracks and features are now also available in JBrowse. Please use [JBrowse](#) for your queries.

# Outreach: Fly Lab List



## FlyBase:Fly Lab List

The 'Fly Lab List' is an ongoing project aiming to generate an accurate list of all active labs undertaking a substantial fraction of their research using *Drosophila* (any species of *Drosophila*).

The current list can also be downloaded as an [Excel file](#).

If you wish to **revise** your lab's entry, please use this [form](#).

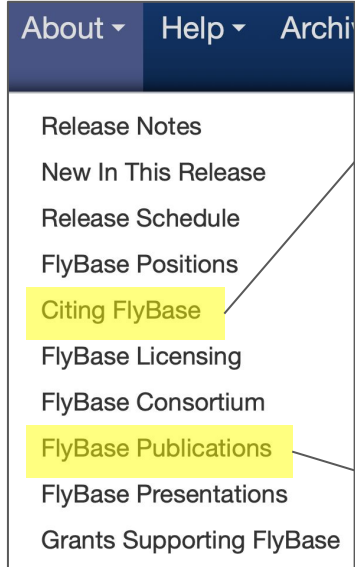
If your lab is not on the list and you wish to **add** it, please use this [form](#). Note the PI/Lab must have a website to be included.

If you wish to **remove** your entry from the list, or do not consider your lab to be an active Fly Lab, please opt out using this [form](#).

*Last updated: 20th October 2023 (1,939 entries)*

Lab head name			Website ↕	Primary affiliation				Additional affiliation(s) ↕
First ↕	Middle ↕	Last ↕		Institution ↕	City ↕	State/Province/Prefecture ↕	Country ↕	
Jessica	K	Abbott	<a href="#">Link</a>	Lund University	Lund		Sweden	
Uri		Abdu	<a href="#">Link</a>	Ben-Gurion University of the Negev	Beer'Sheva		Israel	
Hermann		Aberle	<a href="#">Link</a>	Heinrich Heine University Düsseldorf	Düsseldorf		Germany	
Elizabeth	T	Ables	<a href="#">Link</a>	East Carolina University	Greenville	North Carolina	USA	
Amos	O	Abolaji	<a href="#">Link</a>	University of Ibadan	Ibadan		Nigeria	
John	M	Abrams	<a href="#">Link</a>	University of Texas Southwestern Medical Center	Dallas	Texas	USA	
Jairaj	K	Acharya	<a href="#">Link</a>	National Cancer Institute	Frederick	Maryland	USA	
Takashi		Adachi-Yamada	<a href="#">Link</a>	Gakushuin University	Tokyo		Japan	
Michael	E	Adams	<a href="#">Link</a>	University of California	Riverside	California	USA	
Christof	M	Aegerter	<a href="#">Link</a>	Universität Zürich	Zürich		Switzerland	
Stein		Aerts	<a href="#">Link</a>	Katholieke Universiteit Leuven	Leuven		Belgium	
Markus		Affolter	<a href="#">Link</a>	University of Basel	Basel		Switzerland	

# Outreach: FlyBase publications



## Citing FlyBase

If you have used FlyBase during your research project, we urge you to cite us in your published work. We recommend that you use our latest *Genetics Knowledgebase and Database Resources* publication:

Gramates LS, Agapite J, Attrill H, Calvi BR, Crosby M, dos Santos G Goodman JL, Goutte-Gattat D, Jenkins V, Kaufman T, Larkin A, Matthews B, Millburn G, Strelets VB, and the FlyBase Consortium (2022) **FlyBase: a guided tour of highlighted features**. *Genetics*, Volume 220, Issue 4, April 2022, iyac035 [↗](#)

## FlyBase Publications

### 2023

Gene Ontology Consortium (2023)

**The Gene Ontology knowledgebase in 2023.**

*Genetics* 224(1):iyad031. DOI:10.1093/genetics/iyad031 [↗](#)

Hu Y, Comjean A, Attrill H, Antonazzo G, Thurmond J, Chen W, Li F, Chao T, Mohr SE, Brown NH, Perrimon N. (2023)

**PANGEA: a new gene set enrichment tool for Drosophila and common research organisms.**

*Nucleic Acids Res*:gkad331. DOI:10.1093/nar/gkad331 [↗](#) FBrf0256969 [↗](#)

Marygold SJ and the FlyBase Consortium (2023)

**Exploring FlyBase Data Using QuickSearch (Updated protocol)**

*Current Protocols* 3:e731. DOI:10.1002/cpz1.731 [↗](#) FBrf0256232 [↗](#)

# Coming soon...

- Macromolecular complex reports
- Better representation of split-GAL4 lines & combinations
- ‘Gene toolkit’
- Metabolic pathways
- Further integration with the



**ALLIANCE**  
of GENOME RESOURCES

# Notifications and help



Tools ▾

Downloads ▾

Links ▾

Community ▾

Species ▾

About ▾

Help ▾

Archives ▾

Fast-Track Your Paper

FlyBase Community  
Advisory Group

Fly Lab List

Recent Papers With  
Technical Advances

Gene Snapshots

FlyBase Forum

Newsletter

FlySlack Community

Twitter

Fly Board

NIH Funding

Release Notes

New In This Release

Release Schedule

FlyBase Positions

Citing FlyBase

FlyBase Licensing

FlyBase Consortium

FlyBase Publications

FlyBase Presentations

Grants Supporting FlyBase

**FlyBase Wiki:**

Help Index

Tool help

Report help

FlyBase FAQ

Info for Authors ▶

Linking to/from FlyBase

Nomenclature

Curation documentation ▶

New to Flies?

Video Tutorials ↗

FlyBase For Developers

Contact FlyBase



# Acknowledgements

## **FlyBase @ University of Cambridge:**

Nick Brown (Co-PI)  
Giulia Antonazzo  
Helen Attrill  
Damien Goutte-Gattat  
Aoife Larkin  
Steven Marygold  
Alex McLachlan  
Gillian Millburn  
Arzu Ozturk-Colak  
Clare Pilgrim  
Esther Shackleton\*

## **FlyBase @ University of New Mexico:**

Richard Cripps (Co-PI)  
TyAnna Lovato

## **FlyBase @ Harvard University:**

Norbert Perrimon (PI)  
Susan Russo Gelbart (PD)  
Kris Broll  
Lynn Crosby  
Gil dos Santos  
Kathleen Falls  
L. Sian Gramates  
Victoria Jenkins  
Ian Longden  
Beverley Matthews  
Jolene Seme  
Christopher Tabone  
Pinglei Zhou  
Mark Zytkovicz

## **FlyBase @ Indiana University:**

Thomas Kaufman (Co-PI)  
Brian Calvi (Co-PI)  
Seth Campbell  
Josh Goodman  
Victor Strelets  
Jim Thurmond

## **Collaborators:**

Fly Cell Atlas / SCEA  
FlyAtlas2  
Eric Lai  
DRSC (PANGEA, DIOPT)  
Rhea  
AlphaFold  
RNAcentral  
Ensembl VEP  
OrthoDB

## **Funding:**



**NHGRI** U41HG000739, U24HG010859

**NIDDK:** 1R01DK136945-01

**MRC:** W024233/1

**NSF:** DBI-2035515, 2039324

**BBSRC:** BB/T014008/1

**Wellcome Trust:** PLM13398

**FlyBase Users**