



# New and Classic Features in FlyBase

Gil dos Santos, FlyBase Consortium

Presented at the 17th Canadian Drosophila Research Conference, Canfly 2023

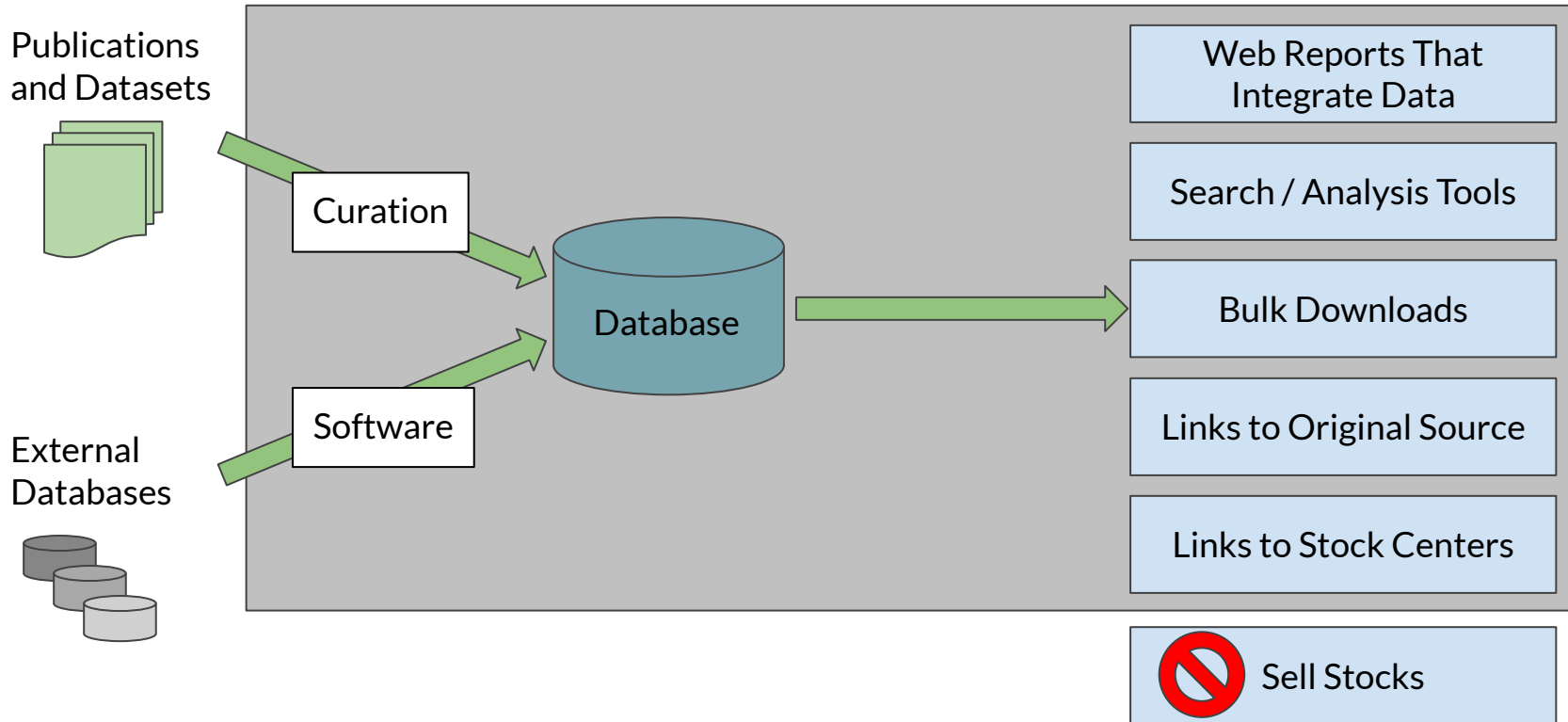


# Overview

1. What does FlyBase do?
2. Finding your favourite gene.
3. FlyBase homepage overview.
4. Navigation bar.
5. Big buttons.
6. QuickSearch tabs.
7. External resources
8. Connecting with FlyBase.

# What does FlyBase do?

FlyBase is the main data warehouse for 100+ years of *Drosophila* research.



# Finding your favourite gene

The screenshot shows the FlyBase homepage with the following elements:

- Header: FlyBase logo, "A Database of Drosophila Genes & Genomes", and navigation menu (Home, Tools, Downloads, Links, Community, Species, About, Help, Archives).
- Search bar: "J2G" selected in the dropdown, "Jump to Gene" text, and a "Go" button.
- Navigation icons: BLAST, JBrowse, New to Files?, PHENOTYPE ANATOMY DISEASE MORE, ID Validator, Batch Download, Fast-Track.
- Left sidebar: "Tweets from @FlyBaseDotOrg" featuring Brian Oliver's tweet about website access fees, and "External Resources" including News and Outreach, Journals and Protocols, Tool/Reagent Resources, Meetings/Courses/Jobs, Disease Research, Inter-species, and Fly Boards.
- Main content: A green banner for website access fees, a "QuickSearch" section with filters (Human Disease, Protein Domains, Gene Groups, Pathways, GO, Data Class) and a search input field, and a "Commentary" section for the "CanFly 2023" conference.
- Footer: "version FB2023\_02, released April 18, 2023", "Contact FlyBase", "FAQ", "Cite FlyBase", and social media icons.

A close-up of the search bar showing the dropdown menu with "J2G" selected, the text "wg", and a "Go" button.

The screenshot shows the FlyBase Gene Detail page for the 'wg' gene. The page includes:

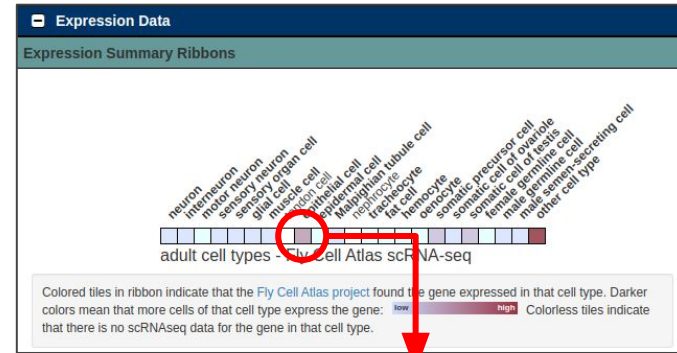
- Gene Summary: Name, description, and accession numbers.
- Gene Location: Genomic coordinates and map.
- Gene Structure: Exon-intron diagram.
- Protein: Amino acid sequence and domain annotations.
- References: List of publications.
- External Links: Links to other databases like UniProt, Ensembl, and NCBI.

# New on gene reports: scRNA-seq

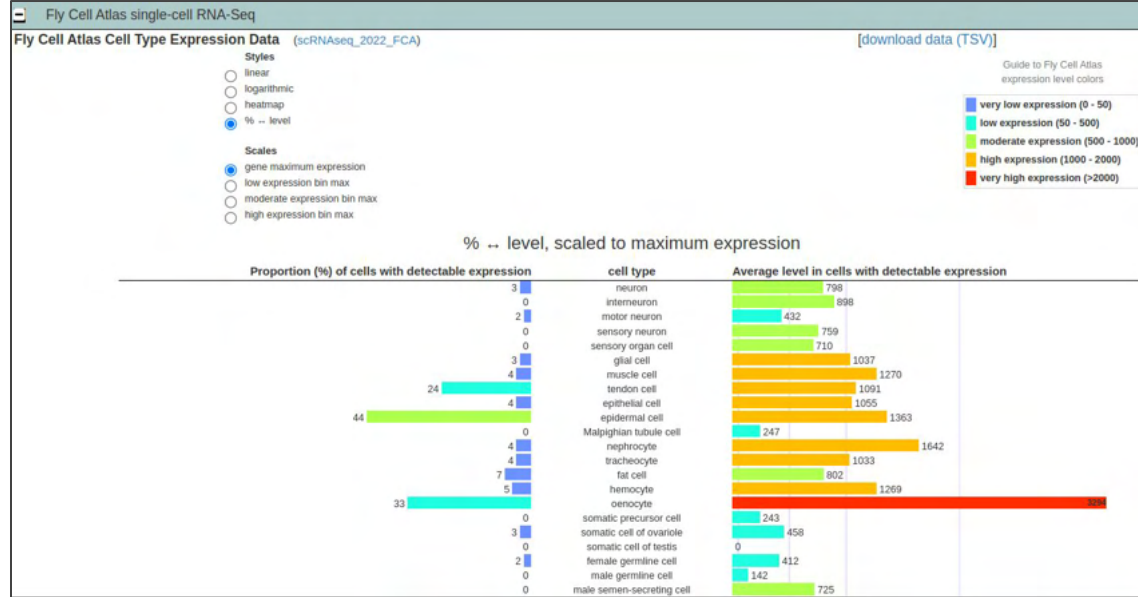
## scRNA-seq expression data

Explore expression in specific cell types.

View expression levels and proportion of cells with expression.



**EPITHELIAL CELL:**  
 cell types + gene expression spread (%) within related cluster :  
 28% : adult proventriculus epithelial cell  
 10% : enterocyte of posterior adult midgut epithelium  
 4% : adult hindgut epithelial cell  
 3% : adult copper cell  
 3% : adult enteroendocrine cell  
 3% : midgut large flat cell  
 2% : enterocyte of anterior adult midgut epithelium  
 1% : adult enterocyte  
 1% : adult midgut enterocyte  
 1% : adult pyloric epithelial cell  
 1% : epithelial cell of crop  
 <1% : 11 cell type(s) not shown



# FlyBase homepage overview

navigation bar

big buttons

Twitter feed

external resources

connecting with FlyBase

The screenshot shows the FlyBase homepage with the following sections and features:

- Navigation Bar:** Includes links for Home, Tools, Downloads, Links, Community, Species, About, Help, and Archives. A search bar is present with the text "J2G - Jump to Gene".
- Release Version:** A small box at the top right indicates "FB2023\_02, released April 18, 2023".
- Big Buttons:** A row of large, colorful buttons for BLAST, JBrowse, New to Files?, GO PHENOTYPE ANATOMY DISEASE MORE, ID Validator, Batch Download, and Fast-Track.
- Twitter Feed:** A section titled "Tweets from @FlyBaseDotOrg" showing a tweet from Brian Oliver.
- External Resources:** A sidebar menu with categories like News and Outreach, Drosophila Journals and Protocols, Tool/Reagent Resources (wiki), Meetings/Courses/Jobs, Drosophila in Disease Research, Inter-species, and Fly Boards.
- QuickSearch:** A search interface with tabs for Human Disease, Protein Domains, Gene Groups, Pathways, GO, and Data Class. It includes a search input field and a search button.
- Commentary:** A section titled "Commentary" featuring a large image of a fly and a text box for "CanFly 2023" registration information.
- Footer:** Contains text about grants supporting FlyBase, copyright information, and social media icons for Twitter, YouTube, and GitHub.

release version

Jump To Gene

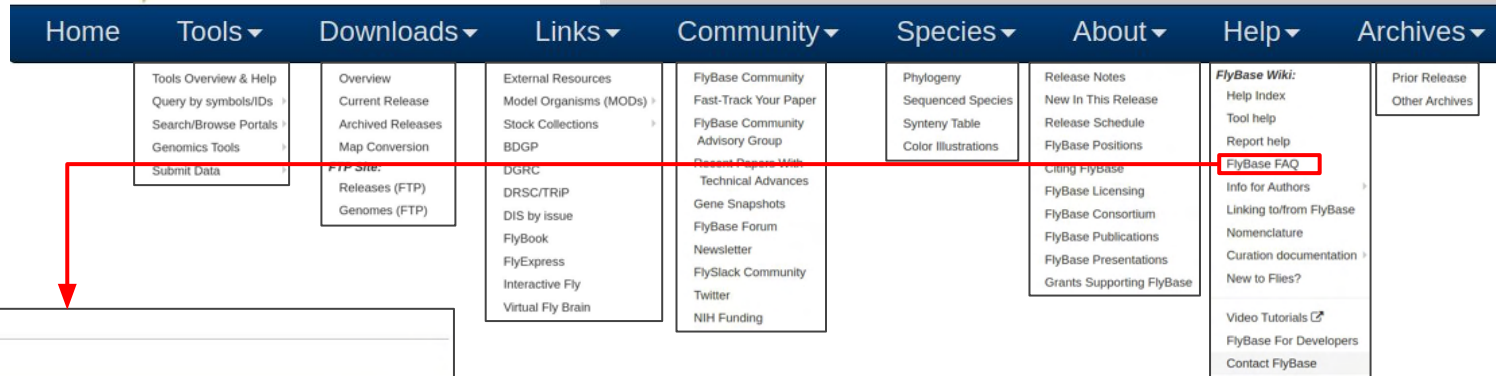
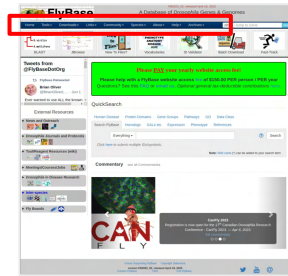
fees

QuickSearch tabs

Commentary

# Navigation bar

Access all of FlyBase tools and resources through organized menus



## 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 [\[Expand\]](#)
5. FlyBase fee [\[Expand\]](#)
6. FlyBase people database [\[Expand\]](#)
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.
  - 7.2. How can I download FASTA sequence from JBrowse?

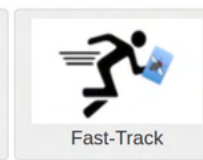
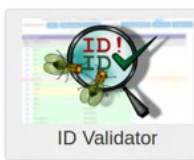
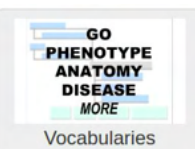
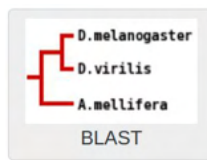
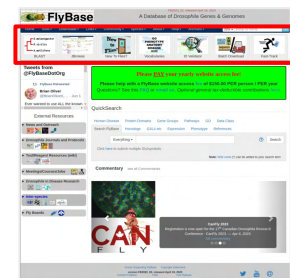
Please find instructions in [this FlyBase tweetorial@](#).

It is currently not possible to download decorated FASTA from JBrowse; however you can download decorated FASTA from the Gene Report of your gene of interest using the "Get Decorated FASTA" button in the "Genomic Location" section (see Question 8.2 for more information).



# Big buttons: New to Flies

Quick access to key FlyBase tools: e.g., resources for newbies.



A screenshot of the FlyBase Wiki page titled 'New to Flies'. The page includes a fly icon, a navigation menu, and a main content area. The main content area has a 'Contents' section with links to 'Learning how to use Drosophila', 'Learning Drosophila genetics', 'External Resources on the FlyBase Wiki', 'How to use FlyBase', and 'Social media and online groups'. Below this is a section titled 'Learning how to use Drosophila' with a list of links to various resources. A red arrow points from the 'New to Flies?' button to this page. At the bottom, there is a note about the Manchester Fly Facility.

FlyBase Discussion Read View source View history Search FlyBase Wiki

## FlyBase:New to Flies

FlyBase is a database of genetic and molecular data for *D. melanogaster* and other *Drosophila* species, targeted to an audience of research professionals. The resources listed here may be especially useful to those new to flies. For a quick look at answers to common issues, check out the [Frequently Asked Questions](#) page.

- 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 Lakhotia 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 Plevrens; 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

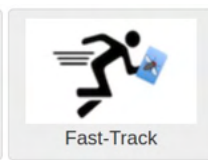
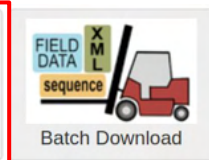
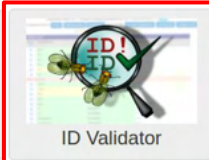
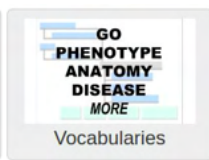
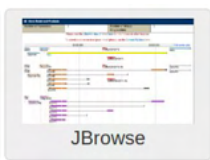
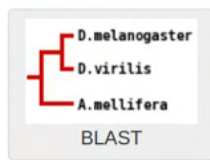
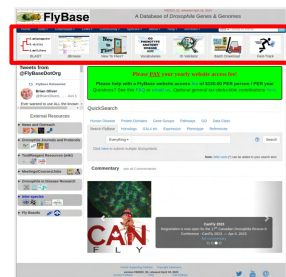


The [Manchester Fly Facility](#) has a large compendium of *Drosophila*-related links for the public, educators, and



# Big buttons: ID Validator

ID Validator accepts many types of IDs and names and finds related FlyBase object(s)



Enter IDs or symbols to validate Help

This tool will accept a list of FlyBase symbols/IDs (for any data type) and, where necessary/possible, update them to their current versions. It will also convert certain external IDs (GenBank nucleotide/protein accessions, UniProt accessions, PubMed IDs) into their equivalent FlyBase IDs. The output is provided as a validation table that can either be downloaded as a file or exported to a FlyBase HitList for further processing (including conversion between data types).

Enter IDs or Symbols:

- watermelon
- CG15086
- CG41367
- CG42619
- CG40930
- P09615

or Upload File of Identifiers:

Return non-melanogaster matches  
 Match synonyms / secondary IDs

Submitted Items: 6 | Unique Validated/Updated IDs: 6 | Unknown IDs: 0 | [Help](#)

Export selected IDs to:

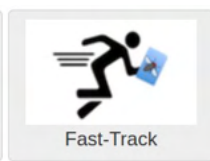
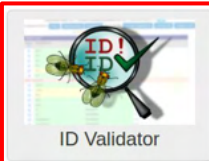
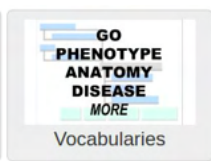
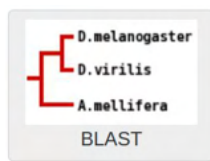
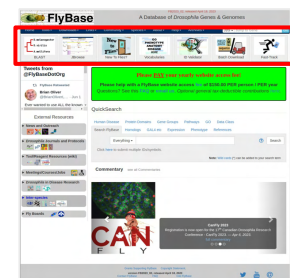
Save as file:

(1) WARNING: SOME ITEMS HAVE MULTIPLE MAPPINGS

Validation report					
Export	Submitted Item	Validated ID	Data Class	Filter in HitList	Related record
watermelon - unknown ID					
<input checked="" type="checkbox"/>	CG15086 !	FBgn0265625	Gene		CG44433
<input checked="" type="checkbox"/>	CG15086 !	FBgn0265626	Gene		CG44434
<input checked="" type="checkbox"/>	CG15086 !	FBgn0265627	Gene		CG44435
<input checked="" type="checkbox"/>	CG41367	FBgn0267489 !	Gene		PRY
<input checked="" type="checkbox"/>	CG42619	FBgn0267489 !	Gene		PRY
<input checked="" type="checkbox"/>	CG40930	FBgn0267489 !	Gene		PRY
<input checked="" type="checkbox"/>	P09615	FBgn0284084 !	Gene		wg
<input checked="" type="checkbox"/>	spd !	FBgn0065045	Gene		spd
<input checked="" type="checkbox"/>	spd !	FBgn0284084 !	Gene		wg

# Big buttons: ID Validator

ID Validator is an entry point to many list filtering and analysis tools



Convert - Export - Analyze - View As

Filter by species

- 4 selected
- D. melanogaster (8)
- H. sapiens (transgenes in flies) (0)
- other Drosophila species (0)
- Other species (0)

Filter by data class

- Gene (4)
- Transgenic Construct (2)

selected items in a FlyBase tool

- Sequence Download
- GO Ribbon Stack Viewer
- Batch Download
- QueryBuilder
- FeatureMapper
- selected items as a file
- ID list (download)
- FlyBase records crossreferencing table
- selected items to an external tool
- PANGEA Enrichment Tool (DRSC C)

Gene model status: Current 94E1-94E1  
Cytogenetic Map: 1 Polypeptide 2999 References

Gene Snapshot >

melanogaster  Gene

**dpp**  wingless (CG4889; FBgn0294084) *D. melanogaster*  Gene

Feature type: protein coding gene  
Sequence Location: 2L:2,428,372..2,459,823 [-]  
350 Alleles 106 Stocks 4 Transcripts 4 Polypeptides 2991 References

Gene model status: Current 22F1-22F3  
Cytogenetic Map: 2 Polypeptides 2991 References

Gene Snapshot >

**wg**  wingless (CG4889; FBgn0294084) *D. melanogaster*  Gene

Feature type: protein coding gene  
Sequence Location: 2L:7,207,159..7,216,265 [-]  
303 Alleles 718 Stocks 1 Transcript 1 Polypeptide 3470 References

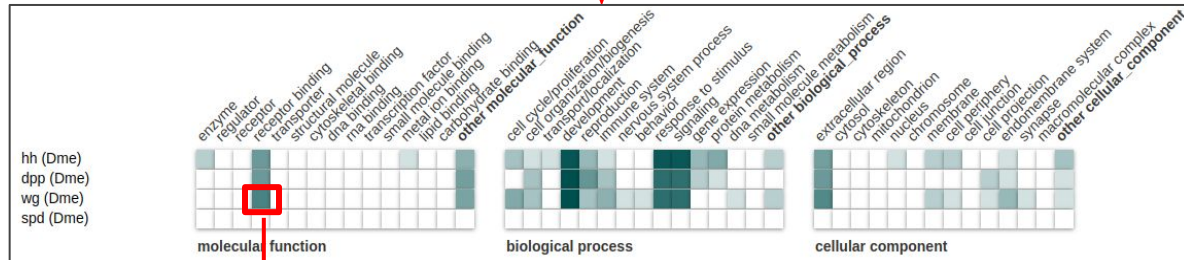
Gene model status: Current 27F1-27F1  
Cytogenetic Map: 1 Polypeptide 3470 References

Gene Snapshot >

**spd**  slow-phase-defective (FBgn0065045) *D. melanogaster*  Gene

Feature type: gene  
Sequence Location: [unlabeled]  
0 Alleles 0 Stocks 0 Transcripts 0 Polypeptides 1 Reference

Gene model status: Unannotated  
Cytogenetic Map: [unlabeled]



wg : "receptor binding"  
16 total annotation(s) using 5 GO term(s) :

- 5 : frizzled binding
- 5 : morphogen activity
- 4 : receptor ligand activity
- 1 : cytokine activity
- 1 : signaling receptor binding

Click cell for more information.

# QuickSearch tabs

## Various ways to search or browse FlyBase data.

Google-esque search of everything.

Human Disease Protein Domains Gene Groups Pathways GO Data Class  
Search FlyBase Homologs GAL4 etc Expression Phenotype References

Everything  Search

Click here to submit multiple IDs/symbols.

Note: Wild cards (\*) can be added to your search term

Data class-specific search is **faster**:  
chemicals are new!

Human Disease Protein Domains Gene Groups Pathways GO Data Class  
Search FlyBase Homologs GAL4 etc Expression Phenotype References

Search:  
 ID/Symbol/Name  Everything

Data Class:  
Chemical

Enter text:

- paclitaxel
- palbociclib
- palmitoleic acid
- pamidronate
- panobinostat
- pantheine
- pantheine 4'-phosphate

multiple IDs/symbols.

Note: Wild cards (\*) can be added to your search term

all Commentaries

Search by phenotype.

Human Disease Protein Domains Gene Groups Pathways GO Data Class  
Search FlyBase Homologs GAL4 etc Expression Phenotype References

Phenotypic class: e.g. 'lethal' Search

refinement: e.g. 'recessive' and (refinements are optional)

Autocomplete coordinates only within each of the two groups of fields.

Tissue/cell affected: e.g. 'fat body'

refinement: e.g. 'somatic clone' and (refinements are optional)

Search for homologs.

Human Disease Protein Domains Gene Groups Pathways GO Data Class  
Search FlyBase Homologs GAL4 etc Expression Phenotype References

Input

Species:

Gene(s):

Enter gene symbol(s) or ID(s), separated by spaces

Output

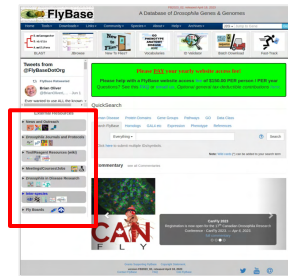
HUMAN AND MODEL ORGANISMS (via DIOP7)

<input checked="" type="checkbox"/> <i>H. sapiens</i> (Human)	<input checked="" type="checkbox"/> <i>D. melanogaster</i> (Fruit fly) PARALOGS
<input checked="" type="checkbox"/> <i>R. norvegicus</i> (Norway rat)	<input checked="" type="checkbox"/> <i>C. elegans</i> (Nematode, roundworm)
<input checked="" type="checkbox"/> <i>M. musculus</i> (Laboratory mouse)	<input checked="" type="checkbox"/> <i>A. thaliana</i> (Thale cress)
<input checked="" type="checkbox"/> <i>X. tropicalis</i> (Western clawed frog)	<input checked="" type="checkbox"/> <i>S. cerevisiae</i> (Brewer's yeast)
<input checked="" type="checkbox"/> <i>D. rerio</i> (Zebrafish)	<input checked="" type="checkbox"/> <i>S. pombe</i> (Fission yeast)





uncheck all:

# External resources





## Featured external resources






▼ News and Outreach

-  FlyBase News
-  Community News
-  FlySlack
-  Public, Teachers, Students



▼ *Drosophila* Journals and Protocols

-  Fly journal
-  microPublication
-  FlyBook
-  Experiments with *Drosophila*




▼ Tool/Reagent Resources (wiki)

-  scRNA-Seq
-  Technical Advances
-  External Resources





▼ Meetings/Courses/Jobs


-  Meetings Courses
-  Positions Available

▼ *Drosophila* in Disease Research



-  Biomedical Research
-  MARRVEL
-  Model Matcher

▼ Inter-species

-  Gene2Function
-  MIST
-  iProteinDB
-  BioLitMine

 Alliance of Genome Resources

▼ Fly Boards

-  Fly Board
-  European Drosophila Society



# The Alliance of Genome Resources

FlyBase is a founding member (alliancegenome.org) ; data is shared both ways.

Data imports into FlyBase.

Gene reports

General Information			
Symbol	DmelCG17807	Species	D. melanogaster
Name		Annotation Symbol	CG17807
Feature Type	protein_coding_gene	FlyBase ID	FBgn0034748
Gene Model Status	Current	Stock	2 publicly available
Enzyme Name (EC)	EC1.1.1.18 (methyltransferase)		
Gene Summary	Predicted to enable tRNA (carboxymethyluridine(34)-5-O)-methyltransferase activity; tRNA (uracil) methyltransferase activity; and tRNA binding activity. Predicted to be involved in tRNA methylation and tRNA wobble uridine modification. Predicted to be active in cytoplasm and nucleus. Human ortholog(s) of this gene implicated in autosomal recessive intellectual developmental disorder 71. Orthologous to human ALKBH8 (alkB homolog 8, tRNA methyltransferase). (Alliance, FBgn0034748)		
Key Links	<a href="#">Alliance</a> <a href="#">Auto summary</a> <a href="#">Gene Group</a>		

Key Links: [Alliance of Genome Resources](#), [NCBI](#), [Ensembl](#), [UniProt](#), [MARRVEL](#)

Allele reports

Report of the Allele			
Type	Location	Additional Notes	References
Allele	3L:1,247,891..1,249,852 (+)	CG17807 is an insertion, resulting from the imprecise excision of P[EP50438] elements from the 3rd L ending insertion.	[Lee et al., 2012]
Variant Molecular Consequences			
Affected Genes	Related Alleles	Affected Transcripts	
DmelCG17807	CG17807	DmelCG17807	

Cross-species data alignment.



Cross-model organism forum.

Screenshot of the Alliance of Genome Resources forum. The forum displays various posts and categories. The top navigation bar includes links for Home, Data and Tools, Members, News, About, Working Groups, Help, Community, and Contact Us. The main content area shows a list of posts with categories, topics, and latest dates. A red box highlights a survey link: "Help us improve Alliance resources by taking a very brief SURVEY". Another red box highlights the "Alliance User Community" link. The forum posts include "Alliance of Genome Resources", "Job Postings", "Positions Wanted", "Meeting Announcements", and "Model Organism: Files".

# Connecting with FlyBase

Ask a question, provide a suggestion, get news and helpful hints.

The image shows a screenshot of the FlyBase website with several elements highlighted by red boxes and arrows:

- Top Navigation Bar:** Home, Tools, Downloads, Links, Community, Species, About, Help, Archives.
- Community Menu:** FlyBase Community, Fast-Track Your Paper, FlyBase Community Advisory Group, Recent Papers With Technical Advances, Gene Snapshots, FlyBase Forum, Newsletter, FlySlack Community, Twitter, NIH Funding.
- Help Menu:** 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 Files?, Video Tutorials, FlyBase For Developers, Contact FlyBase.
- Footer:** Grants Supporting FlyBase, Copyright Statement, version FB2023\_02, released April 18, 2023, Cite FlyBase, Contact FlyBase, FAQ, Twitter, YouTube, @.

Red arrows indicate the following connections:

- From the top navigation bar to the main content area.
- From the 'FlyBase' logo in the top left to the main content area.
- From the 'FlyBase Presentations' link in the Help menu to the email address `helpfb@morgan.harvard.edu`.
- From the 'Contact FlyBase' link in the footer to the email address `helpfb@morgan.harvard.edu`.
- From the 'Contact FlyBase' link in the footer to the 'Contact FlyBase' link in the footer.
- From the 'FAQ' link in the footer to the 'FAQ' link in the footer.
- From the 'Twitter' icon in the footer to the 'Twitter' link in the Community menu.
- From the 'YouTube' icon in the footer to the 'YouTube' icon in the footer.
- From the '@' icon in the footer to the '@' icon in the footer.

# 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

## **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 Zytковicz

## **FlyBase @ Indiana University:**

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

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

Richard Cripps (Co-PI)  
TyAnna Lovato

**We thank FlyBase users  
worldwide for their support.**



**NHGRI U41HG000739  
MRC: MR/N030117/1  
NSF: DBI-2035515, 2039324**