



New and Classic Features in FlyBase

Gil dos Santos, FlyBase Consortium

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

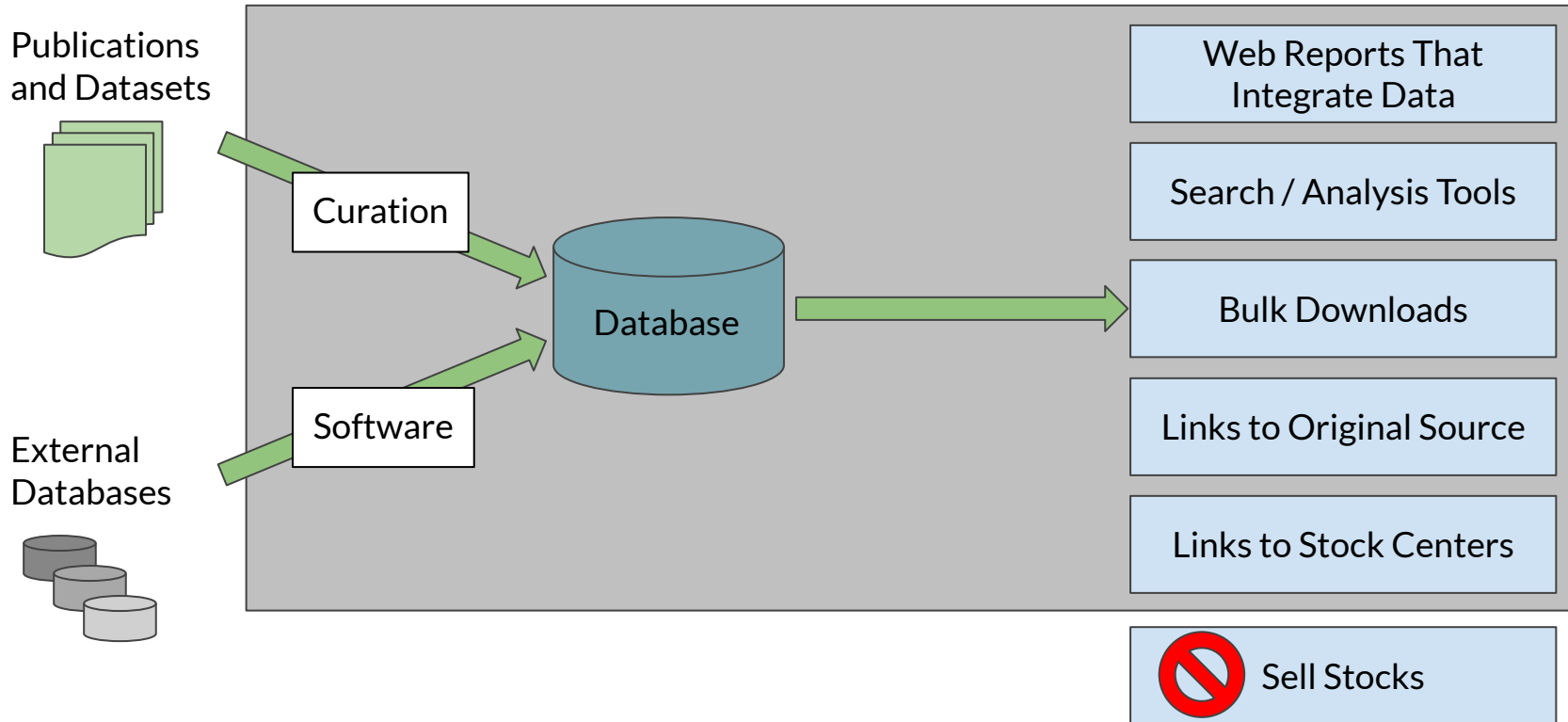


Overview

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. At the top, the navigation bar includes 'Home', 'Tools', 'Downloads', 'Links', 'Community', 'Species', 'About', and 'Help'. A search bar contains 'J2G' and a 'Go' button. Below the navigation bar are several utility icons: 'D. melanogaster', 'D. virilis', 'A. mellifera', 'New to Files?', 'GO PHENOTYPE ANATOMY DISEASE MORE', 'ID Validator', 'Batch Download', and 'Fast-Track'. A sidebar on the left features 'Tweets from @FlyBaseDotOrg' and 'External Resources' such as 'News and Outreach', 'Drosophila Journals and Protocols', 'Tool/Reagent Resources (wiki)', 'Meetings/Courses/Jobs', 'Drosophila in Disease Research', 'Inter-species', and 'Fly Boards'. A central green banner reads 'Please PAY your yearly website access fee!'. Below this is a 'QuickSearch' section with a search bar and filters for 'Human Disease', 'Protein Domains', 'Gene Groups', 'Pathways', 'GO', and 'Data Class'. At the bottom, there is a 'Commentary' section with an image of a fly and a 'CanFly 2023' announcement.

A close-up of the search bar from the FlyBase homepage. The search bar contains the text 'J2G' followed by a dropdown arrow and 'wg'. To the right of the search bar is a 'Go' button. A red arrow points from the search bar to the 'Go' button.

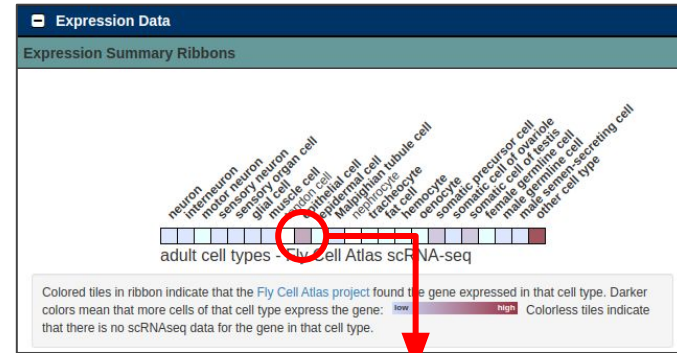
The screenshot shows the FlyBase Gene Detail page for the gene 'J2G'. The page is titled 'Gene Detail' and displays various information about the gene, including its name, location, and associated data. The page is organized into several sections: 'Gene Summary', 'Gene Location', 'Gene Structure', 'Gene Ontology', 'Gene Expression', 'Gene Interactions', and 'Gene History'. The 'Gene Summary' section includes a table with columns for 'Gene Name', 'Gene ID', 'Gene Symbol', 'Gene Type', 'Gene Class', and 'Gene Status'. The 'Gene Location' section shows a map of the gene's location on the chromosome. The 'Gene Structure' section displays a diagram of the gene's structure, including exons and introns. The 'Gene Ontology' section lists the gene's associated GO terms. The 'Gene Expression' section shows a table of gene expression data. The 'Gene Interactions' section displays a network of gene interactions. The 'Gene History' section provides a timeline of the gene's discovery and characterization. The page also includes a 'References' section with a list of publications and a 'Comments' section for user feedback.

New on gene reports: scRNA-seq

scRNA-seq expression data

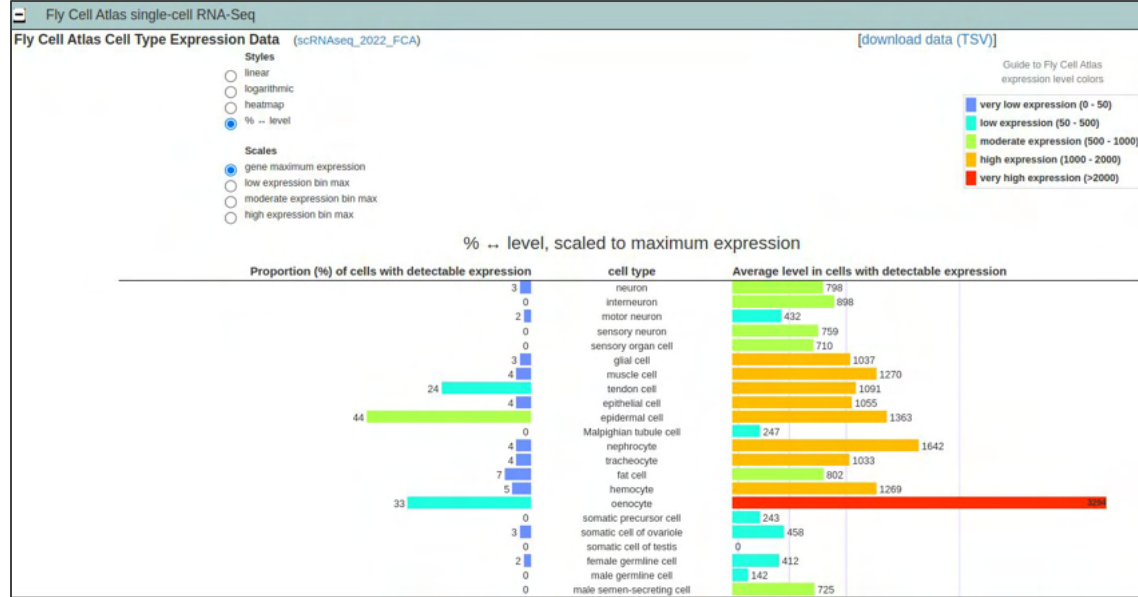
Explore expression in specific cell types.

View expression levels and proportion of cells with expression.

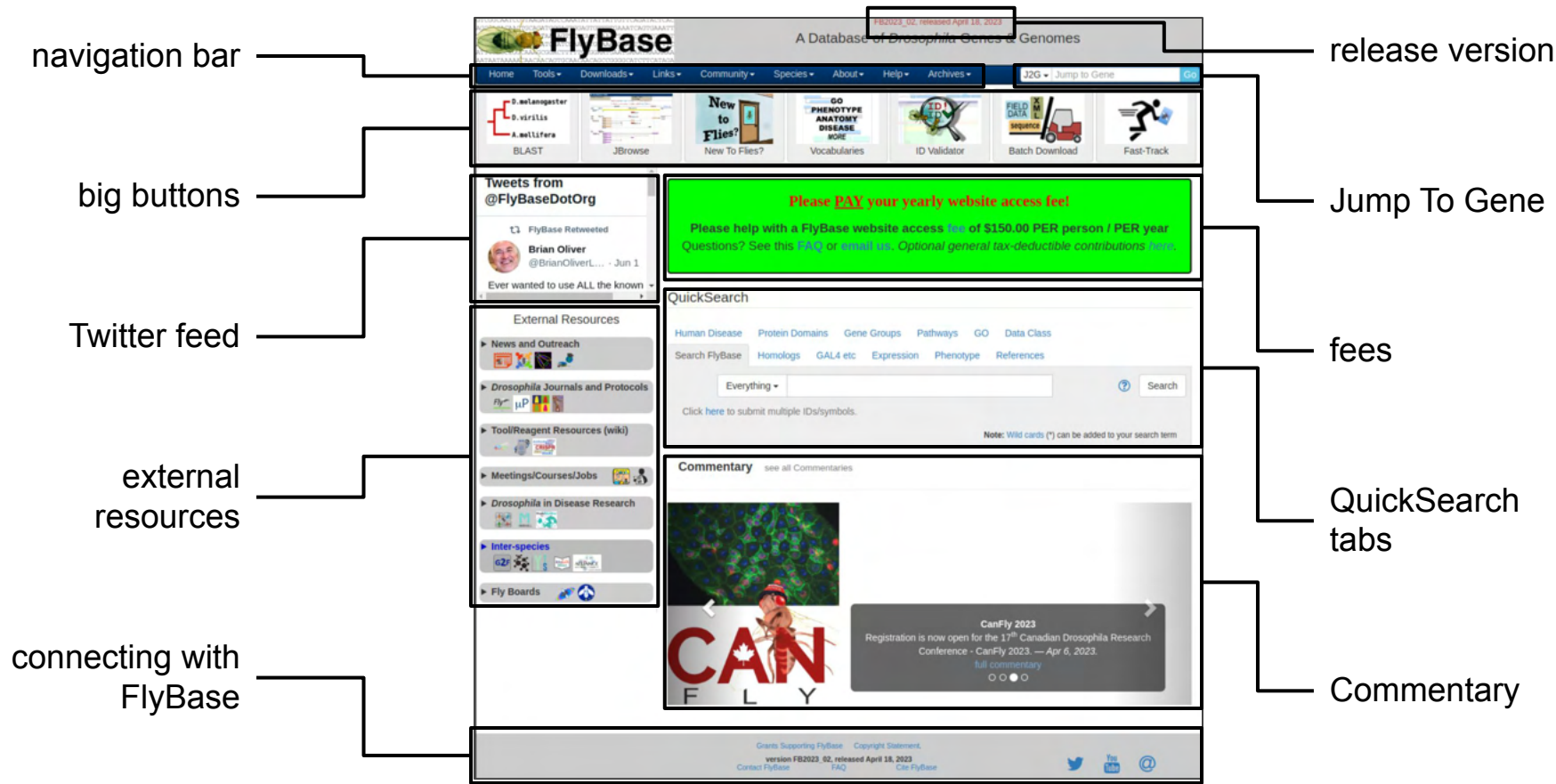


EPITHELIAL CELL:
 cell types + gene expression spread (%) within related cluster :

- 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

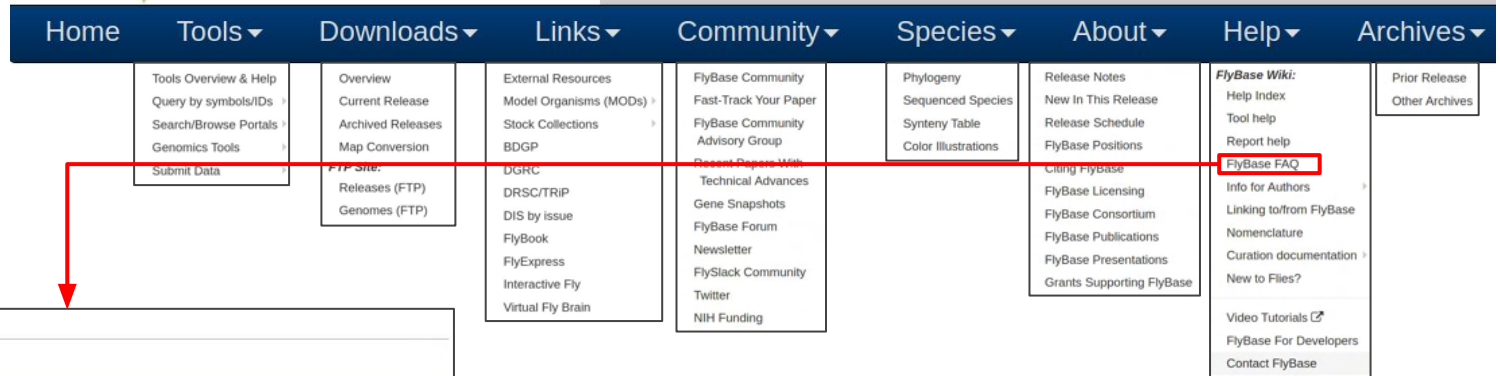
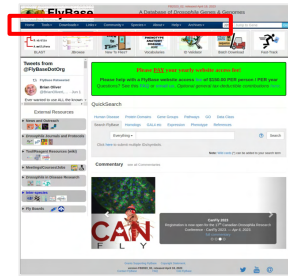


FlyBase homepage overview



Navigation bar

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?

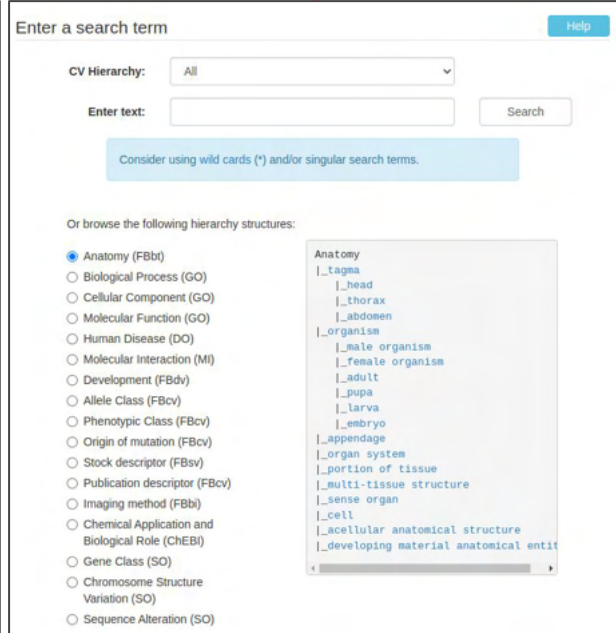
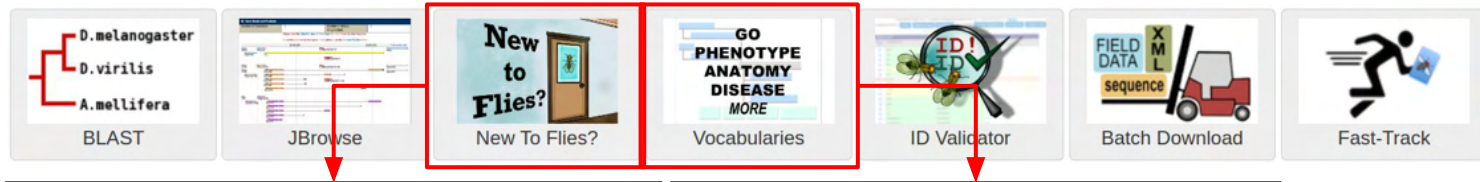
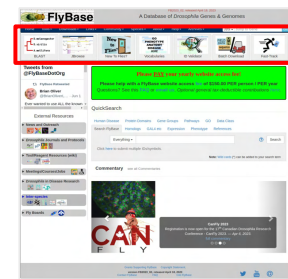
Yes, GBrowse is no longer being maintained, and GBrowse access will be discontinued in FlyBase release FB2202_06. All GBrowse tracks and features are now also available in JBrowse. Please use [JBrowse@](#) for your queries.
 - 7.2. How can I download FASTA sequence from JBrowse?

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).

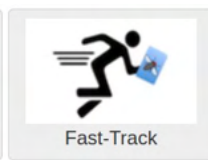
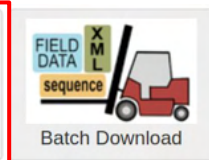
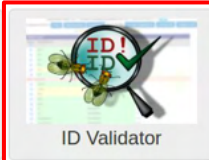
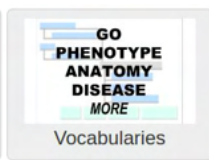
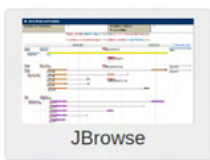
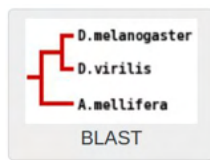
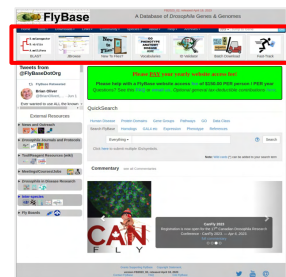
Big buttons: New to Flies

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



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:

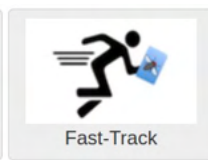
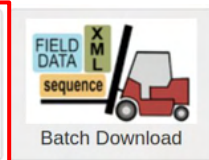
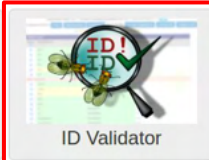
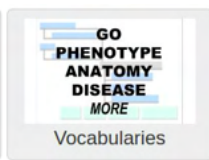
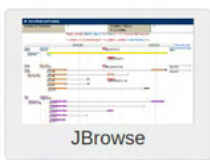
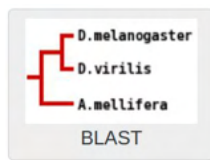
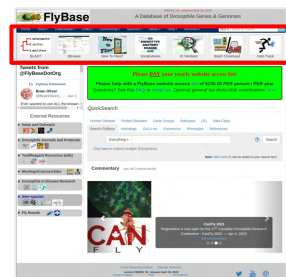
Save as file:

(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



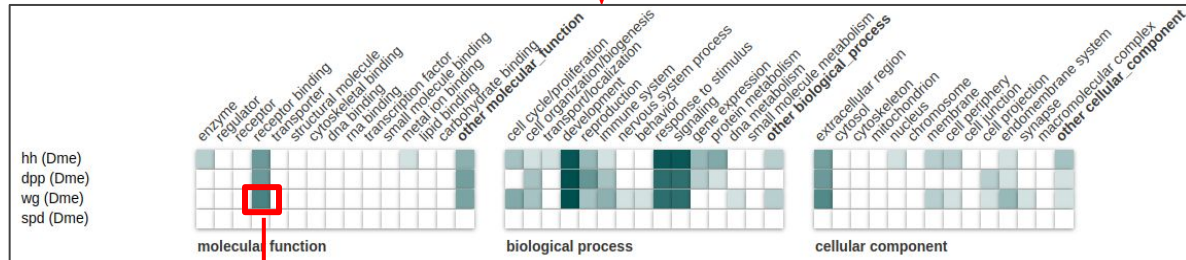
selected items in a FlyBase tool:

- 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)

Filter by species: 4 selected (D. melanogaster, H. sapiens, other Drosophila species, Other species)

Filter by data class: Gene (4), Transgenic Construct (2)

Gene	Feature type	Sequence Location	Gene model status	Cylogenetic Map	References
dpp	protein coding gene	2L:2,428,372..2,459,823 [-]	Current	94E1-94E1	2999
wg	protein coding gene	2L:7,207,159..7,216,265 [-]	Current	27F1-27F1	3470
spd	gene		Unannotated		1

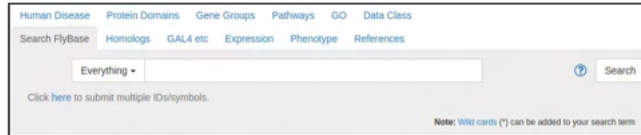


wg : "receptor binding"
16 total annotation(s) using 5 GO term(s) :
5 : frizzled binding
5 : morphogen activity
4 : receptor ligand activity
1 : cytokine activity
1 : signaling receptor binding
Click cell for more information.

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

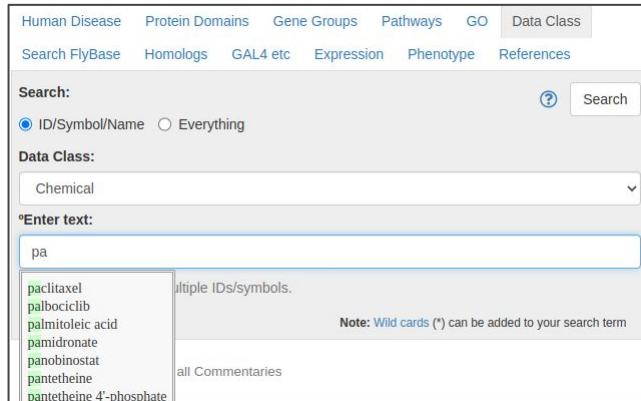
Search FlyBase Homologs GAL4 etc Expression Phenotype References

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: Search

ID/Symbol/Name Everything

Data Class:

Enter text:

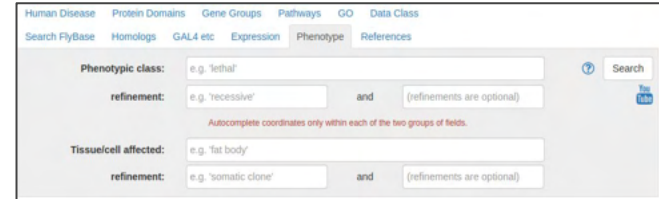
- paclitaxel
- palbociclib
- palmitoleic acid
- pamidronate
- panobinostat
- pantheine
- pantheine 4'-phosphate

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: Search

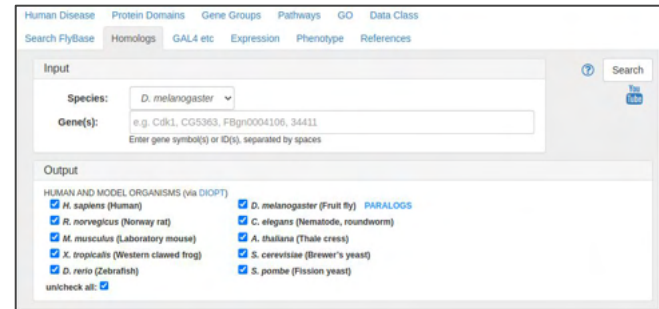
refinement: and

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

Tissue/cell affected:

refinement: and

Search for homologs.



Human Disease Protein Domains Gene Groups Pathways GO Data Class

Search FlyBase Homologs GAL4 etc Expression Phenotype References

Input Search

Species:

Gene(s):

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

Output

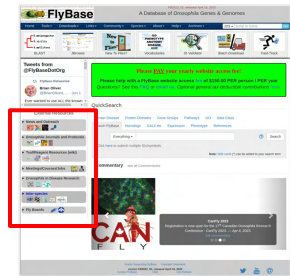
HUMAN AND MODEL ORGANISMS (via DIOPT)

<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	Alliance Auto summary Gene Group		

Allele reports

Type	Location	Additional Notes	References
deletion	3k, 1,247,891..1,249,852 (-)	CG17807 deletion, resulting from the impregne excision of P[GD]50438 on chromosome III of the full-length chromosome.	[Lee et al., 2012]

Cross-species data alignment.



Cross-model organism forum.

Forum post titled "Model Organism: Files". The post discusses the inclusion of files for the Model Organism: Files. The post is dated June 2nd, 2023. The post is part of a discussion about the Alliance of Genome Resources.

Connecting with FlyBase

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

The screenshot shows the FlyBase website interface. The top navigation bar includes: Home, Tools, Downloads, Links, Community, Species, About, Help, and Archives. The 'Community' dropdown menu contains: FlyBase Community, Fast-Track Your Paper, FlyBase Community Advisory Group, Recent Papers With Technical Advances, Gene Snapshots, FlyBase Forum, Newsletter, FlySlack Community, Twitter, and NIH Funding. The 'Help' dropdown menu contains: 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, and Contact FlyBase. The footer contains: Grants Supporting FlyBase, Copyright Statement, version FB2023_02, released April 18, 2023, Cite FlyBase, Contact FlyBase, FAQ, and social media icons for Twitter, YouTube, and a general contact symbol.

helpfb@morgan.harvard.edu

Contact FlyBase

FAQ

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