What's New in FlyBase

527B

Sian Gramates, Gillian Millburn, Josh Goodman, Helen Attrill , Victoria Jenkins, Victor Strelets,
Jim Thurmond, and the FlyBase Consortium

The 'etc' in GAL4 etc

FlyBase has recently redesigned the **GAL4 etc** QuickSearch tab. It is now possible to search for any transgenic driver or reporter, not just the original set of five (GAL4, lexA, QF, lacZ, GFP). We've removed the Driver/Reporter selector from the tab; selection now happens on the results page.

As before, the **GAL4 etc** tab searches for drivers/reporters by temporal-spatial expression pattern. To use this option, fill at least one field with a valid controlled vocabulary tool (FlyBase development or anatomy CV, GO cellular component). Autocomplete guides you in CV term choice. You must use a valid CV term in the search field; this tool does not support use of synonyms. Use the **Vocabularies** tool to find valid CV terms.

We've added a new search option: you can now search for drivers/reporters that have been curated as reflecting the expression pattern of a particular gene, such as *tin*. This won't find drivers/reporters with patterns similar to *tin*, nor will it find *tin* alleles that don't have curated expression data attached.

Searching by gene is offered as an alternative to searching by temporal-spatial expression pattern; you can't do both at once.

Using Experimental Tools to find everything

The hitlist for the GAL4 etc tab displays in a table format, and includes all drivers and reporters that have the temporal-spatial or gene expression pattern you searched for. We've added four new Experimental Tools columns to the hitlist, allowing you to filter by encoded tool (such as **GAL4**), tool use (such as **binary expression system – driver**), tag (such as **Tag:MYC**), or tag use (such as **epitope tag**). Note that the columns are paired, so that the tool use appears next to the tool, giving you a hint of how to filter to find what you want. Every column can be sorted, filtered, moved, or hidden, and you can filter multiple columns at once, allowing you to find exactly the driver or reporter you want in a hitlist that may include thousands of alleles.

Filtering GAL4 etc results with Experimental Tools

This filtered search identifies proteins tagged with yellow fluorescent protein that are localized to larval motor neurons. Filtering the **Tagged with** column with 'YFP' failed to yield a result; however, filtering the **Tag (uses)** column with 'fluorescent' narrows the hitlist, revealing that the YFP tag is named **Venus**.

This filtered search identifies split-GAL4 drivers expressed in the adult mushroom body; the text string 'split' does not appear in **Allele**, **Construct**, or **Encoded tool** symbols, but can be found by filtering the **Encoded tool use(s)** column.

Want to know more? Further details available in the FlyBase Commentary and Tweetorial.

Paralogs

FlyBase has added a new **Paralogs** section to Gene Reports (right under **Orthologs**), reporting *Drosophila melanogaster* paralogs to a given *Drosophila* gene (*ena* paralogs shown here).

Paralogs can be searched using the **Homologs** (previously called **Orthologs**) tab in QuickSearch. Choose *D. melanogaster* as both input and output. We also provide a downloadable file containing *Drosophila melanogaster* paralogs in Current Release **Downloads** (under Homologs).

Revamped genetic reagents lists in Gene Reports

FlyBase has updated the **Alleles, Insertions, and Transgenic Constructs** section of Gene Reports to make it easier to browse and search the genetic reagents that are available for your favorite gene. The update includes new table functionality and the addition of experimental tool data, both of which facilitate finding reagents with the particular characteristics you are interested in.

New columns, containing experimental tool information, have been added. For transgenic constructs, these columns describe the components that make up the construct (*e.g.* regulatory region, encoded product/tool, tags). For alleles caused by an insertion, these columns indicate the type of inserted element (*e.g.* gene trap, misexpression element, enhancer trap), along with any tools encoded by the inserted element

Allele tables can exported to a hitlist, a CSV file, or an Excel file.

Customizing the allele table

Some columns from the previous version of the allele tables are not shown in the default view of the updated table. These columns can be returned to view using the **Show/Hide Columns** button; hidden columns are listed to the left of the button.

As with the **GAL4 etc** hitlist, every column can be sorted and/or filtered, as well as dragged into a different order. Here, the **Stocks** column has been sorted, so that alleles with stocks that have been deposited in a public stock center are sorted to the top of the column; **Stocks** has also been dragged to be the second column. The **Allele class** column has been filtered to display only loss of function alleles. Several columns have been hidden, and the **Mutagen(s)** column has been returned to view.

Further details are available in the FlyBase Commentary and Tweetorial.

Human Disease Model reports: Infectious diseases

Over the past year, FlyBase has increased our coverage of eukaryotic, prokaryotic, and viral pathogenic infections that can be modeled in Drosophila. These reports focus on diseases that infect humans and can either infect Drosophila, or be transmitted by them. This does not include pathogens that infect Drosophila but not humans, such as the Drosophila C virus or the entomopathogenic nematode *S. carpocapsae*.

Unlike most of our Human Disease Models, these reports do not link to a specific causative gene and its ortholog(s). However, they do cover Drosophila genes that are relevant to modeling the infection in the Summary, and provide links to Drosophila genetic reagents that contain genes from pathogens. The reports also connect to available Disease Ontology terms, other Human Disease Models where relevant, and have links to all publications in FlyBase that use the disease model in the References section.

Eukaryotic pathogens	Prokaryotic pathogens	Viral pathogens
Aspergillus species	Escherichia coli	Epstein-Barr virus
Candida species	Listeria monocytogenes	flavivirus (dengue, West Nile, Zika)
	Mycobacterium species	Sindbis virus
	Pseudomonas aeruginosa	
	Serratia marcescens	
	Staphylococcus aureus	
	Vibrio cholerae	

Pathways in FlyBase

Pathway reports list genes that have been experimentally shown to be a member of a pathway or to regulate a pathway, together with supporting research papers. The number of pathway references associated with a gene may also help to weigh the evidence behind its inclusion. Useful links to external resources and publications are provided, as well as options for export and orthology tools.

Fifteen pathways have been curated to date. They can be accessed via the **Pathways** tab QuickSearch tab on the FlyBase homepage.
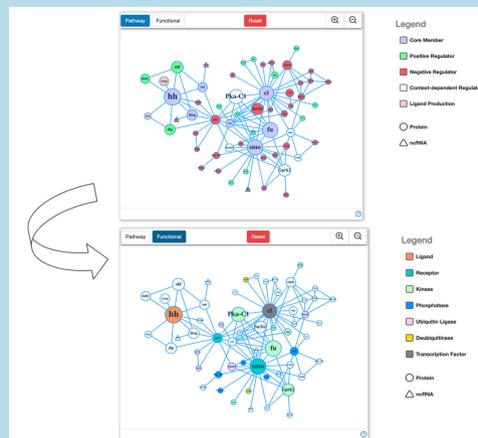
Interactive network diagrams are available for each pathway, generated using experimental data - node size is based on the number of supporting research papers, and the connections from physical interaction data. The networks are interactive, with two viewing options available: a pathway view, which distinguishes regulatory interactors; and a functional view, which colors nodes depending on the functional class of the gene product. More details are available in this Commentary.

Pathway Reports: Members table

Pathway reports provide easy access to linked data and reagents. The members table can be customized to order and display information of interest, such as reagents, functional and disease-relevant information.

Metabolic pathway data

For access to metabolic pathway data, FlyBase now provides **Metabolic Pathway** subsection within the **Pathways** section of the Gene Report. This section includes the names of metabolic pathways in which the gene product is involved, with links to the respective reports at KEGG, Reactome and BioCyc.

FlyBase is supported by a grant from the National Human Genome Research Institute at the U.S. National Institutes of Health #U41 HG000739.
Support is also provided by the British Medical Research Council (#MR/N030117/1) and FlyBase Users all over the world.