FlyBase: A Valuable Source of Molecular Interaction Data

Julie Agapite[1], Gil dos Santos[1], Chris Tabone[1], Phillip Baker[2], Lynn Crosby[1], Kathleen Falls[1], Josh Goodman[3], Andrew Schroeder[1], Victor Strelets[3], Dan Bean[4] and the FlyBase Consortium.

[1]FB-Harvard University, Cambridge, MA, [2]FB-University of New Mexico, Albuquerque, NM, [3]FB-Indiana University, Bloomington, IN, [4]EzyN, University of Cambridge, Cambridge, UK.

Abstract

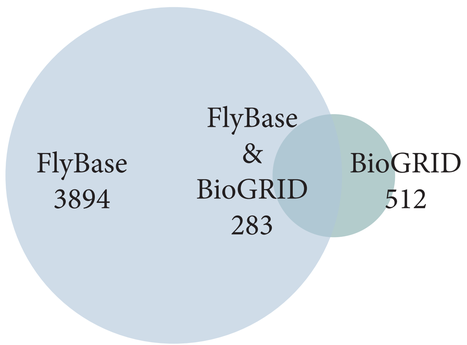
The knowledge of a molecule's binding partners can provide insights into that molecule's function and/or its involvement in a particular biological process. FlyBase curation of molecular interactions is primarily focused on capturing protein-protein, RNA-protein and miRNA-mRNA interactions from low-throughput studies, in which interactions are typically supported by multiple independent forms of evidence. By concentrating on low-throughput studies of *Drosophila* gene products, the FlyBase molecular interactions dataset complements that of interaction databases and consortia that tend to focus on high-throughput studies.

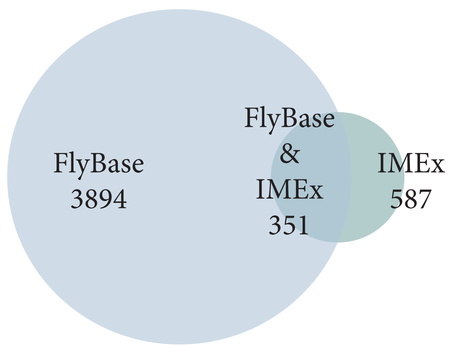
Drosophila Dataset Comparison

	FlyBase	IMEx	BioGRID
total interaction experiments	43,593	50,234	38,900
interspecies	38	508	208
protein-protein	40,900	50,102	na*
protein-RNA	2085	8	na*
RNA-RNA	608	0	na*
protein-DNA	0	104	na*
protein-small molecule	0	18	na*
publications	3894	587	512
unique gene pairs	29498	39628	36801

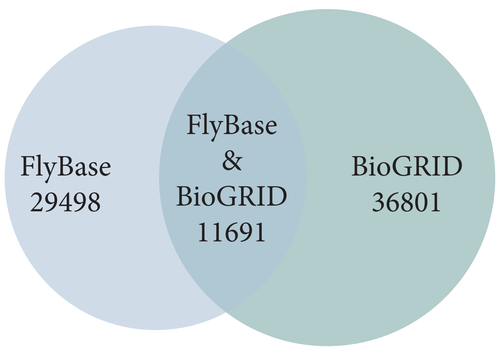
*data not available

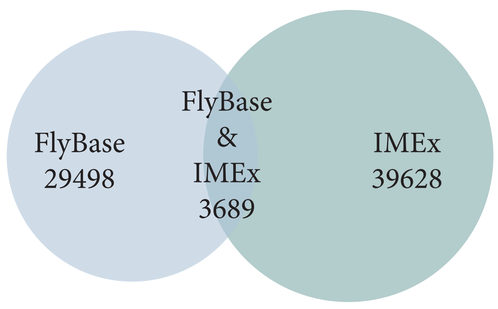
Publications Curated



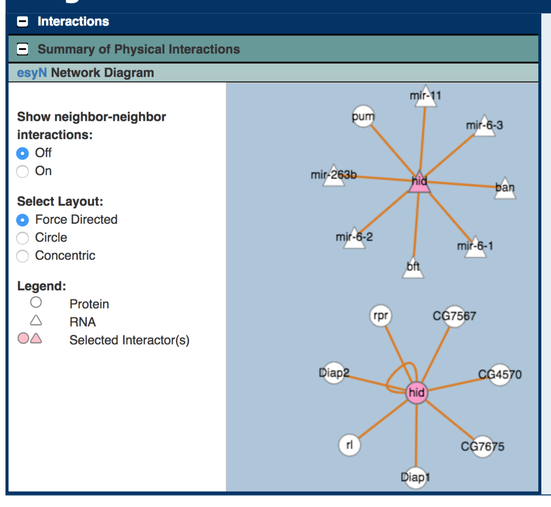


Unique Gene Pairs





FlyBase Interactions Display



FlyBase curated interaction data can be viewed in the "Interactions" section of every FlyBase gene page. The data is summarized as a network diagram where circles represent protein interactors and triangles represent RNA making it possible to distinguish protein-protein, protein-RNA and RNA-RNA interactions at a glance. Network diagrams are generated by EsyN (http://www.esyn.org).

FlyBase Interactions Display

RNA-RNA

Interacting group	Assay	References
hid - ban	immunohistochemistry, luminiscence technology, fluorescence technology, necessary binding region	(*Shen et al., 2015*, *Kertesz et al., 2007*, *Leaman et al., 2005*, *Robins et al., 2005*, *Brennecke et al., 2003*)
hid - bft	quantitative reverse transcription pcr, luminiscence technology, necessary binding region	(*Hilgers et al., 2010*)
hid - mir-11	quantitative reverse transcription pcr, luminiscence technology	(*Truscott et al., 2011*)
hid - mir-263b	luminiscence technology, necessary binding region	(*Hilgers et al., 2010*)
hid - mir-6-1	fluorescence technology, luminiscence technology	(*Leaman et al., 2005*)
hid - mir-6-2	luminiscence technology, fluorescence technology	(*Leaman et al., 2005*)
hid - mir-6-3	fluorescence technology, luminiscence technology	(*Leaman et al., 2005*)

RNA-protein

Interacting group	Assay	References
hid - pum	electrophoretic mobility shift assay, autoradiography	(*Bhogal et al., 2016*)

protein-protein

Interacting group	Assay	References
hid	cross-linking study, molecular weight estimation by staining	(*Sandu et al., 2010*)
hid - CG4570	two hybrid	(*Hu et al., 2017.6.13*)
hid - CG7567	two hybrid	(*Hu et al., 2017.6.13*)
hid - CG7675	two hybrid	(*Hu et al., 2017.6.13*)
hid - Diap1	pull down, anti tag western blot, anti tag coimmunoprecipitation, molecular weight estimation by staining, autoradiography, western blot, far western blotting, colocalization, fluorescence microscopy, inferred by author	(*Morishita et al., 2013*, *Sandu et al., 2010*, *Leulier et al., 2006*, *Ryoo et al., 2002*)

FlyBase also presents interaction data in tabular form where individual interaction experiments are condensed into unique gene pairs, or interacting groups. For each gene pair, assays used to detect interactions and references are listed.

Clicking on a particular interacting group or gene pair opens a new page with more detailed information. The interacting group page displays a network diagram that includes products of both genes and common interactors are highlighted in purple.

General Information

Interaction Type	physical association, colocalization	Interacting Genes	Diap1, hid
FlyBase ID	FBig0000100632		

Interaction Network

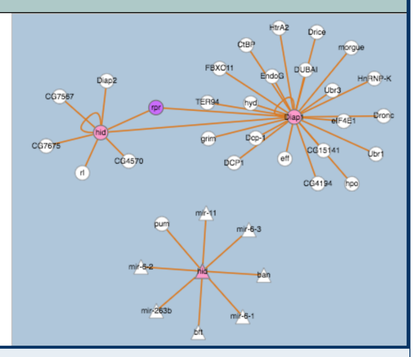
Interactions Browser links	Diap1 network hid network

esyN Network Diagram

Show neighbor-neighbor interactions:
Off
On

Select Layout:
Force Directed
Circle
Concentric

Legend:
Protein
RNA
Selected Interactor(s)
Common Interactor(s)



Reported Interactions

FBrf0149132-2.PD.A

Reference	*Ryoo et al., 2002*
Description	physical association
Assay	pull down autoradiography
Collection	
Source/Stage	
Cell line used	

Participants

Corresponds to	Reported as	Role	Note
hid	Hid	bait	GST tag
Diap1	DIAP1	prey	[35]S label

Experimental entities

Corresponds to	Identifier	Reported	Role	Note

Subregions with role in interaction

Corresponds to	Description	Role	Coordinates	Note
hid	N-terminal region	sufficient binding region	aa 1-37	

Isoform-specific participants

Corresponds to	Description	Role	Note

Comments concerning this interaction

	Interaction in vitro; bait produced as a recombinant fusion protein in bacterial system; prey produced and labeled by in vitro translation.	(*Ryoo et al., 2002*)
	Positive control.	(*Ryoo et al., 2002*)

The interacting group page further presents details for each experiment showing an interaction involving products of the given gene pair. Details include the sources of the molecules, their bait or prey roles, portions of the molecules that are necessary or sufficient for the interaction or mutations that affect the interaction.

External Data

External Data

Subunit Structure (UniProtKB)	Interacts with Diap2 (via BIR2 and BIR3 domains). (UniProt, Q24106)
Linkouts	**BioGRID** - A database of protein and genetic interactions. 65293 **DroID** - A comprehensive database of gene and protein interactions. FBgn0003997 **InterologFinder** - Protein-protein interactions (PPI) from both known and predicted PPI data sets. 40009 **MIST (genetic)** - An integrated Molecular Interaction Database hid **MIST (protein-protein)** - An integrated Molecular Interaction Database hid

General Information

Symbol	DmelNhid
Name	head involution defective
Feature Type	protein_coding_gene
Gene Model Status	Current
Gene Snapshot	In progress. Contributions welcome.
Other Summaries	Auto summary Gene Group
Also Known As	W, l(3)05014
Key Links	ALLIANCE of GENOME RESOURCES NCBI Gene

FlyBase shares the interaction data it curates with other databases likes BioGRID, MIST, DroID and the Alliance of Genome Resources. Linkouts to these databases from FlyBase allow users to quickly explore large and diverse datasets.

The complete corpus of FlyBase interaction data is available for download under the 'Downloads' menu.

Physical Interaction table	TSV	physical_interactions_fb_2019_01.tsv.gz,
	MITAB TSV	physical_interactions_mitab_fb_2019_01.tsv.gz

Conclusion

FlyBase curates interaction data that complements that captured by other databases making FlyBase a valuable source of *Drosophila* interaction data.