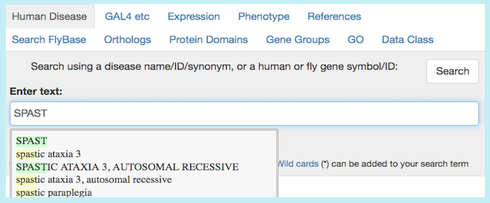
Finding human disease models in FlyBase: You can get there from here

817

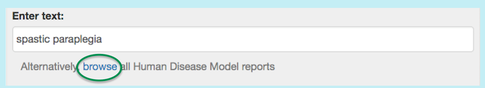
Sian Gramates, Lynn Crosby, Josh Goodman, Steven Marygold, Victor Strelets, Jim Thurmond, and the FlyBase Consortium

Many human diseases have been modeled in *Drosophila*, and a large body of literature has accumulated in recent years. FlyBase has been curating disease models using Disease Ontology (DO) annotation and Human Disease Model reports. We provide multiple methods to access disease model data in FlyBase, including a dedicated 'Human Disease' QuickSearch tab, DO term reports, and disease model information embedded in gene and allele reports. We have organized disease model information in a highly interconnected way, so that a user who has landed on any such information can easily navigate to other related information.

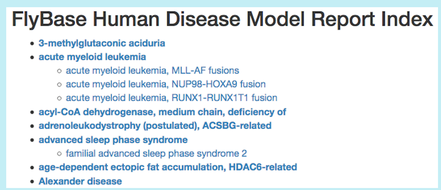
The Human Disease QuickSearch Tab



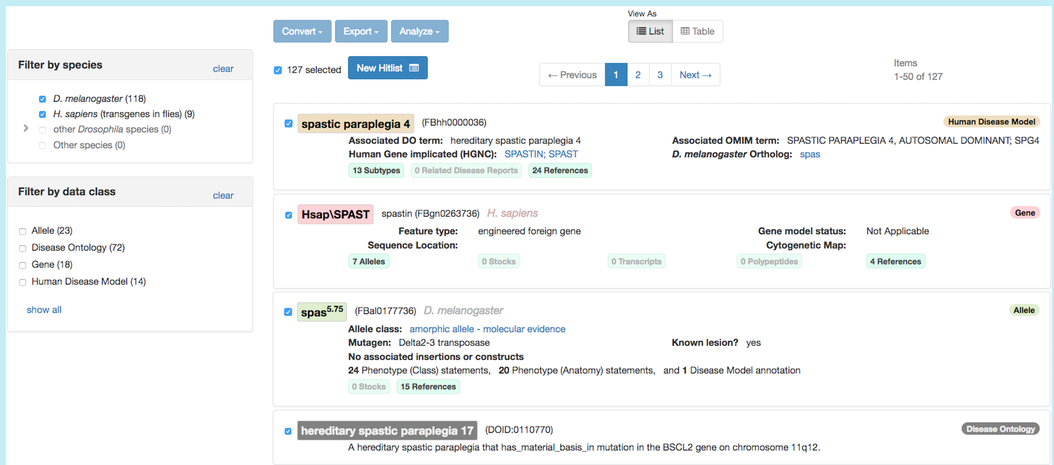
By using the **Human Disease** tab, it is possible to search all FlyBase human disease model data using almost any disease-related search term, including DO terms, Human Disease Models, OMIM phenotypes or genotypes, HGNC symbols, FlyBase genes or alleles, or ID numbers for any of these terms. Autocomplete is enabled for all of these search options; as shown in the example above, the string "SPAST" simultaneously triggers autocompletion of gene symbols, allele symbols, OMIM diseases terms, DO terms, nd Human Disease Model names.



The **Human Disease** tab also includes a link to a browsable index of all Human Disease Model reports; many disease model reports are listed redundantly in the index, , allowing a user to browse to a disease from multiple points. Diseases may be listed as a specific subtype of a disease, by mechanistic cause, by symptomatic group, or as part of a major disease classification.



Human Disease Hit-lists



The Human Disease hit-list features four classes of data: Human Disease Models, Disease Ontology terms, genes, and alleles. This hit-list resulted from the search "spastic paraplegia", which includes the hits the Human Disease Model spastic paraplegia 4, the gene Hsap\SPAST associated with that model, the allele spas$^{5.7}$, and the DO term hereditary spastic paraplegia 17.

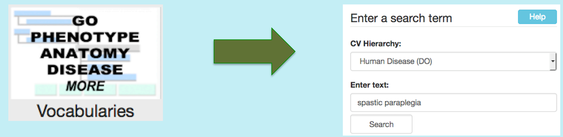
Like all FlyBase faceted hit-lists, a user can choose to display a subset of data classes. Filtering the hit-list to a single data class allows you to further convert the results to another data class, export to a file or to another FlyBase tool, analyze the results in a data class specific manner, or display the hit-list as a table.

Human Disease Model Results

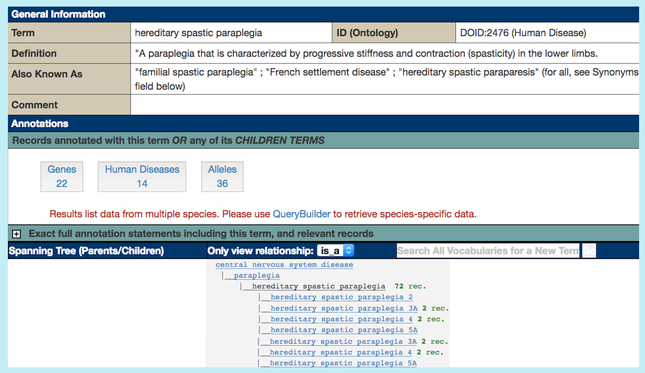
	Name	DO Term	OMIM ID	*D. melanoga…* gene	Human Genes Implicated (HGNC)	# Subtypes	# of Related Disease Terms	# Refs
☑	spastic paraplegia 4	hereditary spastic paraplegia 4	182601	spas	SPASTIN; SPAST	13	0	24
☑	spastic paraplegia 7	hereditary spastic paraplegia 7	607259	Spg7	SPG7 GENE; SPG7	13	0	4
☑	spastic paraplegia 10	hereditary spastic paraplegia 10	604187	Khc	KINESIN FAMILY MEMBER 5A; KIF5A	13	1	9
☑	spastic paraplegia 12	hereditary spastic paraplegia 12	604805	Rtnl2 Rtnl1	RETICULON 2; RTN2	13	0	9
☑	spastic paraplegia 20	Troyer syndrome	275900	spartin	SPARTIN; SPART	13	0	8

This hit-list, resulting from the search "spastic paraplegia", has been filtered to display only Human Disease Models. The table view for this data class highlights the relationships between Human Disease Models, DO terms, and OMIM phenotypes, and displays the associated human disease gene and its orthologous *Drosophila melanogaster* gene.

Vocabularies: Another Path to Disease Models



The **Vocabularies** tool can be accessed using the button on the FlyBase home page, or from the Tools drop-down menu in the FlyBase toolbar. Select the **Human Disease (DO)** CV Hierarchy and enter a search term; autocomplete is enabled for DO terms.

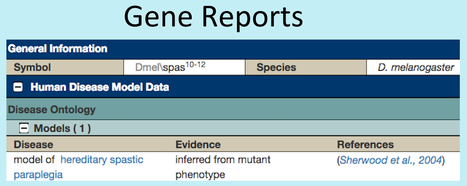


The Disease Ontology term report provides further access to disease model data. The spanning tree allows the user to browse the DO hierarchy, which displays both less specific parental disease categories, and more specific child disease terms. The buttons above the spanning tree lead to hit-lists of genes associated with, Human Disease Model Reports linked to, and alleles annotated with the DO term or its children. The hit-list below includes alleles annotated with the DO term "hereditary spastic paraplegia", or one of its children.

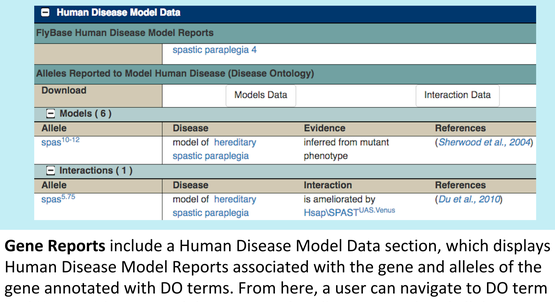
Allele Results

	Symbol	Class	Inserted Elements	# Stocks	Mutagens	Known Lesion?
☑	Arl6IP1^{KK109154}			1	in vitro construct	yes
☑	spas$^{dsRNA.UAS}$			0	in vitro construct	yes
☑	futsch^{EP1419}		P{EP}futsch^{EP1419}	1	P-element activity	yes

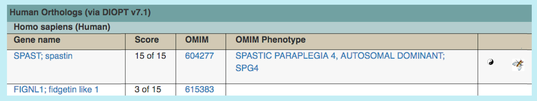
Disease Information in Allele and Gene Reports



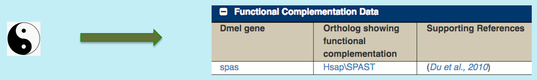
Alleles associated with disease models are annotated with Disease Ontology terms. **Allele reports** link to Disease Ontology term reports.



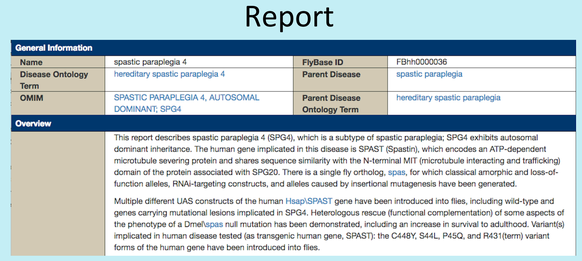
Gene Reports include a Human Disease Model Data section, which displays Human Disease Model Reports associated with the gene and alleles of the gene annotated with DO terms. From here, a user can navigate to DO term and Human Disease Model reports, and to disease-associated alleles.



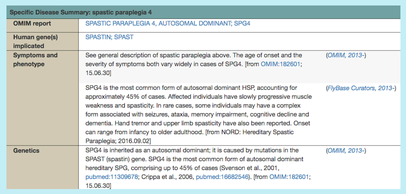
This section also includes DIOPT-identified human orthologous genes. A user can navigate from here to HGNC gene reports, and to OMIM genotype and phenotype reports. The icons in the right column indicate whether the human gene has been expressed in flies, and whether it functionally complements mutations in the *Drosophila* gene.



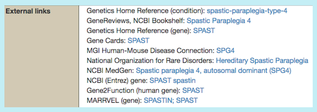
The Integrated Human Disease Model Report



Human Disease Model Reports integrate disease information from many parts of FlyBase. Some, like this one, focus on a single disease, while others are parent reports, describing the general characteristics of groups of related but genetically distinct diseases. We've recently added fly model overview reports that collect methods, fly biology, and reviews pertaining to research in broad classes of diseases, such as hematologic cancer or kidney disease.

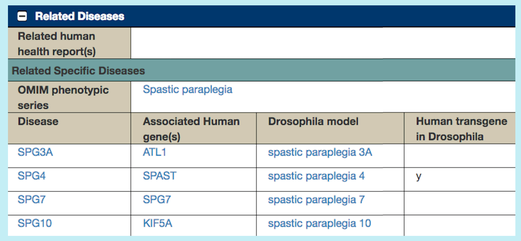


The Disease Summary includes background information, drawn primarily from OMIM (Online Mendelian Inheritance in Man) and research papers. This section includes links to OMIM genotype and phenotype reports, as well as to references.

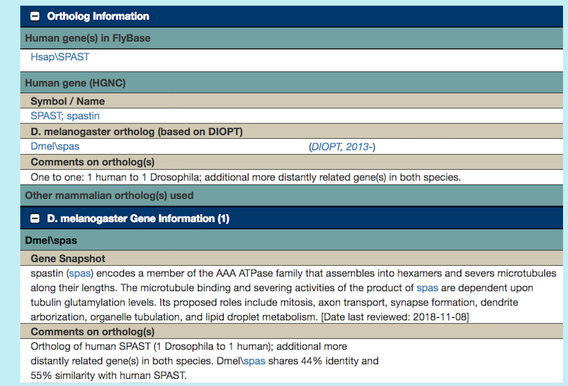


The Disease Summary also includes a collection of external links to other information resources relevant to the disease or its causative gene.

Related Diseases and Orthology

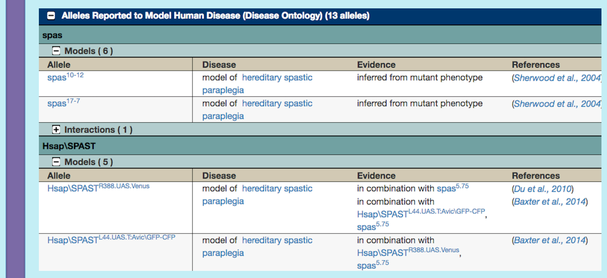


The Related Diseases section allows for easy navigation to Disease Model reports for disease subtypes or related diseases, as well as to OMIM phenotype, genotype, and phenotypic series reports.

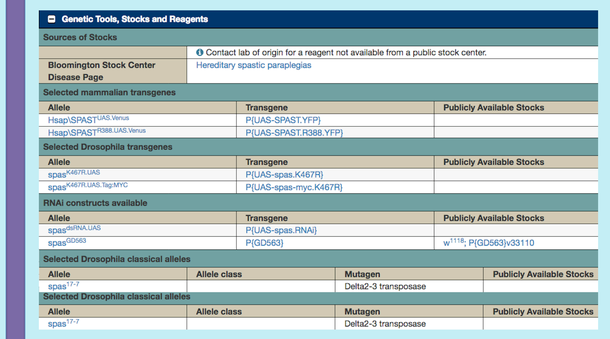


In the Ortholog Information section we associate the human gene identified as the cause of the disease, whether it has been transgenically expressed in flies, and its orthologous fly gene. These gene associations allow us to include data from elsewhere in FlyBase to compute other parts of the report. Links in this section allow the user to navigate to FlyBase and HGNC (HUGO Gene Nomenclature Committee) gene reports.

Alleles and Reagents



The Human Disease Model Report displays DO annotations of alleles of both the transgenically expressed human gene, and orthologous *Drosophila* genes. Links allow the user to navigate to allele reports and DO term reports.



The Genetic Tools, Stocks and Reagents section includes links to allele and transgenic construct reports, to the relevant BDSC disease page, and to stock reports for those available from a public stock repository.

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.