Function-centered approaches for finding and analyzing genes within FlyBase

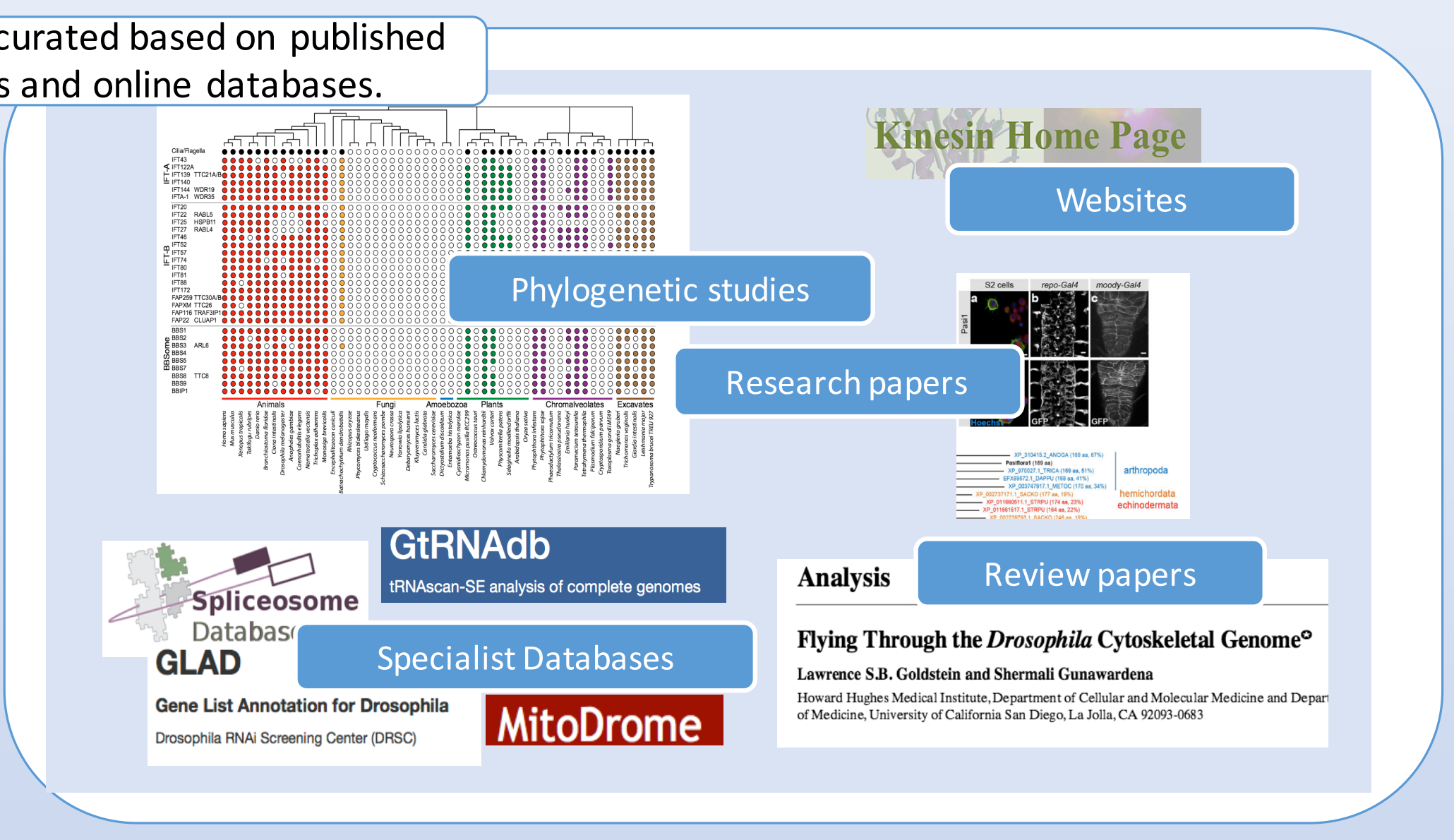
Helen Attrill[1], Giulia Antonazzo[1], Phani Garapati[1], Joshua L. Goodman[2], Steven J. Marygold[1], Alix J. Rey[1], Victor Strelets[2], Jim Thurmond[2] and the FlyBase Consortium

1. Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, CB2 3DY, UK. 2. Dept. of Biology, Indiana University, Bloomington, IN 47405, USA. E-mail H.Attrill: hla28@cam.ac.uk

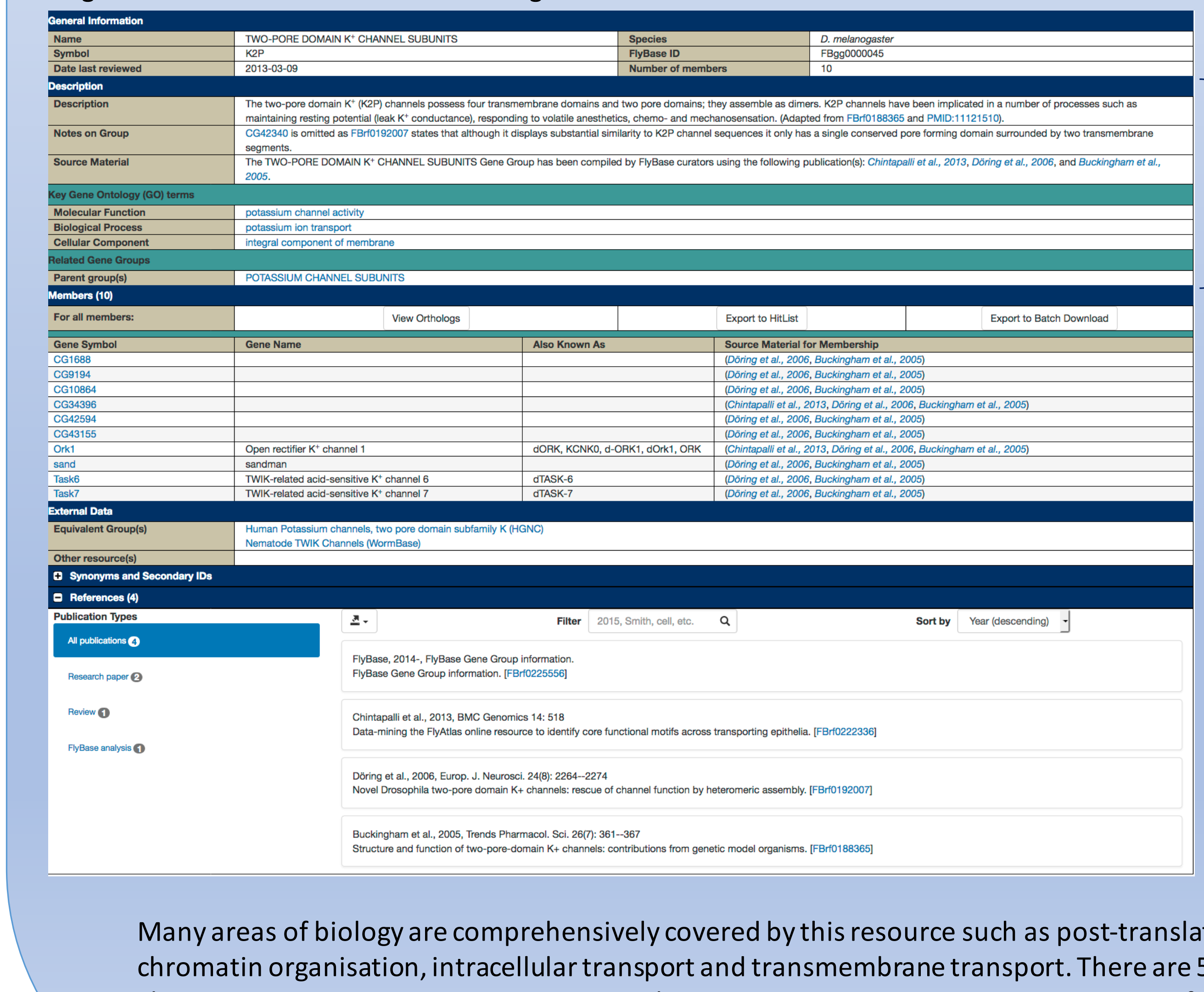
Introduction: It is often desirable to search for and view groups of genes whose products are related in some way, such as their known or predicted function. A list of functionally related genes may provide the starting point for a genetic/molecular screen, or be the basis for in silico analyses using associated data (phenotypes, reagents, genomic data etc.), or allow comparison with equivalent gene sets in other species. FlyBase provides two main ways to search for functionally related *Drosophila melanogaster* genes: via Gene Ontology (GO) annotations and our Gene Group resource. Here we show describe how we compile the Gene Groups resource, assign GO annotations to genes and how these two approaches can be used to find functionally-related genes.

Gene Groups: The FlyBase Gene Group resource is a collection of sets of genes with shared attributes. These groups are primarily focused on well-defined, easily delimited gene sets, such as evolutionary-related gene families (e.g. actins, Wnts), subunits of macromolecular complexes (e.g. ribosome subunits) and sets of genes whose products share a common molecular function (e.g. ubiquitin ligases).

Gene Groups are manually curated based on published research papers, reviews and online databases.


Kinesin Home Page
Websites
Phylogenetic studies
Research papers
GtRNAdb
tRNAscan-SE analysis of complete genomes
Spliceosome Database
GLAD
Gene List Annotation for Drosophila
Drosophila RNAi Screening Center (DRSC)
Specialist Databases
MitoDrome
Analysis
Review papers
Flying Through the Drosophila Cytoskeletal Genome
Lawrence S.B. Goldstein and Shermali Gunawardena
Howard Hughes Medical Institute, Department of Cellular and Molecular Medicine and Department of Medicine, University of California San Diego, La Jolla, CA 92093-0683

Gene Group Reports: Gene Groups are presented in FlyBase report pages. Gene Group reports contain the list of member genes together with additional information organized into sections.



General Information

Name	TWO-PORE DOMAIN K⁺ CHANNEL SUBUNITS	Species	*D. melanogaster*
Symbol	K2P	FlyBase ID	FBgg0000045
Date last reviewed	2013-03-09	Number of members	10

Description

Description	The two-pore domain K⁺ (K2P) channels possess four transmembrane domains and two pore domains; they assemble as dimers. K2P channels have been implicated in a number of processes such as maintaining resting potential (leak K⁺ conductance), responding to volatile anesthetics, chemo- and mechanosensation. (Adapted from FBrf0188365 and PMID:11121510).
Notes on Group	CG42340 is omitted as FBrf0192007 states that although it displays substantial similarity to K2P channel sequences it only has a single conserved pore forming domain surrounded by two transmembrane segments.
Source Material	The TWO-PORE DOMAIN K⁺ CHANNEL SUBUNITS Gene Group has been compiled by FlyBase curators using the following publication(s): *Chintapalli et al., 2013*, *Döring et al., 2006*, and *Buckingham et al., 2005*.

Key Gene Ontology (GO) terms

Molecular Function	potassium channel activity
Biological Process	potassium ion transport
Cellular Component	integral component of membrane

Related Gene Groups

Parent group(s)	POTASSIUM CHANNEL SUBUNITS

Members (10)

For all members: View Orthologs | Export to HitList | Export to Batch Download

Gene Symbol	Gene Name	Also Known As	Source Material for Membership
CG1688			(*Döring et al., 2006*, *Buckingham et al., 2005*)
CG9194			(*Döring et al., 2006*, *Buckingham et al., 2005*)
CG10864			(*Döring et al., 2006*, *Buckingham et al., 2005*)
CG34396			(*Chintapalli et al., 2013*, *Döring et al., 2006*, *Buckingham et al., 2005*)
CG42594			(*Döring et al., 2006*, *Buckingham et al., 2005*)
CG43155			(*Döring et al., 2006*, *Buckingham et al., 2005*)
Ork1	Open rectifier K⁺ channel 1	dORK, KCNK0, d-ORK1, dOrk1, ORK	(*Chintapalli et al., 2013*, *Döring et al., 2006*, *Buckingham et al., 2005*)
sand	sandman		(*Döring et al., 2006*, *Buckingham et al., 2005*)
Task6	TWIK-related acid-sensitive K⁺ channel 6	dTASK-6	(*Döring et al., 2006*, *Buckingham et al., 2005*)
Task7	TWIK-related acid-sensitive K⁺ channel 7	dTASK-7	(*Döring et al., 2006*, *Buckingham et al., 2005*)

External Data

Equivalent Group(s)	Human Potassium channels, two pore domain subfamily K (HGNC) Nematode TWIK Channels (WormBase)
Other resource(s)	

⊞ Synonyms and Secondary IDs

⊟ References (4)

Publication Types: All publications 4 · Research paper 2 · Review 1 · FlyBase analysis 1

Filter: 2015, Smith, cell, etc. · Sort by: Year (descending)

FlyBase, 2014-, FlyBase Gene Group information.
FlyBase Gene Group information. [FBrf0225556]

Chintapalli et al., 2013, BMC Genomics 14: 518
Data-mining the FlyAtlas online resource to identify core functional motifs across transporting epithelia. [FBrf0222336]

Döring et al., 2006, Europ. J. Neurosci. 24(8): 2264--2274
Novel Drosophila two-pore domain K+ channels: rescue of channel function by heteromeric assembly. [FBrf0192007]

Buckingham et al., 2005, Trends Pharmacol. Sci. 26(7): 361--367
Structure and function of two-pore-domain K+ channels: contributions from genetic model organisms. [FBrf0188365]

The 'Description' section gives an overview of the criteria used to compile a group. The 'Notes on Group' may contain justification for the inclusion or exclusion of particular genes. The 'Key Gene Ontology (GO) terms' – are those GO terms (or their children) that are associated with most/all of the member genes and are typical of that group.

The 'Members' table, displays all the genes within the group, frequently used synonyms and the source material for membership of each individual gene within a group. Buttons at the top of this section allow the gene list to be downloaded or exported for analysis.

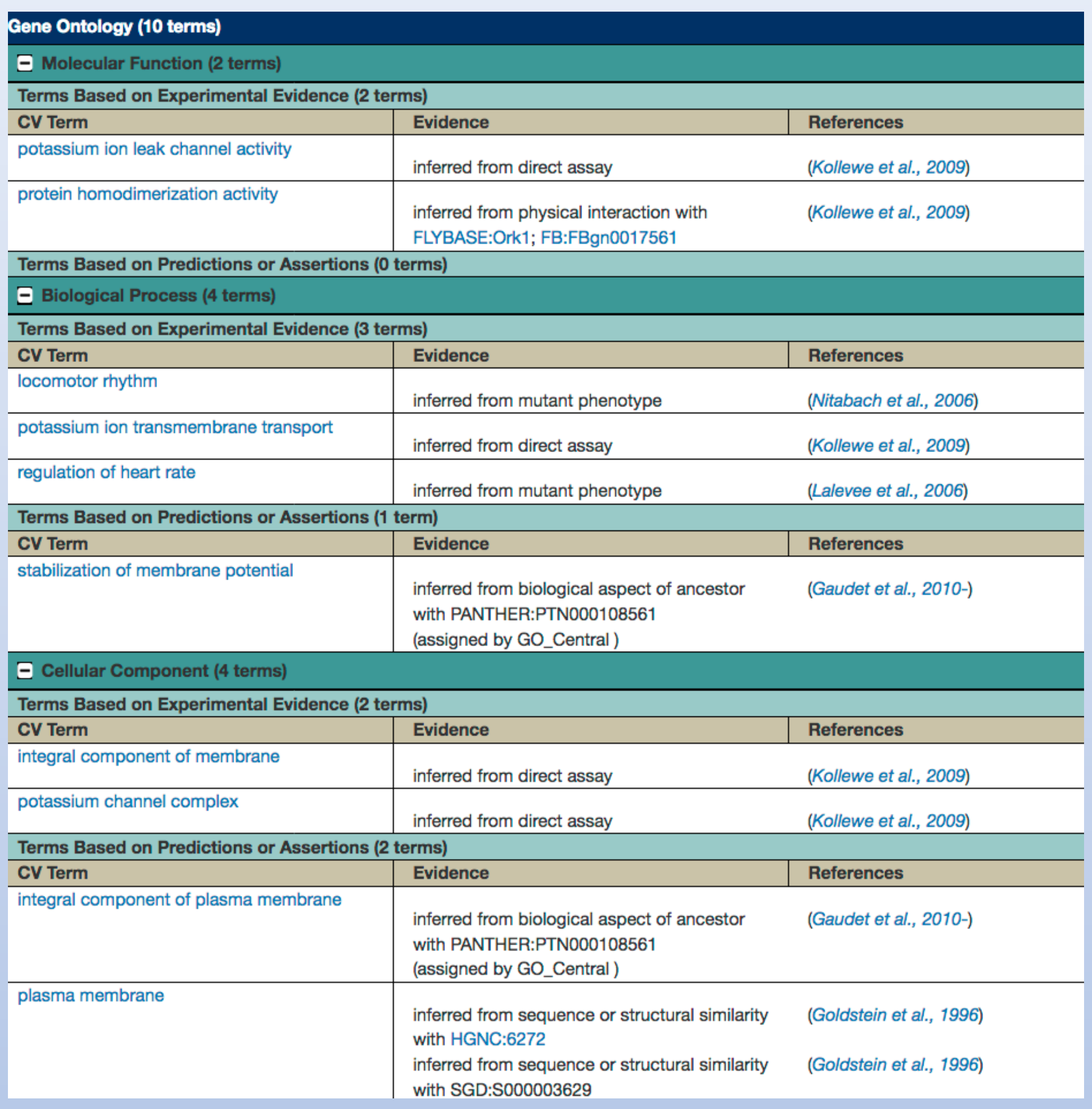
The 'External Data' section provides links to equivalent groups in other databases, facilitating navigation between different species datasets.

All sources used to compile a Gene Group are displayed at the bottom of the page and, for individual genes, in the 'Members' table.

Many areas of biology are comprehensively covered by this resource such as post-translational modification, chromatin organisation, intracellular transport and transmembrane transport. There are 538 gene groups in FlyBase, representing 4487 unique genes. This growing resource now covers over 25% of all of sequence-localized genes and 30% of protein coding genes of the *D. melanogaster* genome.

Gene Ontology (GO) Annotation: The GO is a widely used controlled vocabulary used to label gene products with biological attributes. The GO is arranged in a hierarchical structure, with more specific child terms nested under higher-level parent terms. For example, 'protein kinase activity' is a child of 'kinase activity'. GO annotations are displayed on individual gene pages in FlyBase.

GO annotation for Ork1 (Open rectifier K⁺ channel 1, FBgn0017561)



Gene Ontology (10 terms)

⊟ Molecular Function (2 terms)

Terms Based on Experimental Evidence (2 terms)

CV Term	Evidence	References
potassium ion leak channel activity	inferred from direct assay	(*Kollewe et al., 2009*)
protein homodimerization activity	inferred from physical interaction with FLYBASE:Ork1; FB:FBgn0017561	(*Kollewe et al., 2009*)

Terms Based on Predictions or Assertions (0 terms)

⊟ Biological Process (4 terms)

Terms Based on Experimental Evidence (3 terms)

CV Term	Evidence	References
locomotor rhythm	inferred from mutant phenotype	(*Nitabach et al., 2006*)
potassium ion transmembrane transport	inferred from direct assay	(*Kollewe et al., 2009*)
regulation of heart rate	inferred from mutant phenotype	(*Lalevee et al., 2006*)

Terms Based on Predictions or Assertions (1 term)

CV Term	Evidence	References
stabilization of membrane potential	inferred from biological aspect of ancestor with PANTHER:PTN000108561 (assigned by GO_Central)	(*Gaudet et al., 2010-*)

⊟ Cellular Component (4 terms)

Terms Based on Experimental Evidence (2 terms)

CV Term	Evidence	References
integral component of membrane	inferred from direct assay	(*Kollewe et al., 2009*)
potassium channel complex	inferred from direct assay	(*Kollewe et al., 2009*)

Terms Based on Predictions or Assertions (2 terms)

CV Term	Evidence	References
integral component of plasma membrane	inferred from biological aspect of ancestor with PANTHER:PTN000108561 (assigned by GO_Central)	(*Gaudet et al., 2010-*)
plasma membrane	inferred from sequence or structural similarity with HGNC:6272 inferred from sequence or structural similarity with SGD:S000003629	(*Goldstein et al., 1996*) (*Goldstein et al., 1996*)

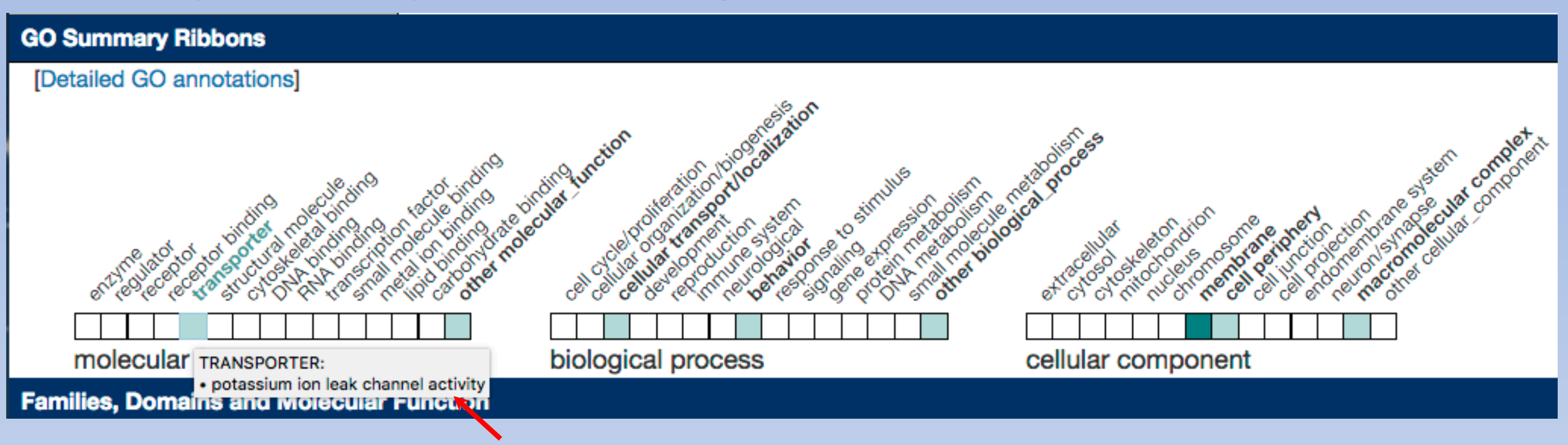
GO annotation
Within the gene report, GO annotations are split into the three aspects: **Molecular Function**, **Biological Process** and **Cellular Component**. These are further subdivided into annotations that have been inferred from **experimental** observations and those that have been inferred from **predictions** or **assertions** made by curators or from automated pipelines.

The combinatorial approach of manual and electronic annotation results in good coverage of GO annotation data over the *D. melanogaster* genome: 73% of sequence-localized genes and 88% of protein-coding genes have associated GO terms.

GO ribbons - a graphical summary of gene function

In the new version of FlyBase, currently available as a beta release, GO summary ribbons are shown at the top of each gene page. The GO summary ribbons use the hierarchical structure of the ontology to group terms under generalized, high-level categories. The detailed annotation can be found in the 'Gene Ontology' section.

The GO summary ribbon for Ork1 (Open rectifier K⁺ channel 1, FBgn0017561)

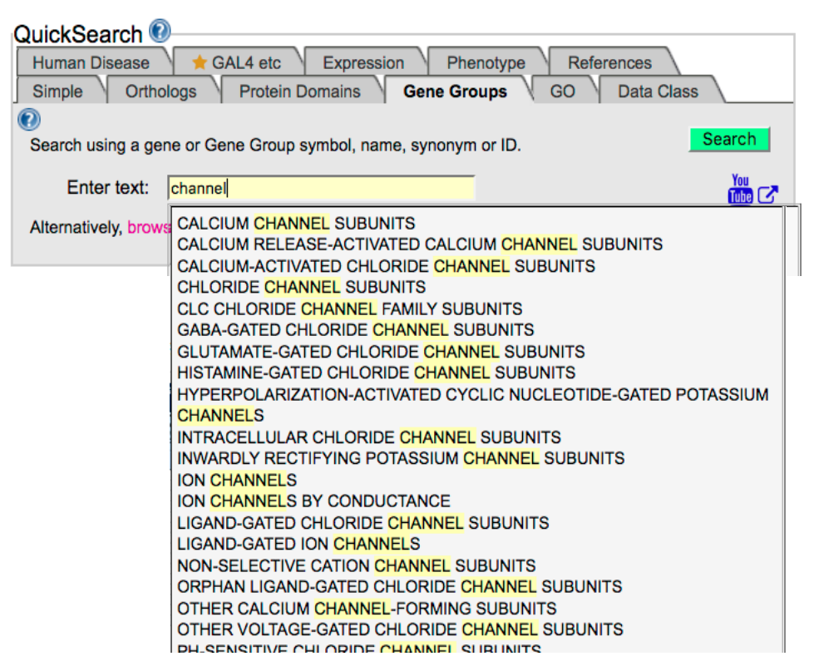


Mousing over cell reveals the term(s) represented.

Finding Gene Groups

Gene Groups can be accessed in three ways:

1. From the FlyBase homepage via QuickSearch

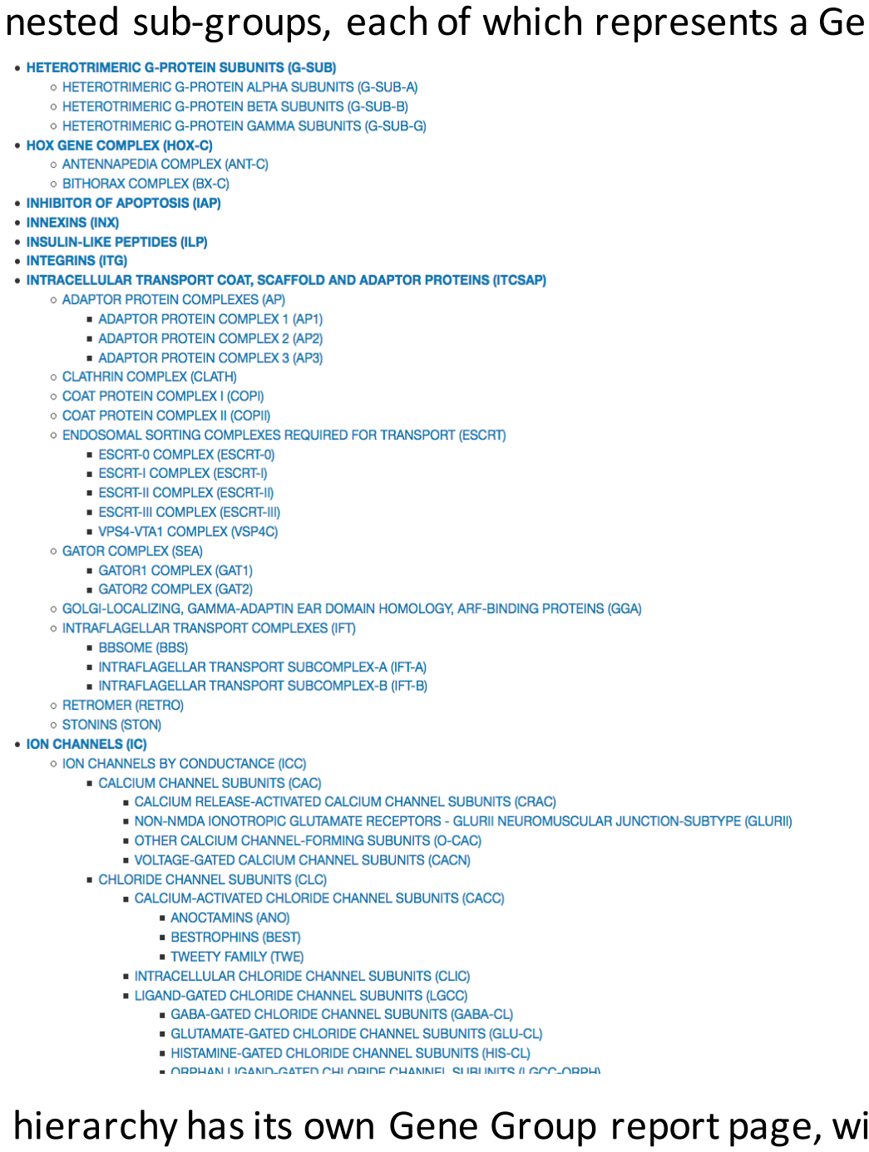


Search using a gene or Gene Group symbol, name, synonym or ID.

Enter text: channel

CALCIUM CHANNEL SUBUNITS
CALCIUM RELEASE-ACTIVATED CALCIUM CHANNEL SUBUNITS
CALCIUM-ACTIVATED CHLORIDE CHANNEL SUBUNITS
CHLORIDE CHANNEL SUBUNITS
CLC CHLORIDE CHANNEL FAMILY SUBUNITS
GABA-GATED CHLORIDE CHANNEL SUBUNITS
GLUTAMATE-GATED CHLORIDE CHANNEL SUBUNITS
HISTAMINE-GATED CHLORIDE CHANNEL SUBUNITS
HYPERPOLARIZATION-ACTIVATED CYCLIC NUCLEOTIDE-GATED POTASSIUM CHANNELS
INTRACELLULAR CHLORIDE CHANNEL SUBUNITS
INWARDLY RECTIFYING POTASSIUM CHANNEL SUBUNITS
ION CHANNELS
ION CHANNELS BY CONDUCTANCE
LIGAND-GATED CHLORIDE CHANNEL SUBUNITS
LIGAND-GATED ION CHANNELS
NON-SELECTIVE CATION CHANNEL SUBUNITS
ORPHAN LIGAND-GATED CHLORIDE CHANNEL SUBUNITS
OTHER CALCIUM CHANNEL-FORMING SUBUNITS
OTHER VOLTAGE-GATED CHLORIDE CHANNEL SUBUNITS
PH-SENSITIVE CHLORIDE CHANNEL SUBUNITS

2. Links within individual Gene reports
3. A browsable, hierarchical list of Gene Groups (example below). Groups may be split into nested sub-groups, each of which represents a Gene Group:



- HETEROTRIMERIC G-PROTEIN SUBUNITS (G-SUB)
  - HETEROTRIMERIC G-PROTEIN ALPHA SUBUNITS (G-SUB-A)
  - HETEROTRIMERIC G-PROTEIN BETA SUBUNITS (G-SUB-B)
  - HETEROTRIMERIC G-PROTEIN GAMMA SUBUNITS (G-SUB-G)
- HOX GENE COMPLEX (HOX-C)
  - ANTENNAPEDIA COMPLEX (ANT-C)
  - BITHORAX COMPLEX (BX-C)
- INHIBITOR OF APOPTOSIS (IAP)
- INNEXINS (INX)
- INSULIN-LIKE PEPTIDES (ILP)
- INTEGRINS (ITG)
- INTRACELLULAR TRANSPORT COAT, SCAFFOLD AND ADAPTOR PROTEINS (ITCSAP)
  - ADAPTOR PROTEIN COMPLEXES (AP)
    - ADAPTOR PROTEIN COMPLEX 1 (AP1)
    - ADAPTOR PROTEIN COMPLEX 2 (AP2)
    - ADAPTOR PROTEIN COMPLEX 3 (AP3)
  - CLATHRIN COMPLEX (CLATH)
  - COAT PROTEIN COMPLEX I (COPI)
  - COAT PROTEIN COMPLEX II (COPII)
  - ENDOSOMAL SORTING COMPLEXES REQUIRED FOR TRANSPORT (ESCRT)
    - ESCRT-0 COMPLEX (ESCRT-0)
    - ESCRT-I COMPLEX (ESCRT-I)
    - ESCRT-II COMPLEX (ESCRT-II)
    - ESCRT-III COMPLEX (ESCRT-III)
    - VPS4-VTA1 COMPLEX (VSP4C)
  - GATOR COMPLEX (SEA)
    - GATOR1 COMPLEX (GAT1)
    - GATOR2 COMPLEX (GAT2)
  - GOLGI-LOCALIZING, GAMMA-ADAPTIN EAR DOMAIN HOMOLOGY, ARF-BINDING PROTEINS (GGA)
  - INTRAFLAGELLAR TRANSPORT COMPLEXES (IFT)
    - BBSOME (BBS)
    - INTRAFLAGELLAR TRANSPORT SUBCOMPLEX-A (IFT-A)
    - INTRAFLAGELLAR TRANSPORT SUBCOMPLEX-B (IFT-B)
  - RETROMER (RETRO)
  - STONINS (STON)
- ION CHANNELS (IC)
  - ION CHANNELS BY CONDUCTANCE (ICC)
    - CALCIUM CHANNEL SUBUNITS (CAC)
      - CALCIUM RELEASE-ACTIVATED CALCIUM CHANNEL SUBUNITS (CRAC)
      - NON-NMDA IONOTROPIC GLUTAMATE RECEPTORS - GLURII NEUROMUSCULAR JUNCTION-SUBTYPE (GLURII)
      - OTHER CALCIUM CHANNEL-FORMING SUBUNITS (O-CAC)
      - VOLTAGE-GATED CALCIUM CHANNEL SUBUNITS (CACN)
    - CHLORIDE CHANNEL SUBUNITS (CLC)
      - CALCIUM-ACTIVATED CHLORIDE CHANNEL SUBUNITS (CACC)
        - ANOCTAMINS (ANO)
        - BESTROPHINS (BEST)
        - TWEETY FAMILY (TWE)
      - INTRACELLULAR CHLORIDE CHANNEL SUBUNITS (CLIC)
      - LIGAND-GATED CHLORIDE CHANNEL SUBUNITS (LGCC)
        - GABA-GATED CHLORIDE CHANNEL SUBUNITS (GABA-CL)
        - GLUTAMATE-GATED CHLORIDE CHANNEL SUBUNITS (GLU-CL)
        - HISTAMINE-GATED CHLORIDE CHANNEL SUBUNITS (HIS-CL)
        - ORPHAN LIGAND-GATED CHLORIDE CHANNEL SUBUNITS (LGCC-ORPH)

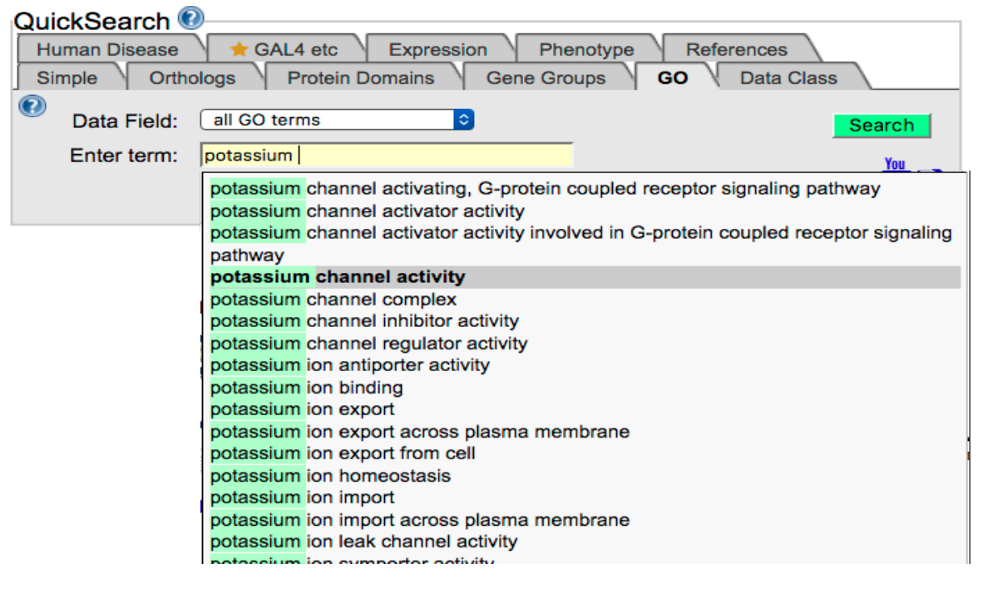
Each level of the hierarchy has its own Gene Group report page, with subgroups and their component member genes displayed.

Gene Group members can be downloaded or exported to analysis tools using the buttons at the top of the 'Members' table within a Gene Group report.

Finding Genes using the GO

The Gene Ontology can be searched in two main ways:

1. From the FlyBase homepage via QuickSearch

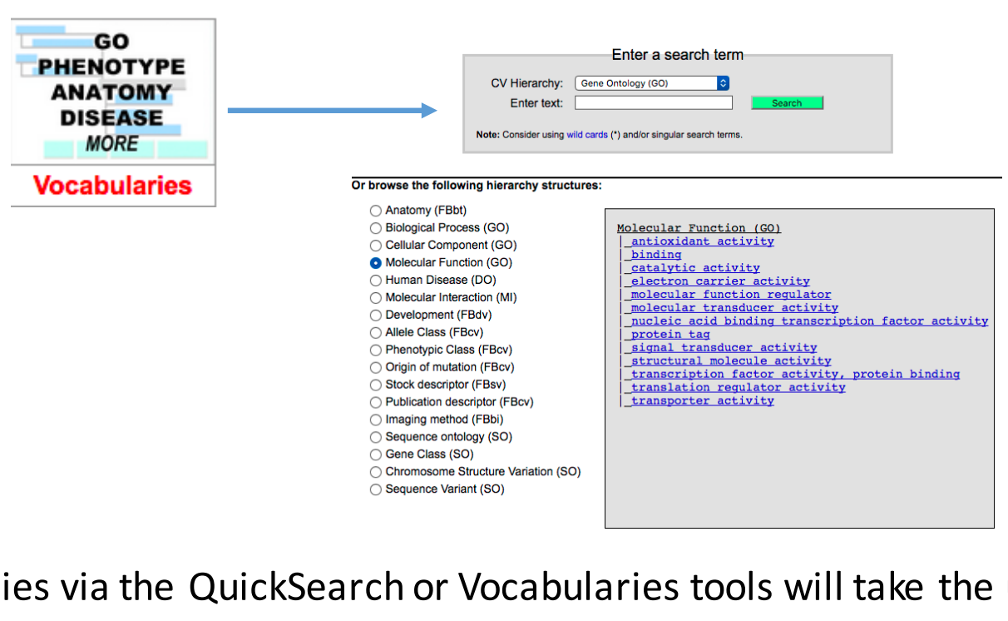


Data Field: all GO terms

Enter term: potassium

potassium channel activating, G-protein coupled receptor signaling pathway
potassium channel activator activity
potassium channel activator activity involved in G-protein coupled receptor signaling pathway
potassium channel activity
potassium channel complex
potassium channel inhibitor activity
potassium channel regulator activity
potassium ion antiporter activity
potassium ion binding
potassium ion export
potassium ion export across plasma membrane
potassium ion export from cell
potassium ion homeostasis
potassium ion import
potassium ion import across plasma membrane
potassium ion leak channel activity
potassium ion symporter activity

2. Or via the Vocabularies link at the top of the homepage.



CV Hierarchy: Gene Ontology (GO)

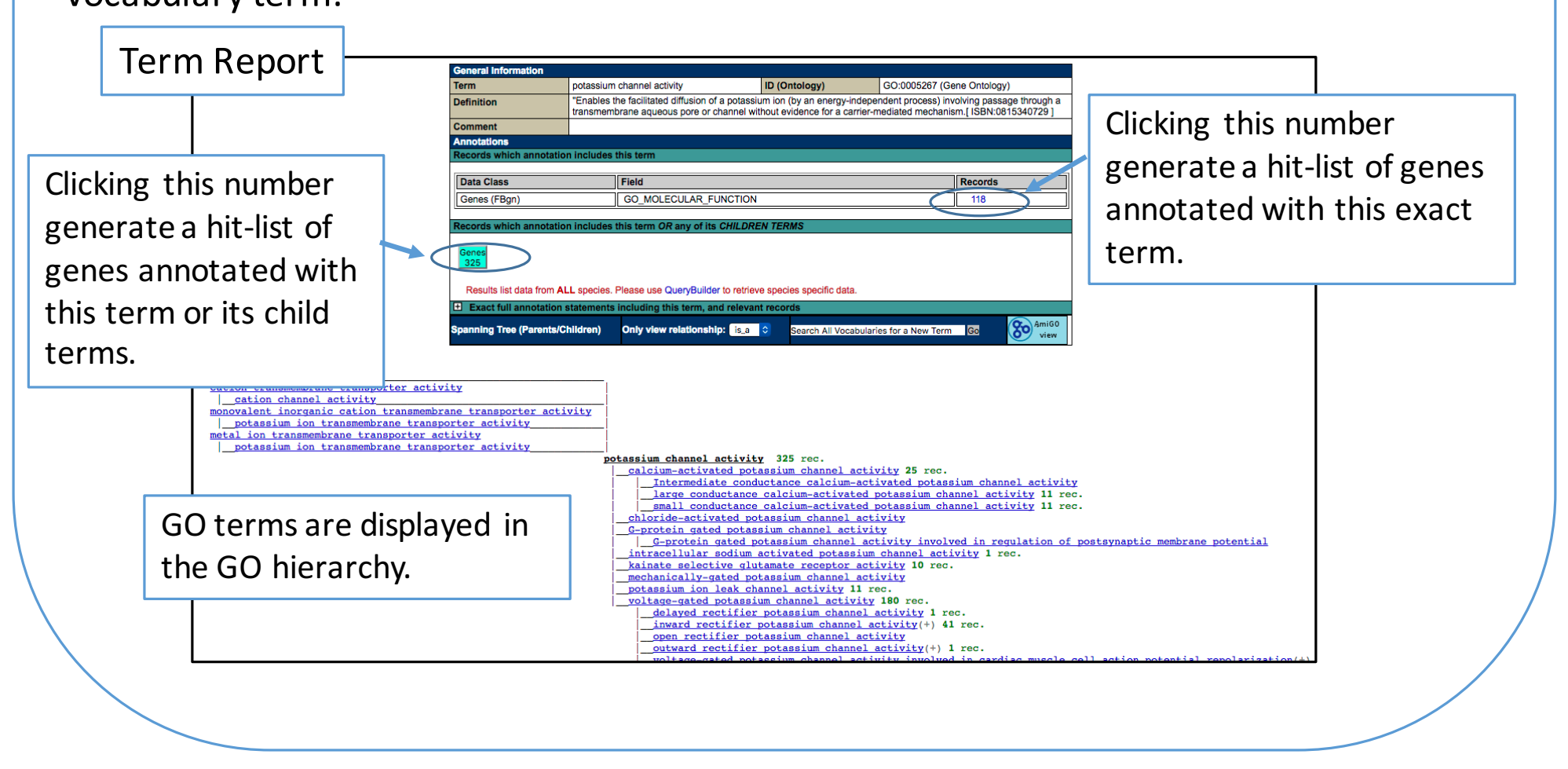
Note: Consider using wild cards (*) and/or singular search terms.

Or browse the following hierarchy structures:

○ Anatomy (FBbt)
○ Biological Process (GO)
○ Cellular Component (GO)
● Molecular Function (GO)
○ Human Disease (DO)
○ Molecular Interaction (MI)
○ Development (FBdv)
○ Allele Class (FBcv)
○ Phenotypic Class (FBcv)
○ Origin of mutation (FBcv)
○ Stock descriptor (FBsv)
○ Publication descriptor (FBcv)
○ Imaging method (FBbi)
○ Sequence ontology (SO)
○ Gene Class (SO)
○ Chromosome Structure Variation (SO)
○ Sequence Variant (SO)

```
Molecular Function (GO)
 |_antioxidant activity
 |_binding
 |_catalytic activity
 |_electron carrier activity
 |_molecular function regulator
 |_molecular transducer activity
 |_nucleic acid binding transcription factor activity
 |_protein tag
 |_signal transducer activity
 |_structural molecule activity
 |_transcription factor activity, protein binding
 |_translation regulator activity
 |_transporter activity
```

GO queries via the QuickSearch or Vocabularies tools will take the user to a a Term Report. A Term Report displays information and data associated with a controlled vocabulary term.



Term Report

General Information

Term	potassium channel activity	ID (Ontology)	GO:0005267 (Gene Ontology)
Definition	"Enables the facilitated diffusion of a potassium ion (by an energy-independent process) involving passage through a transmembrane aqueous pore or channel without evidence for a carrier-mediated mechanism.[ISBN:0815340729]		
Comment			

Annotations

Records which annotation includes this term

Data Class	Field	Records
Genes (FBgn)	GO_MOLECULAR_FUNCTION	118

Records which annotation includes this term OR any of its CHILDREN TERMS

Genes 325

Results list data from ALL species. Please use QueryBuilder to retrieve species specific data.

⊞ Exact full annotation statements including this term, and relevant records

Spanning Tree (Parents/Children) · Only view relationship: is_a · Search All Vocabularies for a New Term · AmiGO view

Clicking this number generate a hit-list of genes annotated with this term or its child terms.

Clicking this number generate a hit-list of genes annotated with this exact term.

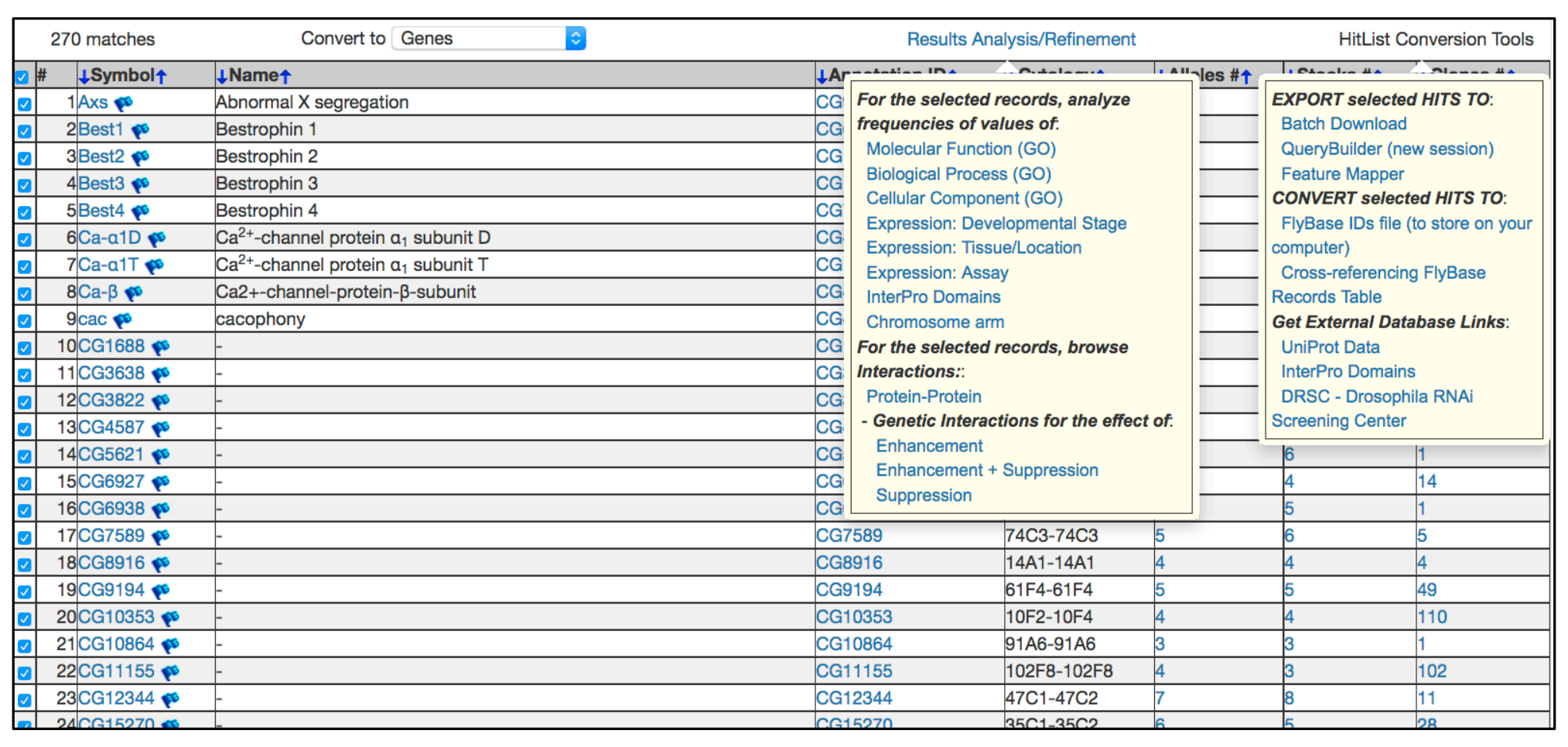
```
cation transmembrane transporter activity
 |_cation channel activity
monovalent inorganic cation transmembrane transporter activity
 |_potassium ion transmembrane transporter activity
metal ion transmembrane transporter activity
 |_potassium ion transmembrane transporter activity

potassium channel activity  325 rec.
 |_calcium-activated potassium channel activity 25 rec.
 |   |_intermediate conductance calcium-activated potassium channel activity
 |   |_large conductance calcium-activated potassium channel activity 11 rec.
 |   |_small conductance calcium-activated potassium channel activity 11 rec.
 |_chloride-activated potassium channel activity
 |_G-protein gated potassium channel activity
 |   |_G-protein gated potassium channel activity involved in regulation of postsynaptic membrane potential
 |_intracellular sodium activated potassium channel activity 1 rec.
 |_kainate selective glutamate receptor activity 10 rec.
 |_mechanically-gated potassium channel activity
 |_potassium ion leak channel activity 11 rec.
 |_voltage-gated potassium channel activity 180 rec.
     |_delayed rectifier potassium channel activity 1 rec.
     |_inward rectifier potassium channel activity(+) 41 rec.
     |_open rectifier potassium channel activity
     |_outward rectifier potassium channel activity(+) 1 rec.
```

GO terms are displayed in the GO hierarchy.

Combinatorial Searches

It may be useful to combine searches of Gene Groups with GO data, or indeed any other data in FlyBase, to generate more refined gene lists. Here we show two ways in which this can be done.

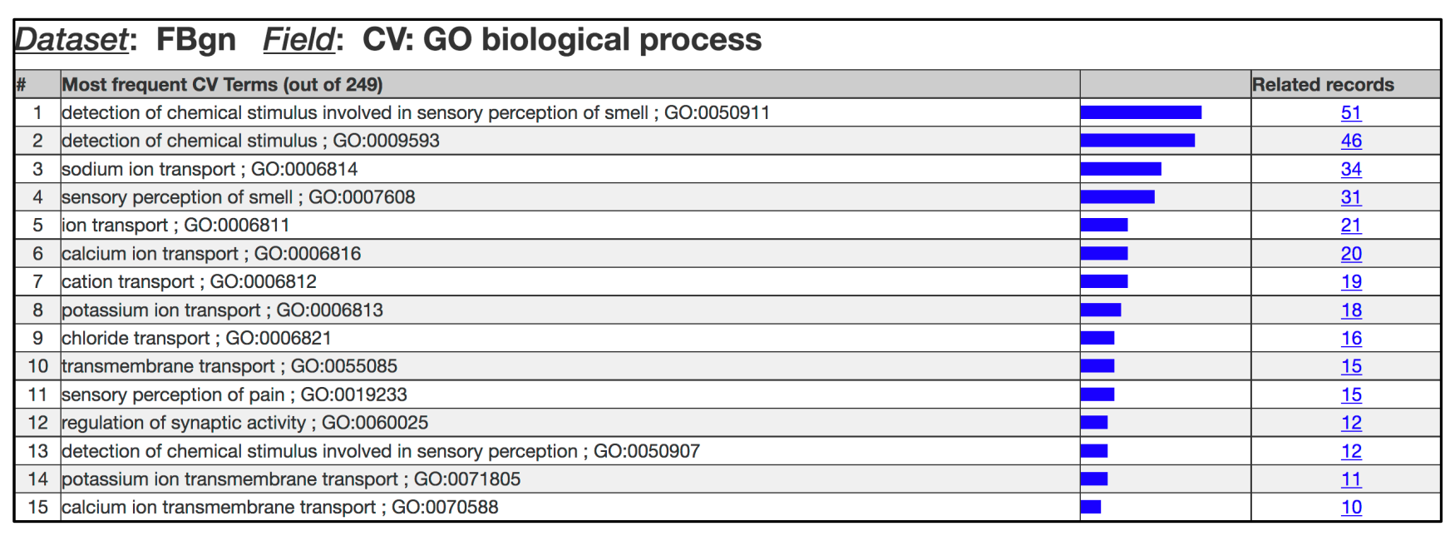


270 matches · Convert to Genes · Results Analysis/Refinement · HitList Conversion Tools

#	Symbol	Name	Annotation ID	Cytology	Alleles #	Stocks #	Clones #
1	Axs	Abnormal X segregation	CG				
2	Best1	Bestrophin 1	CG				
3	Best2	Bestrophin 2	CG				
4	Best3	Bestrophin 3	CG				
5	Best4	Bestrophin 4	CG				
6	Ca-α1D	Ca^{2+}-channel protein $α_1$ subunit D	CG				
7	Ca-α1T	Ca^{2+}-channel protein $α_1$ subunit T	CG				
8	Ca-β	Ca2+-channel-protein-β-subunit	CG				
9	cac	cacophony	CG				
10	CG1688	-	CG				
11	CG3638	-	CG				
12	CG3822	-	CG				
13	CG4587	-	CG				
14	CG5621	-	CG			6	1
15	CG6927	-	CG			4	14
16	CG6938	-	CG			5	1
17	CG7589	-	CG7589	74C3-74C3	5	6	5
18	CG8916	-	CG8916	14A1-14A1	4	4	4
19	CG9194	-	CG9194	61F4-61F4	5	5	49
20	CG10353	-	CG10353	10F2-10F4	4	4	110
21	CG10864	-	CG10864	91A6-91A6	3	3	1
22	CG11155	-	CG11155	102F8-102F8	4	3	102
23	CG12344	-	CG12344	47C1-47C2	7	8	11
24	CG15270	-	CG15270	35C1-35C2	6	5	28

For the selected records, analyze frequencies of values of:
Molecular Function (GO)
Biological Process (GO)
Cellular Component (GO)
Expression: Developmental Stage
Expression: Tissue/Location
Expression: Assay
InterPro Domains
Chromosome arm
For the selected records, browse Interactions::
Protein-Protein
- *Genetic Interactions for the effect of*:
Enhancement
Enhancement + Suppression
Suppression

EXPORT selected HITS TO:
Batch Download
QueryBuilder (new session)
Feature Mapper
CONVERT selected HITS TO:
FlyBase IDs file (to store on your computer)
Cross-referencing FlyBase Records Table
Get External Database Links:
UniProt Data
InterPro Domains
DRSC - Drosophila RNAi Screening Center

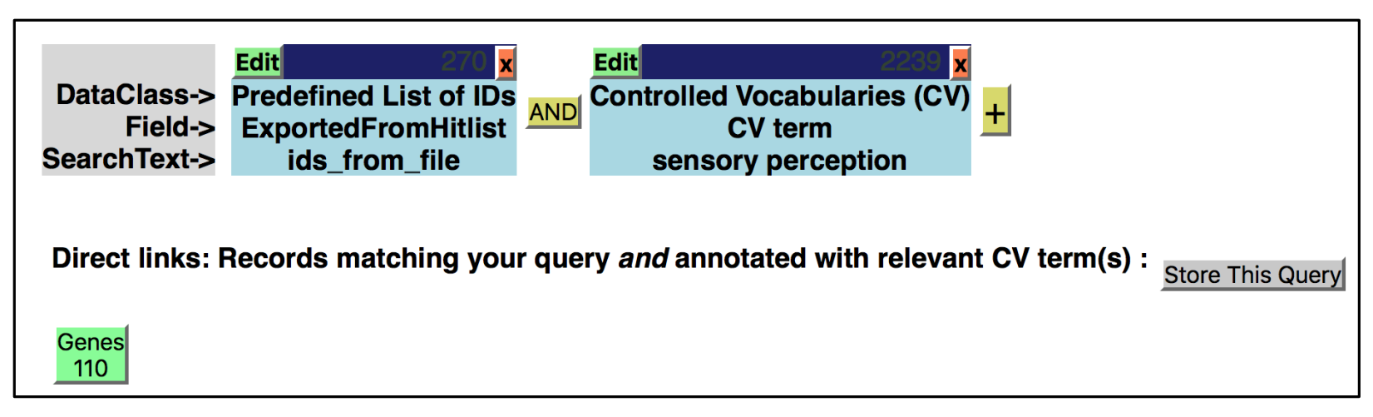


Dataset: FBgn *Field*: CV: GO biological process

#	Most frequent CV Terms (out of 249)	Related records
1	detection of chemical stimulus involved in sensory perception of smell ; GO:0050911	51
2	detection of chemical stimulus ; GO:0009593	46
3	sodium ion transport ; GO:0006814	34
4	sensory perception of smell ; GO:0007608	31
5	ion transport ; GO:0006811	21
6	calcium ion transport ; GO:0006816	20
7	cation transport ; GO:0006812	19
8	potassium ion transport ; GO:0006813	18
9	chloride transport ; GO:0006821	16
10	transmembrane transport ; GO:0055085	15
11	sensory perception of pain ; GO:0019233	15
12	regulation of synaptic activity ; GO:0060025	12
13	detection of chemical stimulus involved in sensory perception ; GO:0050907	12
14	potassium ion transmembrane transport ; GO:0071805	11
15	calcium ion transmembrane transport ; GO:0070588	10

HitList Refinement tools
A list of genes can be downloaded from a Gene Group report page or, when querying a controlled vocabulary, such as the GO, from a Term Report page. The tools under 'Results Analysis/Refinement' button can be used to show simple intersections with other data in FlyBase. In the example shown, the hit-list is populated with genes exported from the ION CHANNELS Gene Group. A results analysis of the top 15 individual Biological Process GO terms associated with genes from the ION CHANNELS Gene Group are shown. These can be selected to generate a refined hit-list.

QueryBuilder
For more complex queries, the QueryBuilder tool can be used to combine individual query segments using Boolean operators (AND, OR, BUT NOT) in order to generate lists that combine or exclude the given criteria. In the example shown here, a 2-leg query has been performed with genes imported from the ION CHANNELS Gene Group and a second query for GO term search for 'sensory perception'.



DataClass->	Predefined List of IDs	AND	Controlled Vocabularies (CV)
Field->	ExportedFromHitlist		CV term
SearchText->	ids_from_file		sensory perception

Direct links: Records matching your query *and* annotated with relevant CV term(s) :

Looking to the future:

- The curation of Gene Groups is often linked with GO annotation curation. When a curator compiles a Gene Group the GO annotation of member genes is quickly reviewed to ensure that it is consistent with the known characteristics of the gene set. There are three main areas into which GO and Gene Group curation will extend in the future:
  - Non-coding RNAs and the components of non-coding RNA biogenesis.
  - Signalling pathways: new pathway and network pages are planned.
  - Protein Complexes with dedicated complex pages linked to physical interaction data.

We would like to thank Mouse Genomic Informatics are sharing the ribbons technology. FlyBase is supported by a grant from the National Human Genome Research Institute at the U.S. National Institutes of Health #P41 HG000739. This work is also supported by the British Medical Research Council (grant number RG83674) **.**
Support is also provided by the Indiana Genomics Initiative, and the National Science Foundation through XSEDE resources provided by Indiana University.

Attrill H, Falls K, Goodman JL, Millburn GH, Antonazzo G, Rey AJ, Marygold SJ, FlyBase consortium (2016). FlyBase: establishing a Gene Group resource for Drosophila melanogaster. *Nucleic Acids Res*. 44(D1):D786-92. doi: 10.1093/nar/gkv1046.
Marygold SJ, Antonazzo G, Attrill H, Costa M, Crosby MA, Dos Santos G, Goodman JL, Gramates LS, Matthews BB, Rey AJ, Thurmond J; FlyBase Consortium. (2016) Exploring FlyBase Data Using QuickSearch. Curr. Protoc. Bioinformatics 56:1.31.1-1.31.23. doi: 10.1002/cpbi.19