

TACACAATCAGTTAGTTTCCACCGACAGTCCGAGAAACCATTCCAGCGG
 GTCCGCAATCC@TAAGATAGCCAAATATATATATGTTTCAGACTCA
 AGCCCAACACTGCAGATCC@T@GAGTGT@CAAATCAGTGAATTC
 ATTTCCCGGC@CAAGCGGACTTTT@G@ATG@ATG@ARA@AAAAA
 AATAATAAAA@CAACACAGTGAACACAGCGGGGCATCTCATAGAT
 AACTTCTGCCT@C@ACTTGGTATATGACTTATACATAGACATATATA



FlyBase

Gene Groups in FlyBase



**UNIVERSITY OF
CAMBRIDGE**

Steven J. Marygold^{1*}, Helen L. Attrill¹, Kathleen Falls², Joshua L. Goodman³, Gillian H. Millburn¹ and the FlyBase Consortium

1. Dept. of Genetics, University of Cambridge, Cambridge, CB2 3EH, UK. 2. Biological Laboratories, Harvard University, Cambridge, MA 02138, USA. 3. Dept. of Biology, Indiana University, Bloomington, IN 47405, USA
 *Email: sjm41@gen.cam.ac.uk

Related sets of genes or gene products, such as gene families (e.g. actins), subunits of protein complexes (e.g. proteasome subunits) or other functional groupings (e.g. Ubiquitin E3 ligases), are frequently described in the literature. Moreover, researchers often wish to obtain a list of related genes from a database and be able to easily download/query data associated with them. We wish to improve the representation of such 'gene groups' in FlyBase through the provision of new 'Gene Group Reports'. These Reports will tabulate the membership of each group, clearly attributed to source references, together with buttons to export the member genes to FlyBase tools to facilitate further analyses. Gene Group Reports will also include a brief description of the group, links to parent/child groups within FlyBase, and links to orthologous groups available at external sites such as HGNC, WormBase and TAIR. Through doing this work and adopting a 'group' approach to curation, we have also been able to improve the consistency and accuracy of Gene Ontology annotations and nomenclature of related genes.

The **General Information** panel displays basic data on the Gene Group, including the number of members in the group and the date the group was last reviewed by FlyBase.

Any 'edge cases' where the inclusion/exclusion of a specific gene is unclear or debated are described in the **Notes on Membership** section.

Key Gene Ontology (GO) terms are those terms that largely define the group and which are used to annotate most/all member genes.

Data downloads and further analyses of the Gene Group are accessible via the 'Export to Hitlist' and 'Export to Batch Download' buttons.

Hyperlinks to equivalent Gene Groups at other databases, such as the HGNC, WormBase or TAIR, are given in the **Orthologous Group(s)** section.

FB2014_02, released March 18th, 2014
 Gene Group Report: CLASS B GPCRs, SUBFAMILY B1

General Information

Name	CLASS B GPCRs, SUBFAMILY B1	Species	<i>D. melanogaster</i>
Symbol	GPCR-B1	FlyBase ID	FBgg0000046
Date last reviewed	2013-08-15	Number of members	5

Description

Description: The B1 subfamily of Class B GPCRs are largely classical hormone receptors. (Adapted from FBrf0147071 and FBrf0221117).

Notes on membership: The CLASS B GPCRs, SUBFAMILY B1 Gene Group has been compiled by FlyBase curators using the membership given in the following publication(s): Hauser et al., 2006, Harmar, 2001, Li et al., 2013, and Hauser et al., 2006.

Key Gene Ontology (GO) terms

Molecular Function	protein-hormone receptor activity
Biological Process	G-protein coupled receptor signaling pathway
Cellular Component	integral component of plasma membrane

Related Gene Groups

Parent group(s): CLASS B GPCRs

Members (5)

Gene Symbol	Gene Name	Annotation ID	Source Material
Dh31-R	Diuretic hormone 31 Receptor	CG32843	(Hauser et al., 2006, Harmar, 2001, Li et al., 2013, Hauser et al., 2006)
Dh44-R1	Diuretic hormone 44 receptor 1	CG8422	(Hauser et al., 2006, Harmar, 2001, Li et al., 2013, Hauser et al., 2006)
Dh44-R2	Diuretic hormone 44 receptor 2	CG12370	(Hauser et al., 2006, Harmar, 2001, Li et al., 2013, Hauser et al., 2006)
hec	hector	CG4395	(Hauser et al., 2006, Harmar, 2001, Li et al., 2013, Hauser et al., 2006)
Pdfr	Pigment-dispersing factor receptor	CG13758	(Hauser et al., 2006, Harmar, 2001, Li et al., 2013, Hauser et al., 2006)

External Data

Orthologous Group(s): Human Class B GPCRs

Synonyms and Secondary IDs

Synonym(s): GPCR-B1, CLASS B GPCRs, SUBFAMILY B1, Family B1 GPCR, Classical hormone receptor, Family B-type neuropeptide receptor

References (5)

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

Research paper: Li et al., 2013, Gene 519(1): 1-12. Comparative genomic analysis and evolution of family-B G protein-coupled receptors from six model insect species. [FBrf0221117]

Review: Hauser et al., 2006, Prog. Neurobiol. 80(1): 1-19. A review of neurohormone GPCRs present in the fruitfly *Drosophila melanogaster* and the honey bee *Apis mellifera*. [FBrf0193587]

Hauser et al., 2006, Brief. Funct. Genomic. Proteomic. 4(4): 321-330. Identifying neuropeptide and protein hormone receptors in *Drosophila melanogaster* by exploiting genomic data. [FBrf0195077]

Harmar, 2001, Genome Biol. 2(12): REVIEWS3013. Family-B G-protein-coupled receptors. [FBrf0147071]

The textual **Description** of the group is written by FlyBase curators, based on the given references.

The **Source Material** section serves to: (i) clearly indicate that the group has been curated manually by FlyBase curators; and (ii) summarizes the primary references used to compile the group.

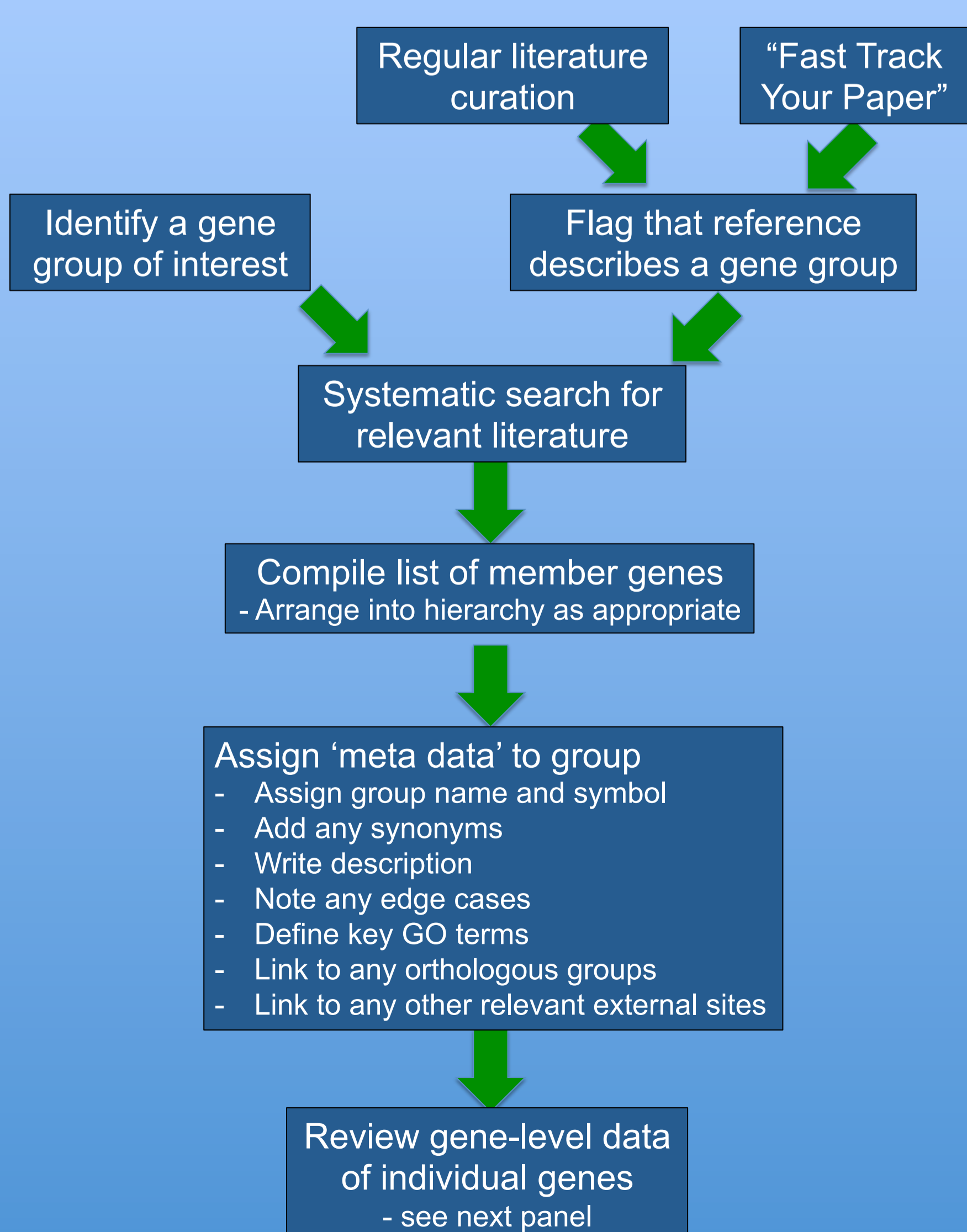
When relevant, parent super-groups or child sub-groups within FlyBase are shown as hyperlinks in the **Related Gene Groups** panel.

The **Members** table lists the member genes of the group and includes:
 - Gene Symbol (linked to Gene Report)
 - Gene Name
 - Annotation ID (linked to GBrowse)
 - Source Material (i.e. the reference(s) that state that this specific gene is a member of this particular group)

Hyperlinks to any relevant specialist websites are provided in the **Other resource(s)** panel.

All citations are given in full in the **References** section, each hyperlinked to its Reference Report

Curation strategy



Added benefits of 'group curation'

Compiling sets of related genes provides the ideal opportunity to review the consistency and accuracy of gene-level data in FlyBase. The following areas are routinely examined:

- Gene Ontology annotations**
 - Add missing terms (~4,500 added to date)
 - Update existing terms (e.g. if the available ontology terms or underlying data have changed)
 - Remove any redundant or incorrect terms
- Gene nomenclature**
 - Check compliance with guidelines
 - Check accuracy and that it reflects usage
 - Add full-name if missing
 - Improve consistency within group where feasible (e.g. related genes that are named for their molecular function/orthology should have the same prefix and format)
- Un-annotated genes**
 - Check if there are any genes not mapped to the sequence that have GO annotations and/or gene symbols that suggest they are members of the current group
 - Merge FlyBase Gene Reports as appropriate

(These improvements mean that GO- or symbol-based searches to identify groups of related genes are more successful *after* review, but such searches will always suffer from false positives/negatives – thus the need for Gene Group Reports.)

Major Groups compiled to date

ACCESSORY GLAND PROTEINS	FLOTILLINS	POTASSIUM CHANNELS
ACTINS	G PROTEIN ALPHA SUBUNITS	PROTEASOME
ACTIN-RELATED PROTEINS	GLUT. S TRANSFERASES-CYTO	RAS FAMILY GAPS
ARP 2/3 COMPLEX	GOLGINS	RAS FAMILY GEFS
ADAM METALLOPROTEASES	GPCR - BIOGENIC_AMINE	RAS FAMILY GTPASEs
ADAPTOR PROTEIN COMPLEX	GPCR - FAMILY_A	RETROMER COMPLEX
ANAPHASE-PROMOTING CPX	GPCR - FAMILY_B	RIBOSOMAL PROTEINS - CYTO
ANNEXINS	GPCR - FAMILY_C	RIBOSOMAL PROTEINS - MITO
APOPTOSIS - APOPTOSOME	GPCR - FAMILY_F	RING FINGER PROTEINS
APOPTOSIS - BCL2-LIKE	GUSTATORY RECEPTORS	RZZ COMPLEX
APOPTOSIS - CASPASE	HALLOWEEN GROUP GENES	SEA COMPLEX
APOPTOSIS - IAP	HETEROCHROMATIN PROTEIN 1	SEPTINS
APOPTOSIS - RHG	IFT-A SUBCOMPLEX	SERPINS
ATAC COMPLEX	IFT-B SUBCOMPLEX	SKP1
AUTOPHAGY-RELATED GENES	INNEXINS	SM PROTEINS
BAR DOMAIN PROTEINS	INTEGRINS	SMC
BBSOME	INWARDLY RECTIFYING K CHLs	SMC 5/6 COMPLEX
BLOC1 COMPLEX	INSULIN-LIKE PEPTIDES	SMC COHESIN COMPLEX
BLOC2 COMPLEX	IONOTROPIC (GLUTAMATE) REC.	SMC CONDENSIN I COMPLEX
BLOC3 COMPLEX	KASH DOMAIN PROTEINS	SMC CONDENSIN II COMPLEX
CADHERINS	KINESIN SUPERFAMILY	SNAPS
CHROM. PASSENGER CPX	LAMINS	SNARES
CLATHRINS	METALLOTHIONEINS	SORTING NEXINS
CNG CHANNELS	METHUSELAH GENES	SYNAPTOTAGMINS
COP9 SIGNALOSOME	MON1-CC21 COMPLEX	TETHERING COMPLEX
COPI COMPLEX	MYOSINS	TFIIH_COMPLEX
COPII COMPLEX	MYSOIN LIGHT CHAINS	TRAFs
CULLINS	NEUROPEP/PROTEIN/HORMONES	TRP CHANNELS
DORSOCROSS GENES	NICOTINIC ACETYLCHOLINE REC.	TUBULINS
DROSOMYCINS	NIMROD GENES	UBIQUITIN-ACTIVATING-E1
DSL1 COMPLEX	NSFs	UBIQUITIN-CONJUGATING-E2
DUBs	NUCLEOPORINS	UBIQUITIN-LIGASE E3 - GOLIATH
ENHANCER OF SPLIT CPX	ODORANT BINDING PROTEINS	V-ATPASES
ESCRT COMPLEX	ODORANT RECEPTORS	WNTs
EXOCYST COMPLEX	P24 TRANSPORTERS	YELLOW GENES
F-BOX PROTEINS	PICKPOCKET GENES	