

Performing annotation queries on the Flannotator

Introduction

Querying the annotation and mapping data on the Flannotator can be performed in a number of different ways, with varying depth.

- Gene search
 - Search sequencing data for insertions
- Simple query
 - Search annotations based on simple high level annotation terms
- Complex query
 - Search annotations based on a custom complex hierarchical tree system
- Table of gene hits
 - Tabulated summary of gene fusions produced by the project

Gene search

1. Enter the name of the gene you wish to search for YFP traps
 1. Please use the official FlyBase gene symbol or CG number

Search for tagged genes

Enter gene symbol or CG number (eg N or CG3936):

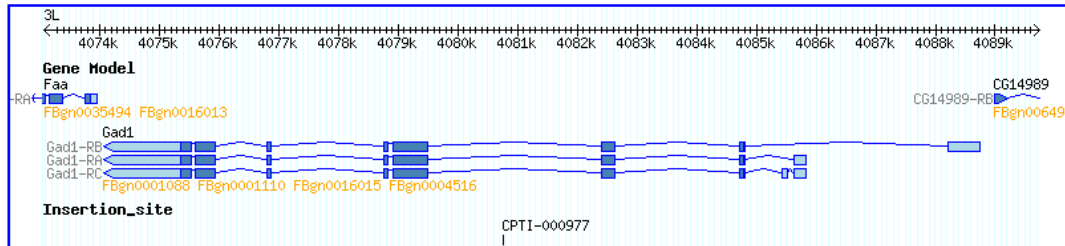
2. Click on *search* to query the database. After a brief moment the results page should appear.

Search results for Gad1

Inserts producing an in-frame YFP-fusion

^ CPTI v	Loc	^ Chr v	^ Gene CG v	^ Gene Name v	Feature ID	^ YFP Status v
CPTI-000977	4080753	3L	CG14994	Gad1	intron_CG14994:3_CG14994:4	Confirmed - Yes

Gene model for CG14994 (Gad1)



3. Click on the stock name to bring up the stock report for the particular insertion
4. Clicking on the gbrowse gene model will load the gbrowse application

Simple query

The simple query system is a single-level query based on either anatomy terms, sub-cellular locations or descriptors. It does not differentiate between any types of relationships that the annotations might have, but it is a good place to start.

High level anatomy terms

1. These are the same as the tissue categories used for the annotation system
2. Choose from the list and click on *choose tissue* to initiate the search
3. Please note that some of the terms on the list are on different levels of the ontology tree to others.
4. For information about a particular term, click on *Term info*, which will display information about the term such as definition and place on the ontology tree.

Choose from a list of high level anatomy terms (this is a good place to start)

Drosophila Anatomy

adult excretory system
adult external head
adult external thorax
adult integumentary system
adult muscle system
adult nervous system
adult peripheral nervous system
adult tracheal system
cell
central nervous system
circulatory system
digestive system

Term info

Select Matching Options: Search

Choose Tissue

Term information

CV term: digestive system

Definition:

The organ system involved in the absorption of nutrients.

Synonyms:

Ontology tree

Drosophila
└─ whole organism
 └─ organ system
 └─ digestive system

All anatomy terms

1. This is similar to the high-level term search, except that all cv anatomy terms are available. This gives much finer control on the search.
2. Click on *choose CV term* to initiate the search on selected terms

Or choose from a list of all anatomy terms...

larval cervical connective
larval cibarium
larval circulatory system
larval corpus allatum
larval corpus cardiacum
larval cuprophilic cell
larval cuticle
larval deutocerebrum
larval digestive system
larval dorsal anastomosis
larval dorsal branch
larval dorsal group branch

Term info

Select Matching Options: Search

Choose CV term

Subcellular terms

1. Instead of searching anatomy terms, you can also search for subcellular localisation of protein traps.
2. Subcellular terms are taken from the GO_slim (<http://www.geneontology.org>) cellular_component ontology.

1. Please note as this is a direct lift from GO, there may be some terms that are not relevant to *Drosophila* (eg cell wall, which is plant cell specific), but are included to keep the integrity of the tree structure. These can be ignored.
3. Click on *choose subcellular* to initiate the search on selected terms.

Or choose from a list of subcellular terms...

cytoplasm	▲
cytoplasmic chromosome	
cytoplasmic membrane-bound vesicle	
cytoskeleton	
cytosol	
endoplasmic reticulum	
endosome	
external encapsulating structure	
extracellular matrix (sensu Metazoa)	
extracellular region	
extracellular space	
Golgi apparatus	▼

Term info

Select Matching Options: Search

Choose subcellular

Descriptor terms

1. Descriptor terms are a subset of the spatial descriptor ontology available from FlyBase, and describe the pattern of expression, rather than where it is expressed.
2. This may be useful for searching for lines with ubiquitous or punctate expression patterns for example.
3. Click on *choose descriptor* to initiate the search on selected terms.

Or choose from a list of descriptor terms...

medial	▲
neurogenic	
posterior	
posterior compartment	
proximal	
punctate	
restricted	
segmentally repeated	
subset	
terminal	
ubiquitous	
uniform	▼

Term info

Select Matching Options: Search

Choose descriptor

Complex queries

Creating complex queries is less straightforward than the simple searches but is much more powerful, allowing for hierarchical questions to be asked.

The main query window is split up into several panes, all providing different information on the query.

The screenshot shows a complex query interface divided into several panes:

- TREE SELECTOR PANE:** A tree view under "Choose a branch from the ontology tree to narrow your search terms". It lists categories like "Whole organism", "organ system?", "adipose system?", "circulatory system?", "digestive system?", "endocrine system?", "excretory system?", "imaginal precursor", "integumentary system?", "muscle system?", "nervous system?", "reproductive system?", "tracheal system?", "stage", and "tagma?".
- TERM SELECTOR PANE:** A list view under "Or choose from a list of tissue types to narrow your search terms". It lists terms like "adult peripheral nervous system", "adult tracheal system", "cell", "central nervous system", "circulatory system", "digestive system", "embryonic adipose system", "embryonic brain", "embryonic central nervous system", and "embryonic circulatory system".
- INFO PANE:** Contains "CV term: optic lobe", a "Definition:" (The thick bundle of nerves branching off from the protocerebrum, innervating the compound eyes, a region of the adult brain that processes the visual information from the compound eyes.), "Synonyms:", and an "Ontology tree" showing "Drosophila" with sub-nodes "whole organism", "stage", "adult", and "adult external".
- QUERY PANE:** Includes "Query Trees:" (Anatomy terms further down in the ontology tree will automatically be included in the search (eg searching for 'testis' will also search for 'spermatocyte)'), "Match ALL these terms:" (central nervous system (s d r n), nucleus (d r n)), "Match ANY of these terms:", a filter "Filter results by ANY developmental stage", and a "Start query" button.
- CV TERM PANE:** A "Cellular Component" list view showing terms like "biological_process", "cell", "cell envelope", "cell wall", "cellular component organization and biogenesis", "cellular process", "cellular_component", "chromosome", "chromosome organization and biogenesis", and "chromosome organization and biogenesis (sensu Eukaryota)". It has "select from list" and "info" buttons.

Queries can be constructed in two main ways, although the end query will be the same

- Using the tree selector
 - Useful if you are unfamiliar with the anatomy ontology tree
 - Slower to load and use than the term selector
- Using the term selector
 - Requires a basic knowledge of the anatomy ontology tree structure (but not much)
 - Much faster operation than the tree selector

Creating a query

As an example we'll create the following query:

“Show all stocks that express in the cytoplasm of the larval salivary gland, and show uniform expression in the cytoplasm of the egg chamber”

The final format of the query will look like this:

```
Match ALL these terms:  
larval salivary gland ( s d r n )  
cytoplasm ( d r n )  
egg chamber ( s d r n )  
cytoplasm ( d r n )  
uniform ( d r n )
```

Adding anatomy terms using the ontology tree selector

1. Find the high-level term from the tree that your search term belongs. In this case the *embryonic/larval digestive system*. Click on the term to populate the CV term pane with the anatomy term sub-tree. This may take a while to load depending on the number of terms in the tree.
2. Find the term you want to add – in this case ‘*larval salivary gland*’ and click on the checkbox. To add multiple terms you can select several checkboxes at once.
3. Click one ‘add all’ to add the selected terms to the *match all* box
4. The selected terms should then appear.

Choose a branch from the ontology tree to narrow your search terms

- ▼ Whole organism
 - ▶ organ system ?
 - ▶ adipose system ?
 - ▶ circulatory system
 - ▶ digestive system ?
 - *embryonic/larval digestive system ← 1
 - *foregut ?
 - *hindgut ?
 - *midgut ?
 - *pupal/adult digestive system
 - ▶ endocrine system
 - ▶ excretory system
 - ▶ imaginal precursor
 - ▶ integumentary system ?
 - ▶ muscle system
 - ▶ nervous system ?

Or choose from a list of tissue types to narrow your search terms

- adipose system
- adult adipose system
- adult brain
- adult central nervous system
- adult circulatory system
- adult endocrine system
- adult excretory system
- adult external head
- adult external thorax
- adult integumentary system

[Get relevant CV terms](#)

[Add a NOT qualifier](#)

CV term information will appear here

Query Trees:

Anatomy terms further down in the ontology tree will automatically be included in the search (eg searching for 'testis' will also search for 'spermatoocyte')

Match ALL these terms:

larval salivary gland (s d r n) ← 4

Match ANY of these terms:

▶ [Filter results by ANY developmental stage](#)

Start query

Expand Collapse All Untick All Tick All

Add to ALL Add to ANY

- ▼ embryonic/larval digestive system
 - ▶ cephalopharyngeal skeleton
 - ▶ embryonic digestive system
 - ▶ embryonic/larval foregut
 - ▶ embryonic/larval hindgut
 - ▶ embryonic/larval midgut
 - ▶ embryonic/larval mouth
 - ▼ embryonic/larval salivary gland
 - ▶ embryonic salivary gland
 - ▶ embryonic/larval salivary gland duct
 - *larval salivary gland ← 2
 - *larval digestive system

5. Repeat for *egg chamber*, which can be found under *reproductive system* -> *female reproductive system*

Choose a branch from the ontology tree to narrow your search terms

- ▼ Whole organism
 - ▶ organ system ?
 - ▶ adipose system ?
 - ▶ circulatory system
 - ▶ digestive system ?
 - ▶ endocrine system
 - ▶ excretory system
 - ▶ imaginal precursor
 - ▶ integumentary system ?
 - ▶ muscle system
 - ▶ nervous system ?
 - ▼ reproductive system
 - *female reproductive system ← 5
 - *gamete
 - *gonad ?
 - *male reproductive system

Add to ALL Add to ANY

- ▼ female reproductive system
 - *calyx
 - ▶ female accessory gland ?
 - *genital chamber ?
 - ▼ ovary ?
 - ▼ ovariole
 - *basal stalk
 - ▶ egg chamber ?
 - ▶ germarium ?
 - *ovarian basement membrane ?
 - ▶ ovarian sheath ?
 - ▶ vitellarium ?

Match ALL these terms:

larval salivary gland (s d r n)

egg chamber (s d r n)

6. We now have the anatomy terms we want to search for. To learn about adding subcellular and descriptor terms, please skip to the section 'adding subcellular and descriptor terms' in this manual.

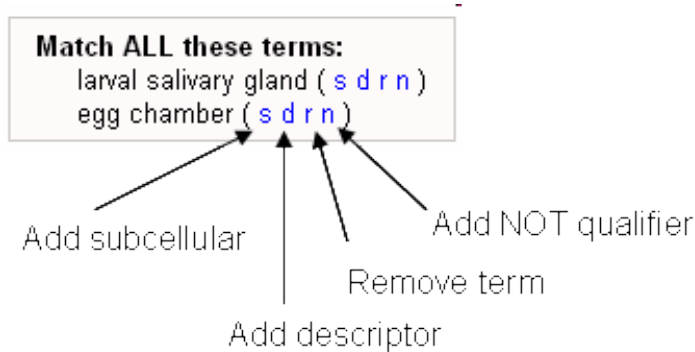
Adding anatomy terms using the ontology term selector

1. Choose the main tissue category which contains your term of interest. In this case *digestive system*. If you don't know which tissue category to use you can use *whole organism*. Click on 'get relevant CV terms' which will populate the CV term pane with a list anatomy terms belonging to the select tissue category. The list loads much faster than the tree view but does not show the relationships between the terms.
2. Find the term you want to add – in this case '*larval salivary gland*' and click on the term to select it.
3. Click one 'add all' to add the selected terms to the *match all* box
4. The selected terms should then appear.
5. Repeat for *egg chamber*, which can be found under *female reproductive system*
6. We now have the anatomy terms we want to search for.

The screenshot illustrates the ontology term selector interface. On the left, a box titled "Choose a branch from the ontology tree to narrow your search terms" shows a tree structure with "Whole organism" selected. In the center, a list of tissue types is shown, with "digestive system" highlighted and an arrow labeled "1" pointing to it. Below this list are buttons for "Get relevant CV terms" and "Add a NOT qualifier". On the right, the "CV term: larval salivary gland" is displayed, along with its definition and synonyms. Below this, an "Ontology tree" shows the hierarchical structure of the term. At the bottom, the "Query Trees" section shows a search query "larval salivary gland (s d r n)" with an arrow labeled "4" pointing to it. A "Start query" button is also visible. On the right side of the bottom section, a "Drosophila CV anatomy terms" list is shown, with "larval salivary gland" highlighted and an arrow labeled "2" pointing to it. Below this list are buttons for "add to ALL", "add to ANY", and "info", with an arrow labeled "3" pointing to the "add to ALL" button. A "Find:" search box is also present.

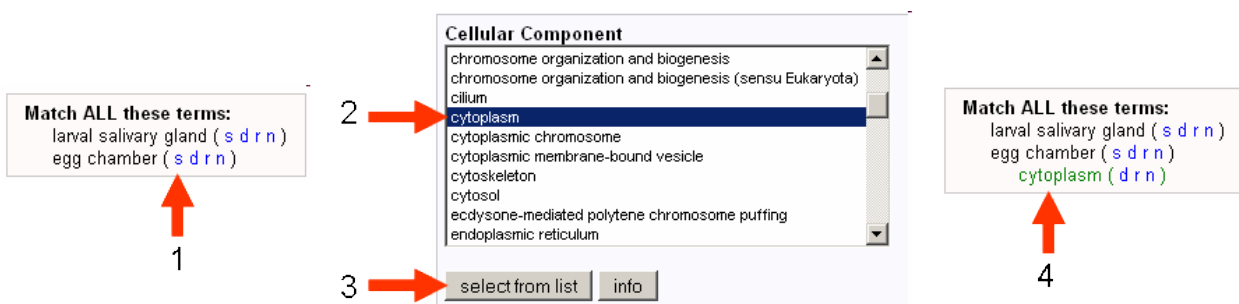
Adding sub cellular and descriptor terms

Adding subcellular, descriptor and NOT terms to your query tree is achieved by clicking on the various letters next to the anatomy term.



Adding a subcellular term

1. Click on the 's' next to *egg chamber* to load up the subcellular menu
2. The cellular component list appears in the CV pane at the bottom right hand side. Click on '*cytoplasm*' to select it
3. To add this term to your query tree, click on '*select from list*'

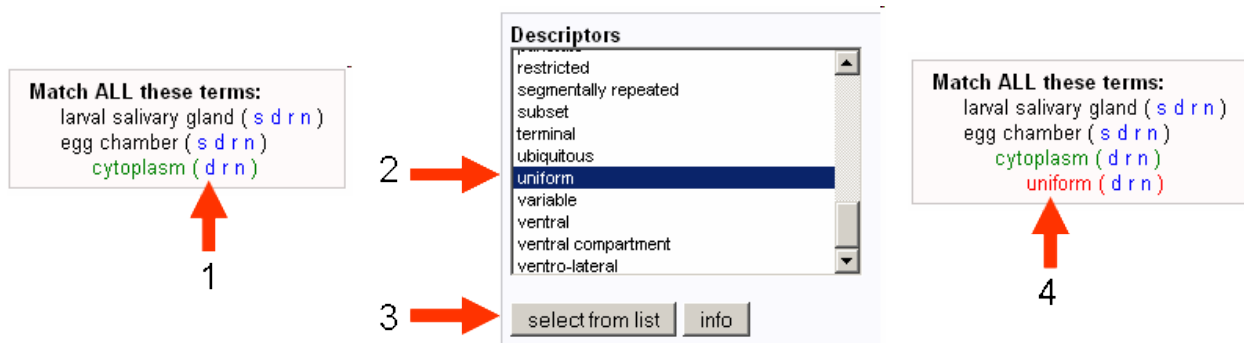


4. Your term now appears as a branch under the egg chamber anatomy term.

Adding a descriptor term

1. Click on the 'd' next to *cytoplasm* to load up the descriptor menu
2. The descriptor list appears in the CV pane at the bottom right hand side. Click on '*uniform*' to select it
3. To add this term to your query tree, click on '*select from list*'

4. Your term now appears as a branch under *egg chamber* -> *cytoplasm*.



After adding *cytoplasm* to *larval salivary gland* your final query should now look like this.

Query Trees:
Anatomy terms further down in the ontology tree will automatically be included in the search (eg searching for 'testis' will also search for 'spermatocyte')

Match ALL these terms:

- larval salivary gland (s d r n)
- cytoplasm (d r n)
- egg chamber (s d r n)
- cytoplasm (d r n)
- uniform (d r n)

Match ANY of these terms:

► Filter results by ANY developmental stage

Start query

Running the query

1. Click on *Start query* to run the query with the selected query tree.
2. After a few moments the results should appear showing the stocks that matched the search term tree. Stock names are shown with any sequencing data and relevant image thumbnails. Stocks and thumbnails are hot-linked to the stock and image reports respectively.

Query results:

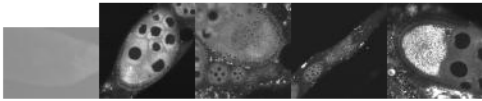

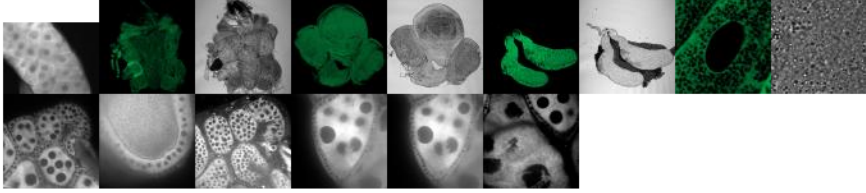
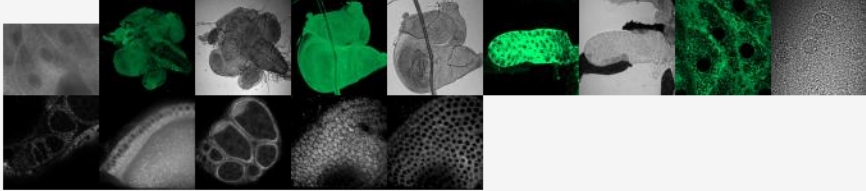
Tree query results

Stocks matching ALL of these CV terms:

larval salivary gland
cytoplasm
egg chamber
cytoplasm
uniform

Stocks matching ANY of these CV terms:

Filtered on ANY of these developmental stages

Stock Name	Sequencing Data	Images
CPTI-000056	Loc: 7584480, Chr: 3R, Gene:CG17342 Loc: 7584482, Chr: 3R, Gene:CG17342	
CPTI-000328	Loc: 18732001, Chr: 2R, Gene:CG17342	
CPTI-000609	Loc: 12437501, Chr: 2L, Gene:CG6382	
CPTI-000688	Loc: 15133815, Chr: 3L, Gene:CG6988	

Congratulations, you should have now completed your first successful query!

Advanced complex queries

The above example gives a taster on how to make complex queries. There are however a few ways you can refine your searches which are discussed below.

ALL and ANY searches.

The ALL and ANY search boxes allow you to do queries in a similar fashion to AND or OR queries in other search engines.

The ALL box looks for matches in stocks that have annotations that match every anatomy branch of the query tree. The ANY box looks for matches in stocks that match at least one of the anatomy branches.

For example:

Match ALL these terms: egg chamber (s d r n) cytoplasm (d r n) uniform (d r n)	Match ANY of these terms: testis (s d r n) hub cell (s d r n) nucleus (d r n)
--	---

...written out in long hand would be *“display stocks that show uniform expression in the cytoplasm of the egg chamber, and also express in either the testis and/or the hub cell nucleus”*.

Filter by developmental stage

It may be that you are only interested in protein expression in a particular developmental stage. This option allows you to filter the results on ANY developmental stage or stages (ALL queries are not yet supported) that you add to your chosen stages of interest.

▼ Filter results by ANY developmental stage

Developmental stages	Your chosen stages
pre-blastoderm stage	adult stage
blastoderm stage	
gastrula stage	
early extended germ band stage	
late extended germ band stage	

Add >> Remove

Using the NOT qualifier

To add a not qualifier click on the ‘n’ next to the term you which to add a NOT branch to. Any sub branch under the NOT branch will be considered as part of the query of where something isn’t expressed.

Example:

“Display all stocks that show uniform but not punctuate expression in the egg chamber cytoplasm, and do not show expression in the nucleus”

Match ALL these terms:
egg chamber (s d r n)
cytoplasm (d r n)
uniform (d r n)
NOT (d r)
punctate (d r n)
NOT (s d r)
nucleus (d r n)

Stocks must have been originally annotated with NOT in order for this option to work properly (ie the annotator definitely did not see any expression there and flagged it as such). Adding a NOT qualifier will not simply retrieve annotations where the annotator didn't look at a particular area or expression pattern. Although this makes the system slightly less flexible you should at least know exactly what you are getting as an answer.

Another example

“Display all stocks that show uniform but not punctuate expression in the egg chamber nucleus”

Match ALL these terms:
egg chamber (s d r n)
nucleus (d r n)
uniform (d r n)
NOT (d r)
punctate (d r n)

General searches.

You can pretty much make queries as detailed as you like, but this relies on the original annotations being as detailed to start with. To make a more general search use high level terms and edit your query appropriately if the number of hits is too high. The ultimate high level term is 'whole organism' which should cover pretty much everything. Your query may take quite a long while to run!

For more general searches you may be much better off using the simple query option instead.