DGIdb walkthrough

Obi Griffith, Malachi Griffith and Nicholas Spies

24 July, 2013

Introduction

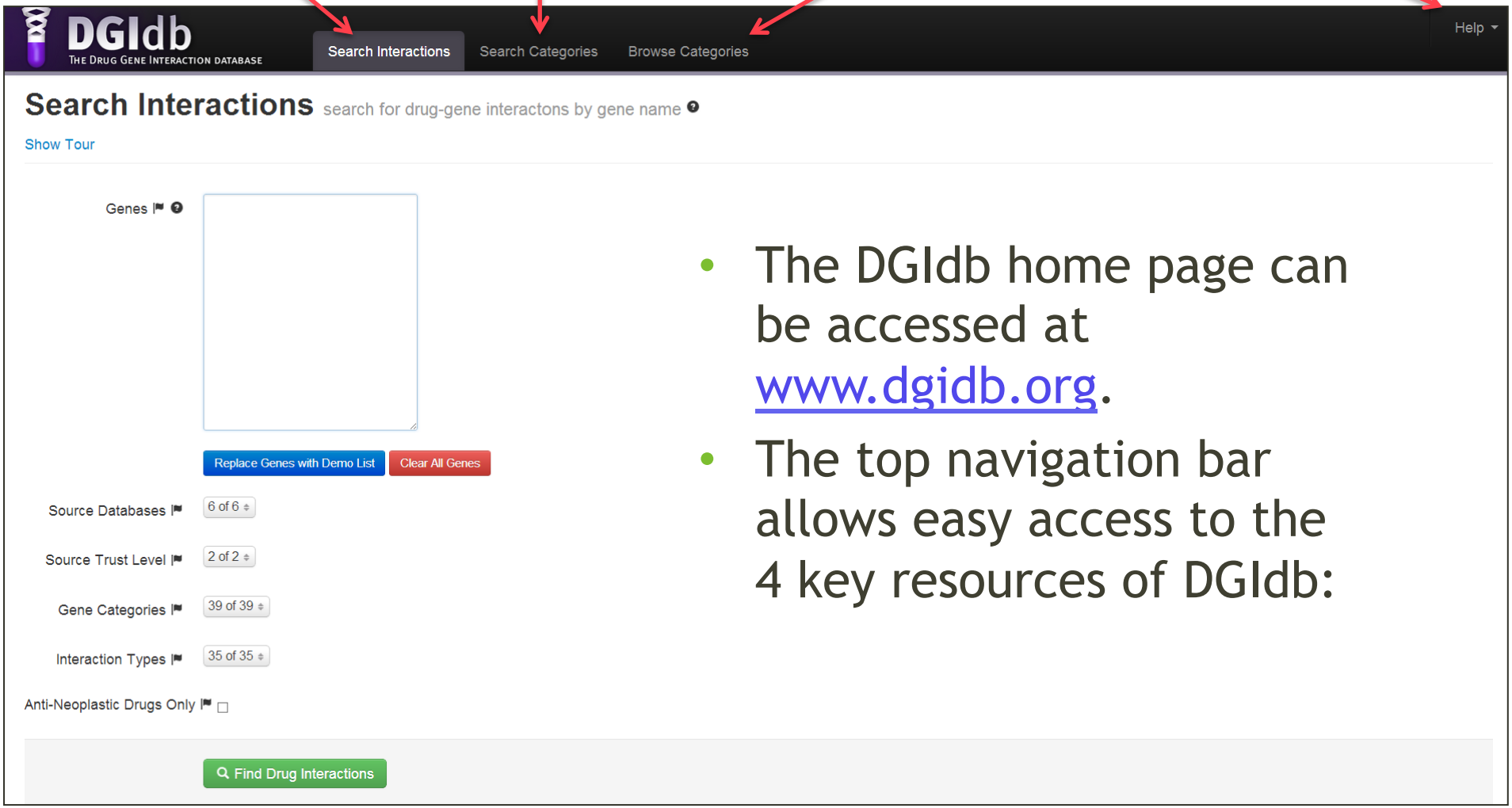
- DGIdb is a database and web interface for identifying known and potential drug-gene relationships.
- Genes are defined initially by Entrez Gene but then expanded and grouped together with other gene definitions such as Ensembl as well as genes from drug-gene interactions and potentially druggable gene categories.
- Drugs are defined initially by PubChem but then expanded and grouped together with drugs from drug-gene interactions.
- Drug-gene interactions come from multiple sources including DrugBank, TTD, PharmGKB, and others.
- Potentially druggable gene categories (e.g., Kinase) are defined by GO, dGene, and published lists such as Russ & Lampel (2005) and Hopkins and Groom (2002).

Outline

- This walkthrough will cover basic functionality of the DGIdb web interface covering the following topics.
  1. General searches for genes of interest.
  2. Identifying known drug-gene interactions for a list of genes of interest.
  3. Identifying potentially druggable genes for a list of genes of interest.
  4. Browsing lists of potentially druggable gene categories.
  5. Getting help.

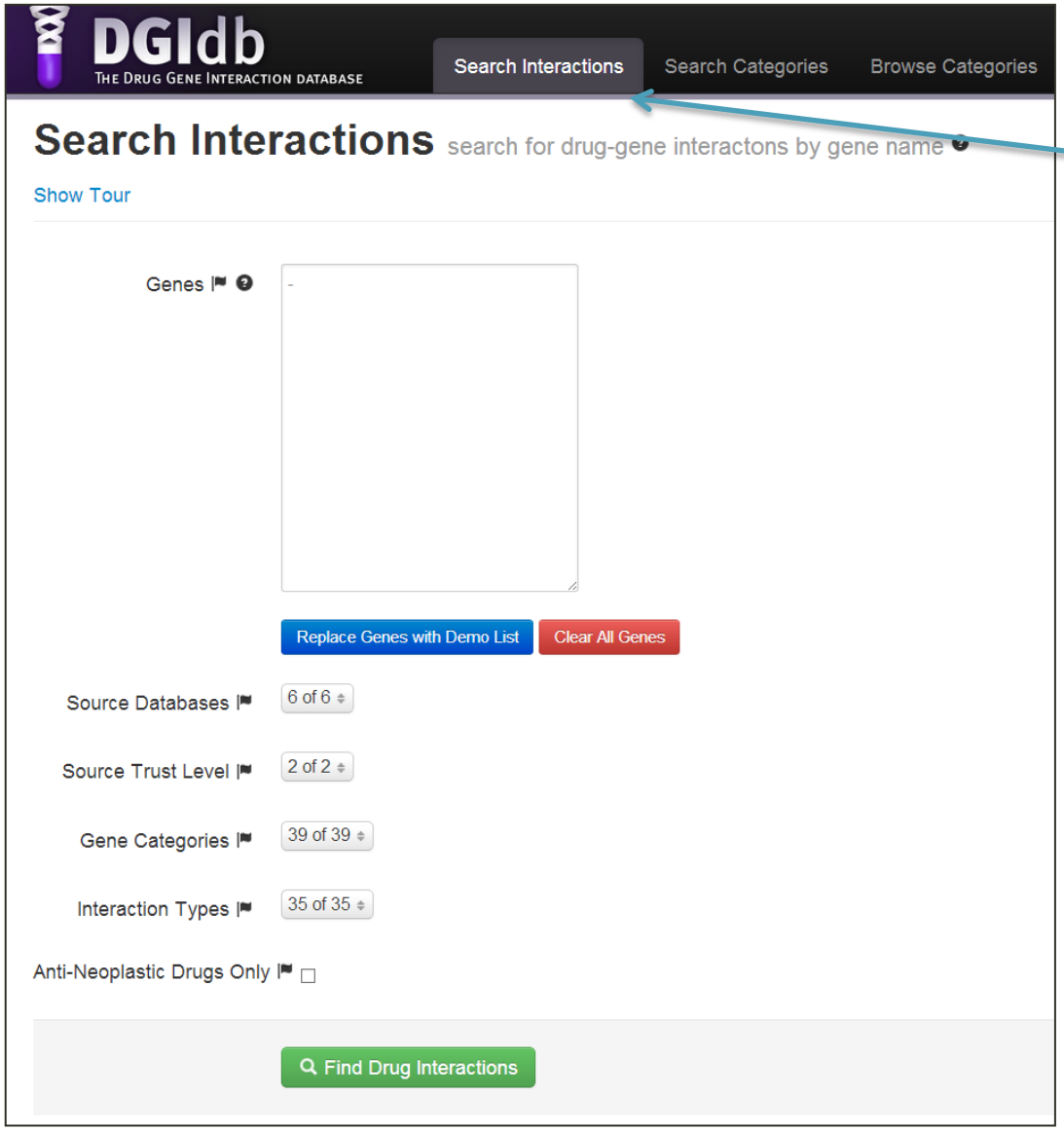
The DGIdb home page

1. Search Interactions
2. Search Categories
3. Browse Categories
4. Help



- The DGIdb home page can be accessed at www.dgidb.org.
- The top navigation bar allows easy access to the 4 key resources of DGIdb:

1. Performing a general search for druggable genes in DGIdb

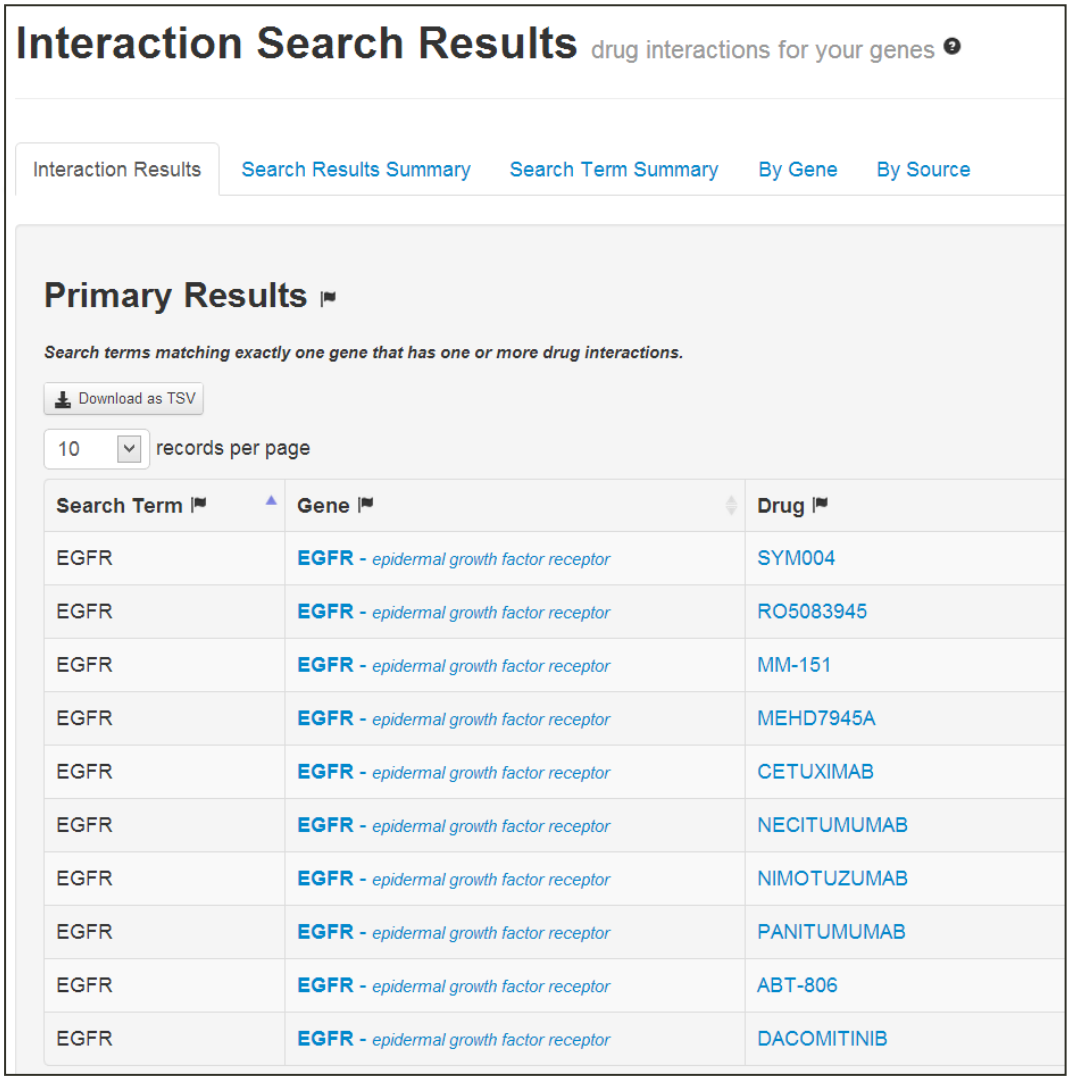


- Select ‘Search Interactions’ from the top navigation bar, enter a gene of interest, then click the green ‘Find Drug Interactions’ button.
- Help flags ( ) are found throughout the web interface. Hover over these icons with your mouse to view definitions of terms and other tips.

1. Performing a general search for druggable genes in DGIdb

- For an in-depth explanation of the various search filters and key features, click the ‘Show Tour’ button.
- The tour explains the 5 available filters that can be used for a search. These are below the search box, here.

Search results



Interaction Search Results drug interactions for your genes

Interaction Results | Search Results Summary | Search Term Summary | By Gene | By Source

Primary Results

Search terms matching exactly one gene that has one or more drug interactions.

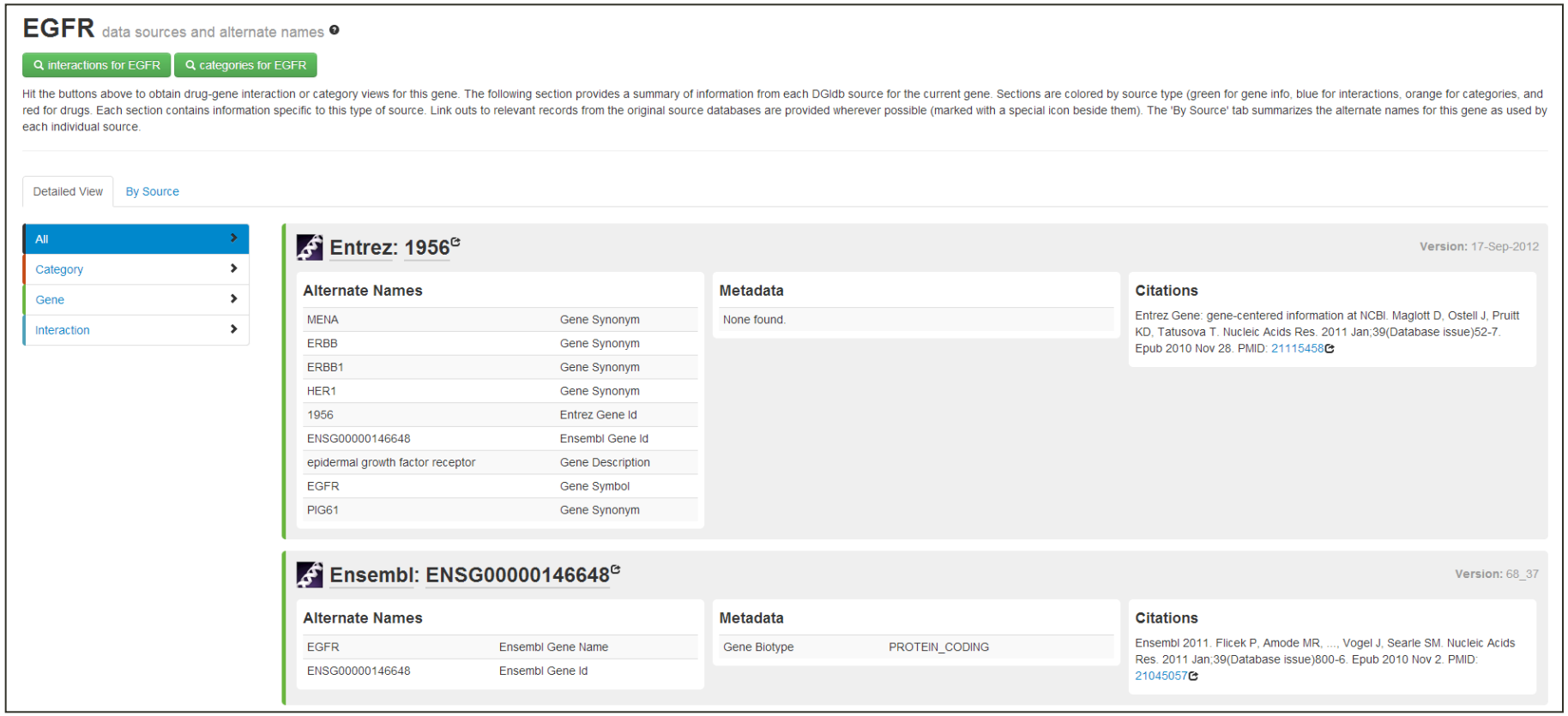
Download as TSV

10 records per page

Search Term	Gene	Drug
EGFR	**EGFR** - *epidermal growth factor receptor*	SYM004
EGFR	**EGFR** - *epidermal growth factor receptor*	RO5083945
EGFR	**EGFR** - *epidermal growth factor receptor*	MM-151
EGFR	**EGFR** - *epidermal growth factor receptor*	MEHD7945A
EGFR	**EGFR** - *epidermal growth factor receptor*	CETUXIMAB
EGFR	**EGFR** - *epidermal growth factor receptor*	NECITUMUMAB
EGFR	**EGFR** - *epidermal growth factor receptor*	NIMOTUZUMAB
EGFR	**EGFR** - *epidermal growth factor receptor*	PANITUMUMAB
EGFR	**EGFR** - *epidermal growth factor receptor*	ABT-806
EGFR	**EGFR** - *epidermal growth factor receptor*	DACOMITINIB

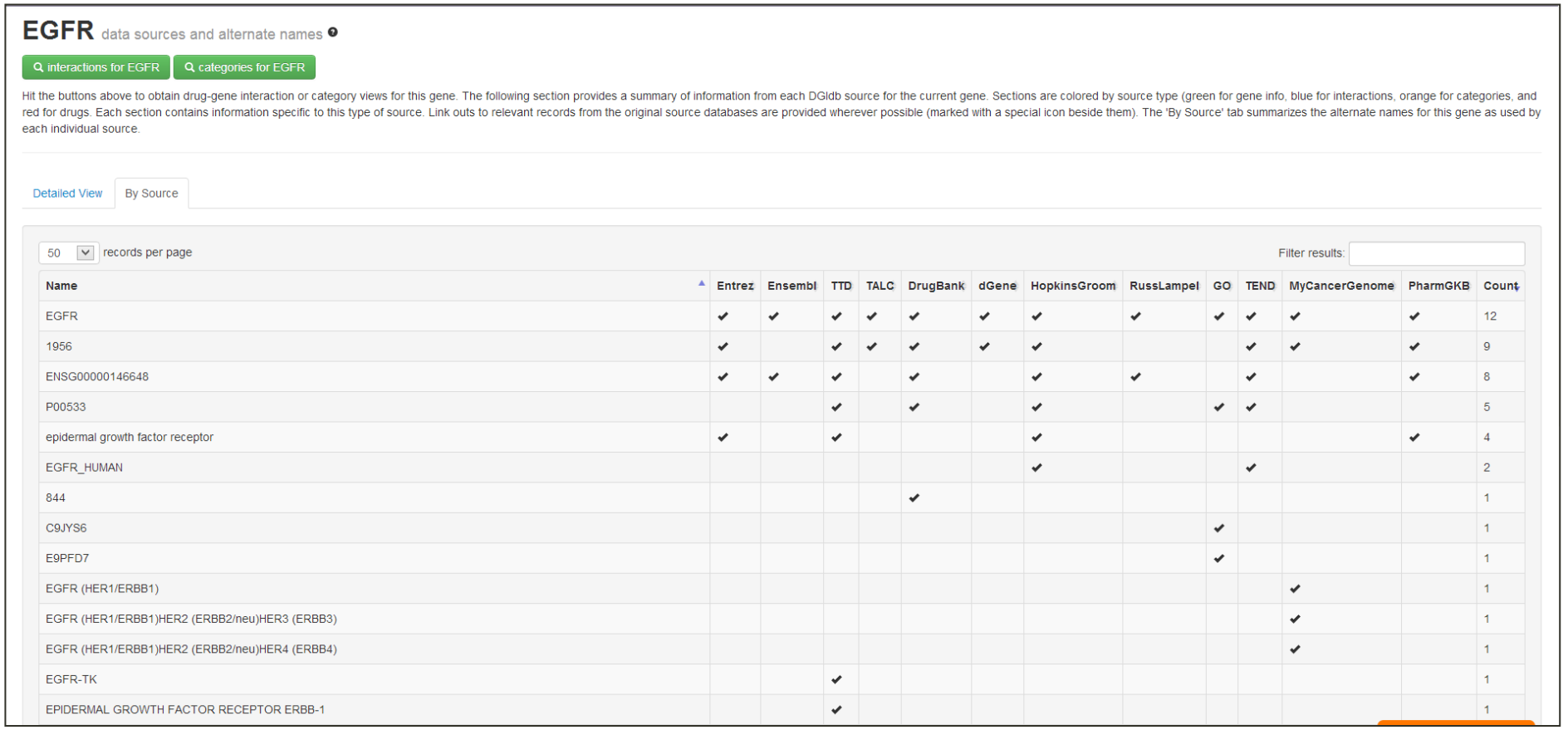
- Search Results display all entities within the database that match the search term
- In this case, the search term matched an exact entry for the gene group‘EGFR’.
- The DGIdb Gene Group for EGFR is the result most likely to be useful. Select the linked text for ‘EGFR-epidermal growth factor receptor’ to see a summary of this gene in DGIdb.

Gene summary page



- The gene summary page shows all information and sources for a gene group in DGIdb.
- A second view/tab summarizes all gene names and synonyms for the gene group by their source (see next slide).

Gene summary by source



EGFR data sources and alternate names

interactions for EGFR | categories for EGFR

Hit the buttons above to obtain drug-gene interaction or category views for this gene. The following section provides a summary of information from each DGIdb source for the current gene. Sections are colored by source type (green for gene info, blue for interactions, orange for categories, and red for drugs. Each section contains information specific to this type of source. Link outs to relevant records from the original source databases are provided wherever possible (marked with a special icon beside them). The 'By Source' tab summarizes the alternate names for this gene as used by each individual source.

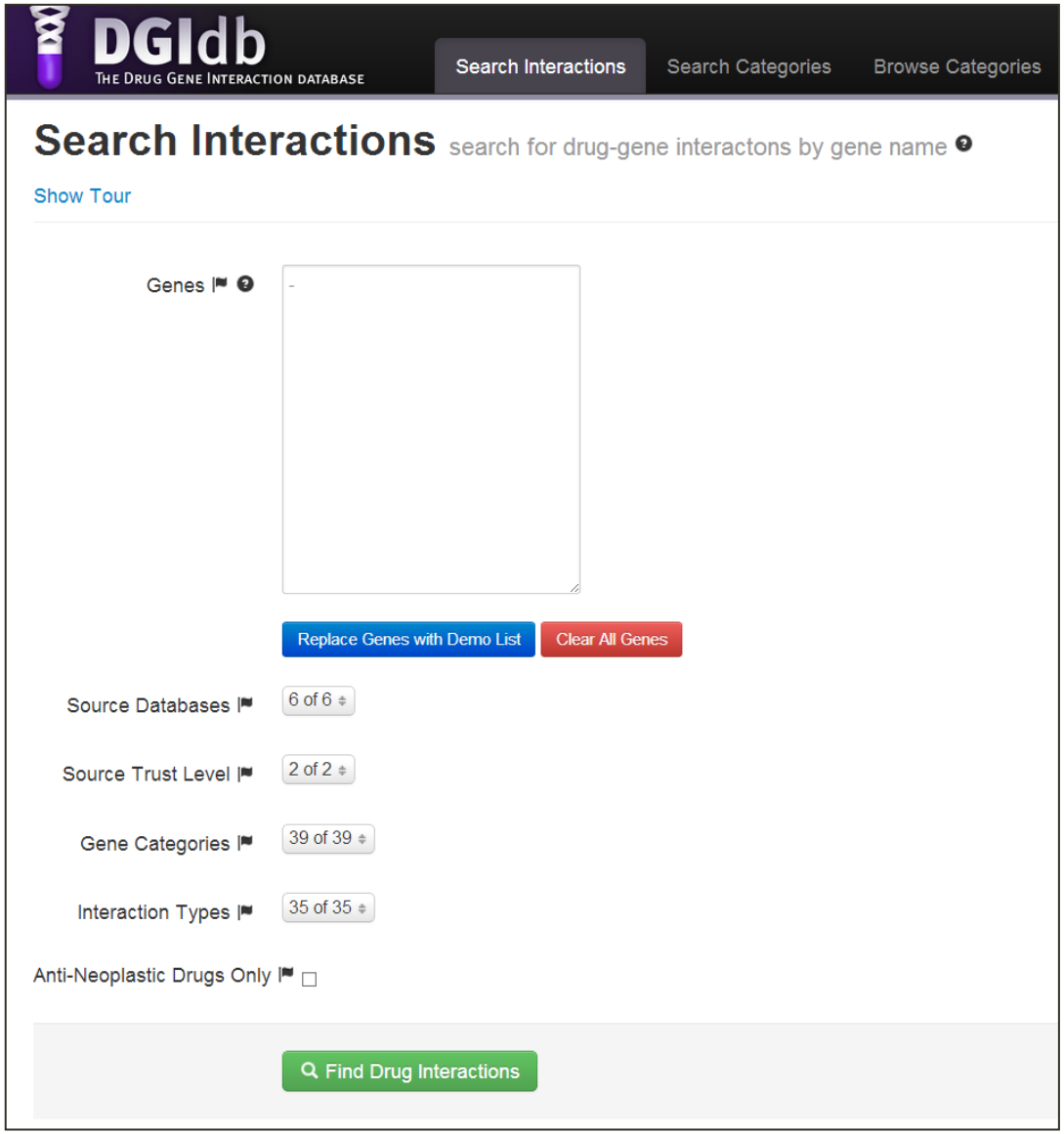
Detailed View | By Source

50 records per page — Filter results:

Name	Entrez	Ensembl	TTD	TALC	DrugBank	dGene	HopkinsGroom	RussLampel	GO	TEND	MyCancerGenome	PharmGKB	Count
EGFR	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	12
1956	✔		✔	✔	✔	✔	✔			✔	✔	✔	9
ENSG00000146648	✔	✔	✔		✔		✔	✔		✔		✔	8
P00533			✔		✔		✔		✔	✔			5
epidermal growth factor receptor	✔		✔				✔					✔	4
EGFR_HUMAN							✔			✔			2
844					✔								1
C9JYS6									✔				1
E9PFD7									✔				1
EGFR (HER1/ERBB1)											✔		1
EGFR (HER1/ERBB1)HER2 (ERBB2/neu)HER3 (ERBB3)											✔		1
EGFR (HER1/ERBB1)HER2 (ERBB2/neu)HER4 (ERBB4)											✔		1
EGFR-TK			✔										1
EPIDERMAL GROWTH FACTOR RECEPTOR ERBB-1			✔										1

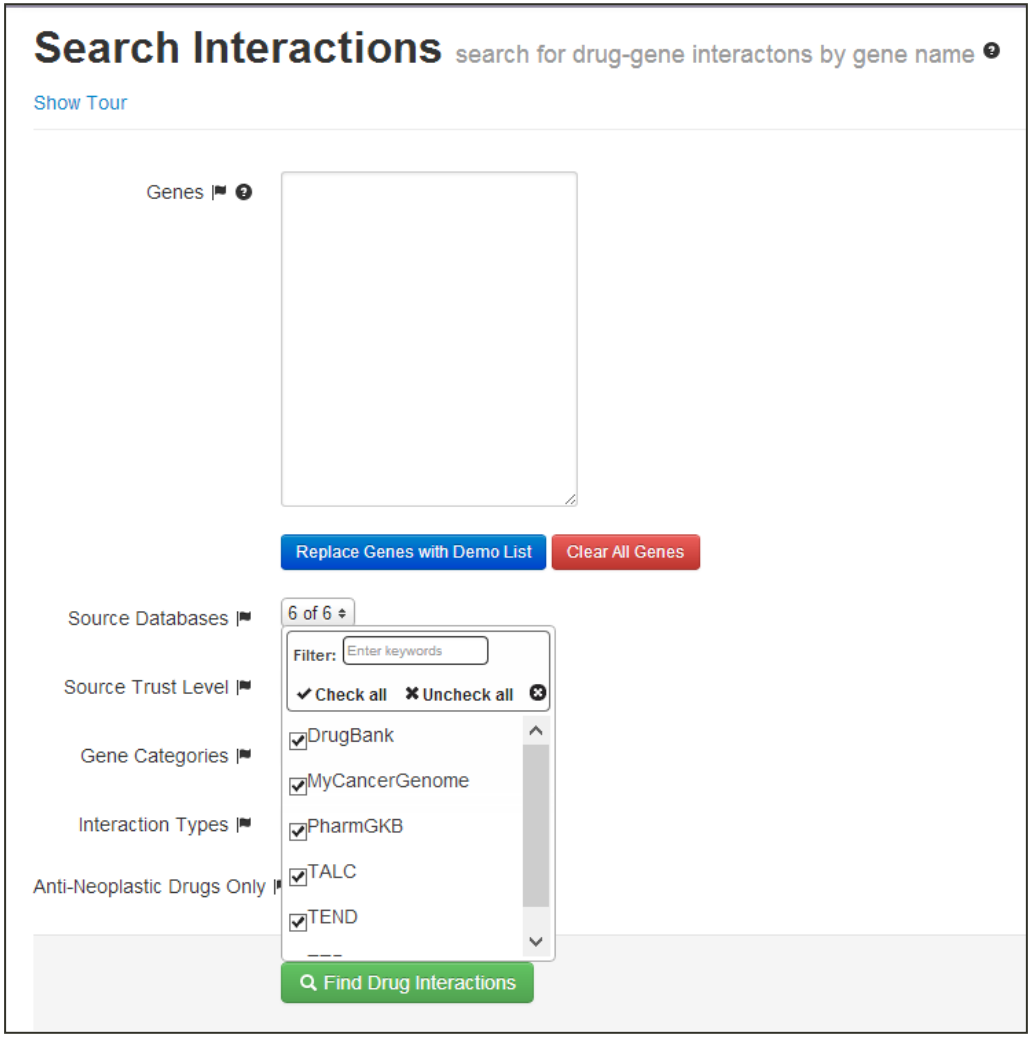
- Gene group (EGFR) summarized by source. This view is accessed by clicking the “By Source” tab next to the “Detailed View” tab.

2. Searching for known drug-gene interactions



- 'Search interactions' allows the user to input one or more genes and determine whether those genes are thought to be targeted by (or interact with) any drugs.
- Genes can be typed/pasted into the 'Genes' box. A set of default genes can also be entered for illustrative purposes.
- Results can be limited to any or all of the interaction source databases and filtered according to several pre-defined criteria, defined on the next slide.
- Once the gene list has been entered or uploaded, select the 'Find Drug Interactions' button to view results of the search.

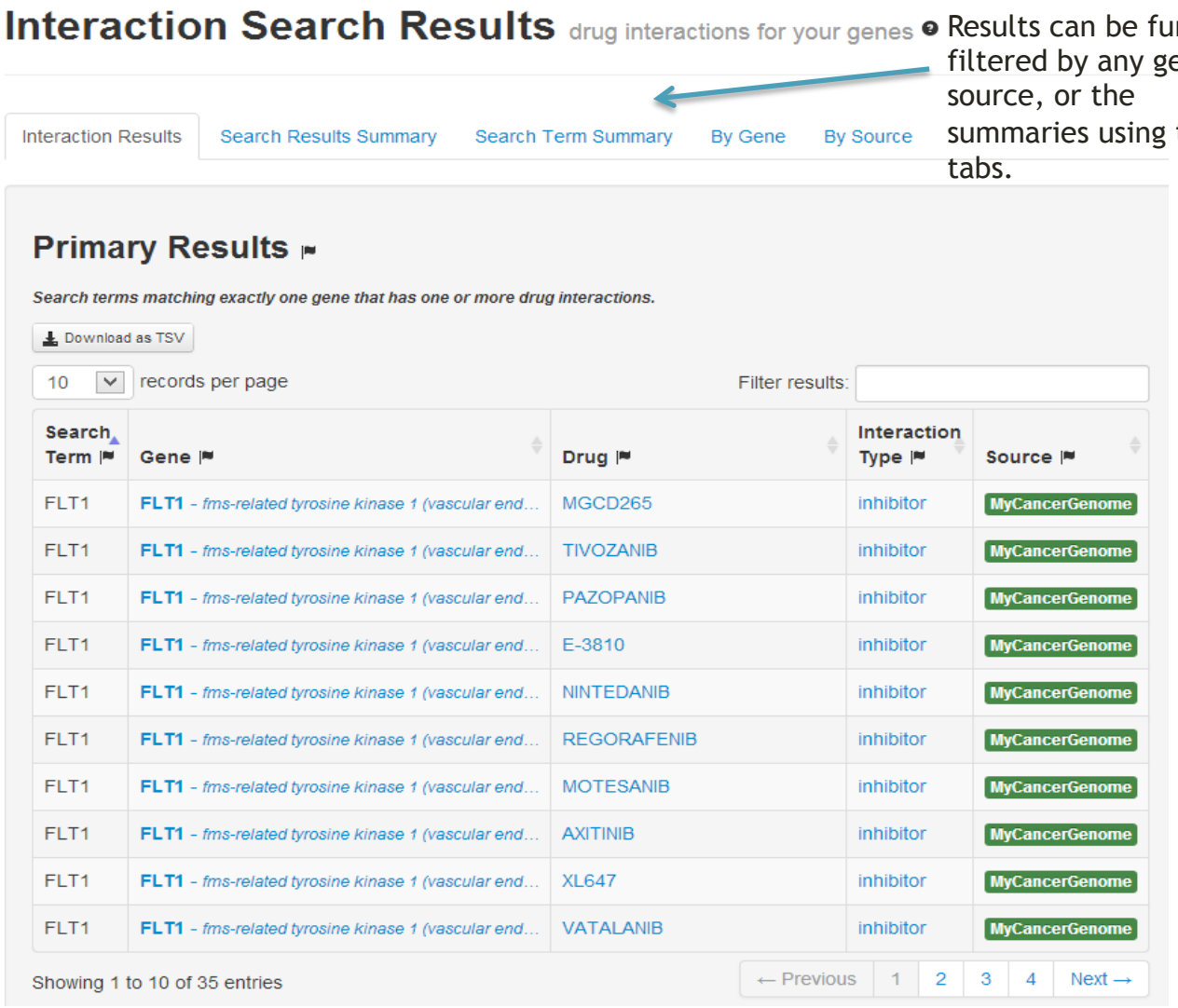
Search Interactions: Filters



- When searching for gene-drug interactions, the results can be put through 5 available filters.
- Upon clicking on the box, a drop-down menu appears with the available options. Shown is the filter for Source Databases. Clicking each option toggles the check, including or excluding that source in the data output.
- The Source Trust Level filter can be used to only included ‘Expert Curated’ or ‘non-curated’ data.
- The Gene Categories Filter allows the results to include only the gene functions desired (ABC Transporters, Cell Signalling.. Etc.)
- The Interaction Types filter allows for the search to include only specific drug action types to be included (inhibitor, allosteric modulator… Etc.)
- Checking the ‘Anti-Neoplastic Drugs Only’ filter will include only drugs proven to be for the treatment of Cancers.

Interaction search results

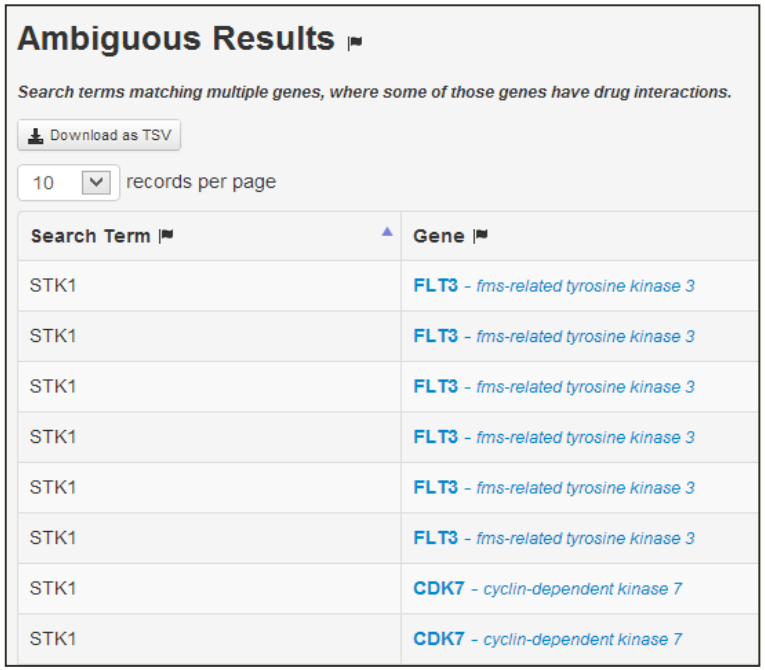
Results are organized according to five different views/tabs. The main view shown here gives a complete detailed result of all the drug-gene interactions matching the search terms (genes) provided. For example, it was found that the search term “FLT1” unambiguously matches a DGIdb gene (FLT1), passed the basic filter, and is thought to be inhibited by a number of drugs according to the MyCancerGenome database.



Results can be further filtered by any gene, source, or the summaries using these tabs.

Interaction search results (cont'd)

(1)



Ambiguous Results

Search terms matching multiple genes, where some of those genes have drug interactions.

Download as TSV

10 records per page

Search Term	Gene
STK1	**FLT3** - *fms-related tyrosine kinase 3*
STK1	**FLT3** - *fms-related tyrosine kinase 3*
STK1	**FLT3** - *fms-related tyrosine kinase 3*
STK1	**FLT3** - *fms-related tyrosine kinase 3*
STK1	**FLT3** - *fms-related tyrosine kinase 3*
STK1	**FLT3** - *fms-related tyrosine kinase 3*
STK1	**CDK7** - *cyclin-dependent kinase 7*
STK1	**CDK7** - *cyclin-dependent kinase 7*

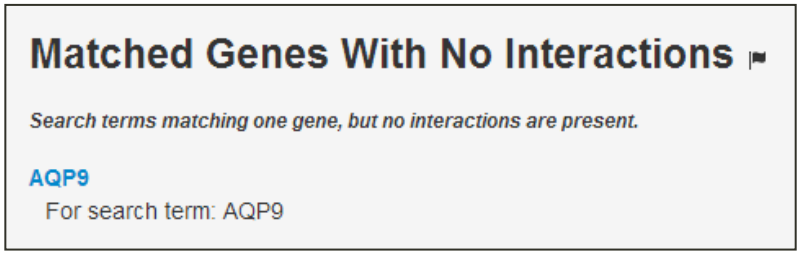
(2)

Ambiguous Search Terms With No Interactions

Search terms matching multiple genes, but none have interactions.

PFDN5
For search term: MM1
PLXNB2
For search term: MM1

(3)



Matched Genes With No Interactions

Search terms matching one gene, but no interactions are present.

AQP9
For search term: AQP9

The main 'Interaction Results' view/tab also summarizes: (1) search terms with ambiguous matches to genes in DGIdb which have interactions that pass filter criteria; (2) search terms with ambiguous matches to genes in DGIdb but with no interactions; (3) search terms with definite matches to genes in DGIdb but with no interactions; (4) search terms with no matches to any genes in DGIdb

NOTE: It is advised that for any ambiguous matches, the input list should be modified to use the official gene symbol to eliminate ambiguity.

Examples shown are generated with the demo list available.

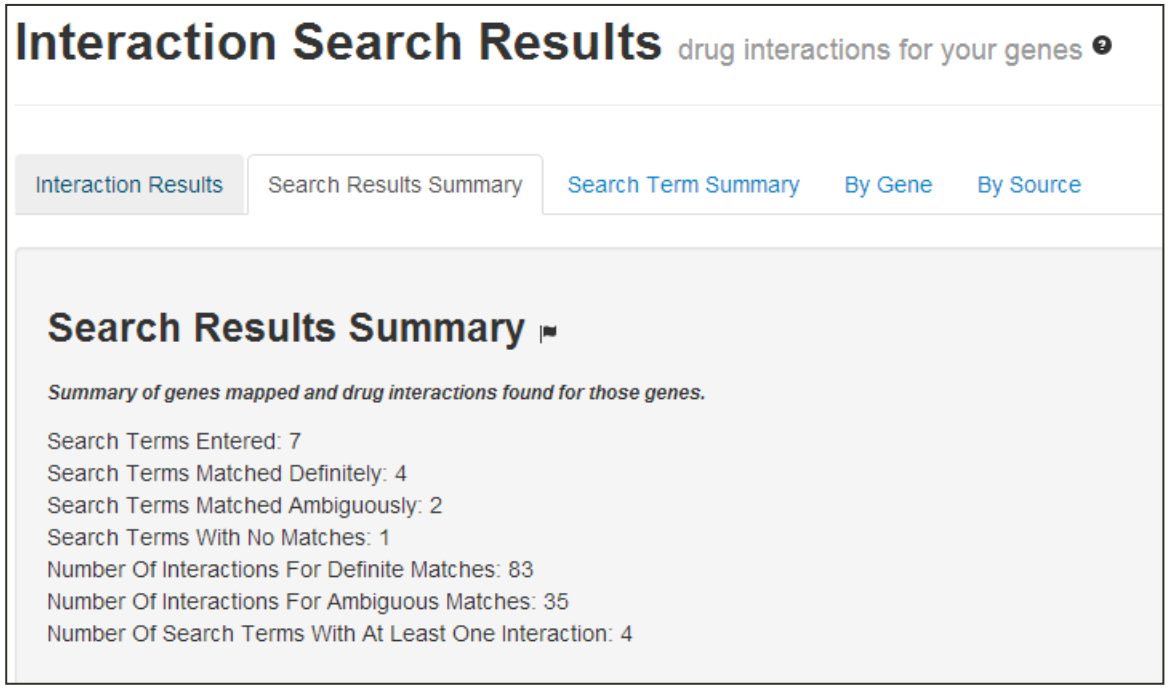
(4)

Search Terms With No Matches

Search terms that could not be matched to a gene

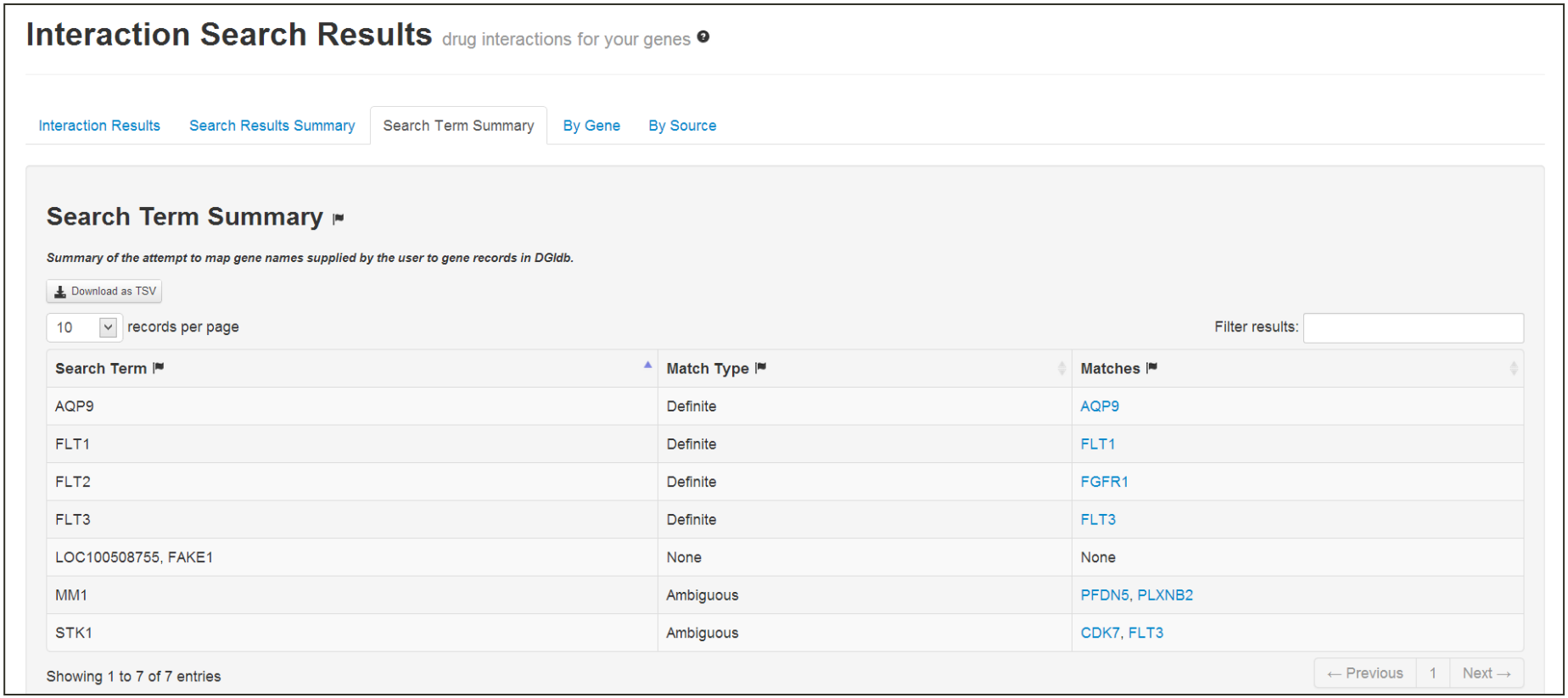
- LOC100508755, FAKE1

Interaction search results summary



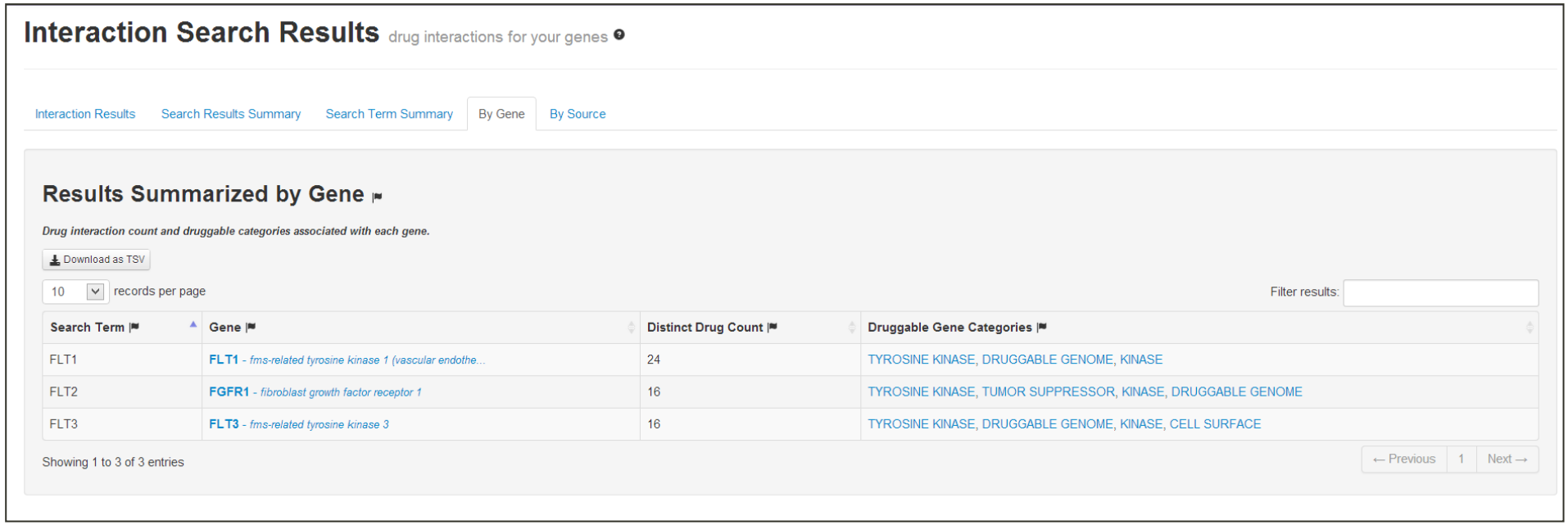
The second view/tab on the Interaction Search Results page provides basic statistics for the search term list such as: Number of search terms provided as input, number of terms with definite matches to genes in DGIdb, number of interactions for definite matches, etc.

Interaction search results term summary



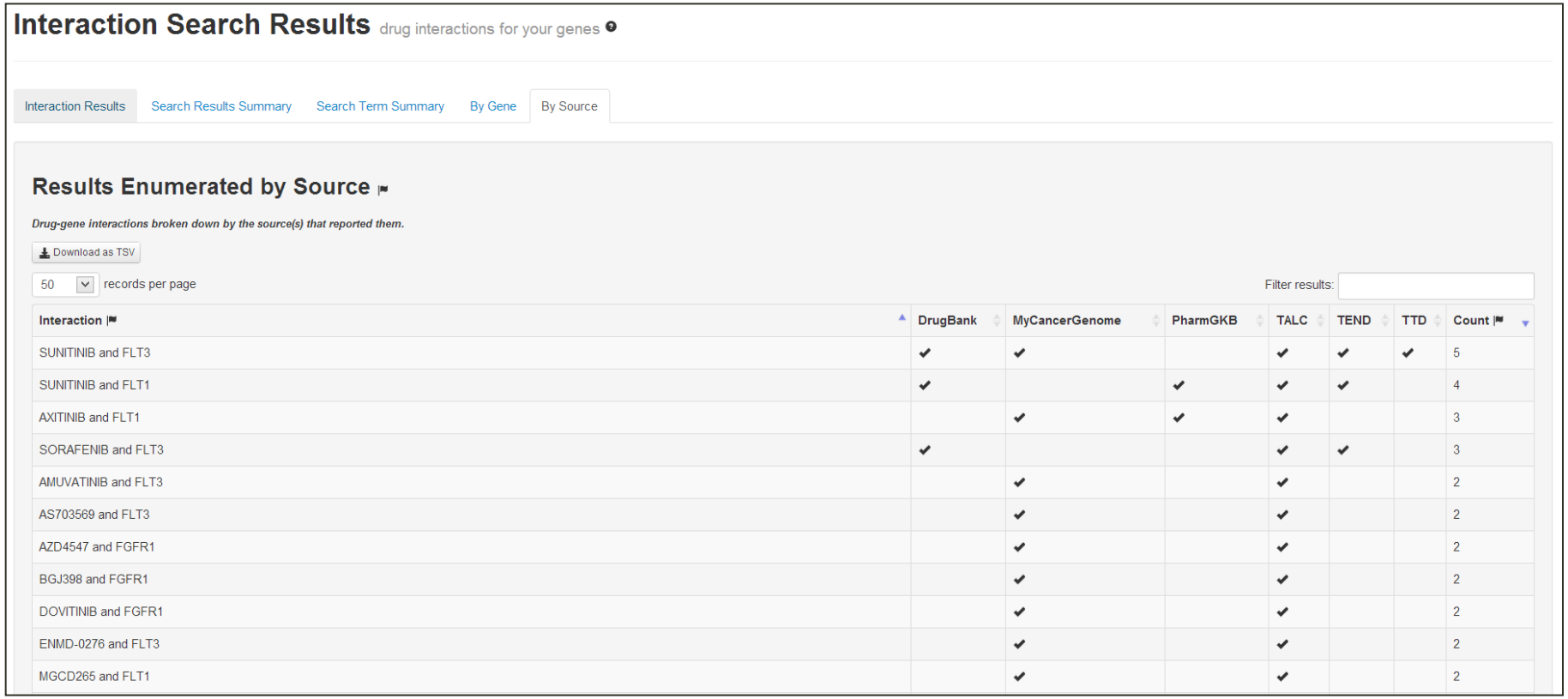
The third view/tab on the Interaction Search Results page provides a summary of how search terms matched results in DGIdb on a term-by-term basis.

Interaction search results gene summary



The Fourth Tab in the menu gives a gene-by-gene breakdown of interaction count for available drugs as well as the catalogged drug categories in the database.

Interaction search results by source



Interaction Search Results drug interactions for your genes

Interaction Results | Search Results Summary | Search Term Summary | By Gene | By Source

Results Enumerated by Source

Drug-gene interactions broken down by the source(s) that reported them.

Download as TSV

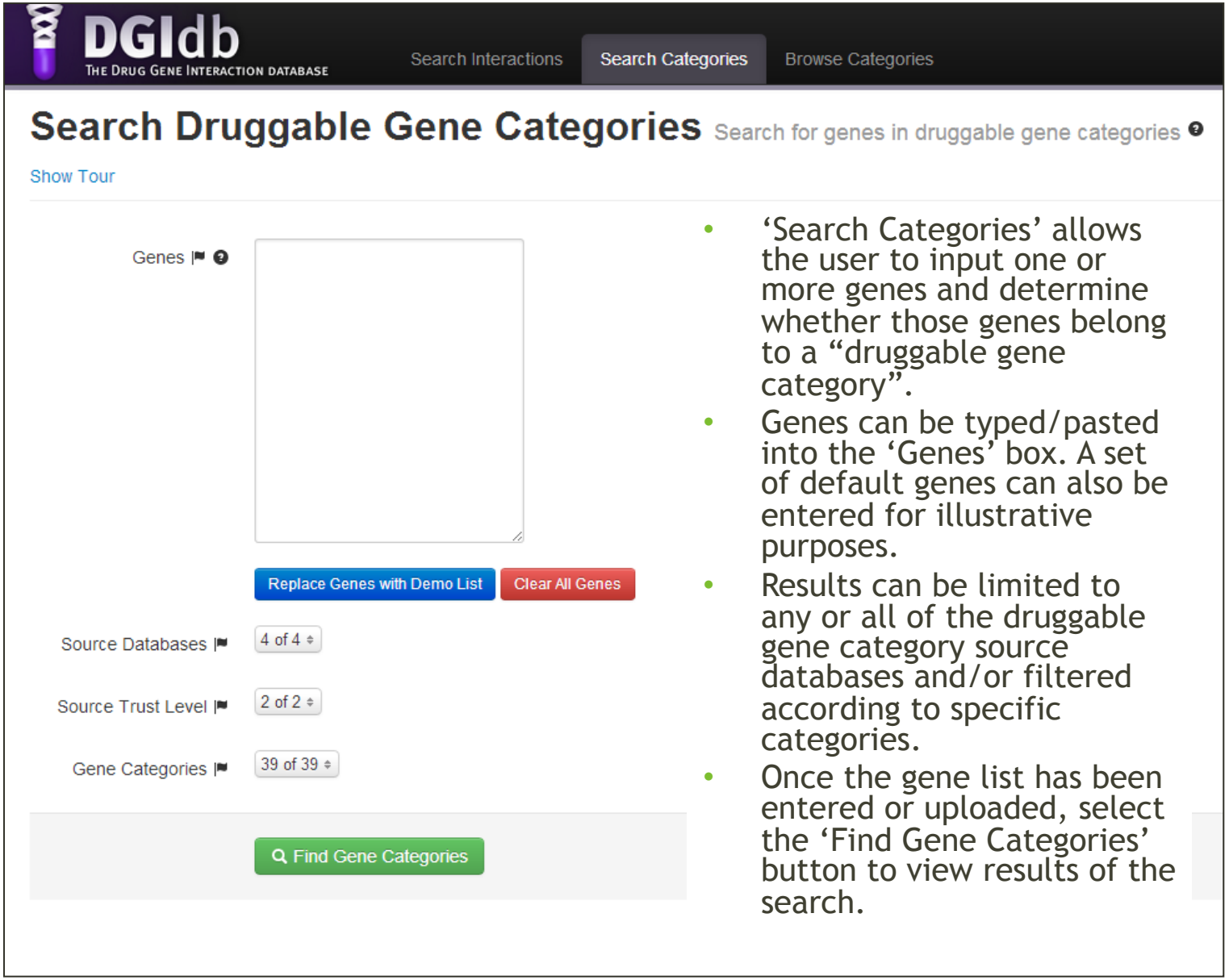
50 records per page

Filter results:

Interaction	DrugBank	MyCancerGenome	PharmGKB	TALC	TEND	TTD	Count
SUNITINIB and FLT3	✔	✔		✔	✔	✔	5
SUNITINIB and FLT1	✔		✔	✔	✔		4
AXITINIB and FLT1		✔	✔	✔			3
SORAFENIB and FLT3	✔			✔	✔		3
AMUVATINIB and FLT3		✔		✔			2
AS703569 and FLT3		✔		✔			2
AZD4547 and FGFR1		✔		✔			2
BGJ398 and FGFR1		✔		✔			2
DOVITINIB and FGFR1		✔		✔			2
ENMD-0276 and FLT3		✔		✔			2
MGCD265 and FLT1		✔		✔			2

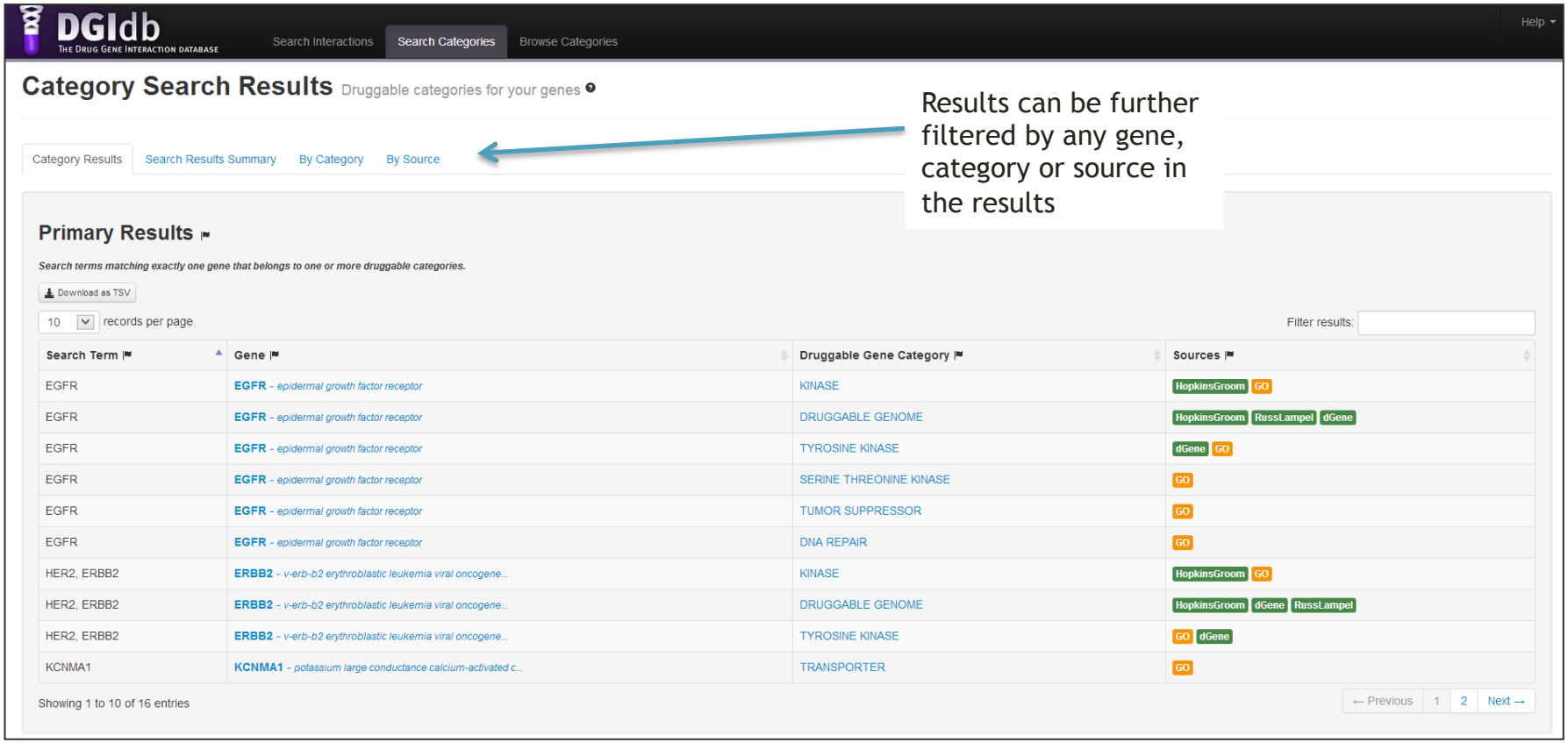
The fifth view/tab on the Interaction Search Results page provides a summary of unambiguous interaction results in DGIdb, grouped by source. To illustrate, the search term “FLT3” was found to interact with Sunitinib by DrugBank, MyCancerGenome, and TALC, TEND, TTD.

3. Searching for potentially druggable genes



- 'Search Categories' allows the user to input one or more genes and determine whether those genes belong to a "druggable gene category".
- Genes can be typed/pasted into the 'Genes' box. A set of default genes can also be entered for illustrative purposes.
- Results can be limited to any or all of the druggable gene category source databases and/or filtered according to specific categories.
- Once the gene list has been entered or uploaded, select the 'Find Gene Categories' button to view results of the search.

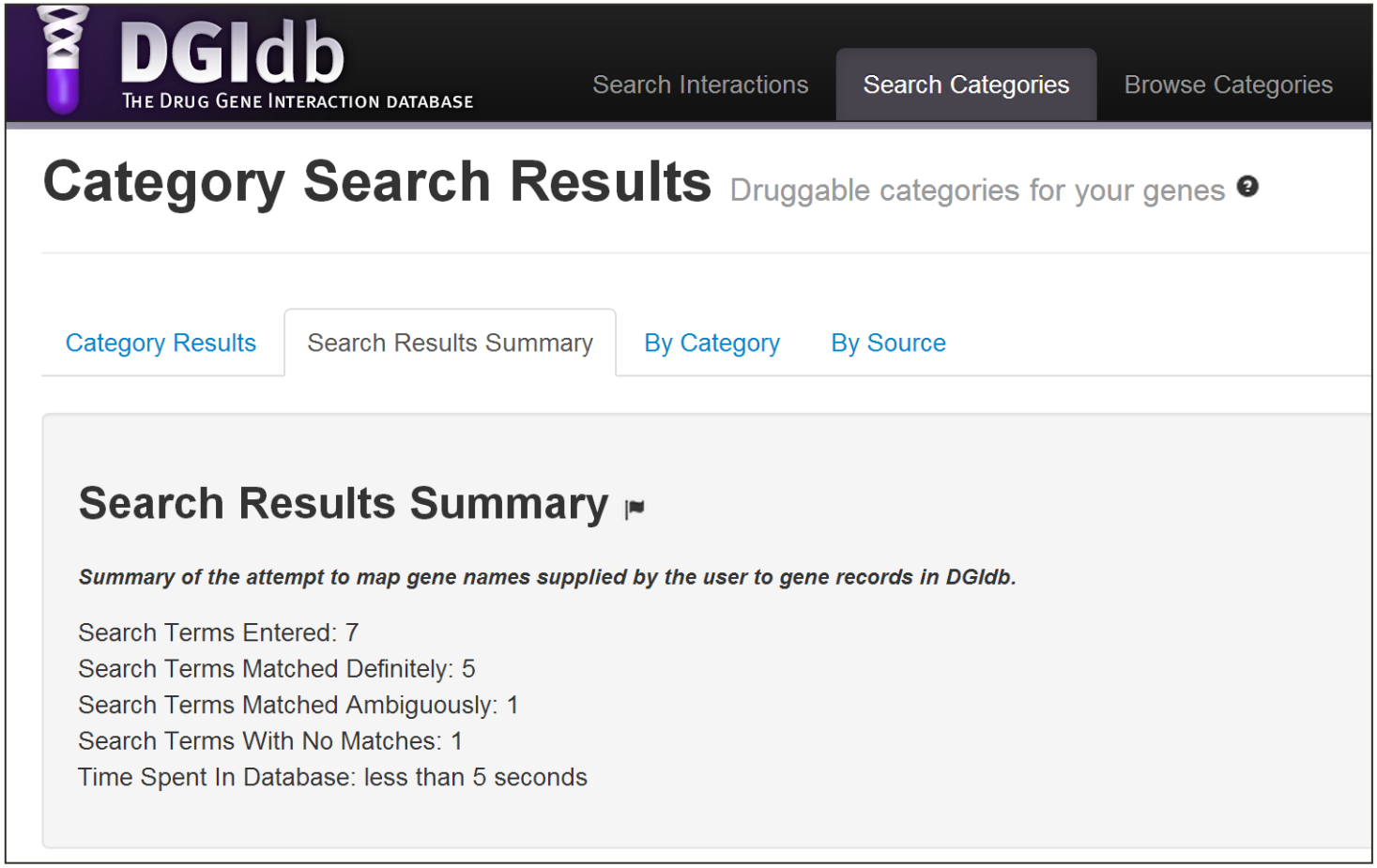
Category search results



Results are organized according to four different views/tabs. The main view shown here gives a complete detailed result of all the druggable gene categories matching the search terms (genes) provided. For example, it was found that the search term “EGFR” belongs to a number of potentially druggable gene categories such as Tyrosine Kinase, DNA repair, etc.

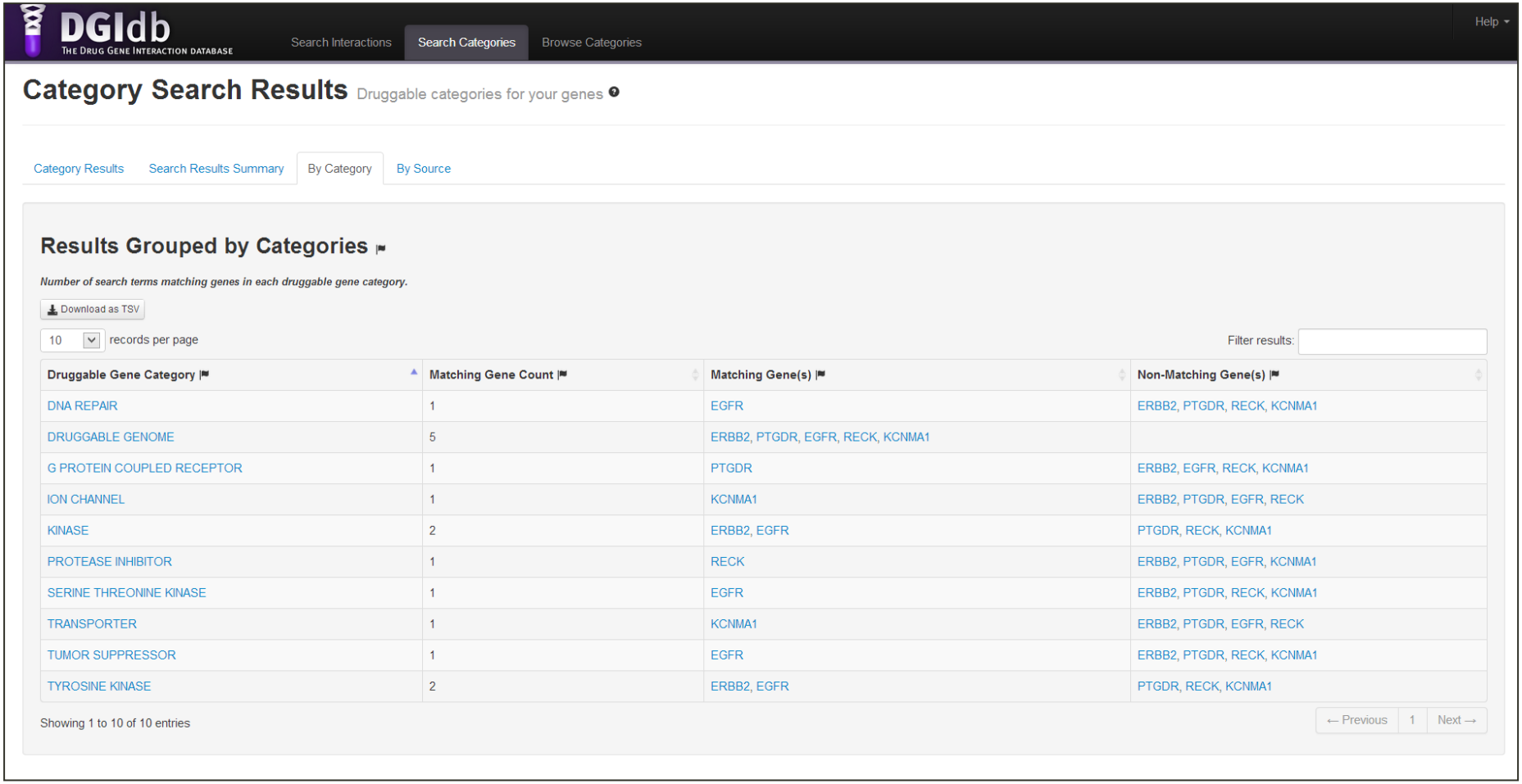
Ambiguous search terms and search terms with no matches are also summarized here

Category search results summary



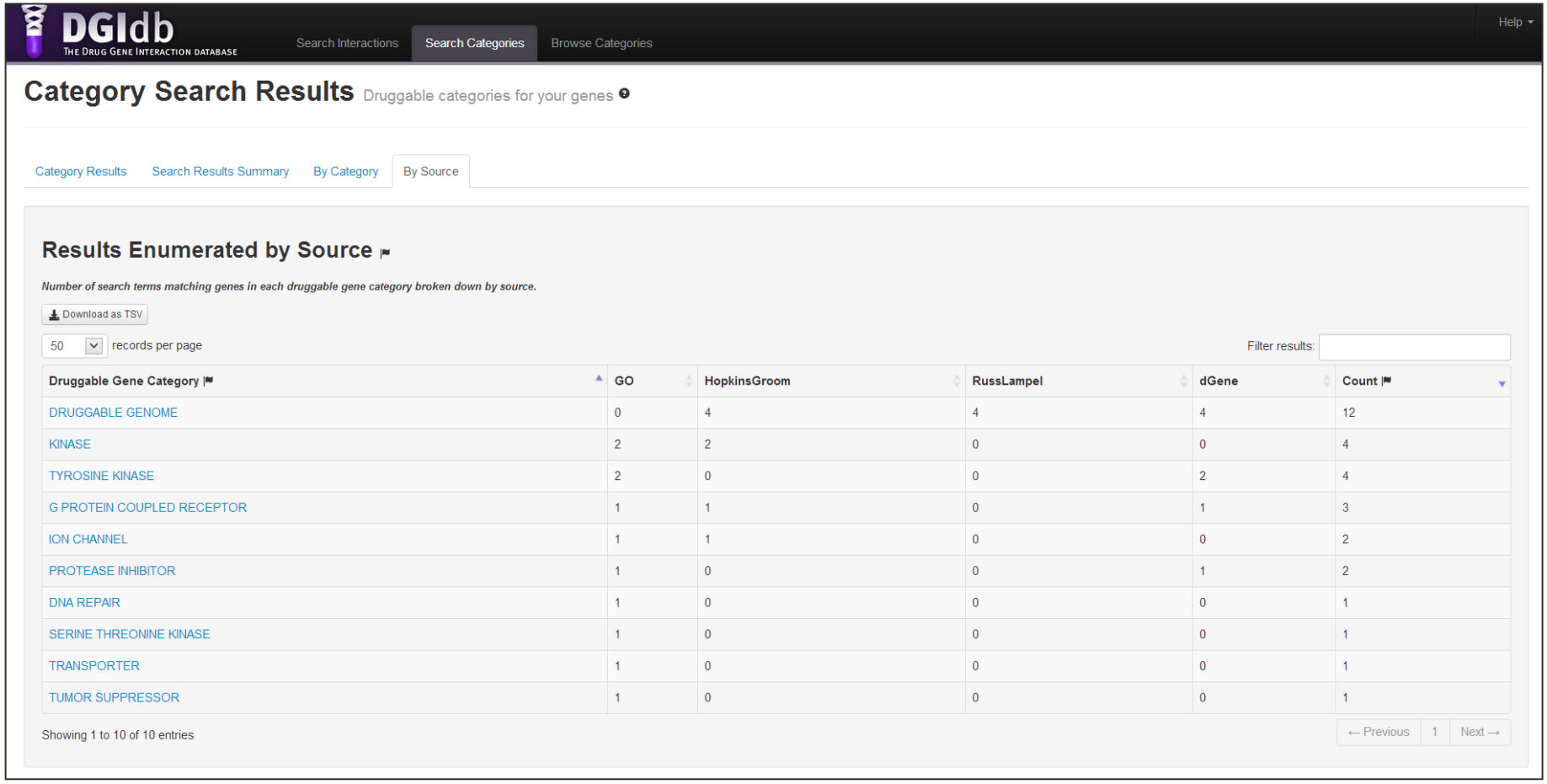
The second view/tab on the Category Search Results page provides basic statistics for the search term list such as: Number of search terms provided as input, number of terms with definite matches to genes in DGIdb, etc.

Category search results by category



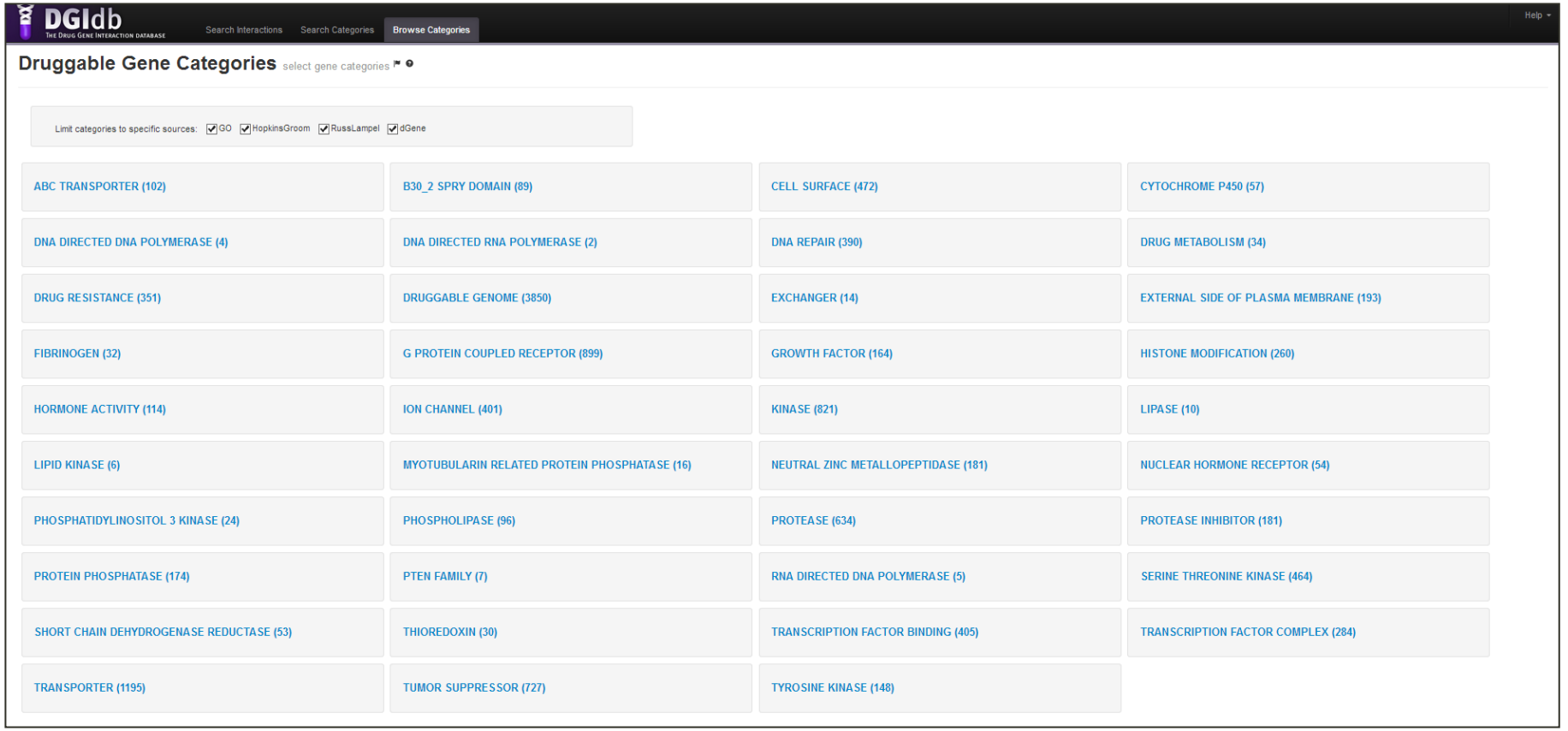
The third view/tab on the Category Search Results page provides a summary of how search terms matched results in DGIdb, grouped by druggable gene category. For example, it was found that two genes (ERBB2 and EGFR) were found in the KINASE category.

Category search results by source



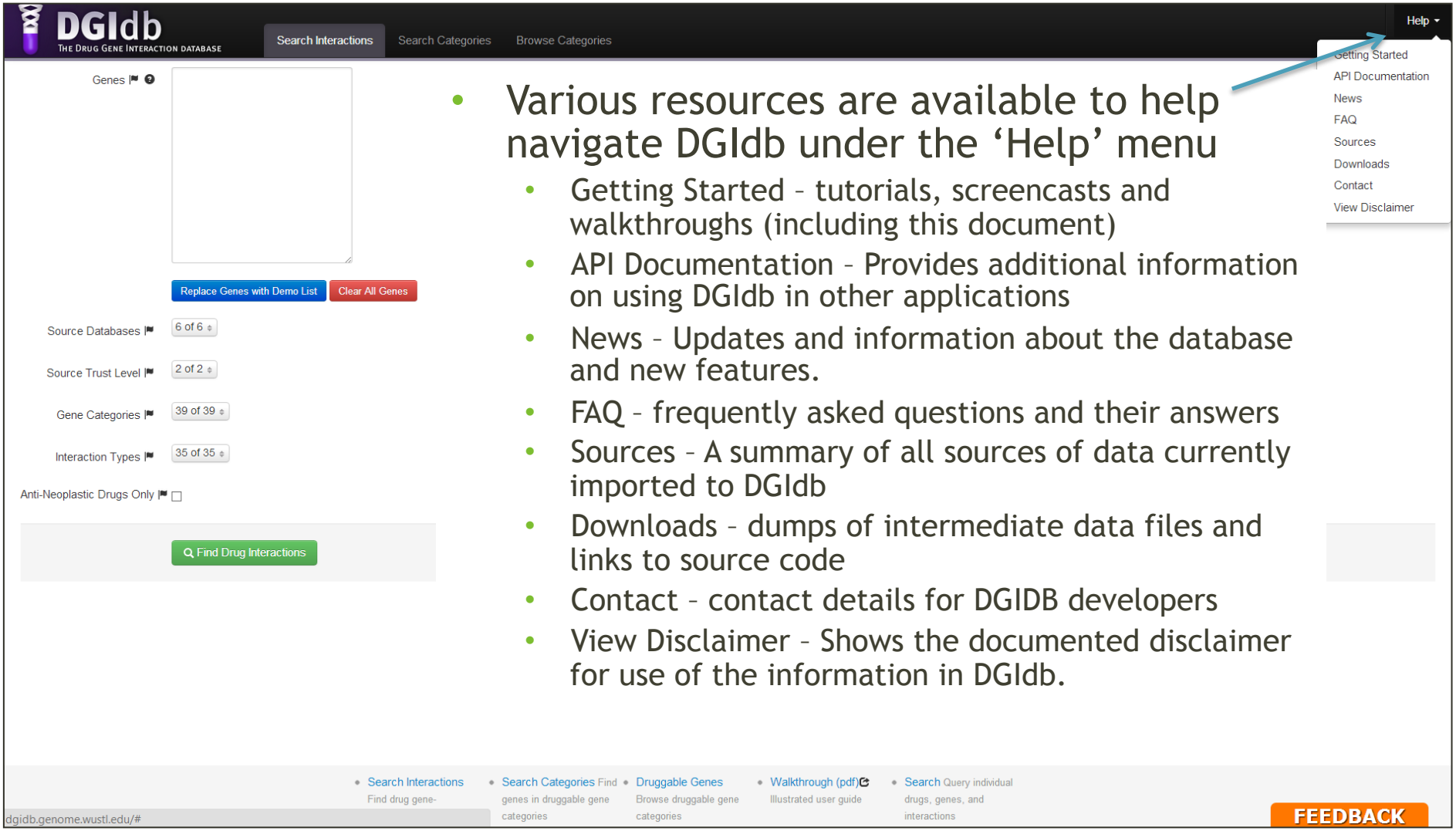
The fourth view/tab on the Category Search Results page provides a summary of how search terms matched results in DGIdb, grouped by interaction source. For example, KINASE category genes have interactions in both the GO and dGene databases.

4. Browsing potentially druggable gene categories



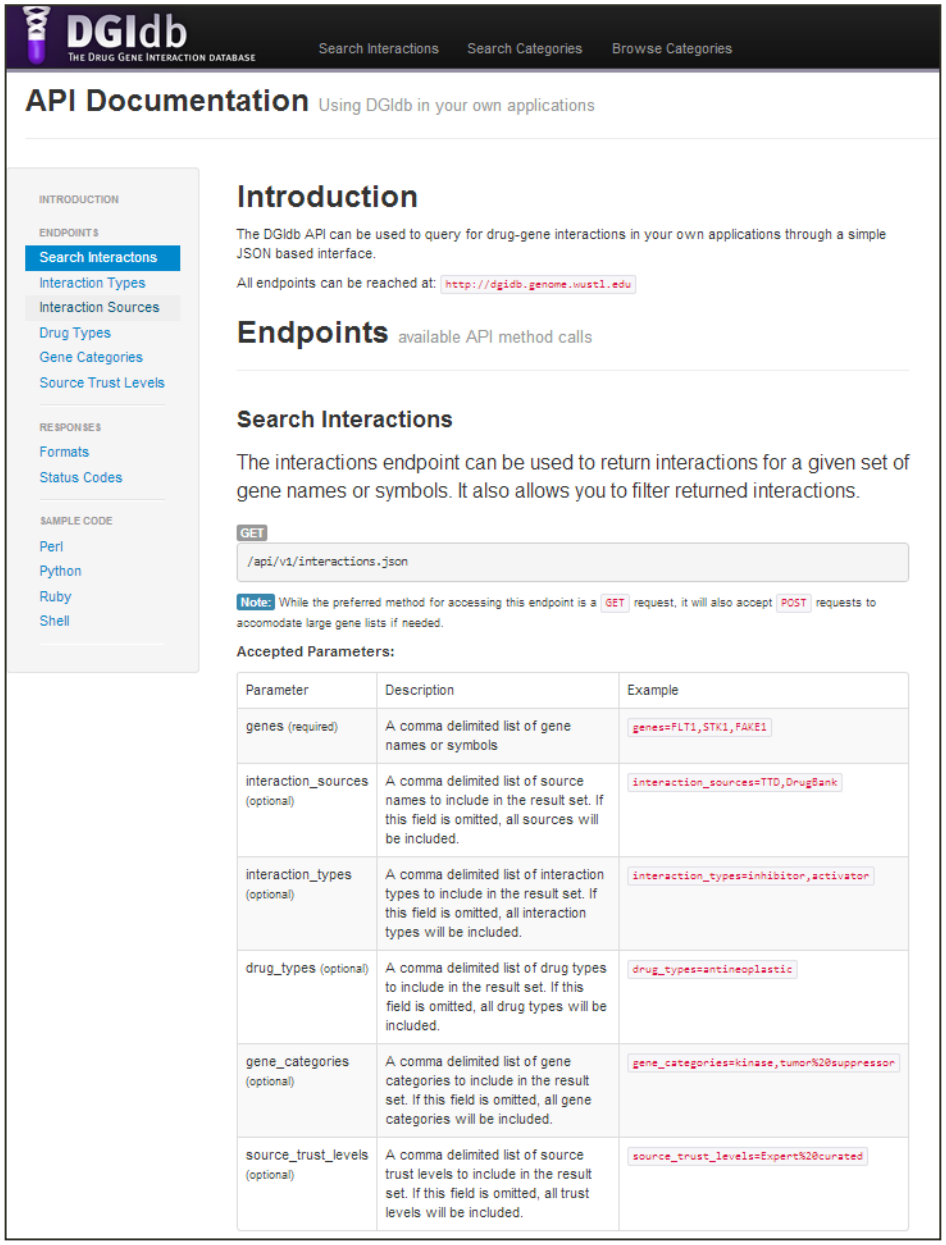
- ‘Browse Categories’ allows the user to explore the druggable gene categories defined in DGIdb.
- Each category displayed links to a complete list of genes found in the category and the source that placed them there, the user can also filter the data by source using the filter menu at the top.
- The number is brackets is the total number of unique genes (from all sources) placed in the category
- It is these lists that are used in the ‘Search Categories’ tool.

5. Finding help



- Various resources are available to help navigate DGIdb under the 'Help' menu
  - Getting Started – tutorials, screencasts and walkthroughs (including this document)
  - API Documentation – Provides additional information on using DGIdb in other applications
  - News – Updates and information about the database and new features.
  - FAQ – frequently asked questions and their answers
  - Sources – A summary of all sources of data currently imported to DGIdb
  - Downloads – dumps of intermediate data files and links to source code
  - Contact – contact details for DGIDB developers
  - View Disclaimer – Shows the documented disclaimer for use of the information in DGIdb.

API Documention: Using DGIdb in your own Applications



DGIdb
THE DRUG GENE INTERACTION DATABASE

Search Interactions | Search Categories | Browse Categories

API Documentation Using DGIdb in your own applications

INTRODUCTION

ENDPOINTS
- Search Interactons
- Interaction Types
- Interaction Sources
- Drug Types
- Gene Categories
- Source Trust Levels

RESPONSES
- Formats
- Status Codes

SAMPLE CODE
- Perl
- Python
- Ruby
- Shell

Introduction

The DGIdb API can be used to query for drug-gene interactions in your own applications through a simple JSON based interface.

All endpoints can be reached at: `http://dgidb.genome.wustl.edu`

Endpoints available API method calls

Search Interactions

The interactions endpoint can be used to return interactions for a given set of gene names or symbols. It also allows you to filter returned interactions.

GET

`/api/v1/interactions.json`

Note: While the preferred method for accessing this endpoint is a `GET` request, it will also accept `POST` requests to accomodate large gene lists if needed.

Accepted Parameters:

Parameter	Description	Example
genes (required)	A comma delimited list of gene names or symbols	`genes=FLT1,STK1,FAKE1`
interaction_sources (optional)	A comma delimited list of source names to include in the result set. If this field is omitted, all sources will be included.	`interaction_sources=TTD,DrugBank`
interaction_types (optional)	A comma delimited list of interaction types to include in the result set. If this field is omitted, all interaction types will be included.	`interaction_types=inhibitor,activator`
drug_types (optional)	A comma delimited list of drug types to include in the result set. If this field is omitted, all drug types will be included.	`drug_types=antineoplastic`
gene_categories (optional)	A comma delimited list of gene categories to include in the result set. If this field is omitted, all gene categories will be included.	`gene_categories=kinase,tumor%20suppressor`
source_trust_levels (optional)	A comma delimited list of source trust levels to include in the result set. If this field is omitted, all trust levels will be included.	`source_trust_levels=Expert%20curated`

- The API Documentation can be accessed by clicking on the Help menu in the top right corner of the home screen (Shown on the previous slide).
- This page provides in-depth instructions to running your own drug-gene interaction queries without the web interface.
- There is sample code provided for a Perl script, while samples of Python, Ruby and Shell are currently in the works.