

# BioGRID to RDF?

Janos G. Hajagos  
Stony Brook University  
School of Medicine  
April 7, 2011

## Welcome to the Biological General Repository for Interaction Datasets

BioGRID is an online interaction repository with data compiled through comprehensive curation efforts. Our current index is version **3.1.75** and searches **26,523** publications for **385,884** raw protein and genetic interactions from major model organism species. All interaction data are **freely** provided through our search index and available via download in a wide variety of standardized formats.

[INTERACTION STATISTICS](#)

[LATEST DOWNLOADS](#)

### Search the BioGRID

Search by identifiers, keywords, and gene names...

All Organisms

[SUBMIT GENE SEARCH](#)



Advanced Search



Search Tips



Featured Datasets

By Gene

By Publication

### AREAS OF INTEREST TO HELP YOU GET STARTED



#### Build and Download Interaction Datasets

Create custom interaction datasets by protein or by publication. You can also download our entire dataset in a wide variety of standard formats.



#### Link To Us or Submit Interactions

Send us your datasets or link to the BioGRID directly from your own website or database. Full details on how to contribute are available here.



#### Online Tools and Resources

We've developed tools that make use of BioGRID data. Check out the list of tools to see if we can help you work with our data.



#### View Our Interaction Statistics

Find out how many organisms, proteins, publications, and interactions are available in the current release of the BioGRID.

### BIOGRID FUNDING AND PARTNERS




[more partners](#)

# BioGRID Downloads

Gene / Identifier Search


















GO

All Organisms 

BioGRID interaction data are 100% freely available to both commercial and academic users and are provided **WITHOUT ANY WARRANTY**. Publications that make use of this data are requested to please cite the contributing authors and : [Stark C](#), [Breitkreutz BJ](#), [Reguly T](#), [Boucher L](#), [Breitkreutz A](#), [Tyers M](#). [Biogrid: A General Repository for Interaction Datasets](#). *Nucleic Acids Res.* Jan1; 34:D535-9 where applicable.

## BioGRID Dataset Downloads

### Current Release

-  [BIOGRID-ALL-3.1.75.mitab.zip](#)
-  [BIOGRID-ALL-3.1.75.psi.zip](#)
-  [BIOGRID-ALL-3.1.75.psi25.zip](#)
-  [BIOGRID-ALL-3.1.75.tab.zip](#)
-  [BIOGRID-ALL-3.1.75.tab2.zip](#)
-  [BIOGRID-IDENTIFIERS-3.1.75.tab.zip](#)
-  [BIOGRID-ORGANISM-3.1.75.mitab.zip](#)
-  [BIOGRID-ORGANISM-3.1.75.psi.zip](#)
-  [BIOGRID-ORGANISM-3.1.75.psi25.zip](#)
-  [BIOGRID-ORGANISM-3.1.75.tab.zip](#)
-  [BIOGRID-ORGANISM-3.1.75.tab2.zip](#)
-  [BIOGRID-OSPREY\\_DATASETS-3.1.75.osprey.zip](#)
-  [BIOGRID-SYSTEM-3.1.75.mitab.zip](#)
-  [BIOGRID-SYSTEM-3.1.75.psi.zip](#)
-  [BIOGRID-SYSTEM-3.1.75.psi25.zip](#)
-  [BIOGRID-SYSTEM-3.1.75.tab.zip](#)
-  [BIOGRID-SYSTEM-3.1.75.tab2.zip](#)

### Previous Release

### Release Archive

### Cytoscape Plugins

### External Database Builds

### Other Datasets

### Published Datasets

## BioGRID Release 3.1.75

This download directory contains the most recent data release from the BioGRID. This release was compiled on **March 25th, 2011** and contains all curated interaction data processed prior to this date and reflects the most recent data available via our search engine. If you are starting a new project using our data, it is **HIGHLY** recommended that you use these data files as they are the most up to date versions of our interaction dataset.

For more information about each of the available files, simply mouse over to see a description.

To download each file, simply **left click** on it to start the download process.

# Pathway Commons uses BioPAX Ontology



Search and visualize public biological pathway information. Single point of access. [\[more...\]](#)

[Home](#) [Data Sources](#) [Download](#) [FAQ](#) [Web Service](#) [About](#)

Send us your [feedback](#). Sign up for Pathway Commons [announcements](#). [RSS Feed](#)

## Search Pathway Commons:

[Find Pathways](#) [Find Molecules](#)

For example, if you enter: [BRCA1](#), you will **get back the list of pathways** containing the keyword "BRCA1", and the list of pathways that contain the BRCA1 gene.

Current filter settings: All Organisms, All Data Sources. [Set filters](#).

## What's New:

- **NEW** February 4, 2011:
  - Added MetaCyc data set (December 7, 2010 Version 14.6).
  - BioGRID data set (December 15, 2010 Version 3.1.72).
  - HumanCyc data set (October 7, 2010 Version 14.6).
  - IntAct data set (December 15, 2010).
  - Mint data set (December 21, 2010).
  - Nature PID (September 16, 2010).
  - Reactome data set (December 17, 2010 Version 35).
- September 7, 2010:
  - BioGRID data set (July 31, 2010 Version 30.0.67).
  - HPRD data set (April 13, 2010 Version 9).
  - HumanCyc data set (June 16, 2010 Version 14.1).
  - IntAct data set (August 8, 2010).
  - MINT data set (July 28, 2010).
  - NCI/Nature Pathway Interaction Database (August 10, 2010).
  - Reactome data set (June 18, 2010 Version 33).
- October 15, 2009:
  - Improved search functionality.
- July 2, 2009:
  - [Batch Download](#) of all Pathway Commons data in multiple file formats is now available.
  - Systems Biology Center New York - IMID data set (December 17, 2008 Version 27).
  - Latest Reactome data set (June 24, 2009 Version 29).
  - Latest HumanCyc data set (June 22, 2009 Version 13.1).
  - All yeast proteins are now annotated with UniProt functional annotation.

## Using Pathway Commons:

- Biologists:** Browse and search pathways across multiple valuable public pathway databases.
- Computational biologists:** Download an integrated set of pathways in BioPAX format for global analysis.
- Software developers:** Build software on top of Pathway Commons using our [web service API](#). Download and install the [cPath software](#) to create a local mirror.

## Current Data Sources:

Pathway Commons currently contains the following data sources ([batch download](#)):



## Pathway Commons Quick Stats:

Number of Pathways:	3,436
Number of Interactions:	565,977
Number of Physical Entities:	116,073
Number of Organisms:	1,295

Integration of additional data sources is planned in the near future. For a comprehensive directory of interaction and pathway databases, please refer to [Pathguide](#).

## Citing Pathway Commons:

**To cite the Pathway Commons Project:** Cerami et al. Pathway Commons, a web resource for biological pathway data. Nucl. Acids Res. (2010) [doi:10.1093/nar/gkq1039](#)

**To cite the cPath Software:** Cerami et al. cPath: open source software for collecting, storing, and querying biological pathways. BMC Bioinformatics. (2006) [doi:10.1186/1471-2105-7-497](#)

# Serialization of BioPAX OWL as ntriples

Gruff - An AllegroGraph Browser - biogrid (read / write - 12,659,705 triples)

File View Display Link Remove Layout Select Edit Inclusion Options Drawing Options Layout Options Help

**Legend:**

- COMMENT
- DATA-SOURCE
- DB
- EVIDENCE
- EVIDENCE-CODE
- EXPERIMENTAL-FORM
- ID
- NAME
- ORGANISM
- PARTICIPANTS
- PHYSICAL-ENTITY
- SHORT-NAME
- SYNONYMS
- Type
- XREF

**Graph Elements:**

- Nodes:** CPATH-LOCAL-18012658, CPATH-LOCAL-18012656, CPATH-LOCAL-18012657, CPATH-LOCAL-18012653, CPATH-LOCAL-18012646, CPATH-LOCAL-18012647, CPATH-LOCAL-18012650, CPATH-LOCAL-18012659, CPATH-LOCAL-18012660, CPATH-LOCAL-18258688, CPATH-LOCAL-18276794, CPATH-1072567, CPATH-1092014, CPATH-1084170, CPATH-LOCAL-18171127, CGS099, msi, msi-CG9106, DMSIDNA, SIDNA, anon-EST:Liang-2.35, Dme|\_CG5099, Dme|\_CG9106, clone 2.35, msi-CG9106, Dme|\_CG9106, A protein interaction map of Drosophila melanogaster.
- Relationships:** Data Source, Unification Xref, Evidence, Physical Interaction, Sequence Participant, Experimental Form, Publication Xref, Unification Xref, CPATH, XREF.
- Literals:** The BioGRID: A General Repository for Interaction Datasets, msi-CG9106, A protein interaction map of Drosophila melanogaster.


# PSI-MITAB formatted BioGRID file

	A	B	C
1	#ID Interactor A	ID Interactor B	Alt IDs Interactor A
2	entrez gene/locuslink:6416   GRID:112315	entrez gene/locuslink:2318   GRID:108607	entrez gene/locuslink:MAP2K4
3	entrez gene/locuslink:84665   GRID:124185	entrez gene/locuslink:88   GRID:106603	entrez gene/locuslink:MYPN
4	entrez gene/locuslink:90   GRID:106605	entrez gene/locuslink:2339   GRID:108625	entrez gene/locuslink:ACVR1
5	entrez gene/locuslink:2624   GRID:108894	entrez gene/locuslink:5371   GRID:111384	entrez gene/locuslink:GATA2
6	entrez gene/locuslink:6118   GRID:112038	entrez gene/locuslink:6774   GRID:112651	entrez gene/locuslink:RPA2   entrez gene/locuslink:RP4-547C9.3
7	entrez gene/locuslink:375   GRID:106870	entrez gene/locuslink:23163   GRID:116775	entrez gene/locuslink:ARF1
8	entrez gene/locuslink:377   GRID:106872	entrez gene/locuslink:23647   GRID:117174	entrez gene/locuslink:ARF3
9	entrez gene/locuslink:377   GRID:106872	entrez gene/locuslink:27236   GRID:118084	entrez gene/locuslink:ARF3
10	entrez gene/locuslink:10327   GRID:115610	entrez gene/locuslink:54512   GRID:120007	entrez gene/locuslink:AKR1A1   entrez gene/locuslink:RP4-697E16.2
11	entrez gene/locuslink:54464   GRID:119970	entrez gene/locuslink:226   GRID:106728	entrez gene/locuslink:XRN1
12	entrez gene/locuslink:351   GRID:106848	entrez gene/locuslink:10513   GRID:115769	entrez gene/locuslink:APP
13	entrez gene/locuslink:333   GRID:106830	entrez gene/locuslink:1600   GRID:107970	entrez gene/locuslink:APLP1
14	entrez gene/locuslink:10370   GRID:115649	entrez gene/locuslink:7020   GRID:112878	entrez gene/locuslink:CITED2
15	entrez gene/locuslink:2033   GRID:108347	entrez gene/locuslink:7020   GRID:112878	entrez gene/locuslink:EP300   entrez gene/locuslink:RP1-85F18.1
16	entrez gene/locuslink:338   GRID:106835	entrez gene/locuslink:4547   GRID:110641	entrez gene/locuslink:APOB
17	entrez gene/locuslink:409   GRID:106902	entrez gene/locuslink:5900   GRID:111836	entrez gene/locuslink:ARRB2
18	entrez gene/locuslink:1436   GRID:107823	entrez gene/locuslink:2885   GRID:109142	entrez gene/locuslink:CSF1R
19	entrez gene/locuslink:7916   GRID:113646	entrez gene/locuslink:2885   GRID:109142	entrez gene/locuslink:BAT2   entrez gene/locuslink:DADB-70P7.5
20	entrez gene/locuslink:27257   GRID:118104	entrez gene/locuslink:4677   GRID:110758	entrez gene/locuslink:LSM1
21	entrez gene/locuslink:6521   GRID:112412	entrez gene/locuslink:22950   GRID:116605	entrez gene/locuslink:SLC4A1

# Parsing of PSI-MITAB formatted file

	A	B	C	D	E	
1	field_name	identifier	sequence_id	db_identifier_type	db_identifier	qualifier
1730	Aliases Interactor A	GRID:2721	2	entrez gene/locuslink	DCOH	gene name synonym
1731	Aliases Interactor A	GRID:2721	3	entrez gene/locuslink	PHS	gene name synonym
1732	Aliases Interactor A	GRID:2721	4	entrez gene/locuslink	PCD	gene name synonym
1733	Aliases Interactor A	GRID:2765	1	entrez gene/locuslink	AD3L	gene name synonym
1734	Aliases Interactor A	GRID:2765	2	entrez gene/locuslink	AD4	gene name synonym
1735	Aliases Interactor A	GRID:2765	3	entrez gene/locuslink	PS2	gene name synonym
1736	Aliases Interactor A	GRID:2765	4	entrez gene/locuslink	STM2	gene name synonym
1737	Aliases Interactor A	GRID:2785	1	entrez gene/locuslink	nCL-1	gene name synonym
1738	Aliases Interactor A	GRID:2785	2	entrez gene/locuslink	MGC14344	gene name synonym
1739	Aliases Interactor A	GRID:2785	3	entrez gene/locuslink	MGC11121	gene name synonym
1740	Aliases Interactor A	GRID:2785	4	entrez gene/locuslink	MGC4403	gene name synonym
1741	Aliases Interactor A	GRID:2785	5	entrez gene/locuslink	p94	gene name synonym
1742	Aliases Interactor A	GRID:2785	6	entrez gene/locuslink	LGMD2A	gene name synonym
1743	Aliases Interactor A	GRID:2785	7	entrez gene/locuslink	LGMD2	gene name synonym
1744	Aliases Interactor A	GRID:2785	8	entrez gene/locuslink	CANPL3	gene name synonym
1745	Aliases Interactor A	GRID:2785	9	entrez gene/locuslink	CANP3	gene name synonym
1746	Aliases Interactor A	GRID:2785	10	entrez gene/locuslink	MGC10767	gene name synonym
1747	Aliases Interactor A	GRID:2827	1	entrez gene/locuslink	Insp3r1	gene name synonym
1748	Aliases Interactor A	GRID:2827	2	entrez gene/locuslink	DKFZp313E1334	gene name synonym
1749	Aliases Interactor A	GRID:2827	3	entrez gene/locuslink	SCA16	gene name synonym
1750	Aliases Interactor A	GRID:2827	4	entrez gene/locuslink	DKFZp313N1434	gene name synonym
1751	Aliases Interactor A	GRID:2827	5	entrez gene/locuslink	IP3R1	gene name synonym
1752	Aliases Interactor A	GRID:2827	6	entrez gene/locuslink	SCA15	gene name synonym
1753	Aliases Interactor A	GRID:2827	7	entrez gene/locuslink	IP3R	gene name synonym

# Using the NCBO Annotator

 BioPortal

[Browse](#) [Search](#) [Projects](#) [Annotate](#) [All Mappings](#) [All Resources](#)

[Online Mendelian Inheritance in Man](#) [HUGO](#) [SNOMED Clinical Terms](#) [COSTART](#)

The NCBO BioPortal Annotator processes text submitted by users, recognizes relevant biomedical ontology terms in the text and returns the annotations to the user. Use the interface below to submit sample text to get ontology-based annotations. Hover the mouse pointer on any button to see what it does. Click on the (?) to see a detailed help panel. Try typing this sample text: "Melanoma is a malignant tumor of melanocytes which are found predominantly in skin but also in the bowel and the eye". Choose SNOMEDCT as your ontology of interest. The annotator---when used as a Web service---allows users to utilize ontologies for annotation of biomedical data on their sites in order to facilitate interoperation, search and translational discoveries. You can use the interface below to test different parameters and settings before using the Annotator programmatically.

Subscribe to the [NCBO Annotator Users Google group](#) to learn more about who and how the Annotator is being used in different projects.

### Annotator

Ontologies  [Choose...](#)

Semantic Types  [Choose...](#)

Options [Change...](#)

The Annotator user interface is currently limited to **300 words**. Please use the [NCBO Annotator web service](#) for more advanced options.

Text

[Annotate](#)

---

[Annotation Term List](#) [Annotation Tag Cloud](#)

### Terms

Filter:

[Select All](#) [Select None](#)

- CALPAIN 3 (1)
- Calpain 3, (p94) (1)

### Annotations

Filter:  (filters by Term or Ontology)

Term: [CALPAIN 3](#)

Ontology: [Online Mendelian Inheritance in Man](#)

Context: **P94**

---

Term: [calpain 3, \(p94\)](#)

Ontology: [HUGO](#)

Context: **P94**



# Aligning Gene Symbols to UMLS CUIs

## Find Meaning

MGC4403  
p94  
LGMD2A



Verbose

Find Meaning

Reset

## Meanings Found

CUI	Label	Type	Definition	Synonyms
<a href="#">C1869123</a>	lgmd2a	Disease or Syndrome		CALPAINOPATHY MUSCULAR DYSTROPHY, LIMB-GIRDLE, TYPE 2A LGMD2A <a href="#">Show All</a>
<a href="#">C1413113</a>	p94	Gene or Genome		CAPN3 gene CANP3 p94 <a href="#">Show All</a>

[Close](#)

```
<http://link.informatics.stonybrook.edu/umls/CUI/C1413113> <http://link.informatics.stonybrook.edu/umls/hasOrder> "1"^^<http://www.w3.org/2001/XMLSchema#long> .
<http://link.informatics.stonybrook.edu/umls/CUI/C1413113> <http://link.informatics.stonybrook.edu/umls/hasSTN> "Gene or Genome" .
<http://link.informatics.stonybrook.edu/umls/CUI/C1413113> <http://www.w3.org/2000/01/rdf-schema#label> "p94" .
<http://link.informatics.stonybrook.edu/umls/CUI/C1869123> <http://link.informatics.stonybrook.edu/umls/hasOrder> "0"^^<http://www.w3.org/2001/XMLSchema#long> .
<http://link.informatics.stonybrook.edu/umls/CUI/C1869123> <http://link.informatics.stonybrook.edu/umls/hasSTN> "Disease or Syndrome" .
<http://link.informatics.stonybrook.edu/umls/CUI/C1869123> <http://www.w3.org/2000/01/rdf-schema#label> "lgmd2a" .
```

[RDF/XML](#)

# Faceted Browser View of P94

**About:** <http://link.informatics.stonybrook.edu/u/mls/AUI/A11683652> (calpain 3, (p94))

predicate	object
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#CHROMOSOME">http://link.informatics.stonybrook.edu/u/mls/ATN#CHROMOSOME</a>	15q15.1-q21.1
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#EZ">http://link.informatics.stonybrook.edu/u/mls/ATN#EZ</a>	3.4.22.17
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#MAPPED_REFSEQ_ID">http://link.informatics.stonybrook.edu/u/mls/ATN#MAPPED_REFSEQ_ID</a>	NM_000070
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#GENESYMBOL">http://link.informatics.stonybrook.edu/u/mls/ATN#GENESYMBOL</a>	CAPN3
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#MAPPED_GDB_ID">http://link.informatics.stonybrook.edu/u/mls/ATN#MAPPED_GDB_ID</a>	GDB:119751
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#LOCUS_TYPE">http://link.informatics.stonybrook.edu/u/mls/ATN#LOCUS_TYPE</a>	gene with protein product
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#ACCESSION_NO">http://link.informatics.stonybrook.edu/u/mls/ATN#ACCESSION_NO</a>	X85030
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#PMID">http://link.informatics.stonybrook.edu/u/mls/ATN#PMID</a>	<ul style="list-style-type: none"> <li>• 2555341</li> <li>• 7720071</li> </ul>
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#SWP">http://link.informatics.stonybrook.edu/u/mls/ATN#SWP</a>	P20807
<a href="http://link.informatics.stonybrook.edu/u/mls/hasTermType">http://link.informatics.stonybrook.edu/u/mls/hasTermType</a>	<a href="http://link.informatics.stonybrook.edu/u/mls/TTY/PT">http://link.informatics.stonybrook.edu/u/mls/TTY/PT</a>
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#MAPPED_RGD_ID">http://link.informatics.stonybrook.edu/u/mls/ATN#MAPPED_RGD_ID</a>	RGD:2269
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#DB_XR">http://link.informatics.stonybrook.edu/u/mls/ATN#DB_XR</a>	MEROPS: <b>C02.004</b> Orphanet: <b>15398</b>
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#CCDS_ID">http://link.informatics.stonybrook.edu/u/mls/ATN#CCDS_ID</a>	<ul style="list-style-type: none"> <li>• CCDS42027.1</li> <li>• CCDS10086.1</li> <li>• CCDS10085.1</li> <li>• CCDS32207.1</li> </ul>
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#DATE_LAST_MODIFIED">http://link.informatics.stonybrook.edu/u/mls/ATN#DATE_LAST_MODIFIED</a>	2007-12-14
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://link.informatics.stonybrook.edu/u/mls/AUI">http://link.informatics.stonybrook.edu/u/mls/AUI</a>
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#OMIM_ID">http://link.informatics.stonybrook.edu/u/mls/ATN#OMIM_ID</a>	114240
<a href="http://link.informatics.stonybrook.edu/u/mls/AUI/AUI">http://link.informatics.stonybrook.edu/u/mls/AUI/AUI</a>	A11683652
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#PREV_SYMBOL">http://link.informatics.stonybrook.edu/u/mls/ATN#PREV_SYMBOL</a>	<ul style="list-style-type: none"> <li>• LGMD2A</li> <li>• LGMD2</li> </ul>
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#ENTREZGENE_ID">http://link.informatics.stonybrook.edu/u/mls/ATN#ENTREZGENE_ID</a>	825
<a href="http://link.informatics.stonybrook.edu/u/mls/ATN#DB_XR_ID">http://link.informatics.stonybrook.edu/u/mls/ATN#DB_XR_ID</a>	<ul style="list-style-type: none"> <li>• MEROPS: C02.004</li> <li>• Orphanet: 15398</li> </ul>