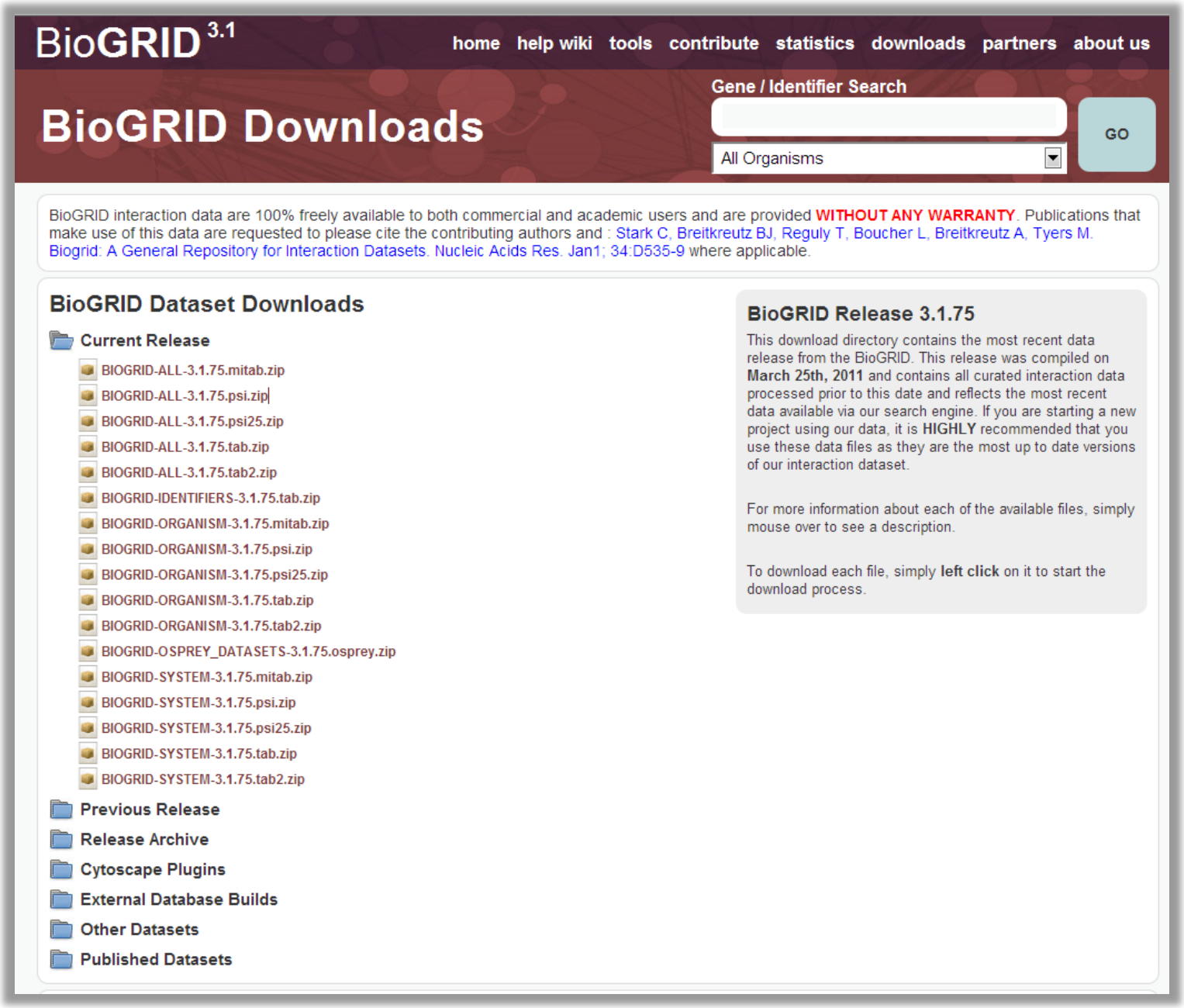
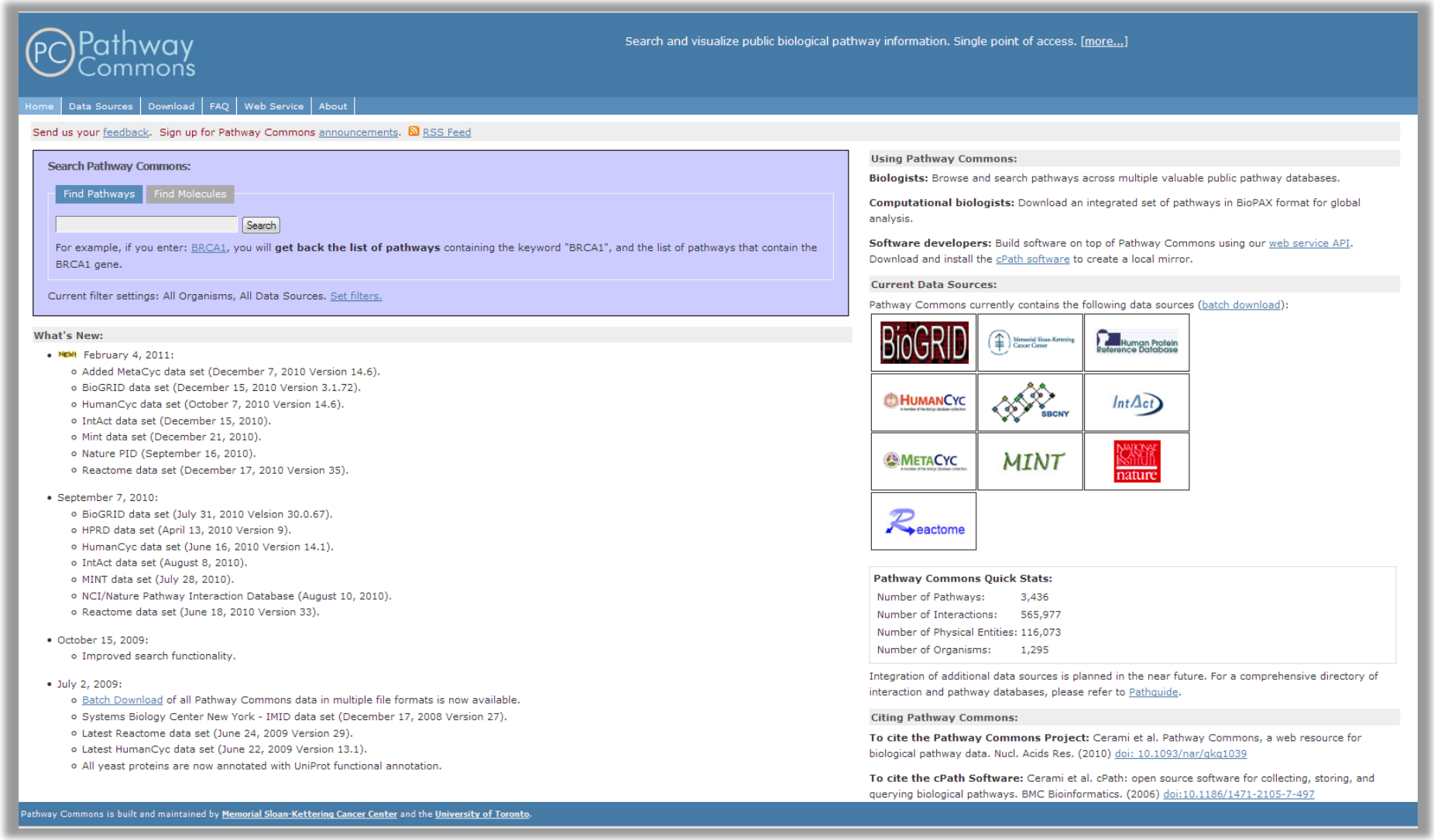
BioGRID to RDF?

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

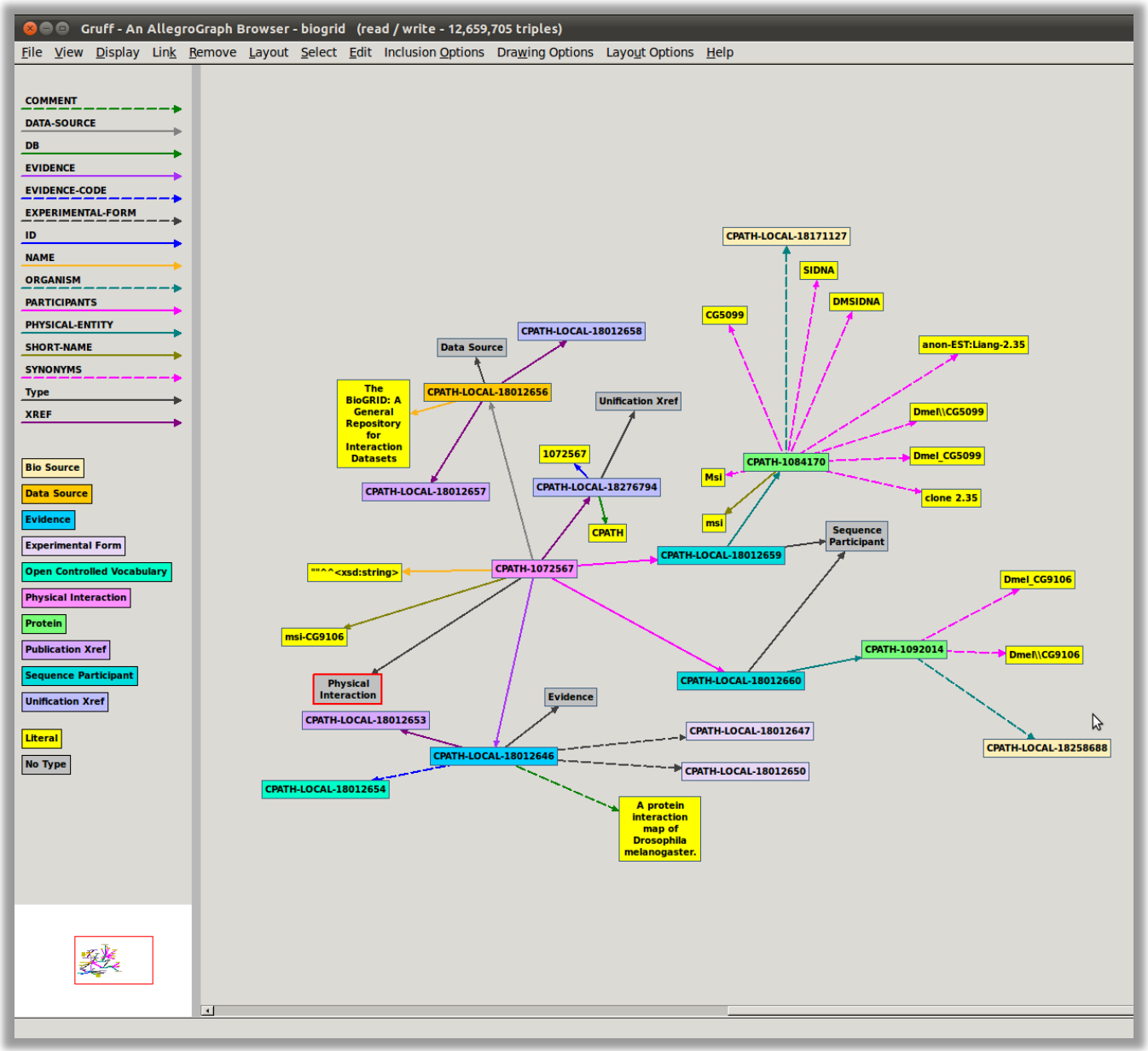

BioGRID 3.1
home help wiki tools contribute statistics downloads partners about us
Welcome to the Biological General Repository for Interaction Datasets
BioGRID is an online interaction repository with data compiled through comprehensive curation efforts. Our current index is version 3.1.75 and searches 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
By Gene
By Publication
Advanced Search
Search Tips
Featured Datasets
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
NATIONAL INSTITUTES OF HEALTH
CIHR IRSC
BBSRC bioscience for the future
MOUNT SINAI HOSPITAL
PRINCETON UNIVERSITY
University of Edinburgh
SGD
LAMHDI
IMEx
more partners


BioGRID 3.1
home help wiki tools contribute statistics downloads partners about us
BioGRID Downloads
Gene / Identifier Search
All Organisms
GO
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



Serialization of BioPAX OWL as ntriples



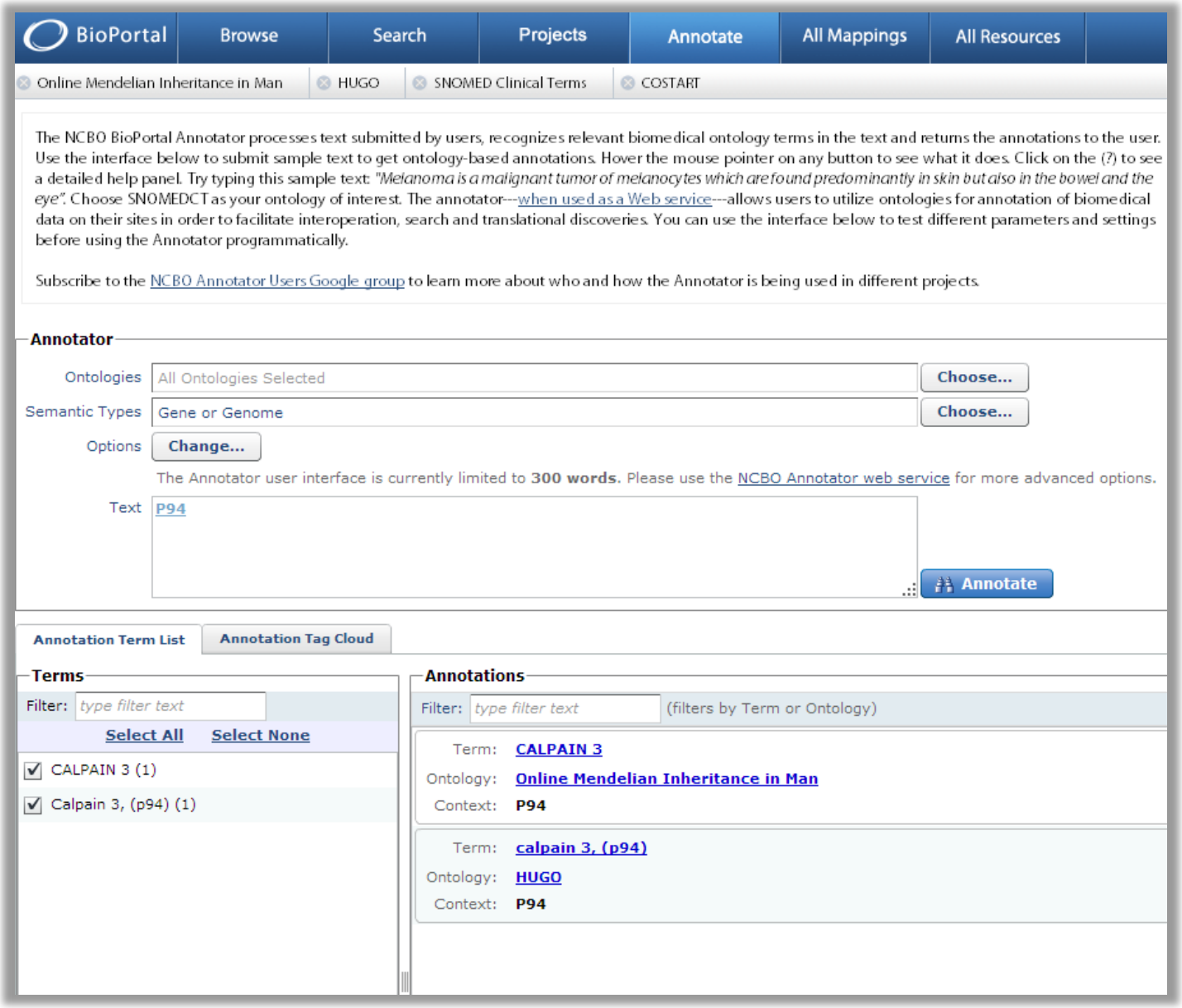
PSI-MITAB formatted BioGRID file

	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

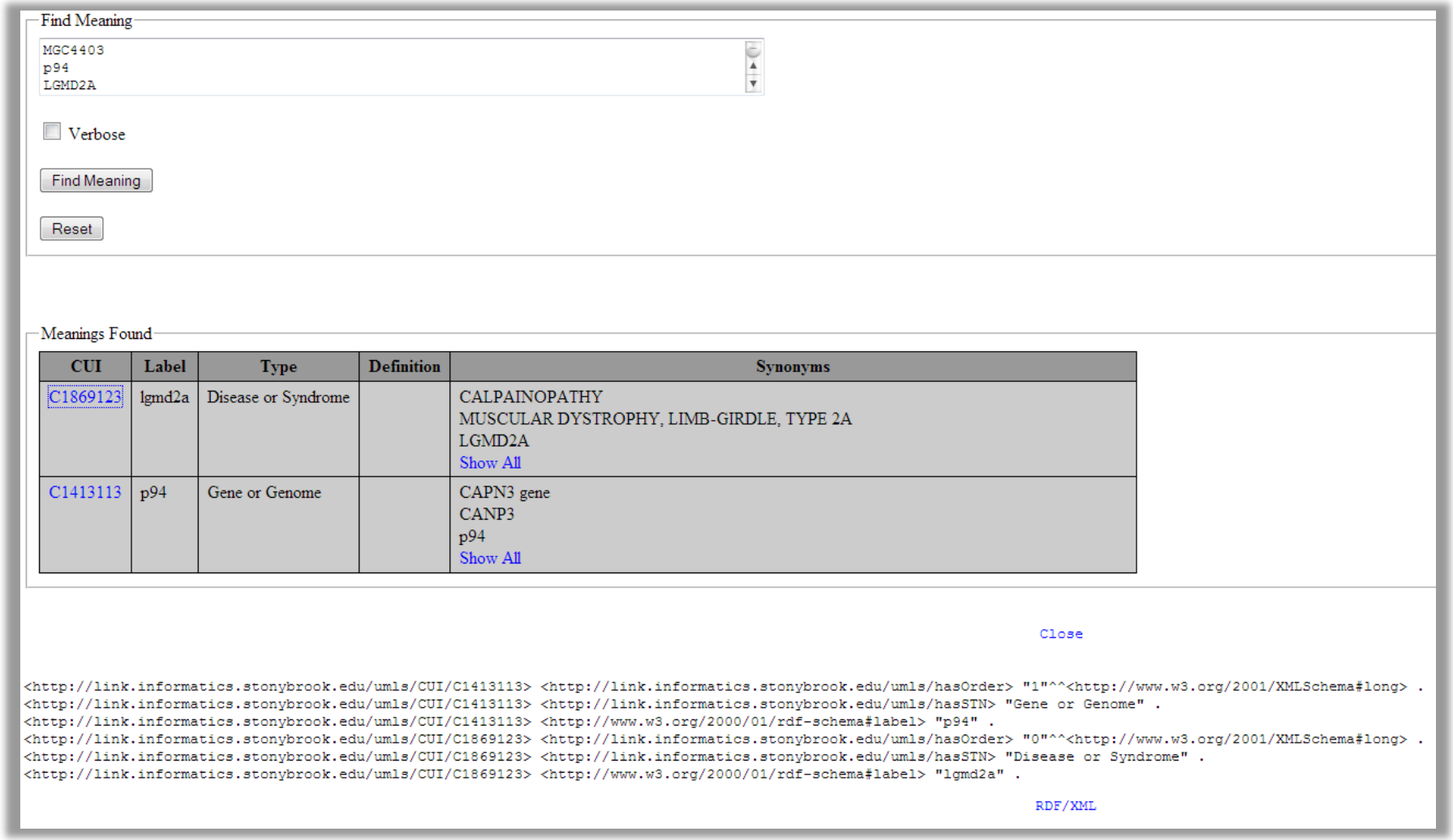
Parsing of PSI-MITAB formatted file

	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



Aligning Gene Symbols to UMLS CUIs



Find Meaning

MGC4403
p94
LGMD2A

Verbose

Find Meaning

Reset

Meanings Found

CUI	Label	Type	Definition	Synonyms
C1869123	lgmd2a	Disease or Syndrome		CALPAINOPATHY MUSCULAR DYSTROPHY, LIMB-GIRDLE, TYPE 2A LGMD2A Show All
C1413113	p94	Gene or Genome		CAPN3 gene CANP3 p94 Show All

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/umls/AUI/A11683652 (calpain 3, (p94))

predicate	object
http://link.informatics.stonybrook.edu/umls/ATN#CHROMOSOME	15q15.1-q21.1
http://link.informatics.stonybrook.edu/umls/ATN#EZ	3.4.22.17
http://link.informatics.stonybrook.edu/umls/ATN#MAPPED_REFSEQ_ID	NM_000070
http://link.informatics.stonybrook.edu/umls/ATN#GENESYMBOL	CAPN3
http://link.informatics.stonybrook.edu/umls/ATN#MAPPED_GDB_ID	GDB:119751
http://link.informatics.stonybrook.edu/umls/ATN#LOCUS_TYPE	gene with protein product
http://link.informatics.stonybrook.edu/umls/ATN#ACCESSION_NO	X85030
http://link.informatics.stonybrook.edu/umls/ATN#PMID	• 2555341 • 7720071
http://link.informatics.stonybrook.edu/umls/ATN#SWP	P20807
http://link.informatics.stonybrook.edu/umls/hasTermType	http://link.informatics.stonybrook.edu/umls/TTY/PT
http://link.informatics.stonybrook.edu/umls/ATN#MAPPED_RGD_ID	RGD:2269
http://link.informatics.stonybrook.edu/umls/ATN#DB_XR	MEROPS:**C02.004** Orphanet:**15398**
http://link.informatics.stonybrook.edu/umls/ATN#CCDS_ID	• CCDS42027.1 • CCDS10086.1 • CCDS10085.1 • CCDS32207.1
http://link.informatics.stonybrook.edu/umls/ATN#DATE_LAST_MODIFIED	2007-12-14
http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://link.informatics.stonybrook.edu/umls/AUI
http://link.informatics.stonybrook.edu/umls/ATN#OMIM_ID	114240
http://link.informatics.stonybrook.edu/umls/AUI/AUI	A11683652
http://link.informatics.stonybrook.edu/umls/ATN#PREV_SYMBOL	• LGMD2A • LGMD2
http://link.informatics.stonybrook.edu/umls/ATN#ENTREZGENE_ID	825
http://link.informatics.stonybrook.edu/umls/ATN#DB_XR_ID	• MEROPS: C02.004 • Orphanet: 15398