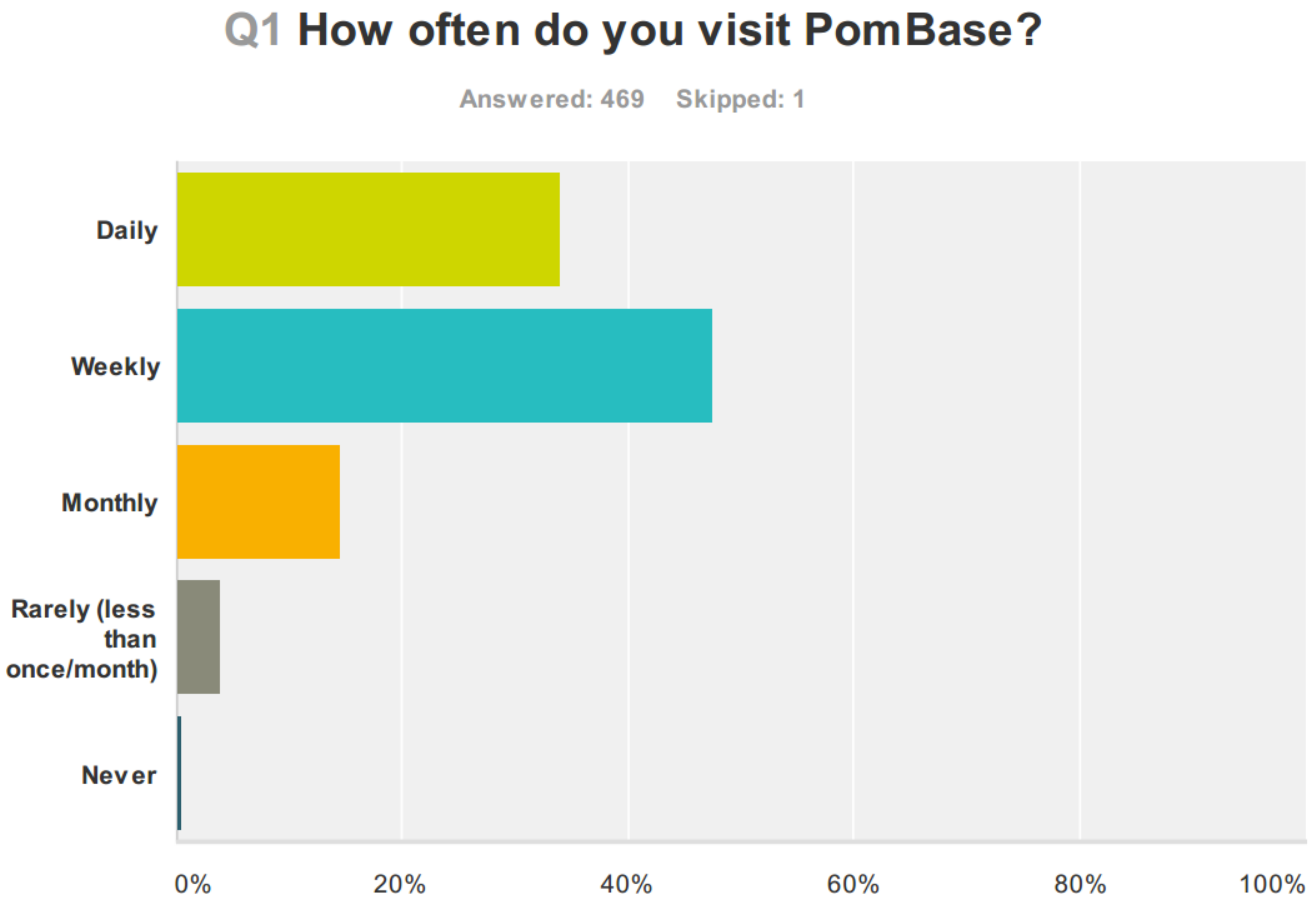
How to get more from PomBase
PomBase workshop, pombe 2015

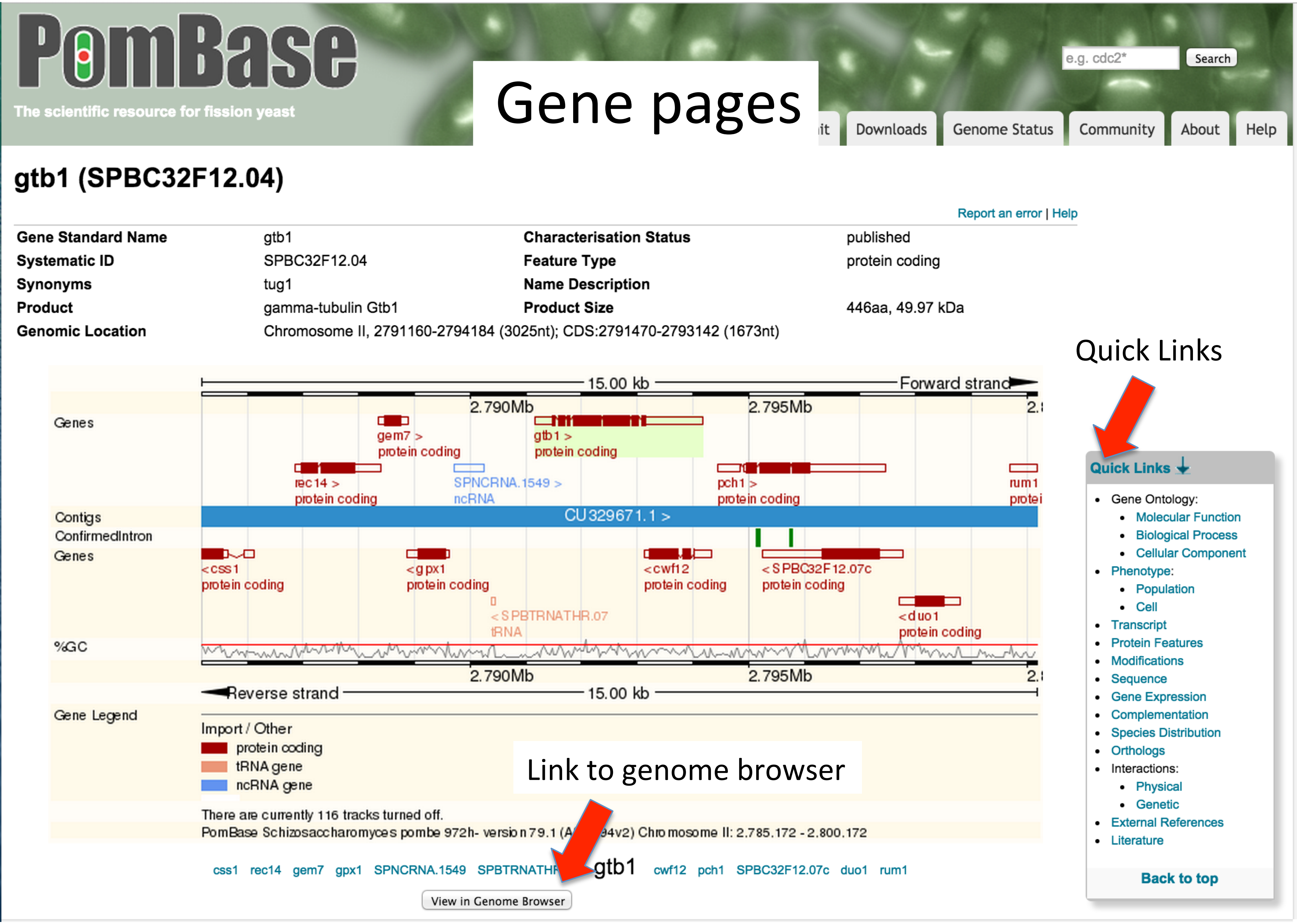
Introductions

PomBase Overview

PomBase User Survey 2013



Gene pages



gtb1 (SPBC32F12.04)

Report an error | Help

Gene Standard Name	gtb1	**Characterisation Status**	published
Systematic ID	SPBC32F12.04	**Feature Type**	protein coding
Synonyms	tug1	**Name Description**	
Product	gamma-tubulin Gtb1	**Product Size**	446aa, 49.97 kDa
Genomic Location	Chromosome II, 2791160-2794184 (3025nt); CDS:2791470-2793142 (1673nt)		

There are currently 116 tracks turned off.

PomBase Schizosaccharomyces pombe 972h- version 79.1 (A 94v2) Chromosome II: 2.785.172 - 2.800.172

Link to genome browser

Quick Links

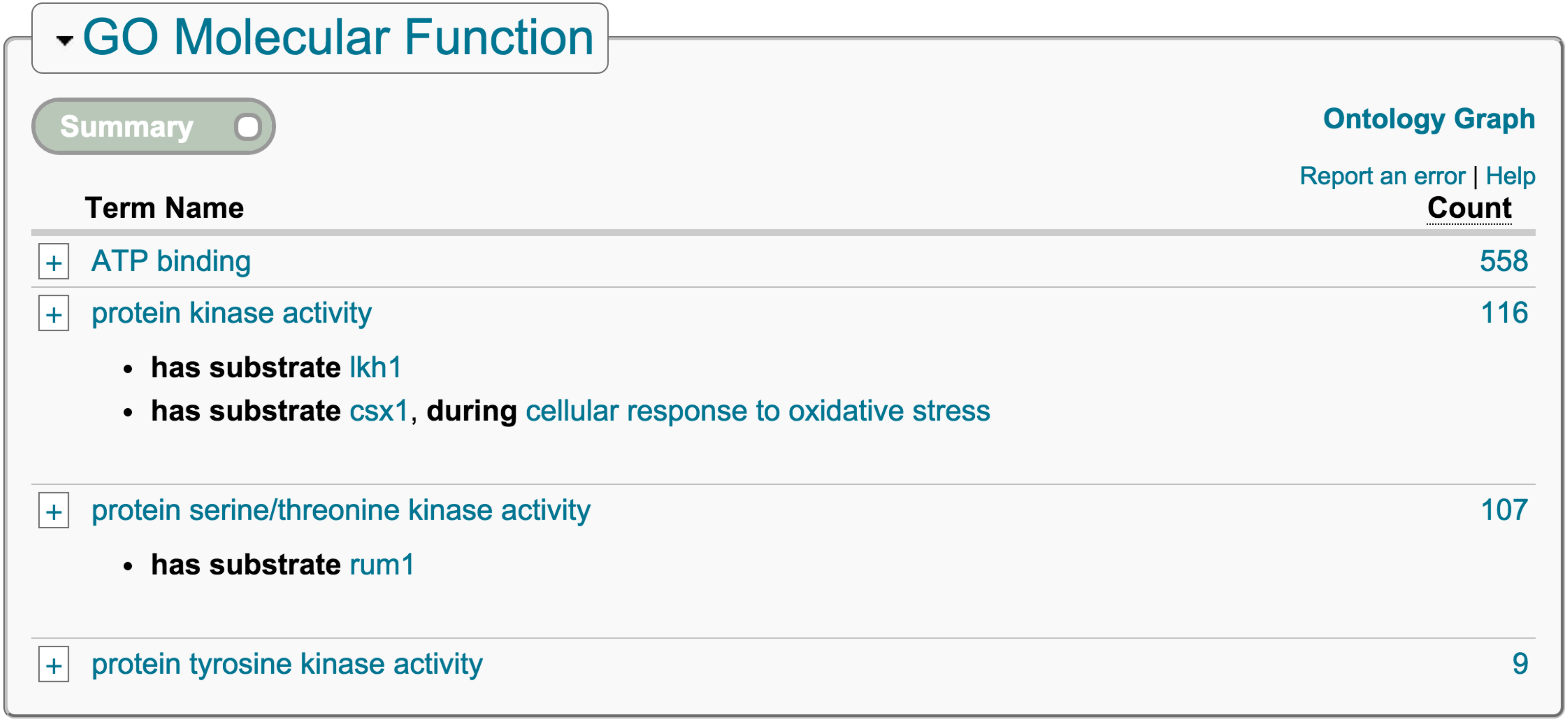
Quick Links

- Gene Ontology:
  - Molecular Function
  - Biological Process
  - Cellular Component
- Phenotype:
  - Population
  - Cell
- Transcript
- Protein Features
- Modifications
- Sequence
- Gene Expression
- Complementation
- Species Distribution
- Orthologs
- Interactions:
  - Physical
  - Genetic
- External References
- Literature

Back to top

lkh1

Gene Ontology Annotation



GO Molecular Function

Term Name	Count
ATP binding	558
protein kinase activity • **has substrate** lkh1 • **has substrate** csx1, **during** cellular response to oxidative stress	116
protein serine/threonine kinase activity • **has substrate** rum1	107
protein tyrosine kinase activity	9

In order to use ontology based annotation effectively it is useful to know some basic concepts of GO, and ontologies in general……….

GO, The basics

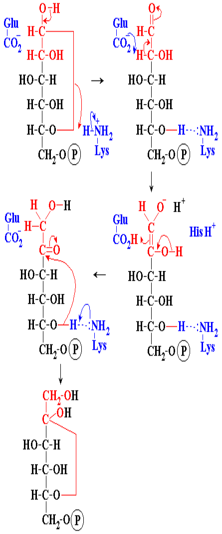
- An evolving vocabulary of terms (names for concepts) to describe attributes of gene products consistently
- Definitions (unambiguous) We annotate based on the definition
- Defined logical relationships to each other

The 3 aspects of GO

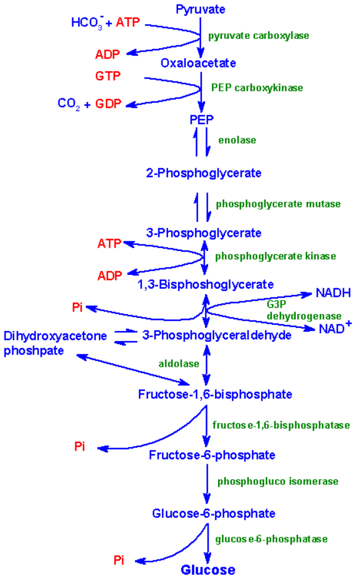
MF Molecular Function (activity, GTPase, transporter, receptor, binding)

BP Biological Process (cell division transcription, gluconeogenesis…
a collection of molecular functions in a specific context)

CC Cellular Component (location or complex)



glucose-6-phosphate
isomerase activity



gluconeogenesis

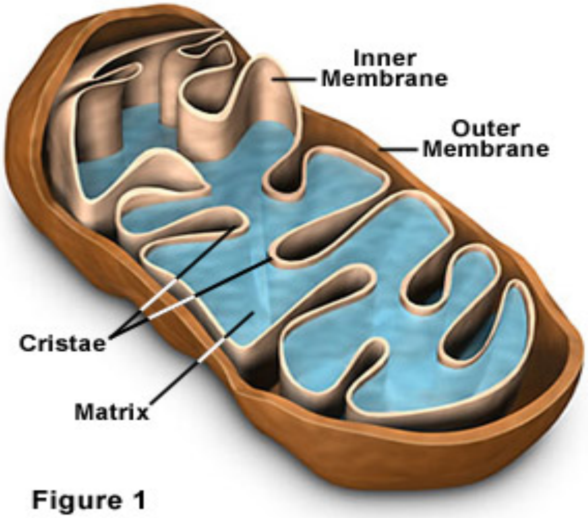


Figure 1

Biological Processes are assemblies of Molecular Functions…..

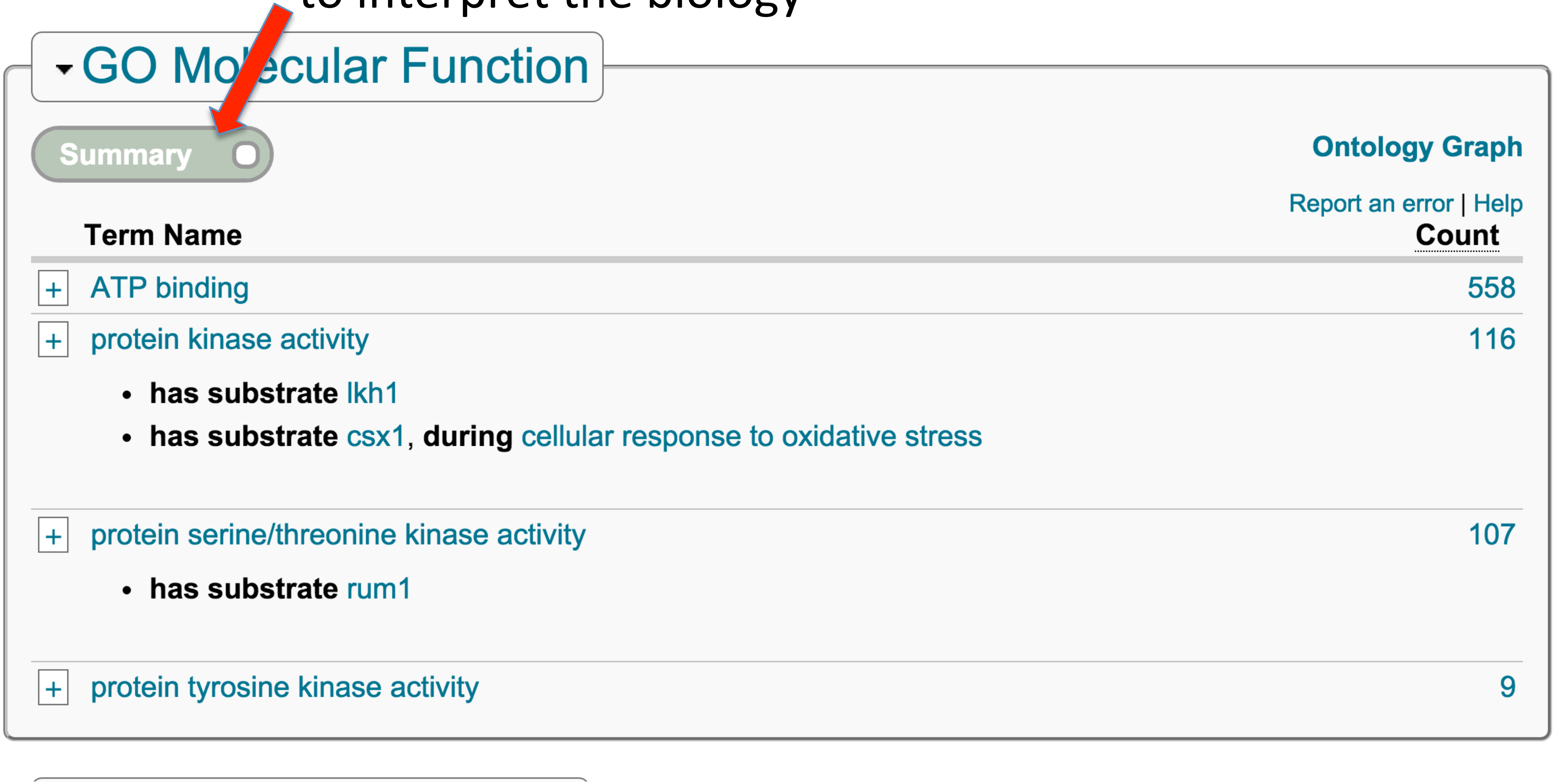
- Glycolysis BIOLOGICAL PROCESS
    1. glucokinase activity
    2. glucose-6-phosphate isomerase activity
    3. 6-phosphofructokinase activity
    4. fructose-bisphosphate aldolase activity
    5. triose-phosphate isomerase activity
    6. glyceraldehyde-3-phosphate dehydrogenase (NAD+) (phosphorylating) activity
    7. phosphoglycerate kinase activity
    8. phosphoglycerate mutase activity
    9. phosphopyruvate hydratase activity
    10. pyruvate kinase activity

MOLECULAR FUNCTIONS

Ideally we would be able to describe all processes in terms of the collection of functions….

lkh1

“Summary” view shows only the information to interpret the biology

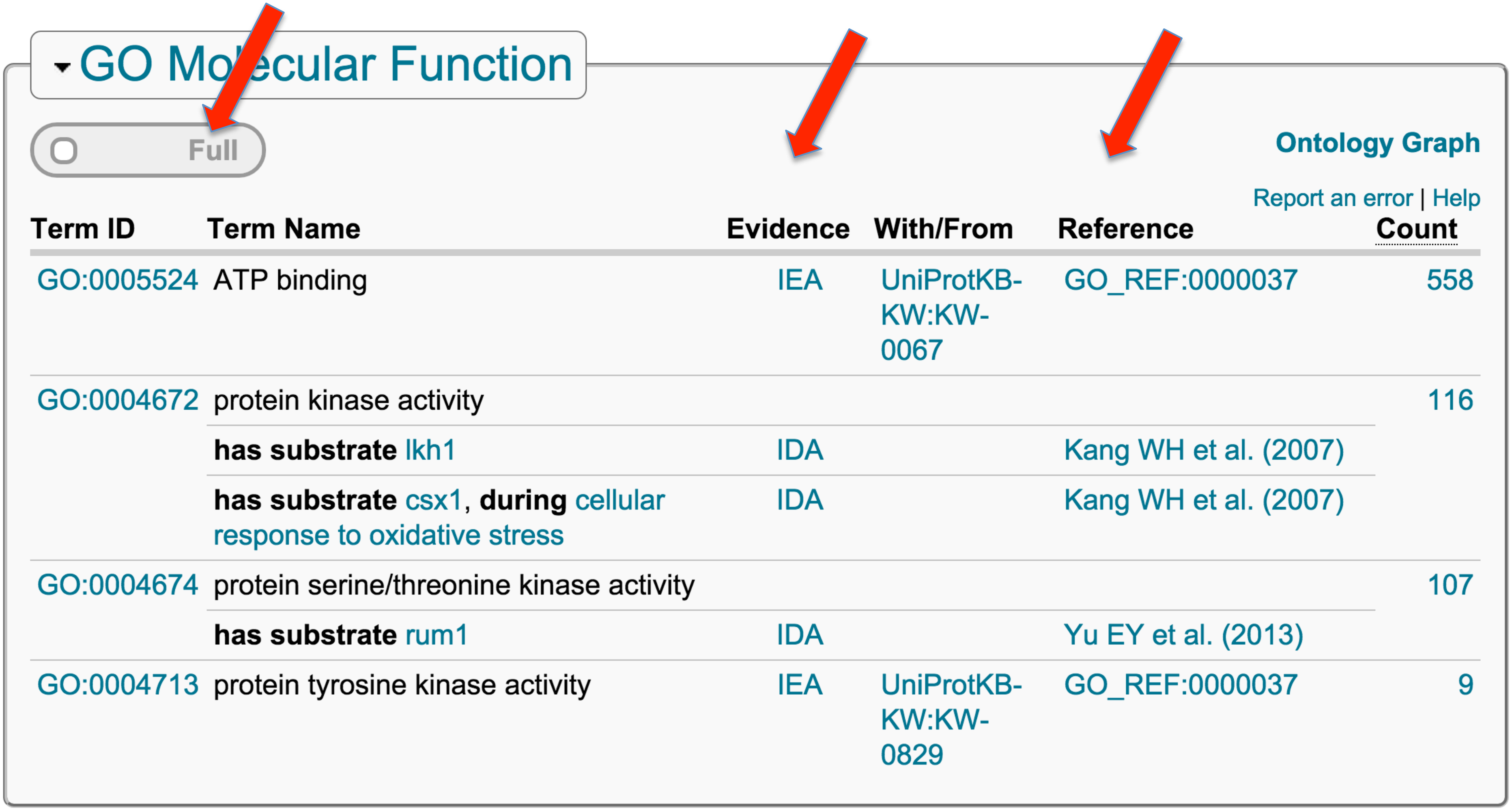


Term Name	Count
+ ATP binding	558
+ protein kinase activity • **has substrate** lkh1 • **has substrate** csx1, **during** cellular response to oxidative stress	116
+ protein serine/threonine kinase activity • **has substrate** rum1	107
+ protein tyrosine kinase activity	9

lkh1

“Full” view shows evidence code, and source (reference)

- we can capture redundant experiments but hide in the summary view



Term ID	Term Name	Evidence	With/From	Reference	Count
GO:0005524	ATP binding	IEA	UniProtKB-KW:KW-0067	GO_REF:0000037	558
GO:0004672	protein kinase activity				116
	has substrate lkh1	IDA		Kang WH et al. (2007)	
	has substrate csx1, **during** cellular response to oxidative stress	IDA		Kang WH et al. (2007)	
GO:0004674	protein serine/threonine kinase activity				107
	has substrate rum1	IDA		Yu EY et al. (2013)	
GO:0004713	protein tyrosine kinase activity	IEA	UniProtKB-KW:KW-0829	GO_REF:0000037	9

Evidence codes

Experimental/manual

- IDA inferred from direct assay
- IMP inferred from mutant phenotype
- IPI inferred from physical interaction
- IGI inferred from genetic interaction

Manual

- ISO inferred from sequence orthology (from characterized ortholog)

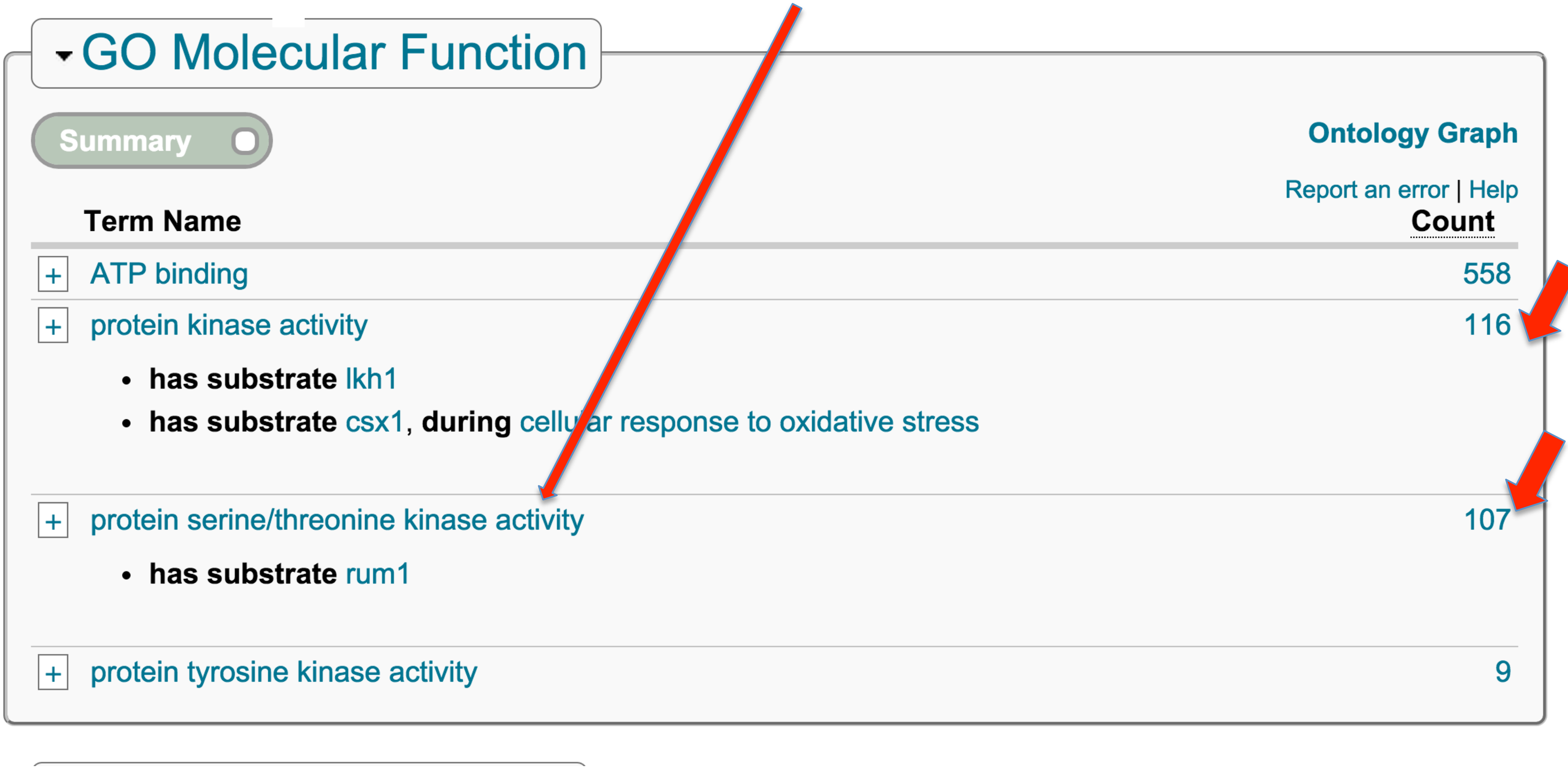
Computational

- IEA inferred from electronic annotation

(for e.g. from the presence of a domain)

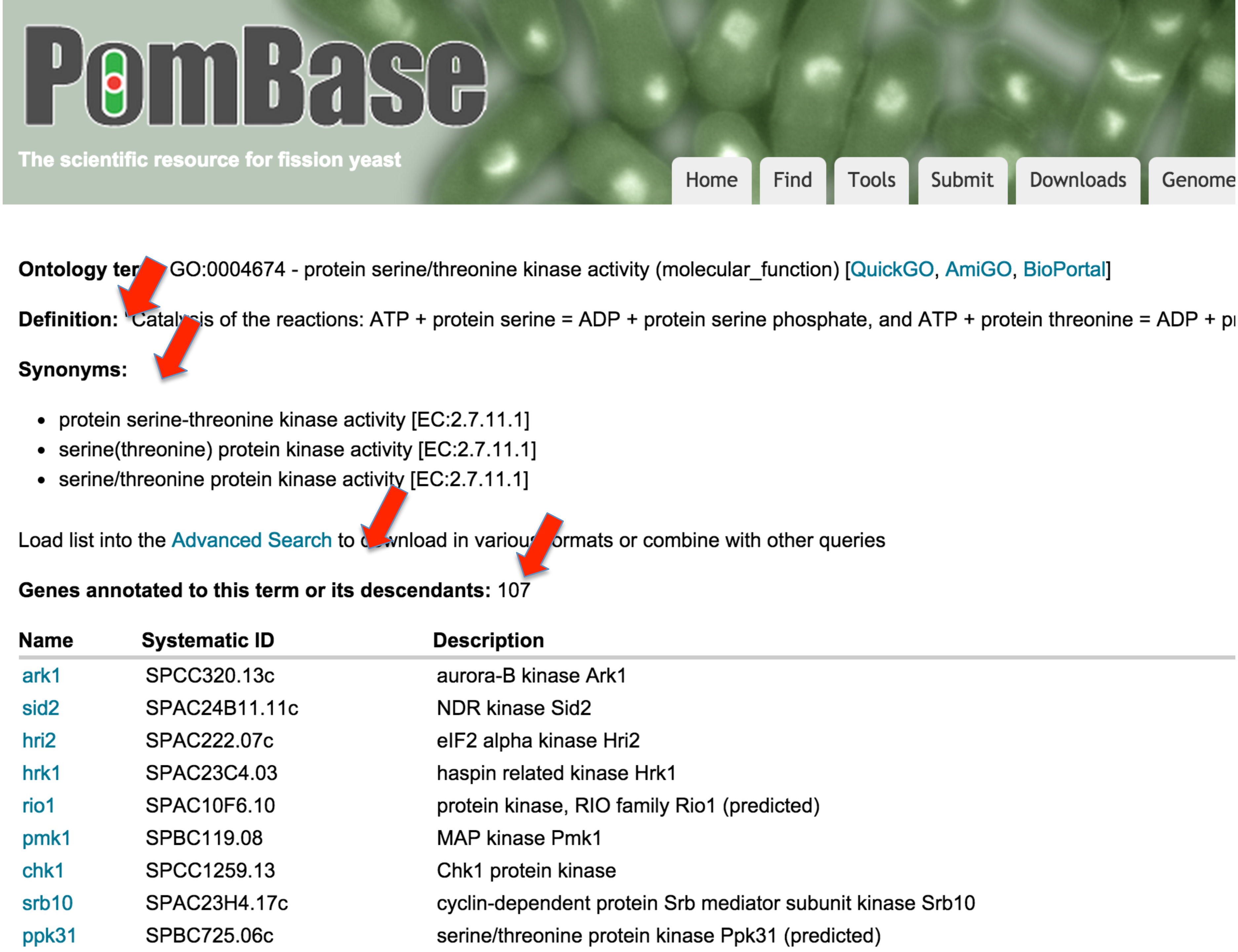
lkh1

The "term name" or "count" link takes you to the list of annotations to a term



GO Molecular Function

Term Name	Count
+ ATP binding	558
+ protein kinase activity • **has substrate** lkh1 • **has substrate** csx1, **during** cellular response to oxidative stress	116
+ protein serine/threonine kinase activity • **has substrate** rum1	107
+ protein tyrosine kinase activity	9



Ontology term: GO:0004674 - protein serine/threonine kinase activity (molecular_function) [QuickGO, AmiGO, BioPortal]

Definition: "Catalysis of the reactions: ATP + protein serine = ADP + protein serine phosphate, and ATP + protein threonine = ADP + pr

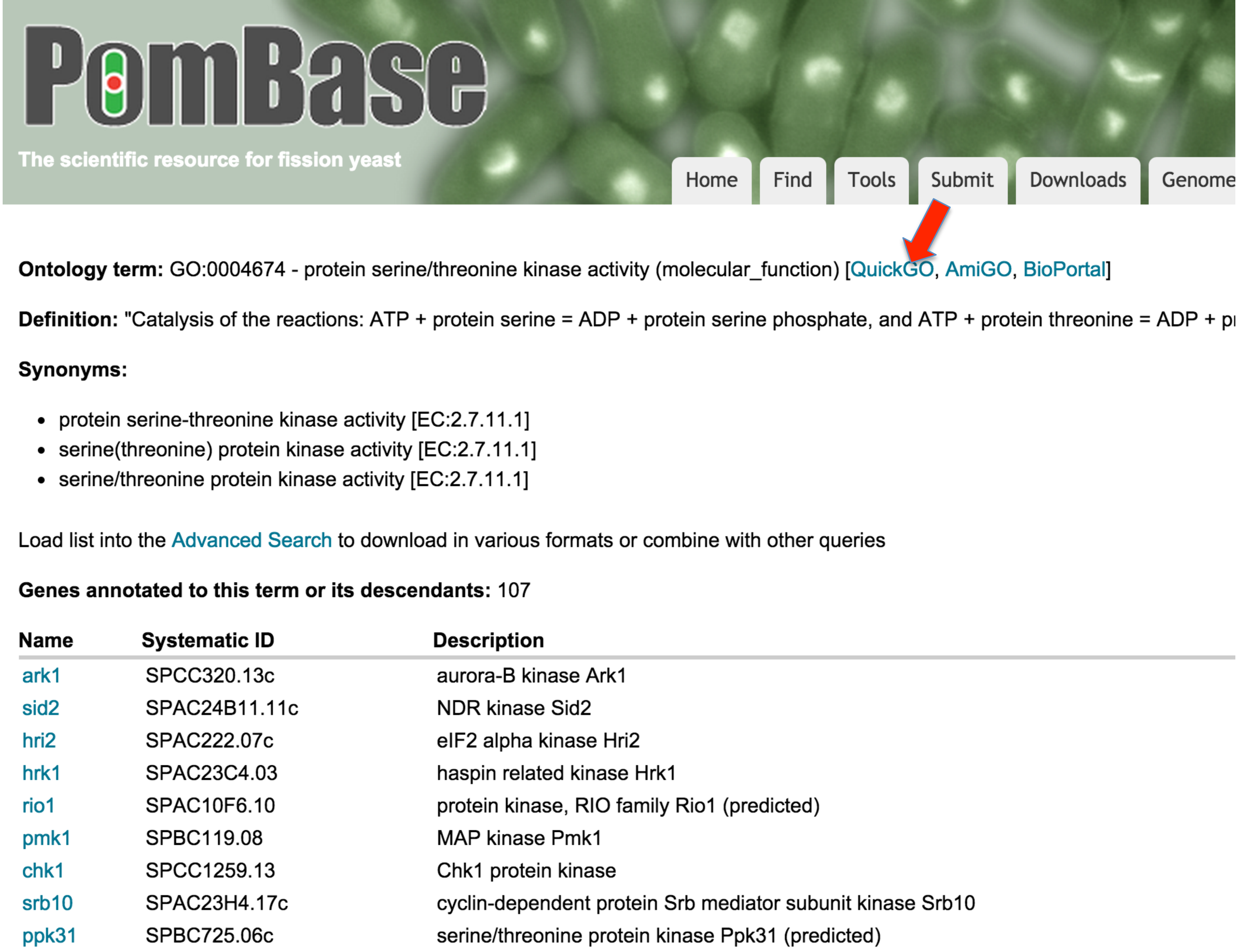
Synonyms:

- protein serine-threonine kinase activity [EC:2.7.11.1]
- serine(threonine) protein kinase activity [EC:2.7.11.1]
- serine/threonine protein kinase activity [EC:2.7.11.1]

Load list into the Advanced Search to download in various formats or combine with other queries

Genes annotated to this term or its descendants: 107

Name	Systematic ID	Description
ark1	SPCC320.13c	aurora-B kinase Ark1
sid2	SPAC24B11.11c	NDR kinase Sid2
hri2	SPAC222.07c	eIF2 alpha kinase Hri2
hrk1	SPAC23C4.03	haspin related kinase Hrk1
rio1	SPAC10F6.10	protein kinase, RIO family Rio1 (predicted)
pmk1	SPBC119.08	MAP kinase Pmk1
chk1	SPCC1259.13	Chk1 protein kinase
srb10	SPAC23H4.17c	cyclin-dependent protein Srb mediator subunit kinase Srb10
ppk31	SPBC725.06c	serine/threonine protein kinase Ppk31 (predicted)



Ontology term: GO:0004674 - protein serine/threonine kinase activity (molecular_function) [QuickGO, AmiGO, BioPortal]

Definition: "Catalysis of the reactions: ATP + protein serine = ADP + protein serine phosphate, and ATP + protein threonine = ADP + p

Synonyms:

- protein serine-threonine kinase activity [EC:2.7.11.1]
- serine(threonine) protein kinase activity [EC:2.7.11.1]
- serine/threonine protein kinase activity [EC:2.7.11.1]

Load list into the Advanced Search to download in various formats or combine with other queries

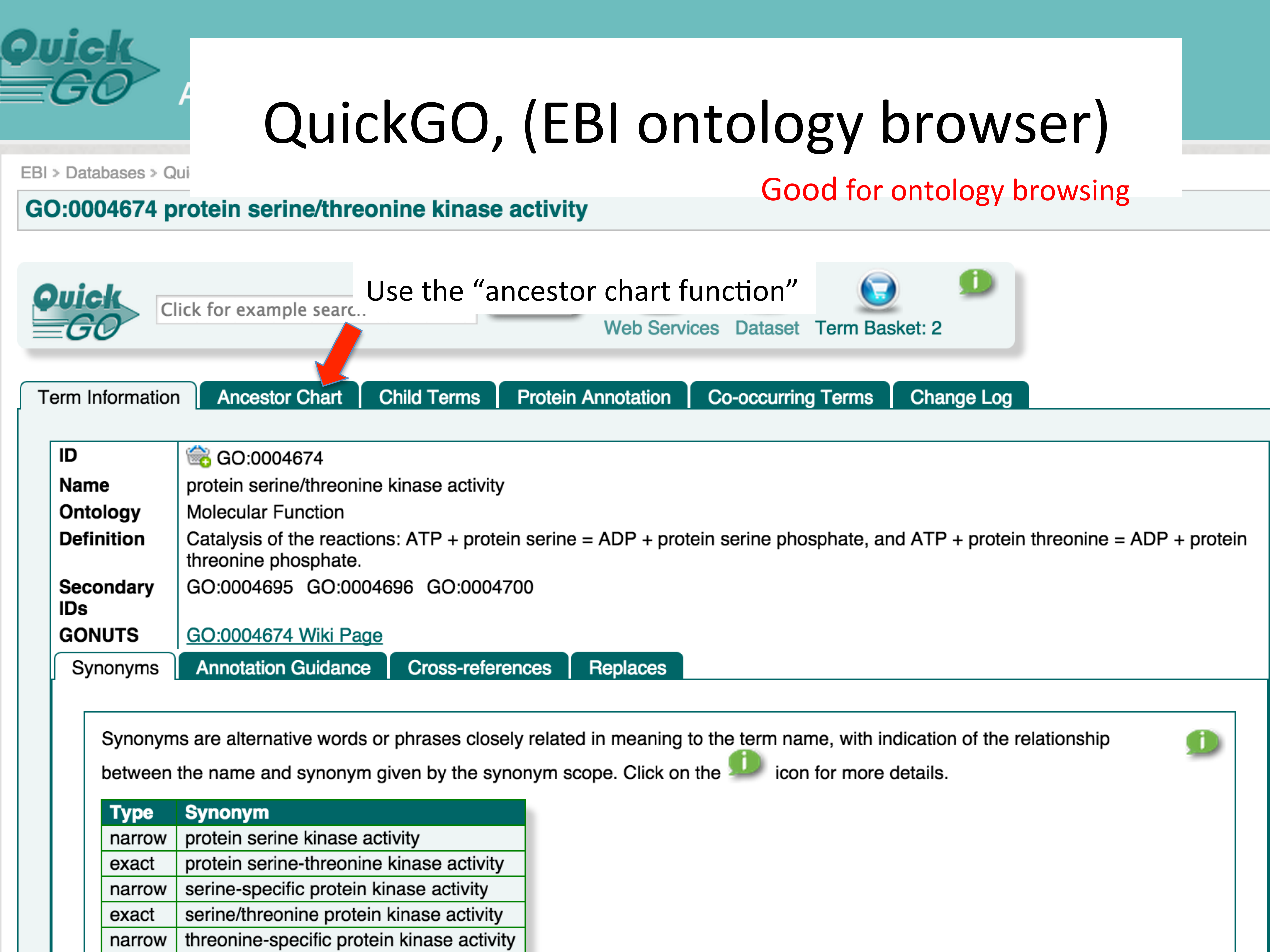
Genes annotated to this term or its descendants: 107

Name	Systematic ID	Description
ark1	SPCC320.13c	aurora-B kinase Ark1
sid2	SPAC24B11.11c	NDR kinase Sid2
hri2	SPAC222.07c	eIF2 alpha kinase Hri2
hrk1	SPAC23C4.03	haspin related kinase Hrk1
rio1	SPAC10F6.10	protein kinase, RIO family Rio1 (predicted)
pmk1	SPBC119.08	MAP kinase Pmk1
chk1	SPCC1259.13	Chk1 protein kinase
srb10	SPAC23H4.17c	cyclin-dependent protein Srb mediator subunit kinase Srb10
ppk31	SPBC725.06c	serine/threonine protein kinase Ppk31 (predicted)

QuickGO, (EBI ontology browser)

Good for ontology browsing

Use the “ancestor chart function”

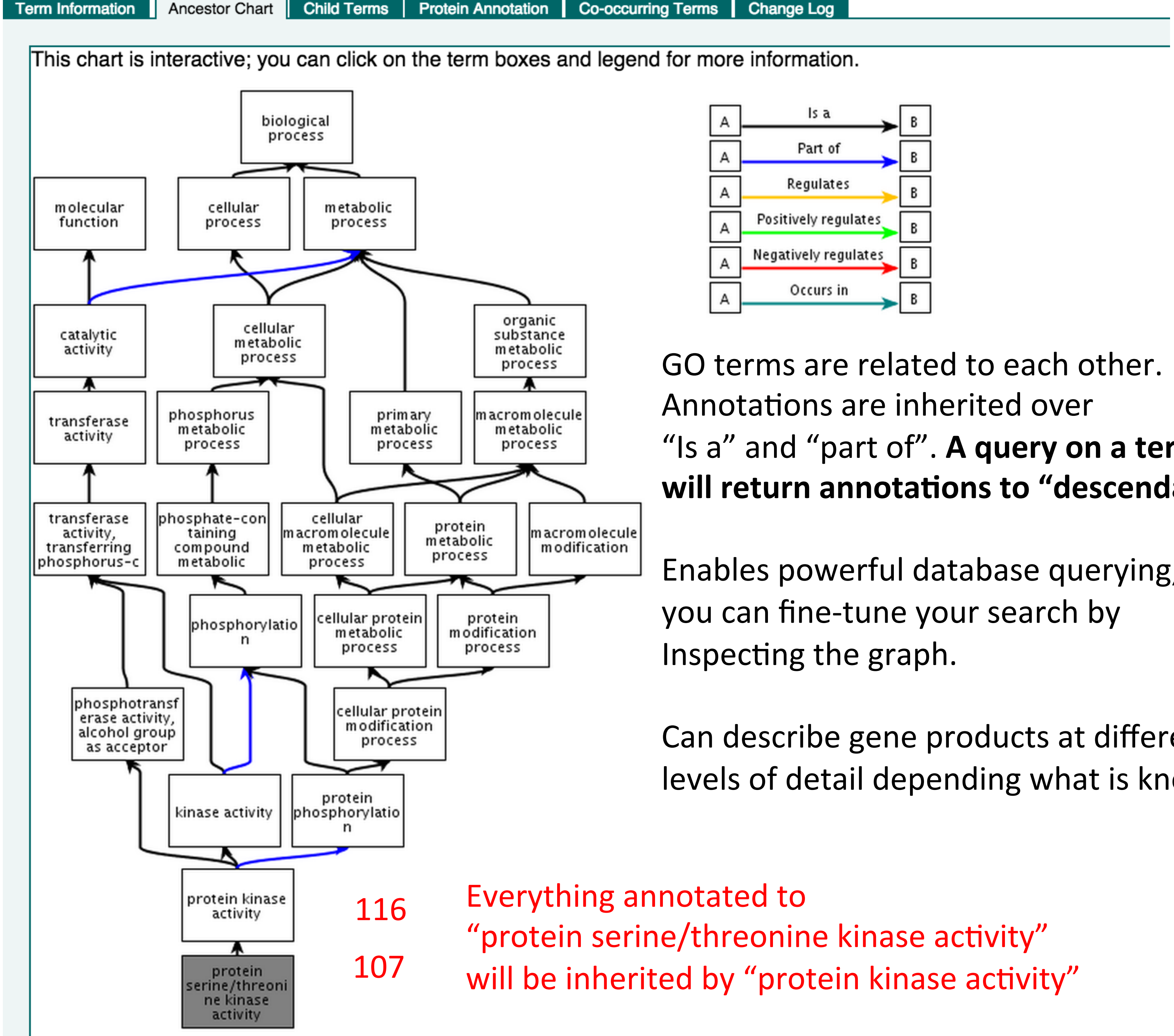


GO:0004674 protein serine/threonine kinase activity

ID	GO:0004674
Name	protein serine/threonine kinase activity
Ontology	Molecular Function
Definition	Catalysis of the reactions: ATP + protein serine = ADP + protein serine phosphate, and ATP + protein threonine = ADP + protein threonine phosphate.
Secondary IDs	GO:0004695 GO:0004696 GO:0004700
GONUTS	GO:0004674 Wiki Page

Synonyms are alternative words or phrases closely related in meaning to the term name, with indication of the relationship between the name and synonym given by the synonym scope. Click on the icon for more details.

Type	Synonym
narrow	protein serine kinase activity
exact	protein serine-threonine kinase activity
narrow	serine-specific protein kinase activity
exact	serine/threonine protein kinase activity
narrow	threonine-specific protein kinase activity



This chart is interactive; you can click on the term boxes and legend for more information.

GO terms are related to each other. Annotations are inherited over "Is a" and "part of". **A query on a term will return annotations to "descendants"**

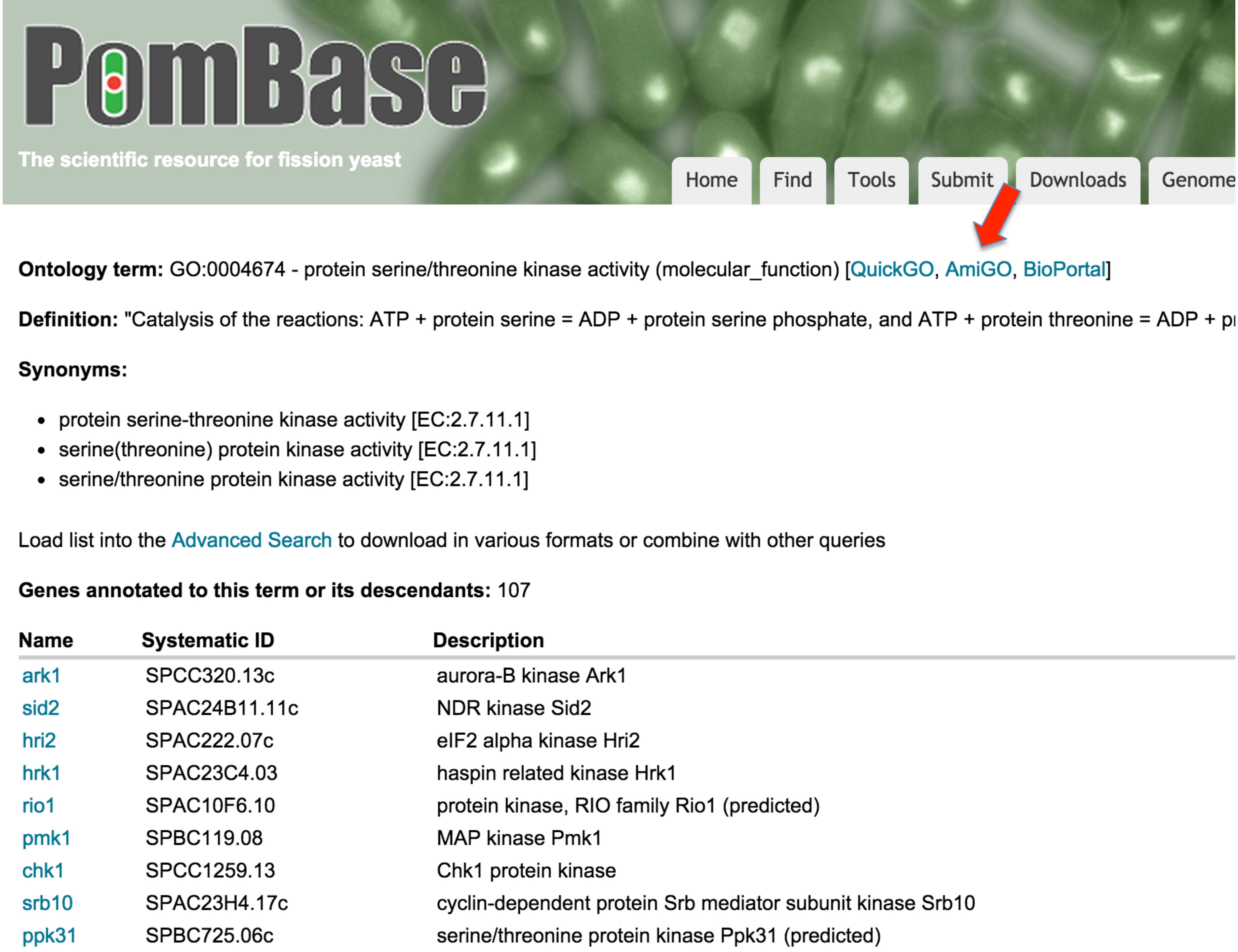
Enables powerful database querying, you can fine-tune your search by Inspecting the graph.

Can describe gene products at different levels of detail depending what is known

116

107

Everything annotated to
"protein serine/threonine kinase activity"
will be inherited by "protein kinase activity"



Ontology term: GO:0004674 - protein serine/threonine kinase activity (molecular_function) [QuickGO, AmiGO, BioPortal]

Definition: "Catalysis of the reactions: ATP + protein serine = ADP + protein serine phosphate, and ATP + protein threonine = ADP + p

Synonyms:

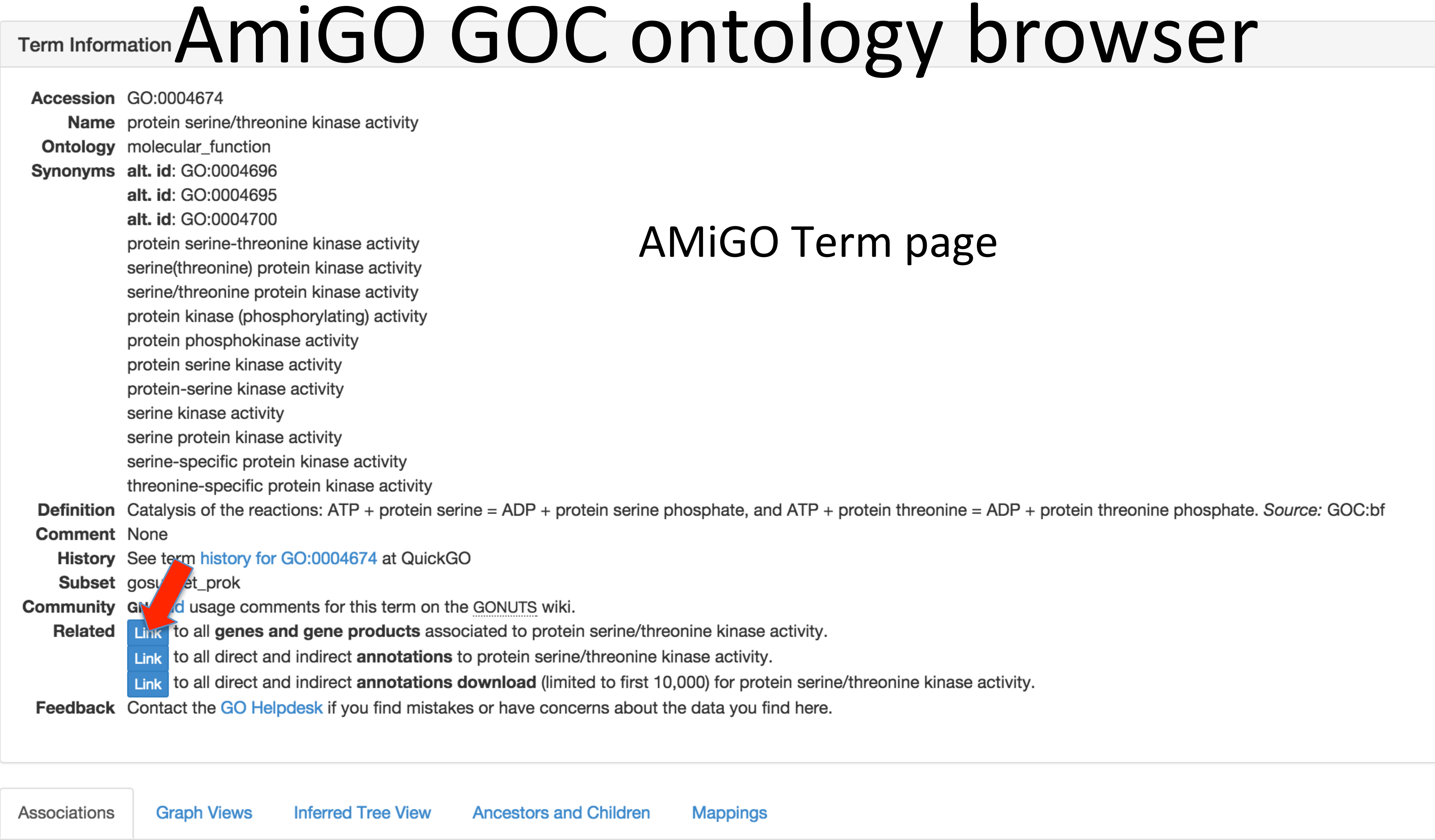
- protein serine-threonine kinase activity [EC:2.7.11.1]
- serine(threonine) protein kinase activity [EC:2.7.11.1]
- serine/threonine protein kinase activity [EC:2.7.11.1]

Load list into the Advanced Search to download in various formats or combine with other queries

Genes annotated to this term or its descendants: 107

Name	Systematic ID	Description
ark1	SPCC320.13c	aurora-B kinase Ark1
sid2	SPAC24B11.11c	NDR kinase Sid2
hri2	SPAC222.07c	eIF2 alpha kinase Hri2
hrk1	SPAC23C4.03	haspin related kinase Hrk1
rio1	SPAC10F6.10	protein kinase, RIO family Rio1 (predicted)
pmk1	SPBC119.08	MAP kinase Pmk1
chk1	SPCC1259.13	Chk1 protein kinase
srb10	SPAC23H4.17c	cyclin-dependent protein Srb mediator subunit kinase Srb10
ppk31	SPBC725.06c	serine/threonine protein kinase Ppk31 (predicted)

AmiGO GOC ontology browser



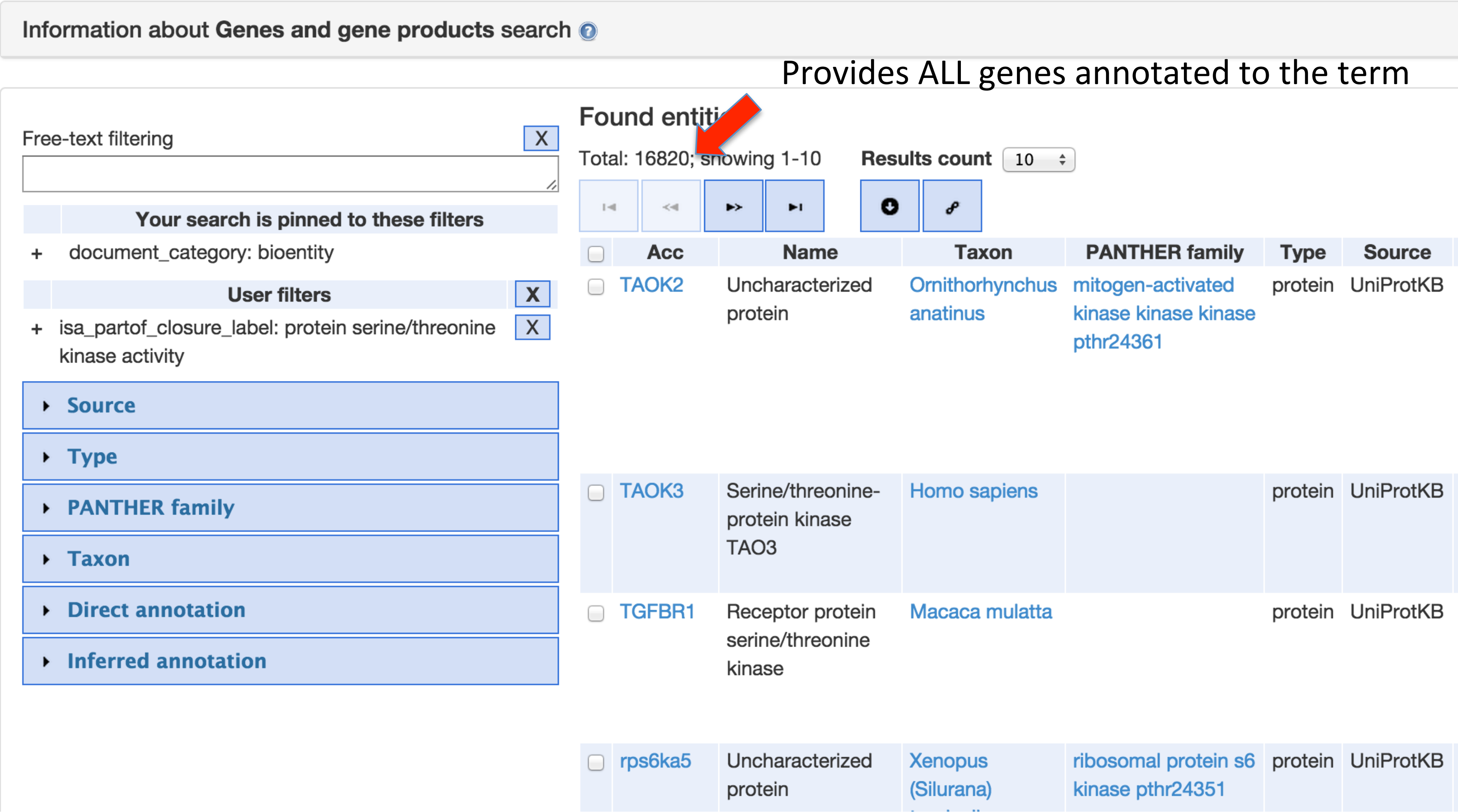
Term Information

Accession GO:0004674
Name protein serine/threonine kinase activity
Ontology molecular_function
Synonyms **alt. id**: GO:0004696
alt. id: GO:0004695
alt. id: GO:0004700
protein serine-threonine kinase activity
serine(threonine) protein kinase activity
serine/threonine protein kinase activity
protein kinase (phosphorylating) activity
protein phosphokinase activity
protein serine kinase activity
protein-serine kinase activity
serine kinase activity
serine protein kinase activity
serine-specific protein kinase activity
threonine-specific protein kinase activity
Definition Catalysis of the reactions: ATP + protein serine = ADP + protein serine phosphate, and ATP + protein threonine = ADP + protein threonine phosphate. *Source:* GOC:bf
Comment None
History See term history for GO:0004674 at QuickGO
Subset gosu et_prok
Community G d usage comments for this term on the GONUTS wiki.
Related Link to all **genes and gene products** associated to protein serine/threonine kinase activity.
Link to all direct and indirect **annotations** to protein serine/threonine kinase activity.
Link to all direct and indirect **annotations download** (limited to first 10,000) for protein serine/threonine kinase activity.
Feedback Contact the GO Helpdesk if you find mistakes or have concerns about the data you find here.

AMiGO Term page

AmiGO

Useful to access annotations, especially in other species



Provides ALL genes annotated to the term

Information about Genes and gene products search

Free-text filtering

Your search is pinned to these filters

+ document_category: bioentity

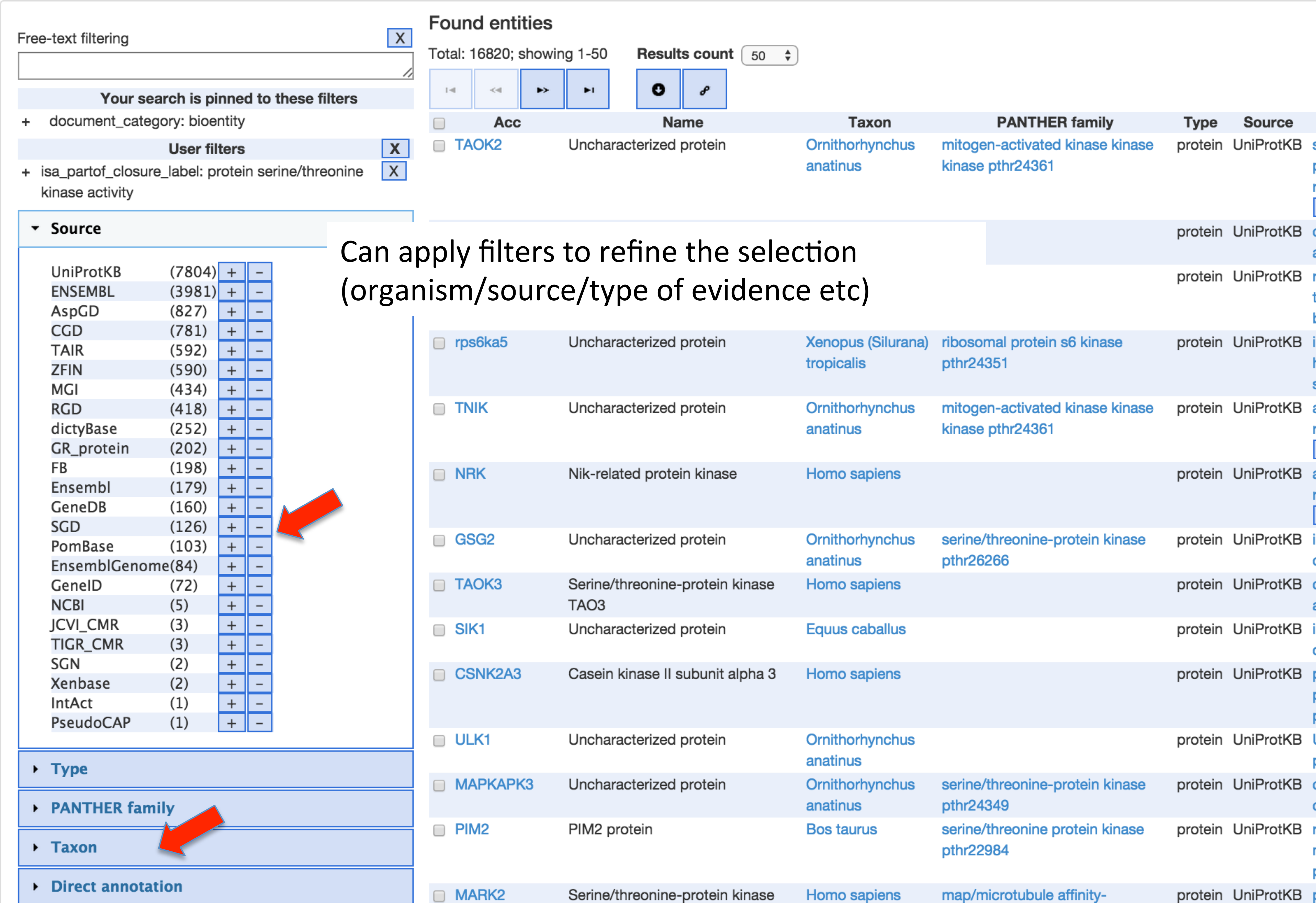
User filters

+ isa_partof_closure_label: protein serine/threonine kinase activity

Found entiti

Total: 16820; showing 1-10 **Results count** 10

☐	Acc	Name	Taxon	PANTHER family	Type	Source
☐	TAOK2	Uncharacterized protein	Ornithorhynchus anatinus	mitogen-activated kinase kinase kinase pthr24361	protein	UniProtKB
☐	TAOK3	Serine/threonine-protein kinase TAO3	Homo sapiens		protein	UniProtKB
☐	TGFBR1	Receptor protein serine/threonine kinase	Macaca mulatta		protein	UniProtKB
☐	rps6ka5	Uncharacterized protein	Xenopus (Silurana)	ribosomal protein s6 kinase pthr24351	protein	UniProtKB



Free-text filtering

+ document_category: bioentity

+ isa_partof_closure_label: protein serine/threonine kinase activity

Source

UniProtKB	(7804)
ENSEMBL	(3981)
AspGD	(827)
CGD	(781)
TAIR	(592)
ZFIN	(590)
MGI	(434)
RGD	(418)
dictyBase	(252)
GR_protein	(202)
FB	(198)
Ensembl	(179)
GeneDB	(160)
SGD	(126)
PomBase	(103)
EnsemblGenome	(84)
GeneID	(72)
NCBI	(5)
JCVI_CMR	(3)
TIGR_CMR	(3)
SGN	(2)
Xenbase	(2)
IntAct	(1)
PseudoCAP	(1)

Type

PANTHER family

Taxon

Direct annotation

Found entities

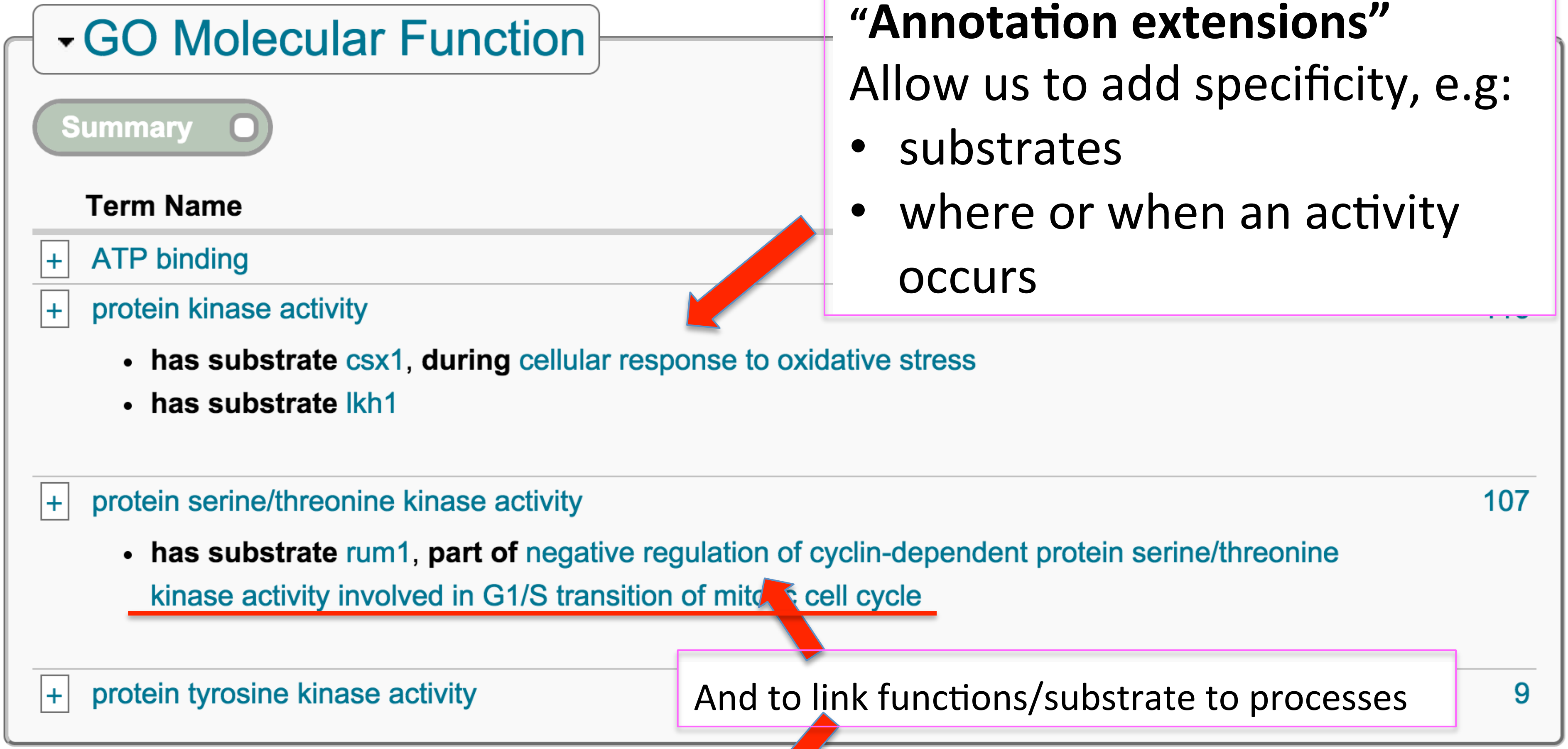
Total: 16820; showing 1-50

Acc	Name	Taxon	PANTHER family	Type	Source
TAOK2	Uncharacterized protein	Ornithorhynchus anatinus	mitogen-activated kinase kinase kinase pthr24361	protein	UniProtKB
				protein	UniProtKB
				protein	UniProtKB
rps6ka5	Uncharacterized protein	Xenopus (Silurana) tropicalis	ribosomal protein s6 kinase pthr24351	protein	UniProtKB
TNIK	Uncharacterized protein	Ornithorhynchus anatinus	mitogen-activated kinase kinase kinase pthr24361	protein	UniProtKB
NRK	Nik-related protein kinase	Homo sapiens		protein	UniProtKB
GSG2	Uncharacterized protein	Ornithorhynchus anatinus	serine/threonine-protein kinase pthr26266	protein	UniProtKB
TAOK3	Serine/threonine-protein kinase TAO3	Homo sapiens		protein	UniProtKB
SIK1	Uncharacterized protein	Equus caballus		protein	UniProtKB
CSNK2A3	Casein kinase II subunit alpha 3	Homo sapiens		protein	UniProtKB
ULK1	Uncharacterized protein	Ornithorhynchus anatinus		protein	UniProtKB
MAPKAPK3	Uncharacterized protein	Ornithorhynchus anatinus	serine/threonine-protein kinase pthr24349	protein	UniProtKB
PIM2	PIM2 protein	Bos taurus	serine/threonine protein kinase pthr22984	protein	UniProtKB
MARK2	Serine/threonine-protein kinase	Homo sapiens	map/microtubule affinity-	protein	UniProtKB

Can apply filters to refine the selection (organism/source/type of evidence etc)

lkh1

“Annotation extensions”
Allow us to add specificity, e.g:
- substrates
- where or when an activity occurs



GO Molecular Function

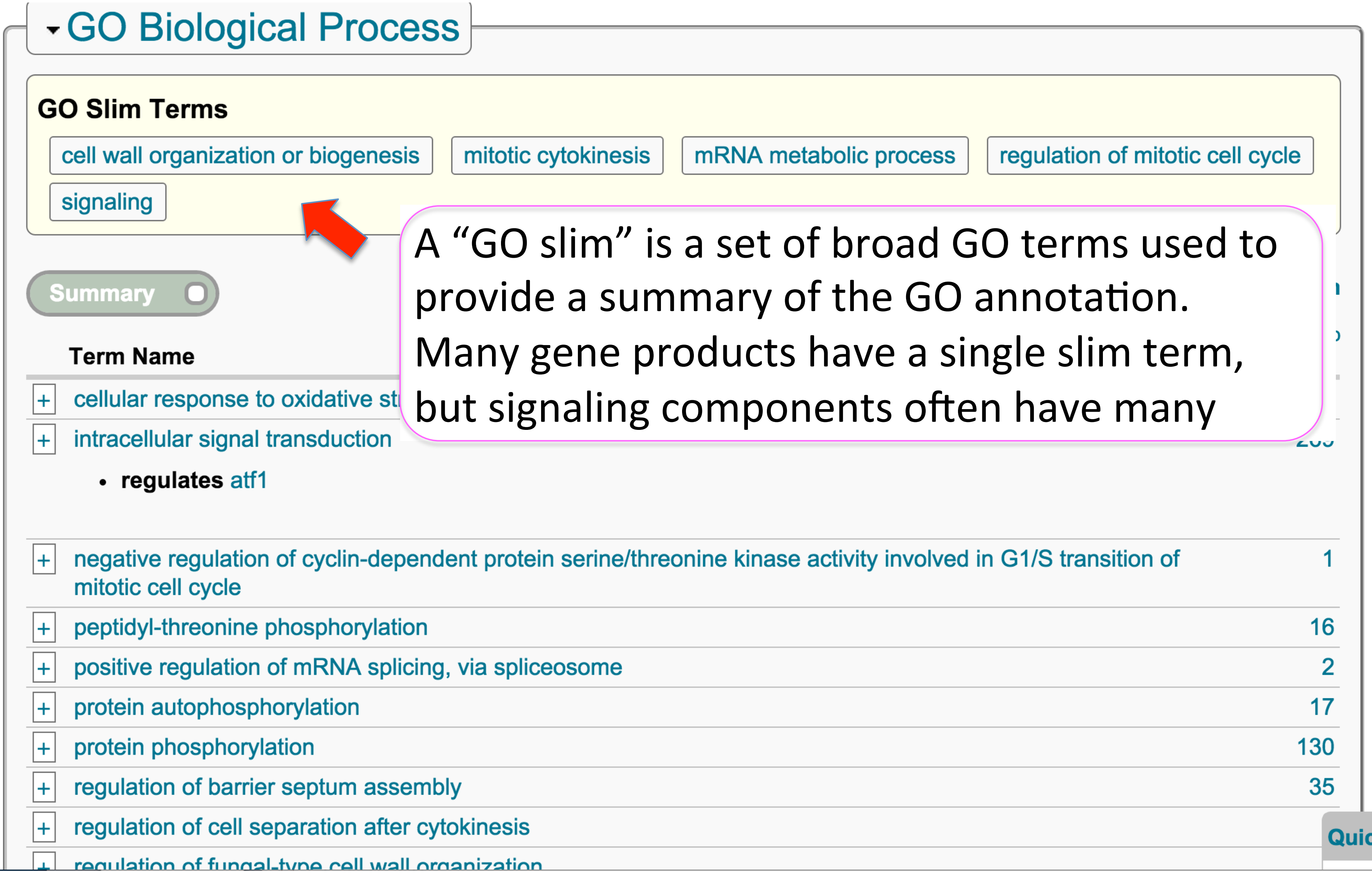
Term Name	
＋ ATP binding	
＋ protein kinase activity • **has substrate** csx1, **during** cellular response to oxidative stress • **has substrate** lkh1	
＋ protein serine/threonine kinase activity • **has substrate** rum1, **part of** negative regulation of cyclin-dependent protein serine/threonine kinase activity involved in G1/S transition of mitotic cell cycle	107
＋ protein tyrosine kinase activity	9

And to link functions/substrate to processes

GO Biological Process

＋ negative regulation of cyclin-dependent protein serine/threonine kinase activity involved in G1/S transition of mitotic cell cycle	1
A splicing, via spliceosome	2
n	17
＋ protein phosphorylation	132
＋ regulation of barrier septum assembly	35
＋ regulation of cell separation after cytokinesis	5
＋ regulation of fungal-type cell wall organization	2

lkh1



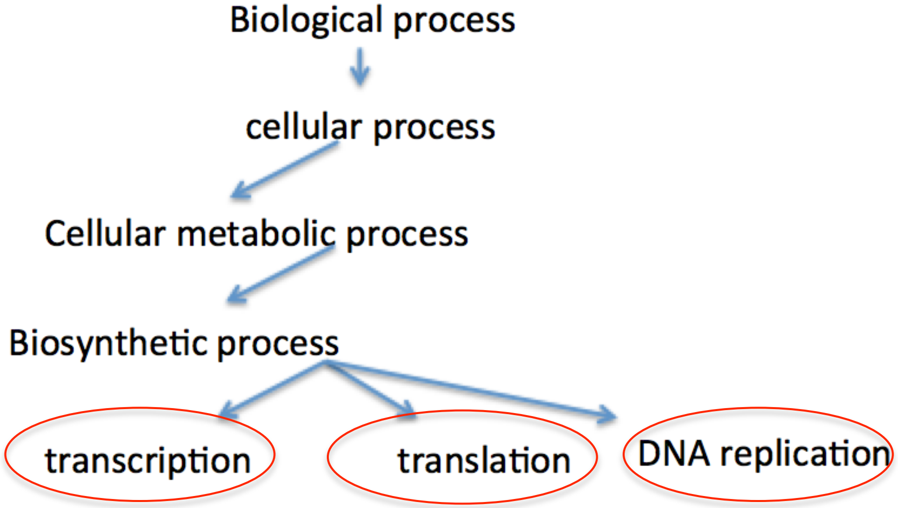
GO Slim Terms

cell wall organization or biogenesis | mitotic cytokinesis | mRNA metabolic process | regulation of mitotic cell cycle | signaling

A “GO slim” is a set of broad GO terms used to provide a summary of the GO annotation. Many gene products have a single slim term, but signaling components often have many

Term Name	
+ cellular response to oxidative st	
+ intracellular signal transduction • **regulates** atf1	
+ negative regulation of cyclin-dependent protein serine/threonine kinase activity involved in G1/S transition of mitotic cell cycle	1
+ peptidyl-threonine phosphorylation	16
+ positive regulation of mRNA splicing, via spliceosome	2
+ protein autophosphorylation	17
+ protein phosphorylation	130
+ regulation of barrier septum assembly	35
+ regulation of cell separation after cytokinesis	
+ regulation of fungal-type cell wall organization	

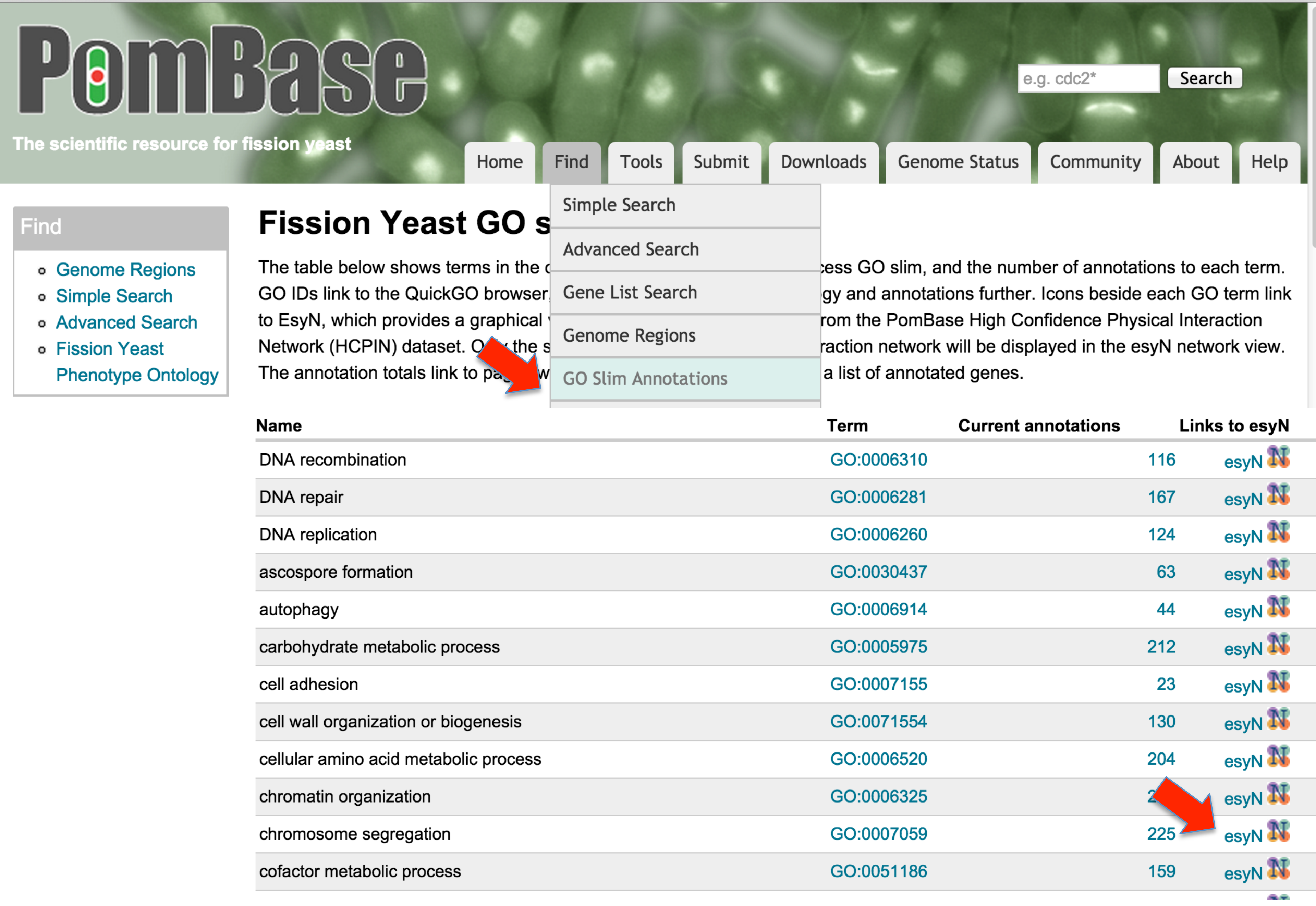
“GO slimming”


Biological process
cellular process
Cellular metabolic process
Biosynthetic process
transcription
translation
DNA replication

"GO slims" are *informative* subsets of the Gene Ontology (GO) that provide a broad overview of annotation distribution. Slims offer a useful overview of a genome or a way of summarizing the results of a large-scale experiment.

Because some gene products can ‘map’ to multiple slim terms it is difficult to generate a non-overlapping slim.

For more information on GO slims, see PomBase Documentation.
Different “slim sets” can be defined for different purposes
A tool to generate user specific “GO slims” http://go.princeton.edu/cgi-bin/GOTermMapper



Find

- Genome Regions
- Simple Search
- Advanced Search
- Fission Yeast Phenotype Ontology

Fission Yeast GO s

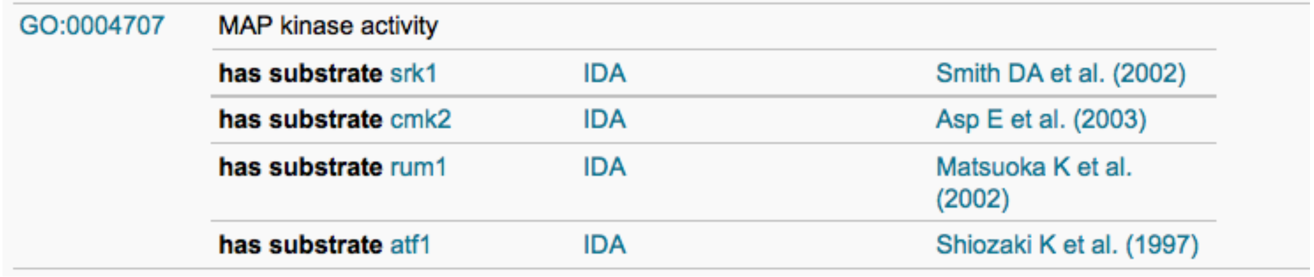
The table below shows terms in the c ess GO slim, and the number of annotations to each term.
GO IDs link to the QuickGO browser gy and annotations further. Icons beside each GO term link
to EsyN, which provides a graphical v rom the PomBase High Confidence Physical Interaction
Network (HCPIN) dataset. O the s raction network will be displayed in the esyN network view.
The annotation totals link to pa w a list of annotated genes.

Name	Term	Current annotations	Links to esyN
DNA recombination	GO:0006310	116	esyN
DNA repair	GO:0006281	167	esyN
DNA replication	GO:0006260	124	esyN
ascospore formation	GO:0030437	63	esyN
autophagy	GO:0006914	44	esyN
carbohydrate metabolic process	GO:0005975	212	esyN
cell adhesion	GO:0007155	23	esyN
cell wall organization or biogenesis	GO:0071554	130	esyN
cellular amino acid metabolic process	GO:0006520	204	esyN
chromatin organization	GO:0006325	2	esyN
chromosome segregation	GO:0007059	225	esyN
cofactor metabolic process	GO:0051186	159	esyN

Using GO Annotation Extensions to build networks

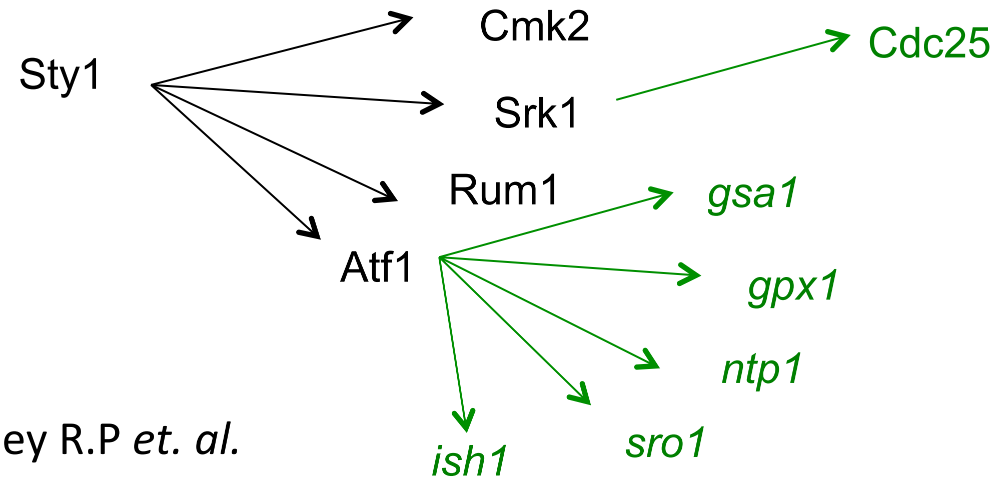
Previously all annotations were 'independent' statements

Annotation extensions allow us to connect annotations



GO:0004707	MAP kinase activity		
	has substrate srk1	IDA	Smith DA et al. (2002)
	has substrate cmk2	IDA	Asp E et al. (2003)
	has substrate rum1	IDA	Matsuoka K et al. (2002)
	has substrate atf1	IDA	Shiozaki K et al. (1997)

Schematically:



Sty1

Cmk2

Srk1

Cdc25

Rum1

Atf1

gsa1

gpx1

ntp1

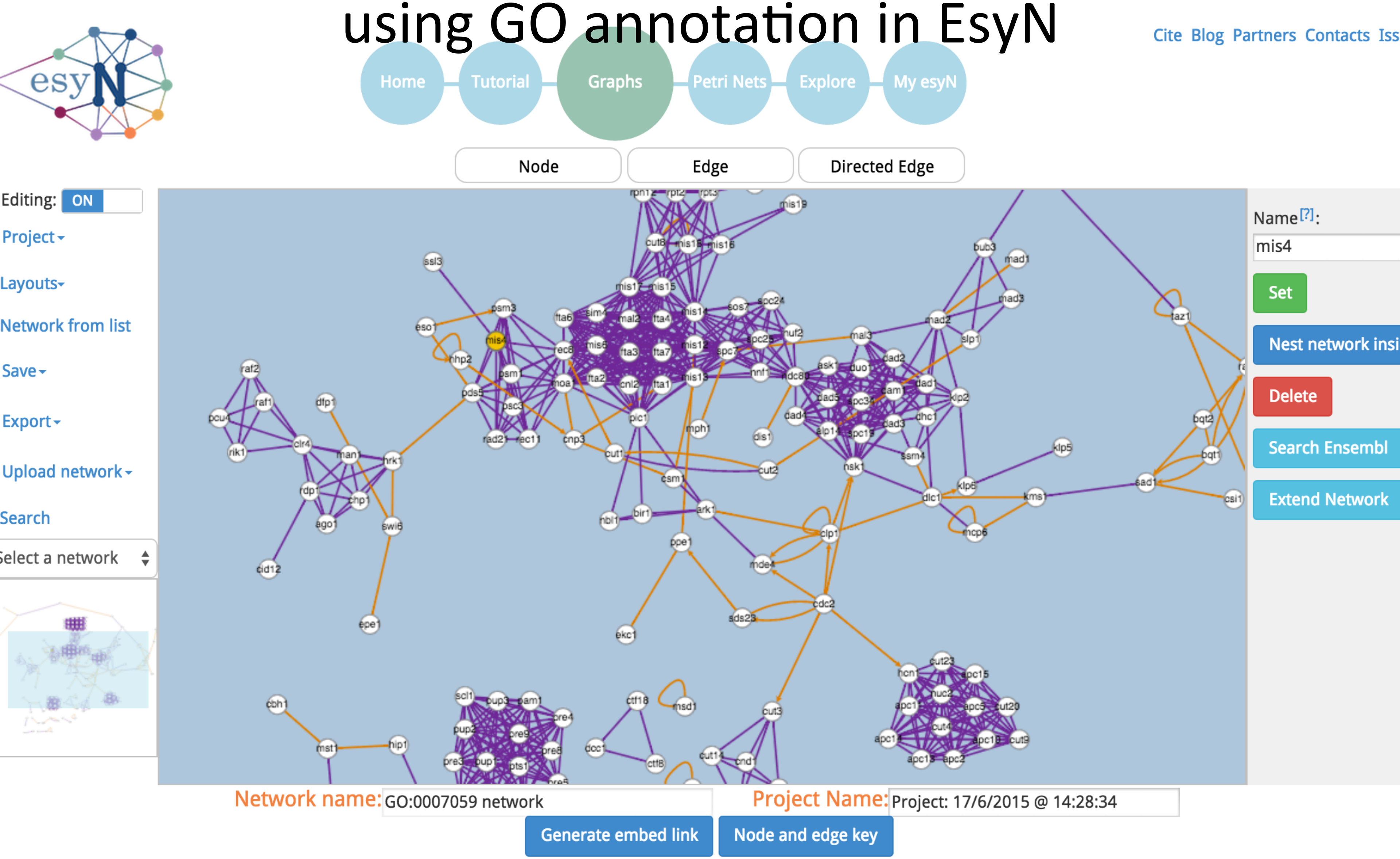
sro1

ish1

PomBase have used GO annotation & extensions to build networks using EsyN

For more about AE's PMID:24885854 Huntley R.P *et. al.* A method for increasing expressivity of Gene Ontology annotations using a compositional approach.

network for chromosome segregation using GO annotation in EsyN


Editing: ON

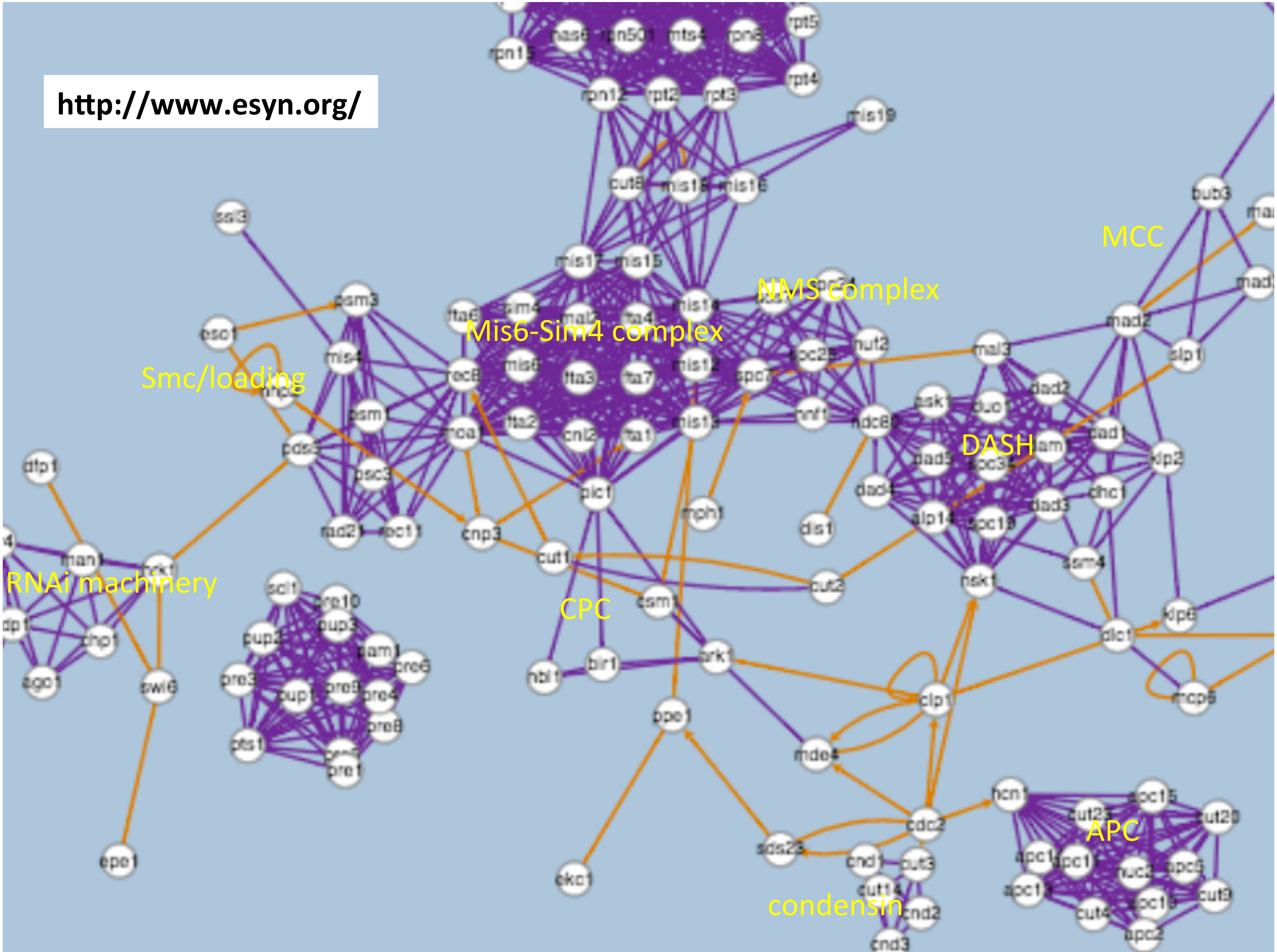
Name[?]: mis4

Network name: GO:0007059 network

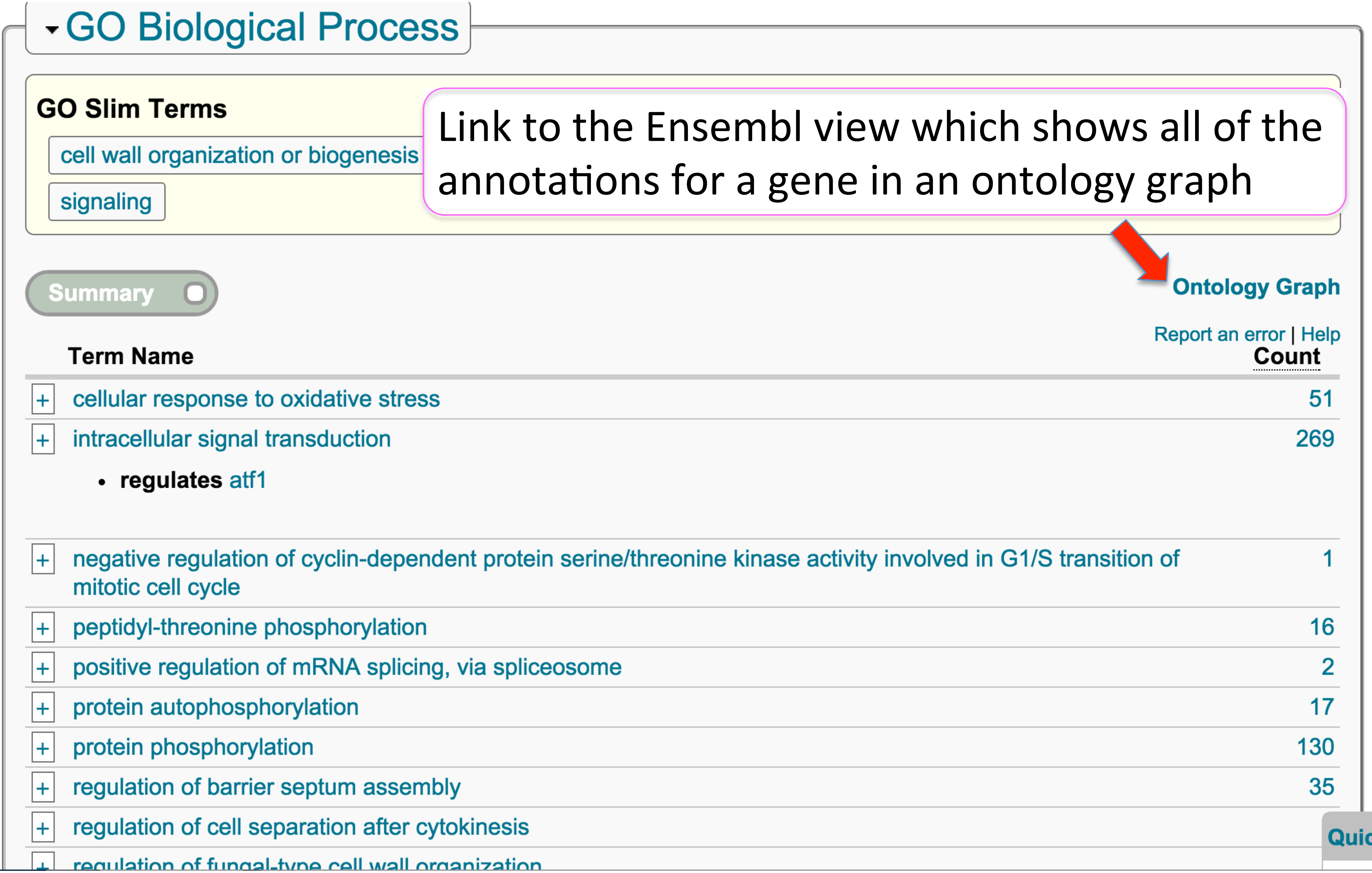
Project Name: Project: 17/6/2015 @ 14:28:34

Login to edit(correct/extend)/save

EsyN developed by Dan Bean and Giorgio Favrin at CSBC) http://www.esyn.org/



lkh1



Link to the Ensembl view which shows all of the annotations for a gene in an ontology graph

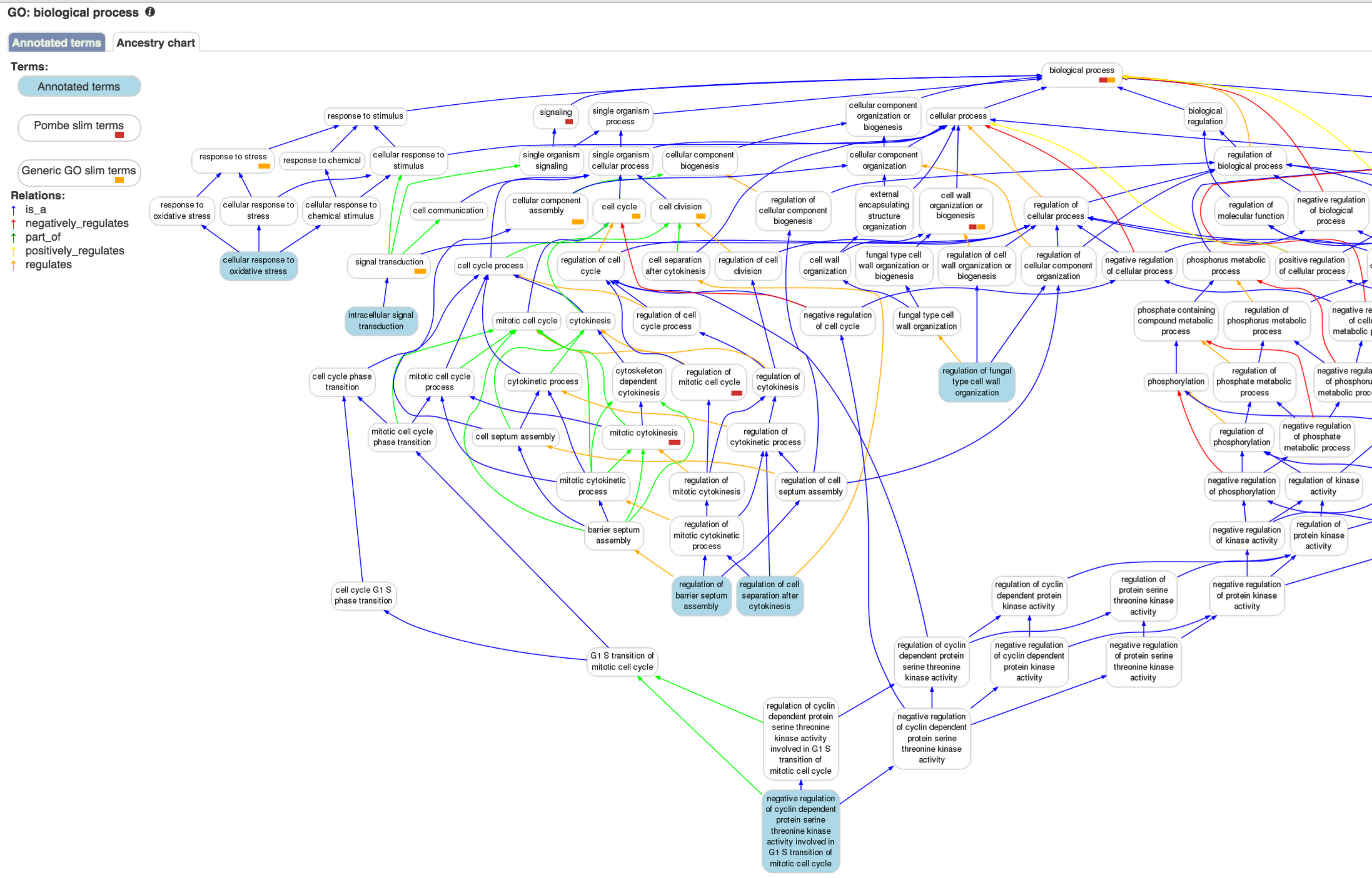
GO Biological Process

GO Slim Terms

cell wall organization or biogenesis

signaling

	Term Name	Count
+	cellular response to oxidative stress	51
+	intracellular signal transduction	269
	- **regulates** atf1	
+	negative regulation of cyclin-dependent protein serine/threonine kinase activity involved in G1/S transition of mitotic cell cycle	1
+	peptidyl-threonine phosphorylation	16
+	positive regulation of mRNA splicing, via spliceosome	2
+	protein autophosphorylation	17
+	protein phosphorylation	130
+	regulation of barrier septum assembly	35
+	regulation of cell separation after cytokinesis	
+	regulation of fungal-type cell wall organization	



GO: biological process

Annotated terms | Ancestry chart

Terms:

- Annotated terms
- Pombe slim terms 🟥
- Generic GO slim terms 🟧

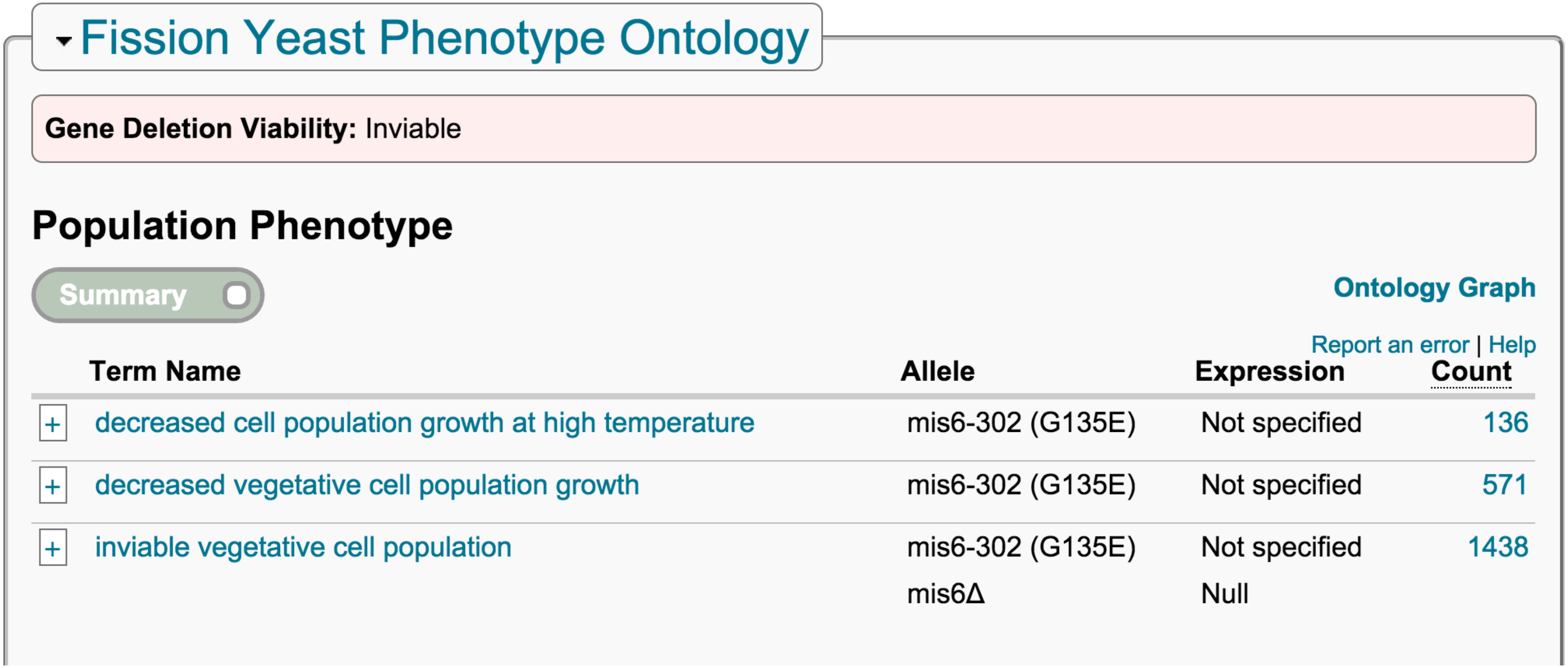
Relations:

- ↑ is_a
- ↑ negatively_regulates
- ↑ part_of
- ↑ positively_regulates
- ↑ regulates

mis6

FYPO (phenotype) annotation

Phenotypes are grouped into ‘population level’ and ‘cell level’
Links (summary, term, ontology) work similarly to the GO section



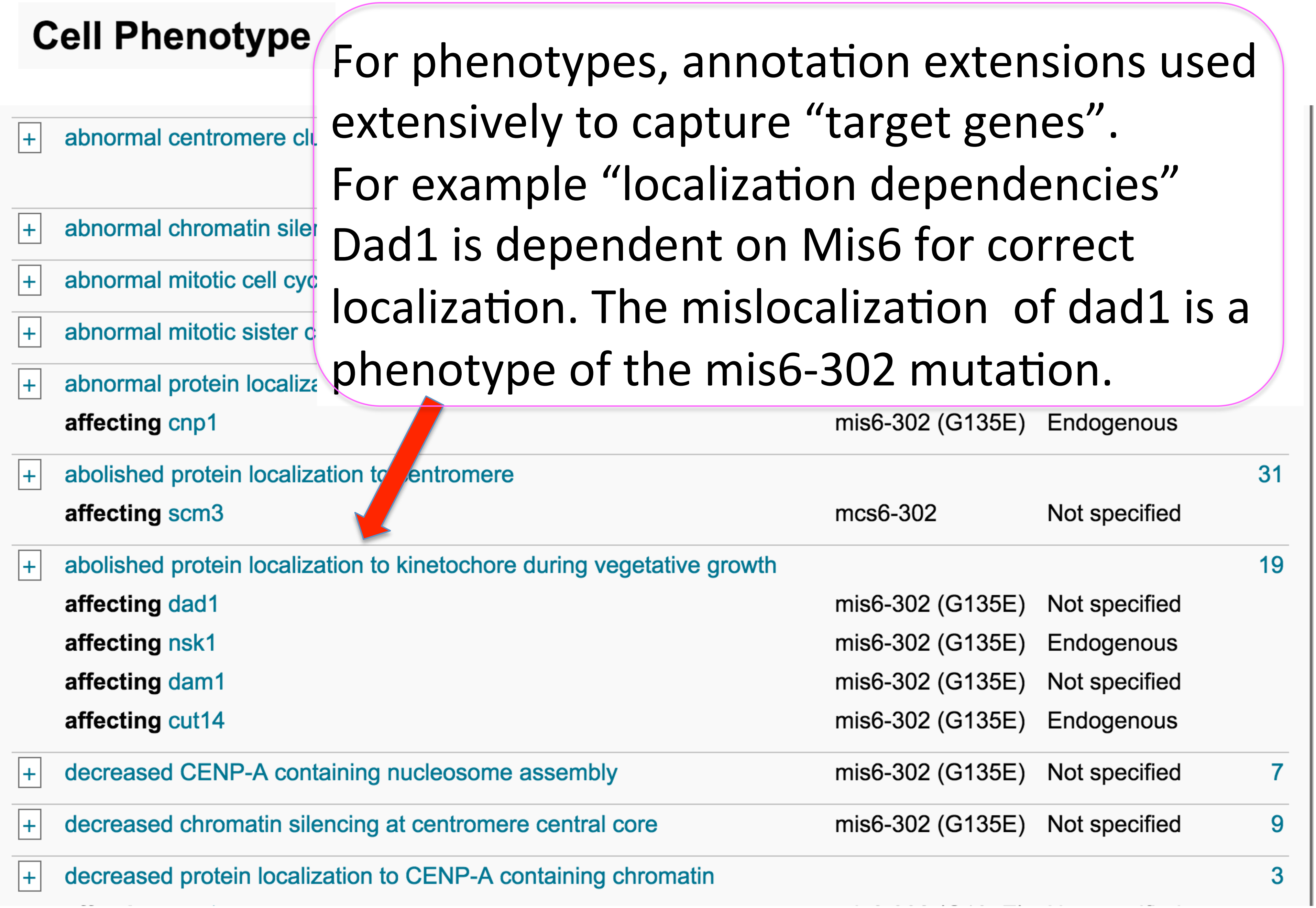
Gene Deletion Viability: Inviable

Population Phenotype

	Term Name	Allele	Expression	Count
+	decreased cell population growth at high temperature	mis6-302 (G135E)	Not specified	136
+	decreased vegetative cell population growth	mis6-302 (G135E)	Not specified	571
+	inviable vegetative cell population	mis6-302 (G135E) mis6Δ	Not specified Null	1438

mis6

For phenotypes, annotation extensions used extensively to capture "target genes".
For example "localization dependencies" Dad1 is dependent on Mis6 for correct localization. The mislocalization of dad1 is a phenotype of the mis6-302 mutation.



Cell Phenotype

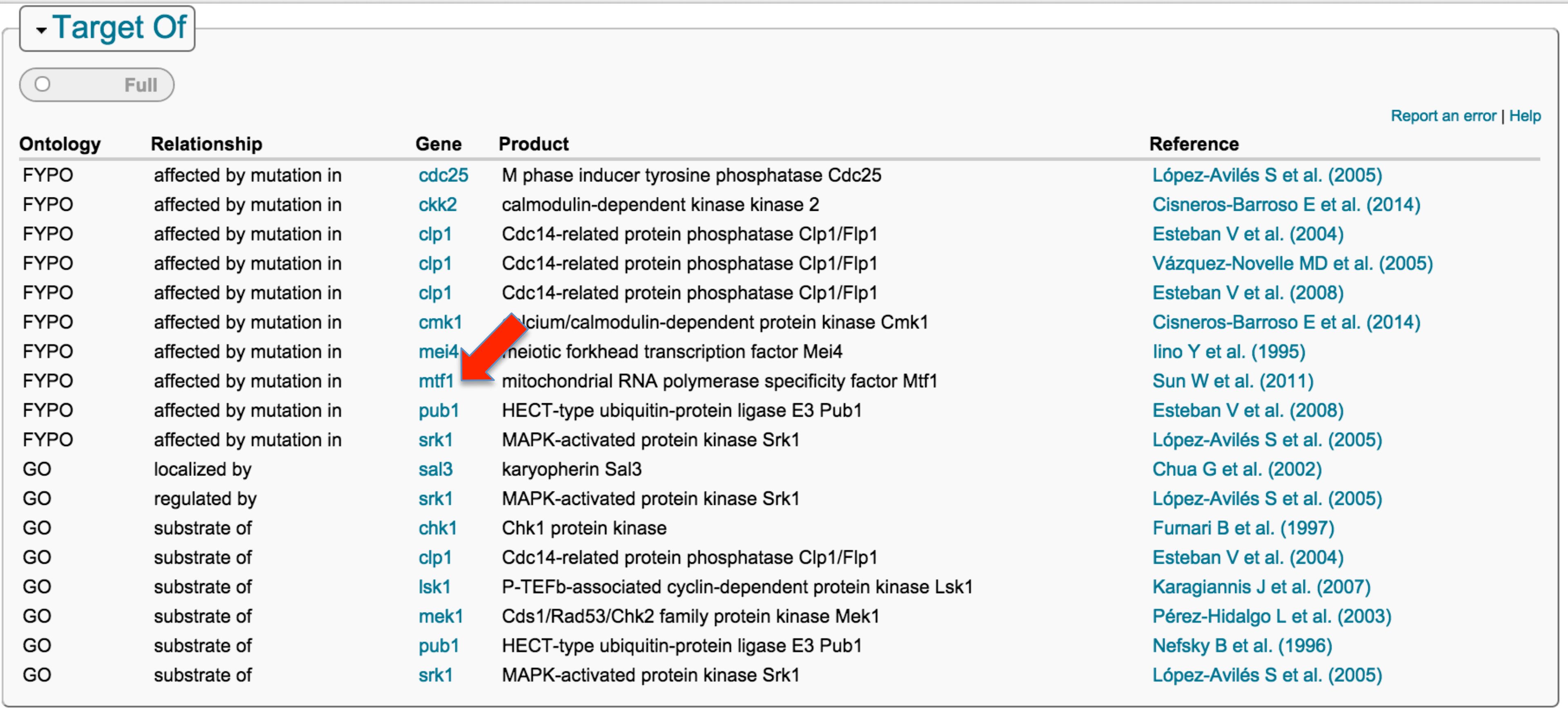
+ abnormal centromere cl			
+ abnormal chromatin sile			
+ abnormal mitotic cell cyc			
+ abnormal mitotic sister c			
+ abnormal protein localiz			
affecting cnp1	mis6-302 (G135E)	Endogenous	
+ abolished protein localization to centromere			31
affecting scm3	mcs6-302	Not specified	
+ abolished protein localization to kinetochore during vegetative growth			19
affecting dad1	mis6-302 (G135E)	Not specified	
affecting nsk1	mis6-302 (G135E)	Endogenous	
affecting dam1	mis6-302 (G135E)	Not specified	
affecting cut14	mis6-302 (G135E)	Endogenous	
+ decreased CENP-A containing nucleosome assembly	mis6-302 (G135E)	Not specified	7
+ decreased chromatin silencing at centromere central core	mis6-302 (G135E)	Not specified	9
+ decreased protein localization to CENP-A containing chromatin			3

GO vs phenotype

- GO Biological Process annotations to indicate that a gene product is ***directly*** involved in a process or its ***regulation***
- FYPO annotations indicate when a mutation in a gene causes a change in a process ***affects (direct or indirect)***
- ***affecting*** is often described in papers as ***"regulating"*** *when a process is not known to be truly regulatory in a Wild Type (*i.e. upstream process is ***broken****)*
- Before FYPO, GO often used for phenotypes
- We still use phenotypes to make GO process annotations, if the phenotype is specific enough to indicate a direct role
- There are many *grey* areas (resolution of these will be helped by having clearly defined beginnings and ends of processes)

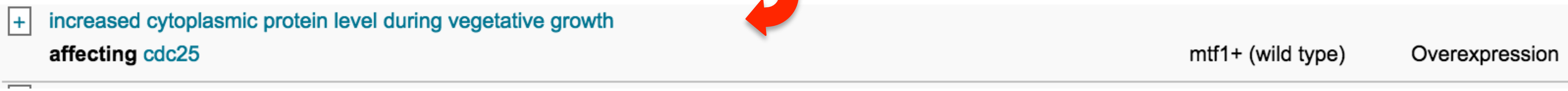
cdc25

“Target of” annotation



Ontology	Relationship	Gene	Product	Reference
FYPO	affected by mutation in	cdc25	M phase inducer tyrosine phosphatase Cdc25	López-Avilés S et al. (2005)
FYPO	affected by mutation in	ckk2	calmodulin-dependent kinase kinase 2	Cisneros-Barroso E et al. (2014)
FYPO	affected by mutation in	clp1	Cdc14-related protein phosphatase Clp1/Flp1	Esteban V et al. (2004)
FYPO	affected by mutation in	clp1	Cdc14-related protein phosphatase Clp1/Flp1	Vázquez-Novelle MD et al. (2005)
FYPO	affected by mutation in	clp1	Cdc14-related protein phosphatase Clp1/Flp1	Esteban V et al. (2008)
FYPO	affected by mutation in	cmk1	calcium/calmodulin-dependent protein kinase Cmk1	Cisneros-Barroso E et al. (2014)
FYPO	affected by mutation in	mei4	meiotic forkhead transcription factor Mei4	Iino Y et al. (1995)
FYPO	affected by mutation in	mtf1	mitochondrial RNA polymerase specificity factor Mtf1	Sun W et al. (2011)
FYPO	affected by mutation in	pub1	HECT-type ubiquitin-protein ligase E3 Pub1	Esteban V et al. (2008)
FYPO	affected by mutation in	srk1	MAPK-activated protein kinase Srk1	López-Avilés S et al. (2005)
GO	localized by	sal3	karyopherin Sal3	Chua G et al. (2002)
GO	regulated by	srk1	MAPK-activated protein kinase Srk1	López-Avilés S et al. (2005)
GO	substrate of	chk1	Chk1 protein kinase	Furnari B et al. (1997)
GO	substrate of	clp1	Cdc14-related protein phosphatase Clp1/Flp1	Esteban V et al. (2004)
GO	substrate of	lsk1	P-TEFb-associated cyclin-dependent protein kinase Lsk1	Karagiannis J et al. (2007)
GO	substrate of	mek1	Cds1/Rad53/Chk2 family protein kinase Mek1	Pérez-Hidalgo L et al. (2003)
GO	substrate of	pub1	HECT-type ubiquitin-protein ligase E3 Pub1	Nefsky B et al. (1996)
GO	substrate of	srk1	MAPK-activated protein kinase Srk1	López-Avilés S et al. (2005)

This section displays the *reciprocal* annotation of an ontology term + annotation extension
For example on the mtf1 gene page



+ increased cytoplasmic protein level during vegetative growth
affecting cdc25 | mtf1+ (wild type) | Overexpression

cdr2

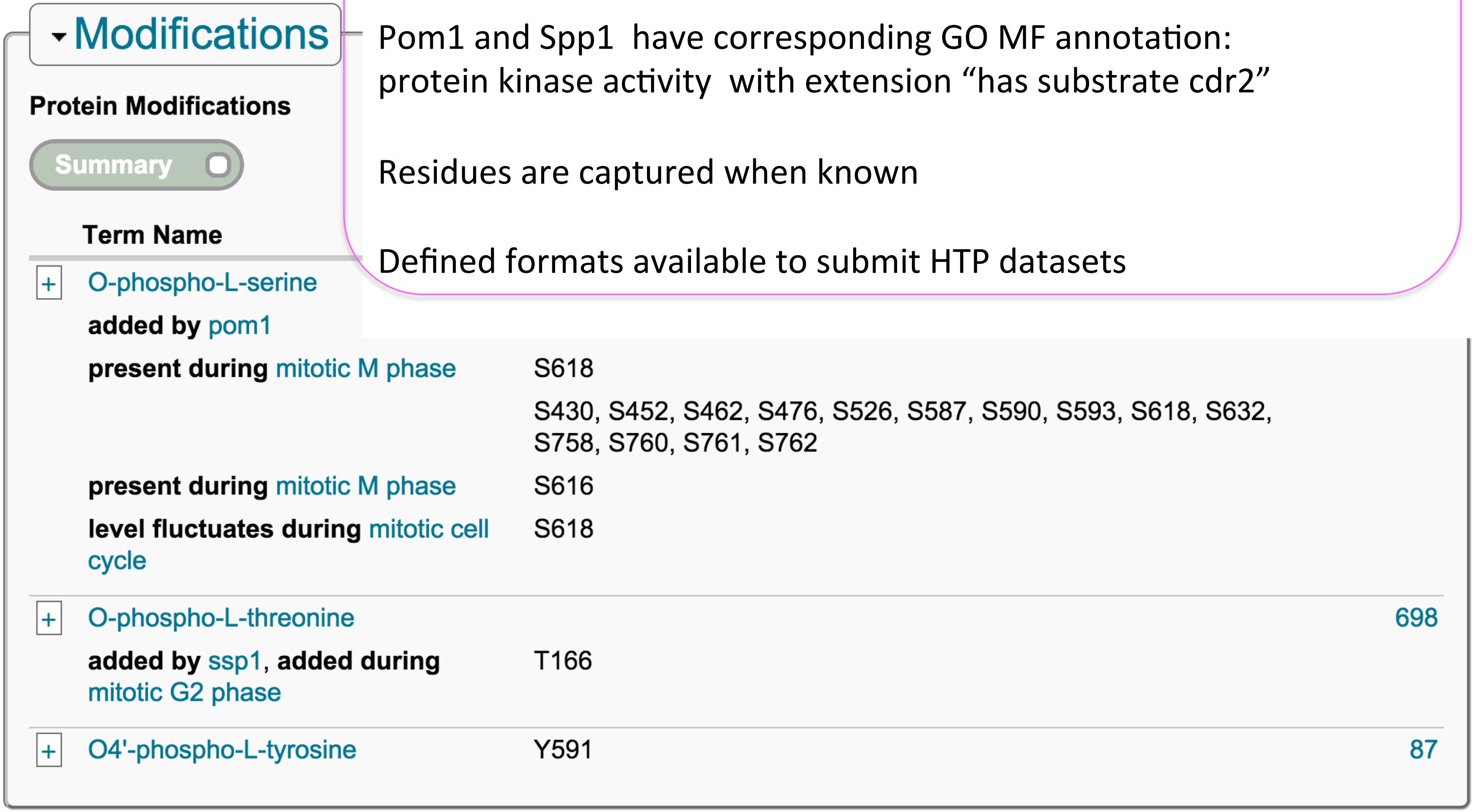
Protein Modifications

Modifications use “annotation extensions” to capture
- modifier
- cell cycle phase(s) when a modification is present/added/removed

Pom1 and Spp1 have corresponding GO MF annotation:
protein kinase activity with extension “has substrate cdr2”

Residues are captured when known

Defined formats available to submit HTP datasets



Protein Modifications

Term Name		
+ O-phospho-L-serine		
added by pom1		
present during mitotic M phase	S618	
	S430, S452, S462, S476, S526, S587, S590, S593, S618, S632, S758, S760, S761, S762	
present during mitotic M phase	S616	
level fluctuates during mitotic cell cycle	S618	
+ O-phospho-L-threonine		698
added by ssp1, **added during** mitotic G2 phase	T166	
+ O4'-phospho-L-tyrosine	Y591	87

Other gene page section

- Protein features: Families & domains (multiple sources) plus “protein sequence features” like NLS, mitochondrial targeting sequence, etc
- Genetic & physical interactions (using BioGRID annotation format)
- Gene expression (qualitative and quantitative)
- Literature
- External references …links added recently

STRING(network inference tool)

pombeA (polyadenylation signals)

SYSGRO (phenotypic data and analysis)

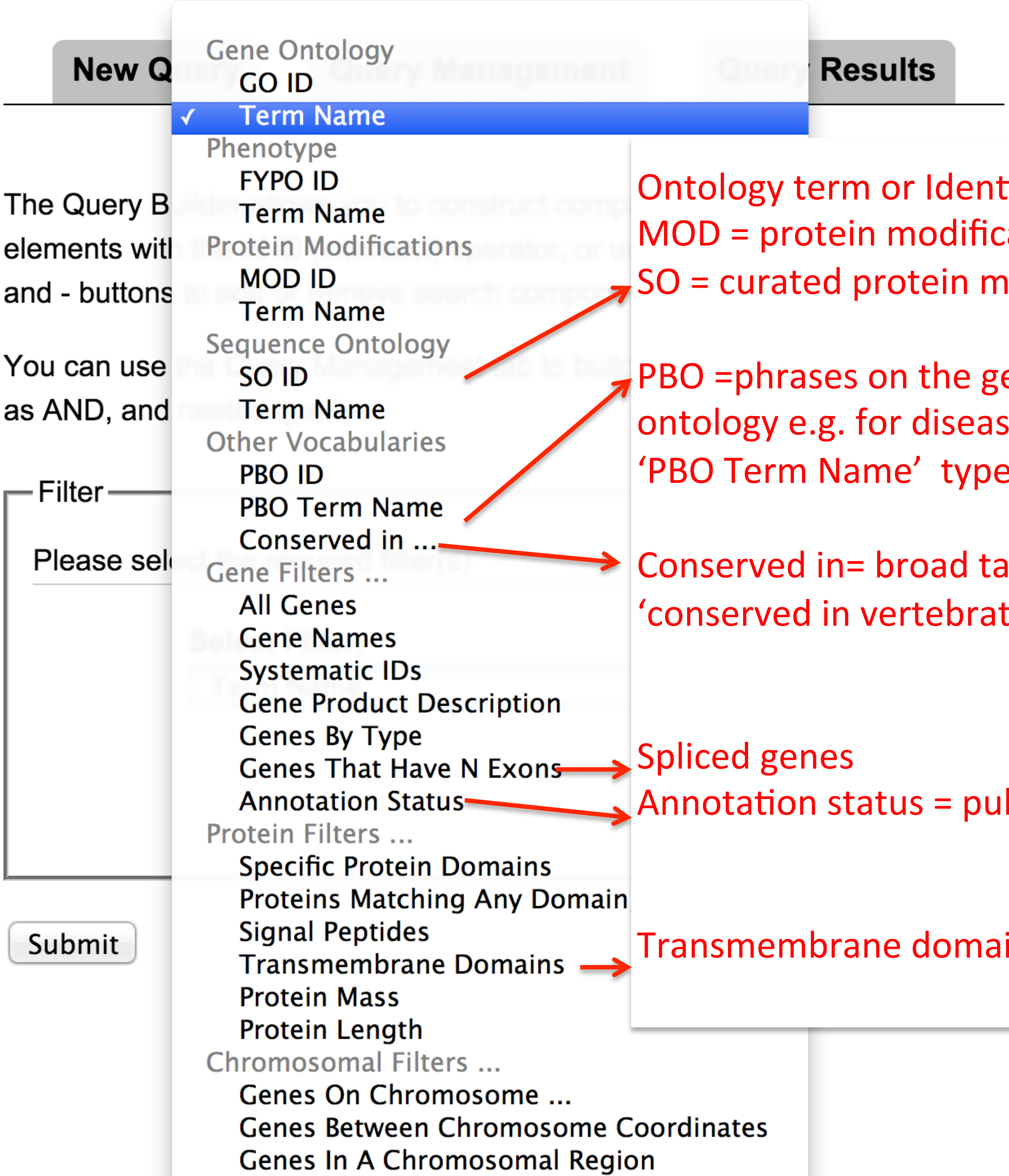
Tools

S. pombe spores and vegetative cells by scanning electron microscopy. Juergen Berger, Maria Langegger and Silke Hauf

PomBase data update

WEDNESDAY, 27TH MAY, 2015

We have updated the data av
include manual curation throu



Ontology term or Identifier

MOD = protein modifications

SO = curated protein motifs i.e "NLS"

PBO =phrases on the gene page which are not part of an ontology e.g. for disease associated genes

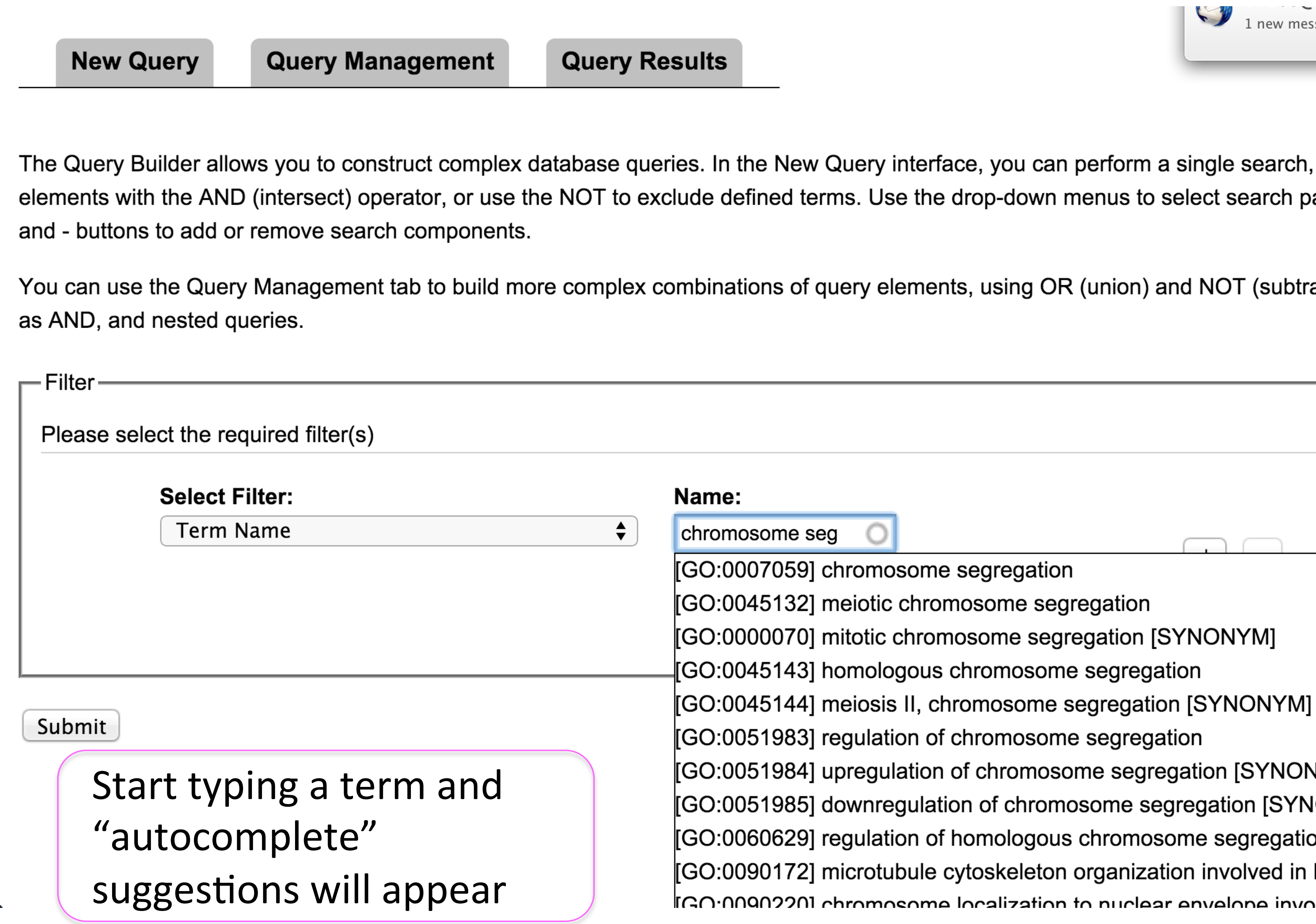
'PBO Term Name' type 'disease associated', or 'cancer'

Conserved in= broad taxonomic distribution, e.g. 'conserved in vertebrate' or 'absent from *S. cerevisiae*'

Spliced genes

Annotation status = published etc.

Transmembrane domain (and number)



The Query Builder allows you to construct complex database queries. In the New Query interface, you can perform a single search, elements with the AND (intersect) operator, or use the NOT to exclude defined terms. Use the drop-down menus to select search pa and - buttons to add or remove search components.

You can use the Query Management tab to build more complex combinations of query elements, using OR (union) and NOT (subtra as AND, and nested queries.

Filter

Please select the required filter(s)

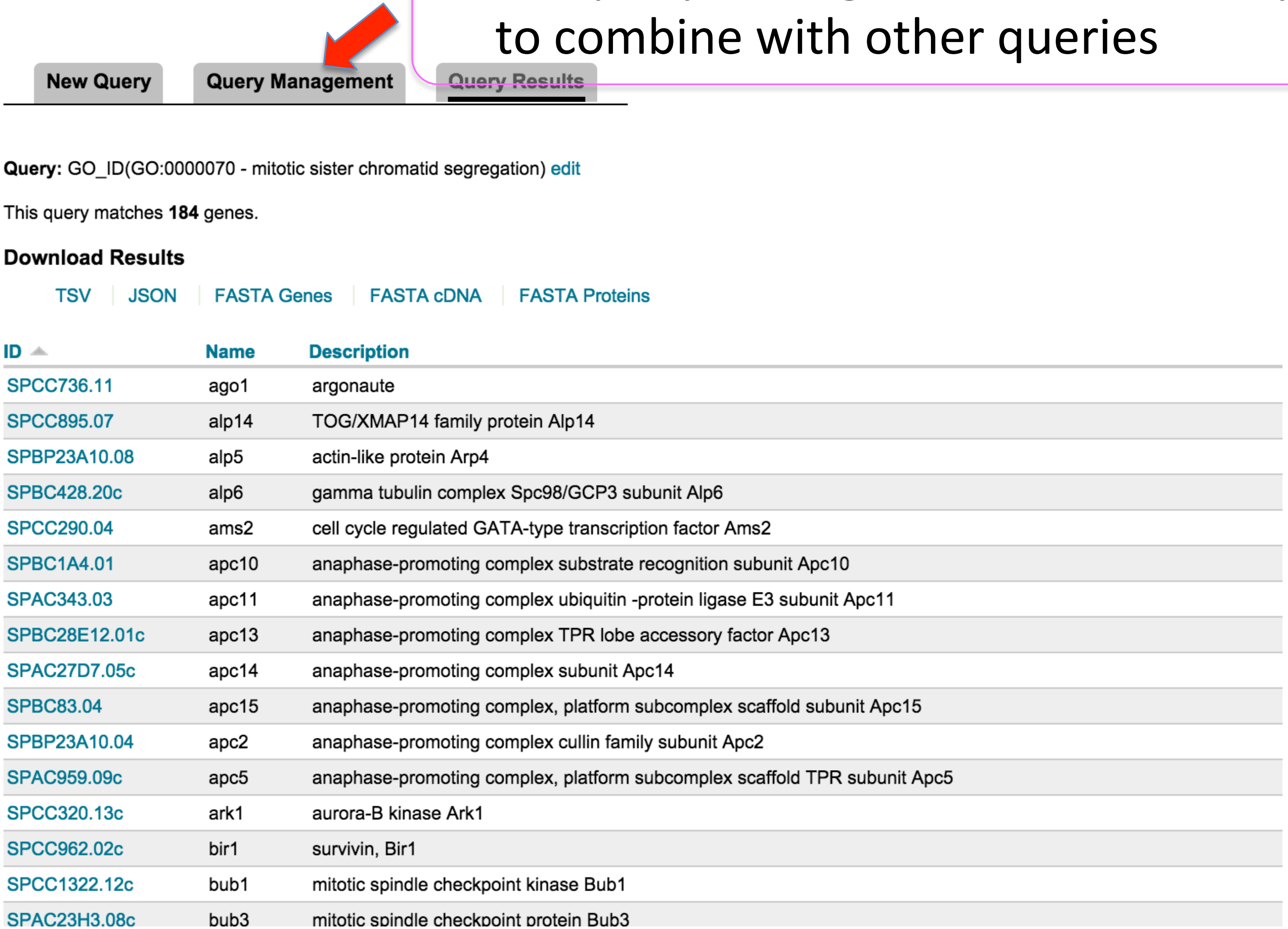
Select Filter: Term Name

Name: chromosome seg

[GO:0007059] chromosome segregation
[GO:0045132] meiotic chromosome segregation
[GO:0000070] mitotic chromosome segregation [SYNONYM]
[GO:0045143] homologous chromosome segregation
[GO:0045144] meiosis II, chromosome segregation [SYNONYM]
[GO:0051983] regulation of chromosome segregation
[GO:0051984] upregulation of chromosome segregation [SYNON
[GO:0051985] downregulation of chromosome segregation [SYN
[GO:0060629] regulation of homologous chromosome segregatio
[GO:0090172] microtubule cytoskeleton organization involved in
[GO:0090220] chromosome localization to nuclear envelope invo

Start typing a term and “autocomplete” suggestions will appear

The query management tabs allows you to combine with other queries



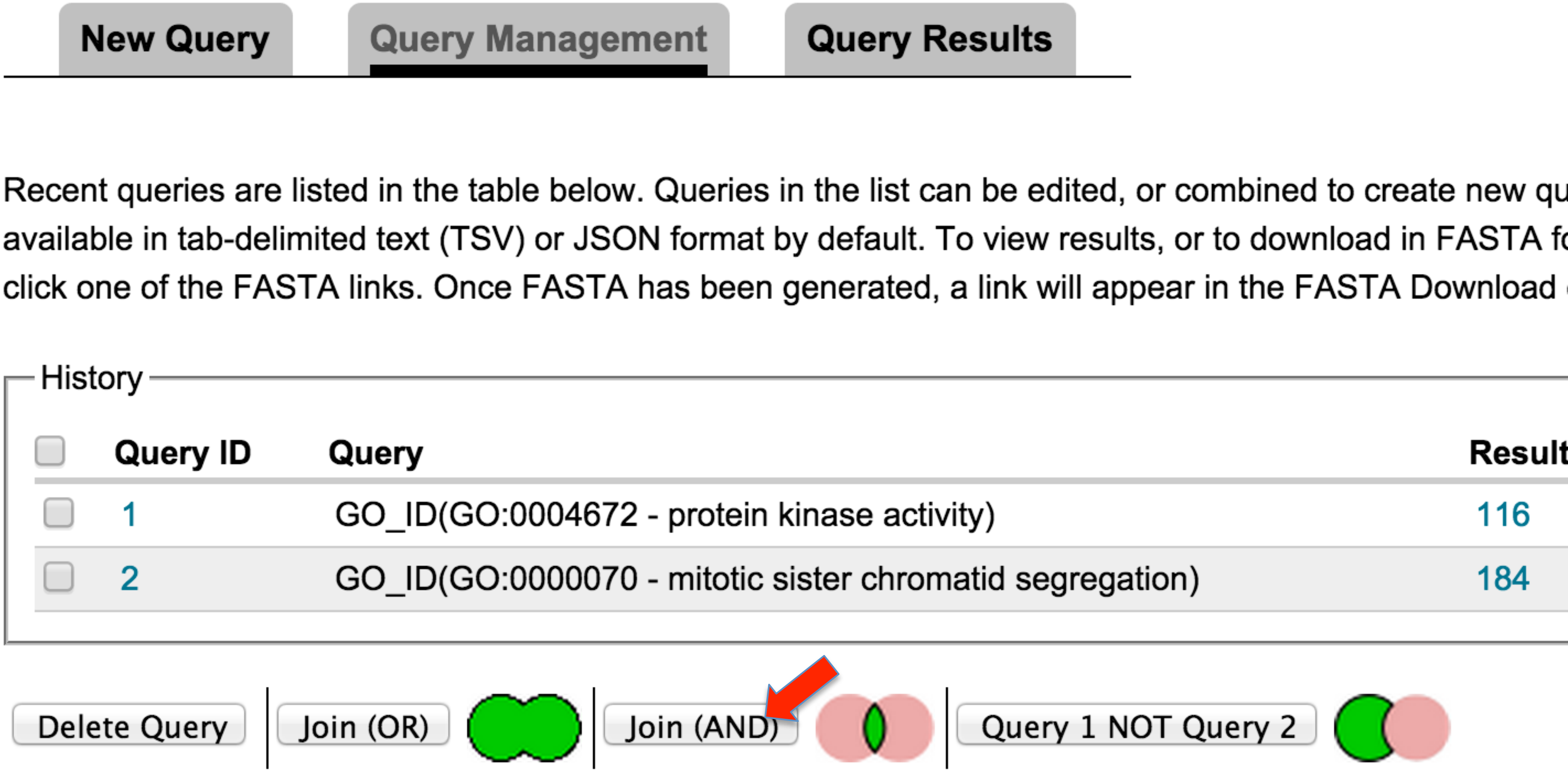
Query: GO_ID(GO:0000070 - mitotic sister chromatid segregation) edit

This query matches **184** genes.

Download Results

TSV | JSON | FASTA Genes | FASTA cDNA | FASTA Proteins

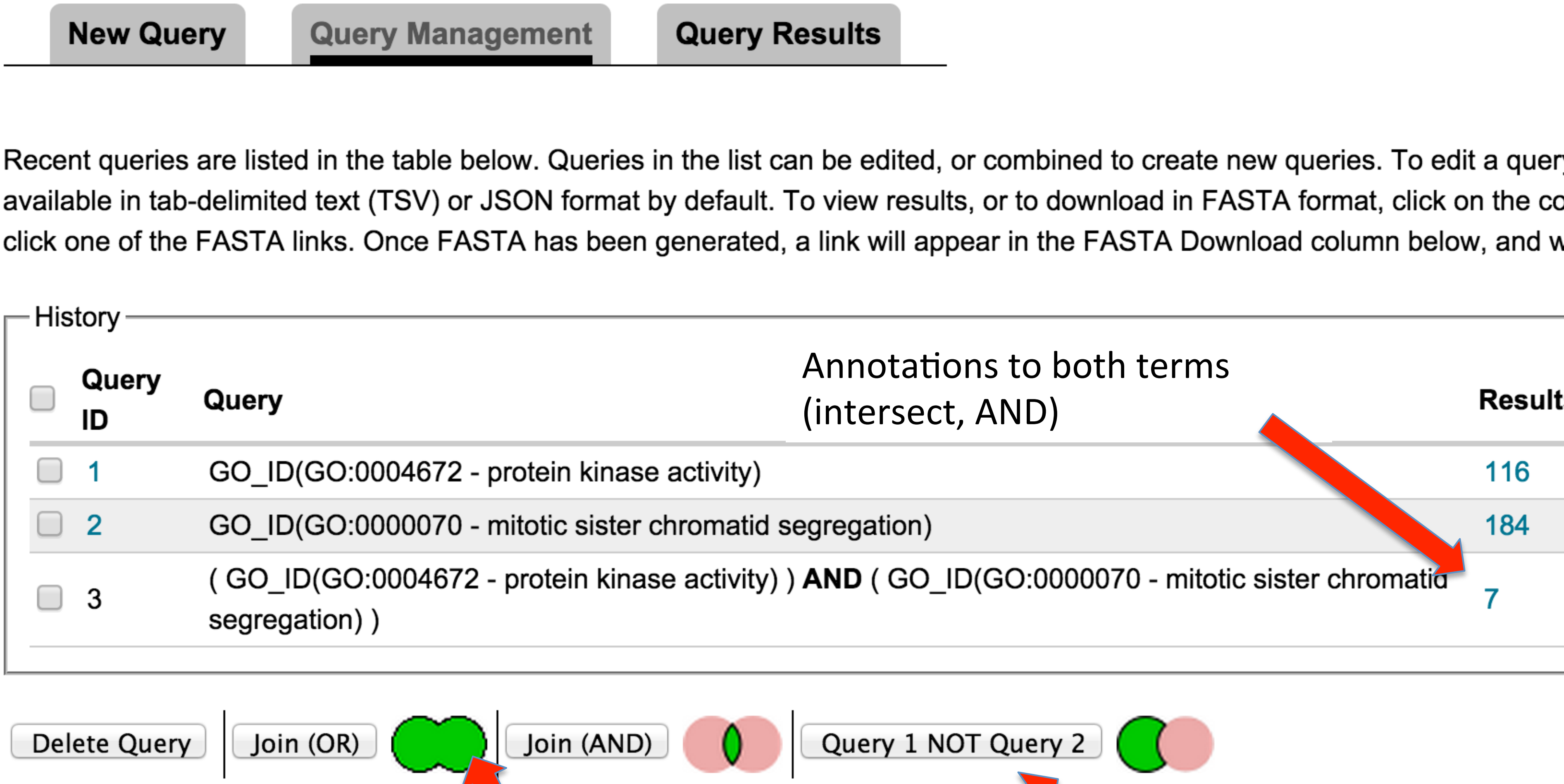
ID	Name	Description
SPCC736.11	ago1	argonaute
SPCC895.07	alp14	TOG/XMAP14 family protein Alp14
SPBP23A10.08	alp5	actin-like protein Arp4
SPBC428.20c	alp6	gamma tubulin complex Spc98/GCP3 subunit Alp6
SPCC290.04	ams2	cell cycle regulated GATA-type transcription factor Ams2
SPBC1A4.01	apc10	anaphase-promoting complex substrate recognition subunit Apc10
SPAC343.03	apc11	anaphase-promoting complex ubiquitin -protein ligase E3 subunit Apc11
SPBC28E12.01c	apc13	anaphase-promoting complex TPR lobe accessory factor Apc13
SPAC27D7.05c	apc14	anaphase-promoting complex subunit Apc14
SPBC83.04	apc15	anaphase-promoting complex, platform subcomplex scaffold subunit Apc15
SPBP23A10.04	apc2	anaphase-promoting complex cullin family subunit Apc2
SPAC959.09c	apc5	anaphase-promoting complex, platform subcomplex scaffold TPR subunit Apc5
SPCC320.13c	ark1	aurora-B kinase Ark1
SPCC962.02c	bir1	survivin, Bir1
SPCC1322.12c	bub1	mitotic spindle checkpoint kinase Bub1
SPAC23H3.08c	bub3	mitotic spindle checkpoint protein Bub3



Recent queries are listed in the table below. Queries in the list can be edited, or combined to create new qu
available in tab-delimited text (TSV) or JSON format by default. To view results, or to download in FASTA fo
click one of the FASTA links. Once FASTA has been generated, a link will appear in the FASTA Download

History

☐	Query ID	Query	Result
☐	1	GO_ID(GO:0004672 - protein kinase activity)	116
☐	2	GO_ID(GO:0000070 - mitotic sister chromatid segregation)	184



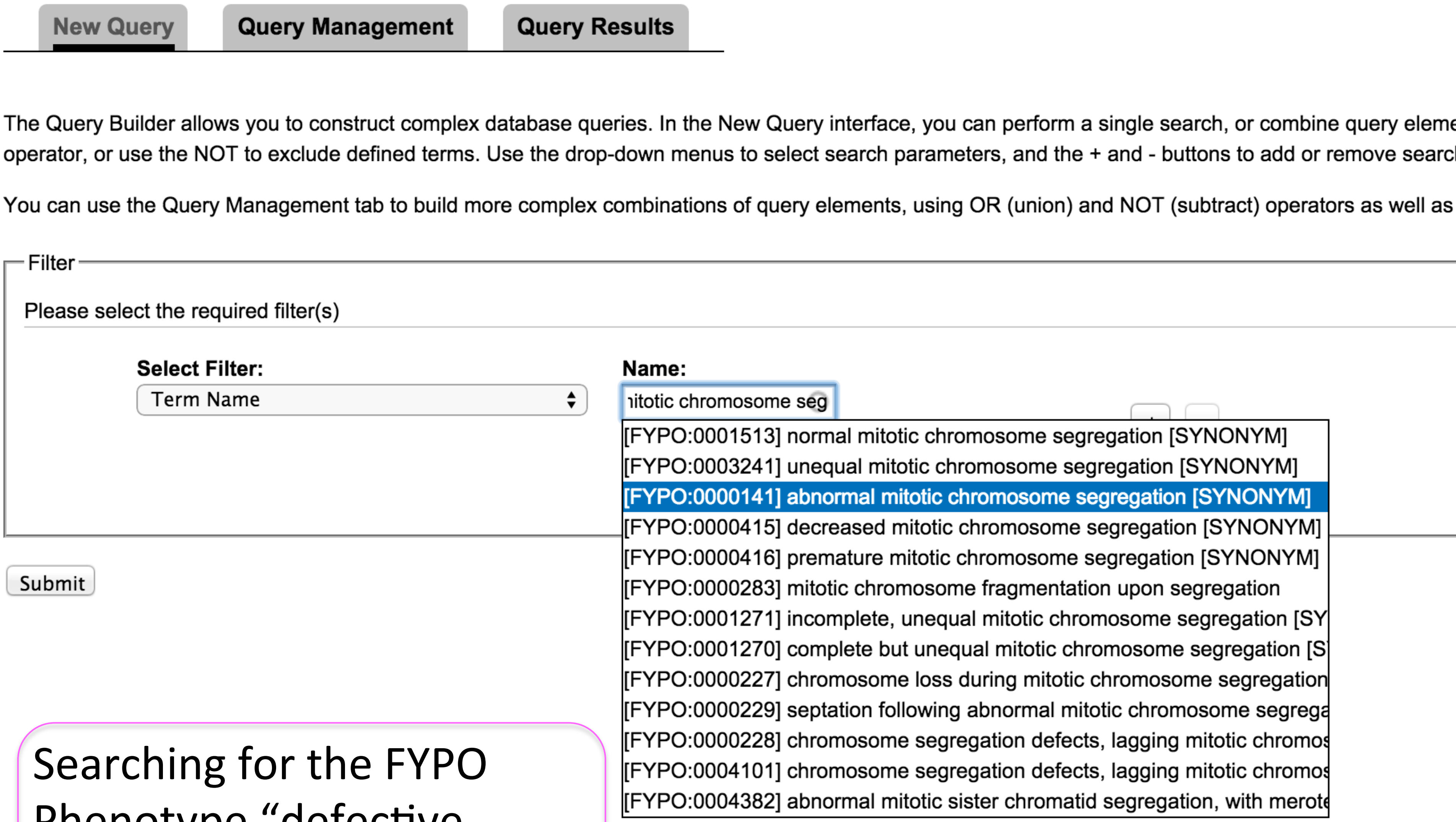
Recent queries are listed in the table below. Queries in the list can be edited, or combined to create new queries. To edit a quer
available in tab-delimited text (TSV) or JSON format by default. To view results, or to download in FASTA format, click on the co
click one of the FASTA links. Once FASTA has been generated, a link will appear in the FASTA Download column below, and w

History

☐	Query ID	Query	Result
☐	1	GO_ID(GO:0004672 - protein kinase activity)	116
☐	2	GO_ID(GO:0000070 - mitotic sister chromatid segregation)	184
☐	3	(GO_ID(GO:0004672 - protein kinase activity)) **AND** (GO_ID(GO:0000070 - mitotic sister chromatid segregation))	7

Annotations to both terms
(intersect, AND)

Can also (union, OR) and subtract lists



The Query Builder allows you to construct complex database queries. In the New Query interface, you can perform a single search, or combine query eleme
operator, or use the NOT to exclude defined terms. Use the drop-down menus to select search parameters, and the + and - buttons to add or remove searc

You can use the Query Management tab to build more complex combinations of query elements, using OR (union) and NOT (subtract) operators as well as

Filter

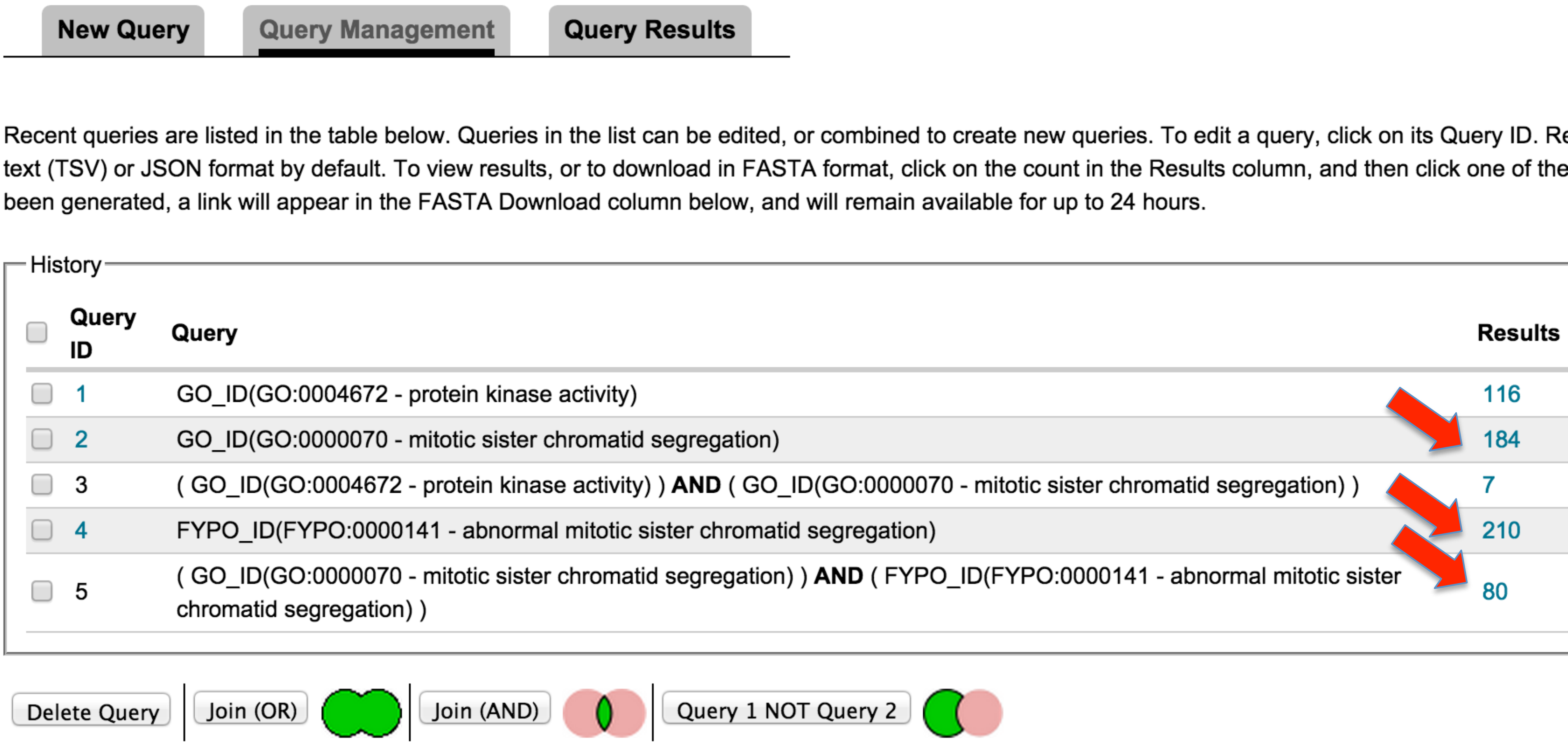
Please select the required filter(s)

Select Filter: Term Name

Name: ıitotic chromosome seg

- [FYPO:0001513] normal mitotic chromosome segregation [SYNONYM]
- [FYPO:0003241] unequal mitotic chromosome segregation [SYNONYM]
- [FYPO:0000141] abnormal mitotic chromosome segregation [SYNONYM]
- [FYPO:0000415] decreased mitotic chromosome segregation [SYNONYM]
- [FYPO:0000416] premature mitotic chromosome segregation [SYNONYM]
- [FYPO:0000283] mitotic chromosome fragmentation upon segregation
- [FYPO:0001271] incomplete, unequal mitotic chromosome segregation [SY
- [FYPO:0001270] complete but unequal mitotic chromosome segregation [S
- [FYPO:0000227] chromosome loss during mitotic chromosome segregation
- [FYPO:0000229] septation following abnormal mitotic chromosome segrega
- [FYPO:0000228] chromosome segregation defects, lagging mitotic chromos
- [FYPO:0004101] chromosome segregation defects, lagging mitotic chromos
- [FYPO:0004382] abnormal mitotic sister chromatid segregation, with merote

Searching for the FYPO Phenotype “defective chromosome segregation”



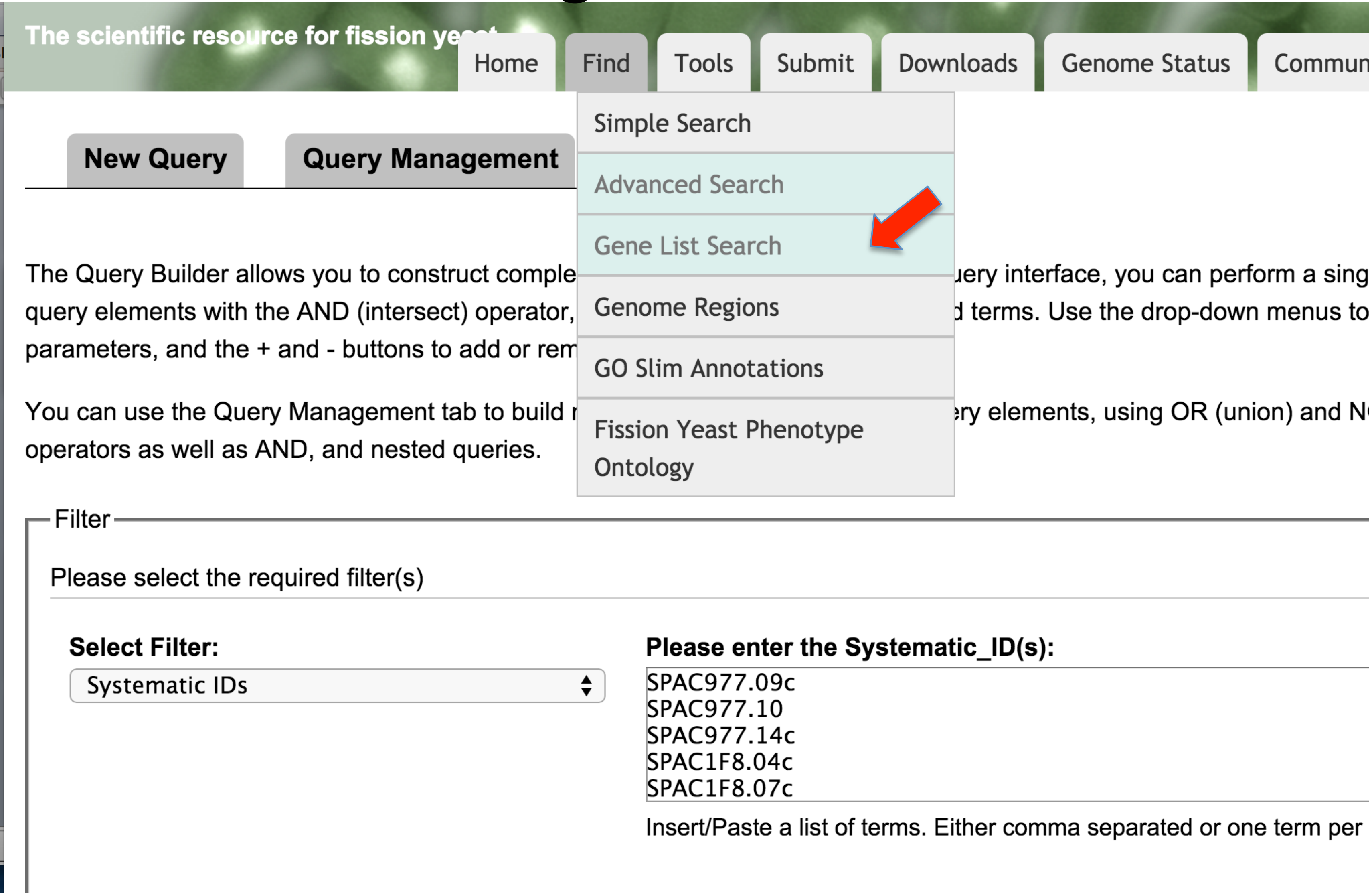
Recent queries are listed in the table below. Queries in the list can be edited, or combined to create new queries. To edit a query, click on its Query ID. Re
text (TSV) or JSON format by default. To view results, or to download in FASTA format, click on the count in the Results column, and then click one of the
been generated, a link will appear in the FASTA Download column below, and will remain available for up to 24 hours.

History

☐	Query ID	Query	Results
☐	1	GO_ID(GO:0004672 - protein kinase activity)	116
☐	2	GO_ID(GO:0000070 - mitotic sister chromatid segregation)	184
☐	3	(GO_ID(GO:0004672 - protein kinase activity)) **AND** (GO_ID(GO:0000070 - mitotic sister chromatid segregation))	7
☐	4	FYPO_ID(FYPO:0000141 - abnormal mitotic sister chromatid segregation)	210
☐	5	(GO_ID(GO:0000070 - mitotic sister chromatid segregation)) **AND** (FYPO_ID(FYPO:0000141 - abnormal mitotic sister chromatid segregation))	80

GO mitotic segregation AND FYPO abnormal chromosome segregation only 80
Many phenotypes indirect (DNA replication)
Many GO have GO annotation but no phenotypic assessment for chromosome segregation defects (e.g. many APC subunits)

Adding a user defined list



The Query Builder allows you to construct comple… …uery interface, you can perform a sing… query elements with the AND (intersect) operator, … terms. Use the drop-down menus to… parameters, and the + and - buttons to add or rem…

You can use the Query Management tab to build … …ry elements, using OR (union) and N… operators as well as AND, and nested queries.

Filter

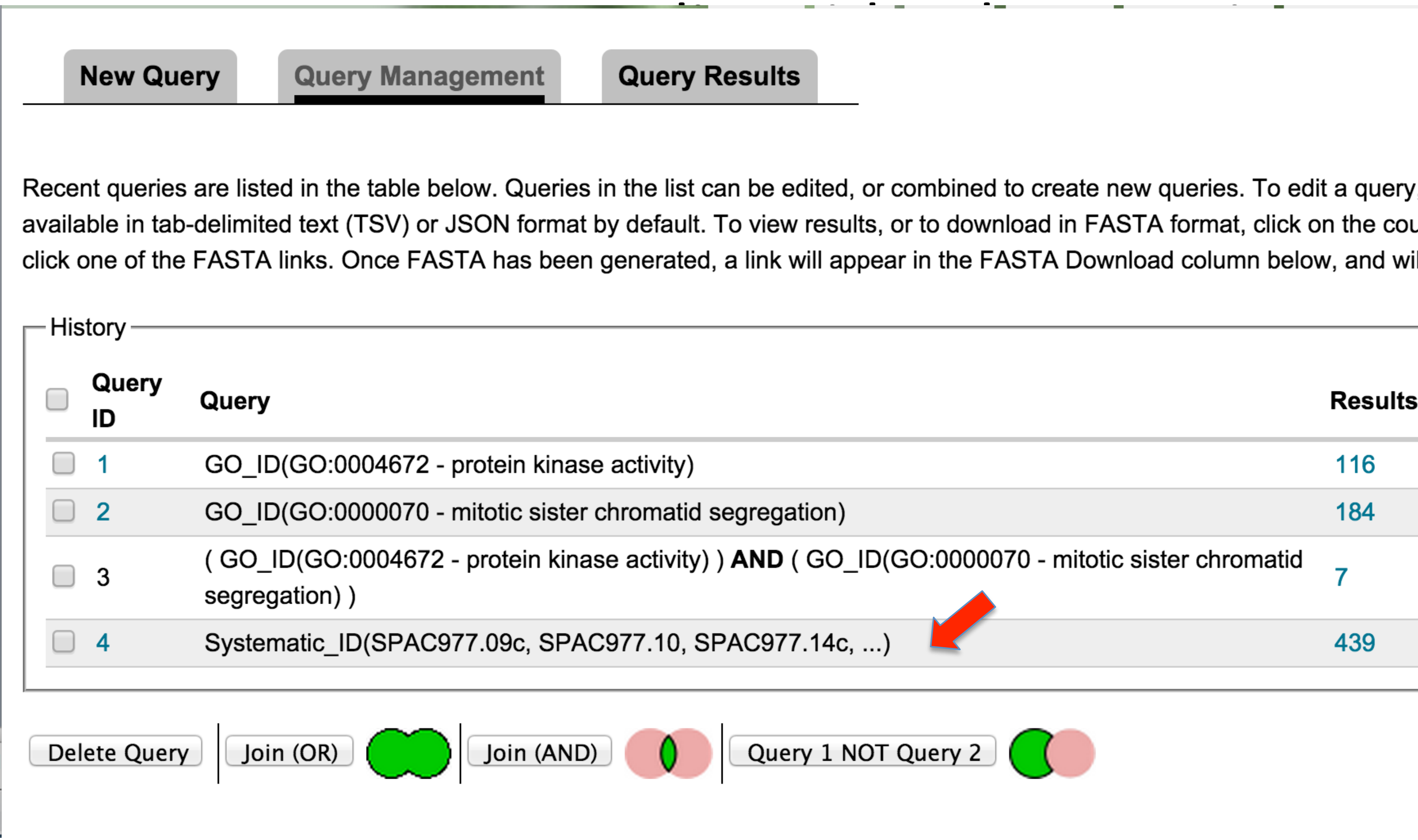
Please select the required filter(s)

Select Filter: Systematic IDs

Please enter the Systematic_ID(s):

SPAC977.09c
SPAC977.10
SPAC977.14c
SPAC1F8.04c
SPAC1F8.07c

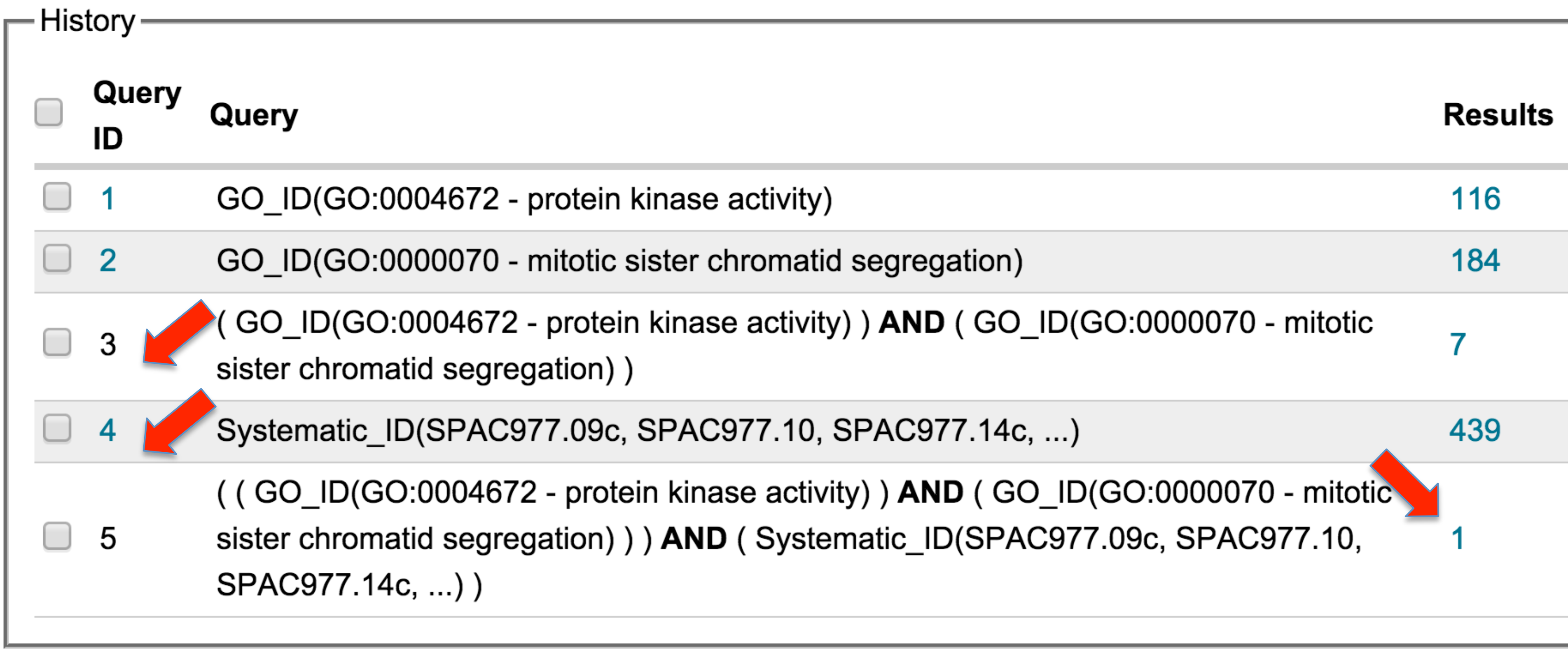
Insert/Paste a list of terms. Either comma separated or one term per



Recent queries are listed in the table below. Queries in the list can be edited, or combined to create new queries. To edit a query available in tab-delimited text (TSV) or JSON format by default. To view results, or to download in FASTA format, click on the cou click one of the FASTA links. Once FASTA has been generated, a link will appear in the FASTA Download column below, and wi

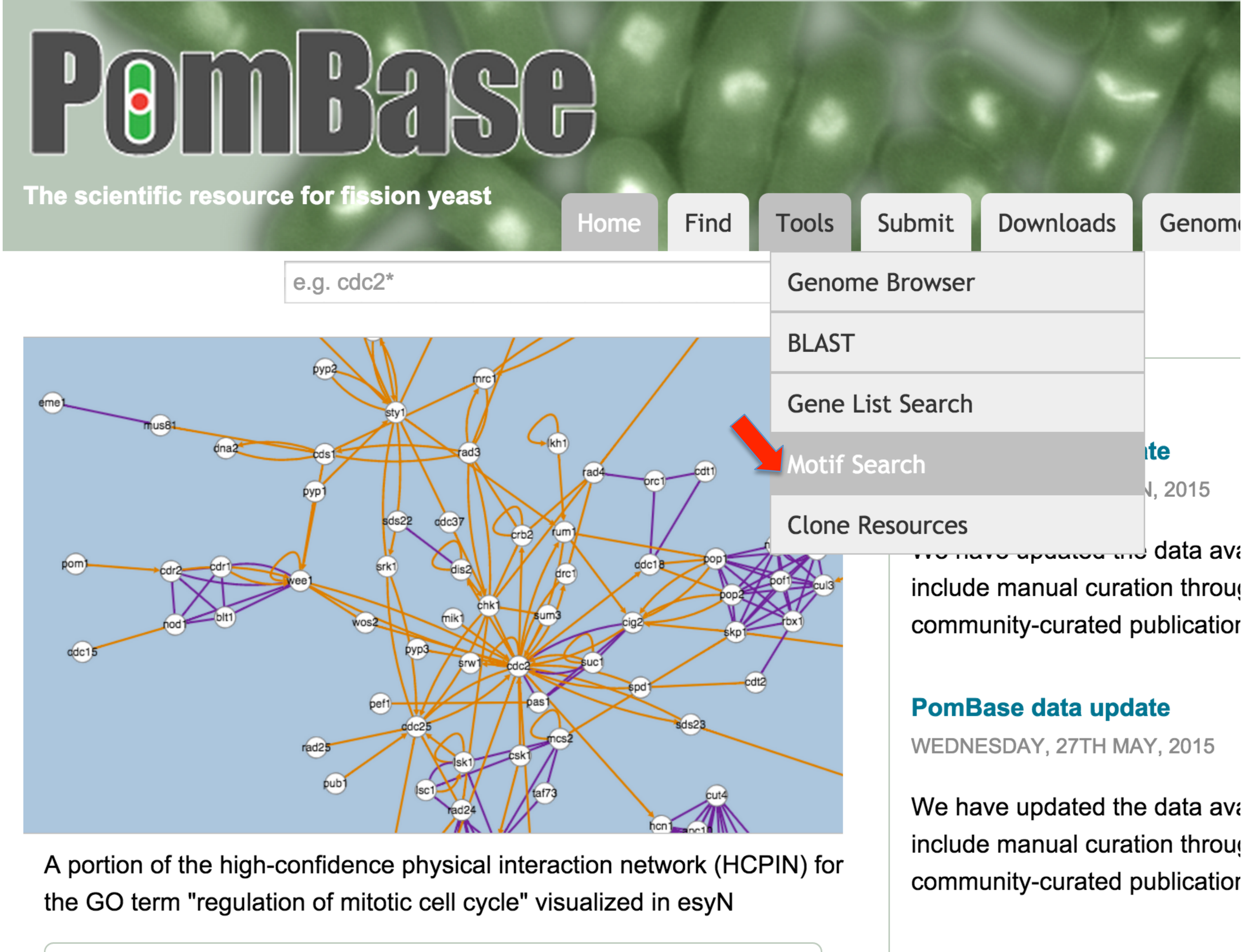
History

☐	Query ID	Query	Results
☐	1	GO_ID(GO:0004672 - protein kinase activity)	116
☐	2	GO_ID(GO:0000070 - mitotic sister chromatid segregation)	184
☐	3	(GO_ID(GO:0004672 - protein kinase activity)) **AND** (GO_ID(GO:0000070 - mitotic sister chromatid segregation))	7
☐	4	Systematic_ID(SPAC977.09c, SPAC977.10, SPAC977.14c, ...)	439



History

☐	Query ID	Query	Results
☐	1	GO_ID(GO:0004672 - protein kinase activity)	116
☐	2	GO_ID(GO:0000070 - mitotic sister chromatid segregation)	184
☐	3	(GO_ID(GO:0004672 - protein kinase activity)) **AND** (GO_ID(GO:0000070 - mitotic sister chromatid segregation))	7
☐	4	Systematic_ID(SPAC977.09c, SPAC977.10, SPAC977.14c, ...)	439
☐	5	((GO_ID(GO:0004672 - protein kinase activity)) **AND** (GO_ID(GO:0000070 - mitotic sister chromatid segregation))) **AND** (Systematic_ID(SPAC977.09c, SPAC977.10, SPAC977.14c, ...))	1



A portion of the high-confidence physical interaction network (HCPIN) for the GO term "regulation of mitotic cell cycle" visualized in esyN

ate

N, 2015

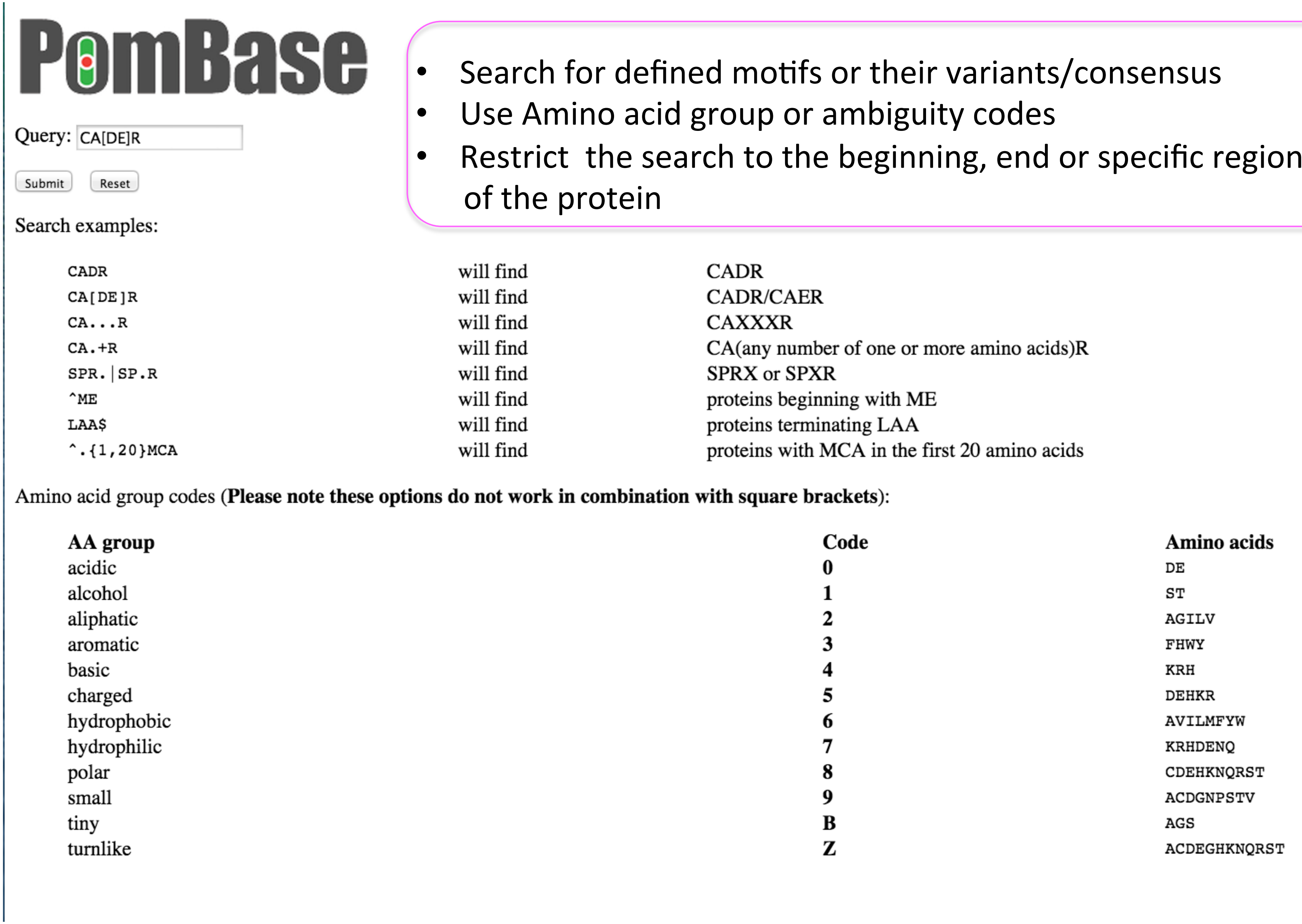
We have updated the data ava
include manual curation throu
community-curated publicatio

PomBase data update

WEDNESDAY, 27TH MAY, 2015

We have updated the data ava
include manual curation throu
community-curated publicatio

- Search for defined motifs or their variants/consensus
- Use Amino acid group or ambiguity codes
- Restrict the search to the beginning, end or specific regions of the protein



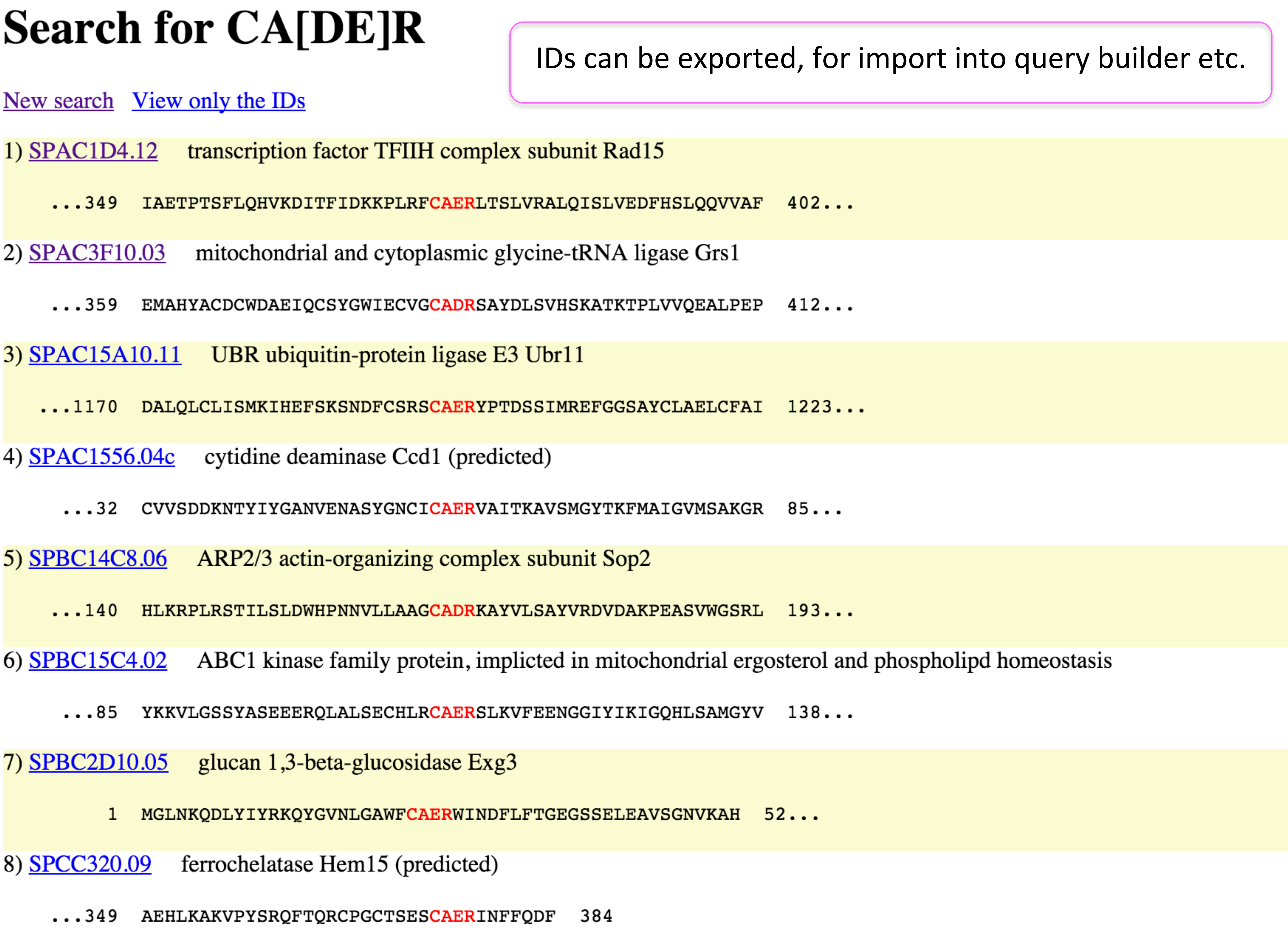
Query: CA[DE]R

Search examples:

`CADR`	will find	CADR
`CA[DE]R`	will find	CADR/CAER
`CA...R`	will find	CAXXXR
`CA.+R`	will find	CA(any number of one or more amino acids)R
`SPR.|SP.R`	will find	SPRX or SPXR
`^ME`	will find	proteins beginning with ME
`LAA$`	will find	proteins terminating LAA
`^.{1,20}MCA`	will find	proteins with MCA in the first 20 amino acids

Amino acid group codes (**Please note these options do not work in combination with square brackets**):

AA group	**Code**	**Amino acids**
acidic	**0**	DE
alcohol	**1**	ST
aliphatic	**2**	AGILV
aromatic	**3**	FHWY
basic	**4**	KRH
charged	**5**	DEHKR
hydrophobic	**6**	AVILMFYW
hydrophilic	**7**	KRHDENQ
polar	**8**	CDEHKNQRST
small	**9**	ACDGNPSTV
tiny	**B**	AGS
turnlike	**Z**	ACDEGHKNQRST



Search for CA[DE]R

IDs can be exported, for import into query builder etc.

New search   View only the IDs

1) SPAC1D4.12   transcription factor TFIIH complex subunit Rad15

```
...349   IAETPTSFLQHVKDITFIDKKPLRFCAERLTSLVRALQISLVEDFHSLQQVVAF   402...
```

2) SPAC3F10.03   mitochondrial and cytoplasmic glycine-tRNA ligase Grs1

```
...359   EMAHYACDCWDAEIQCSYGWIECVGCADRSAYDLSVHSKATKTPLVVQEALPEP   412...
```

3) SPAC15A10.11   UBR ubiquitin-protein ligase E3 Ubr11

```
..1170   DALQLCLISMKIHEFSKSNDFCSRSCAERYPTDSSIMREFGGSAYCLAELCFAI   1223...
```

4) SPAC1556.04c   cytidine deaminase Ccd1 (predicted)

```
...32   CVVSDDKNTYIYGANVENASYGNCICAERVAITKAVSMGYTKFMAIGVMSAKGR   85...
```

5) SPBC14C8.06   ARP2/3 actin-organizing complex subunit Sop2

```
...140   HLKRPLRSTILSLDWHPNNVLLAAGCADRKAYVLSAYVRDVDAKPEASVWGSRL   193...
```

6) SPBC15C4.02   ABC1 kinase family protein, implicted in mitochondrial ergosterol and phospholipd homeostasis

```
...85   YKKVLGSSYASEEERQLALSECHLRCAERSLKVFEENGGIYIKIGQHLSAMGYV   138...
```

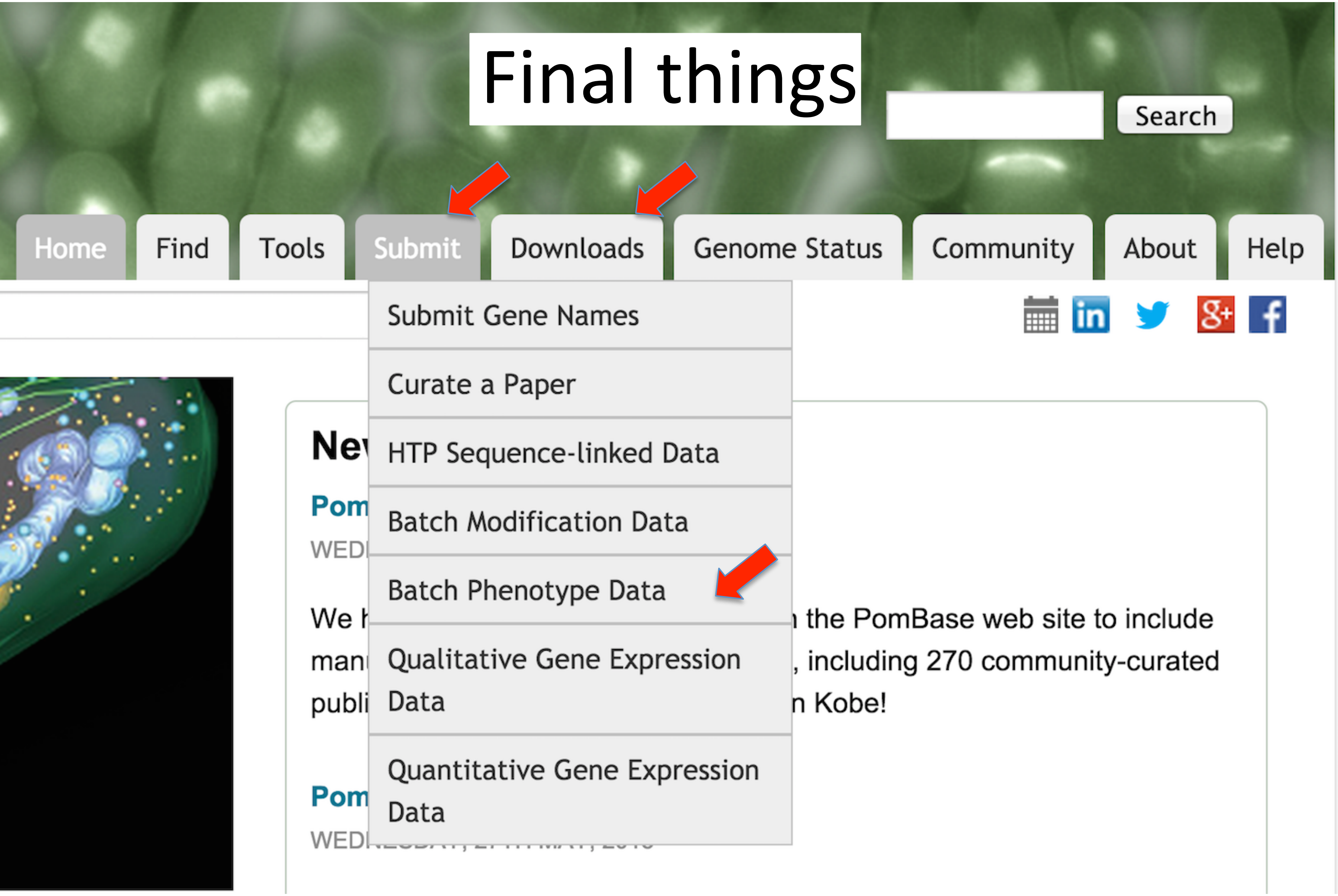
7) SPBC2D10.05   glucan 1,3-beta-glucosidase Exg3

```
1   MGLNKQDLYIYRKQYGVNLGAWFCAERWINDFLFTGEGSSELEAVSGNVKAH   52...
```

8) SPCC320.09   ferrochelatase Hem15 (predicted)

```
...349   AEHLKAKVPYSRQFTQRCPGCTSESCAERINFFQDF   384
```

Final things



Ne

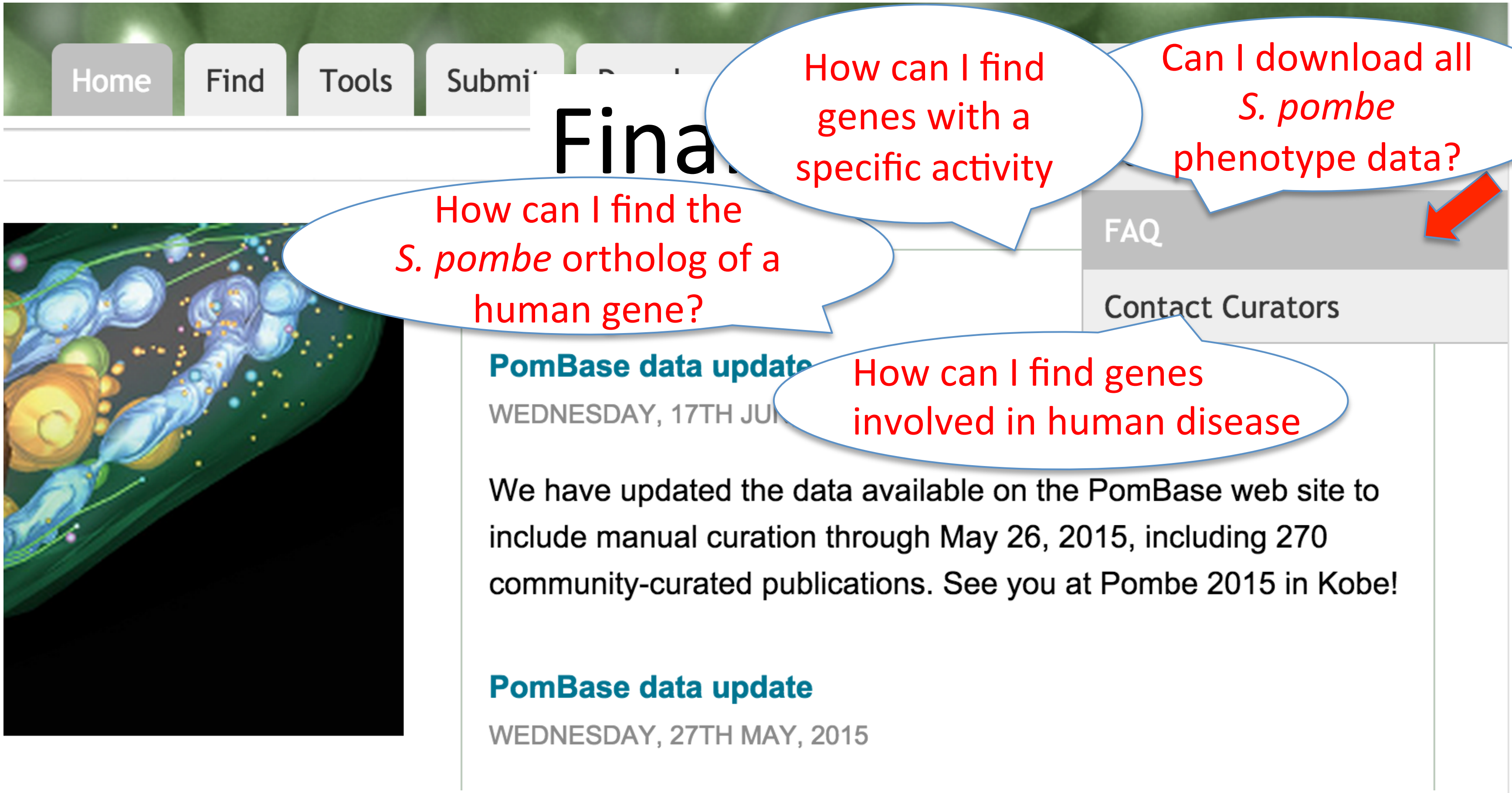
Pom

WED

We h … n the PomBase web site to include
man … , including 270 community-curated
publi … n Kobe!

Pom

WEDNESDAY, 27TH MAY, 2015



PomBase data update

WEDNESDAY, 17TH JU

We have updated the data available on the PomBase web site to include manual curation through May 26, 2015, including 270 community-curated publications. See you at Pombe 2015 in Kobe!

PomBase data update

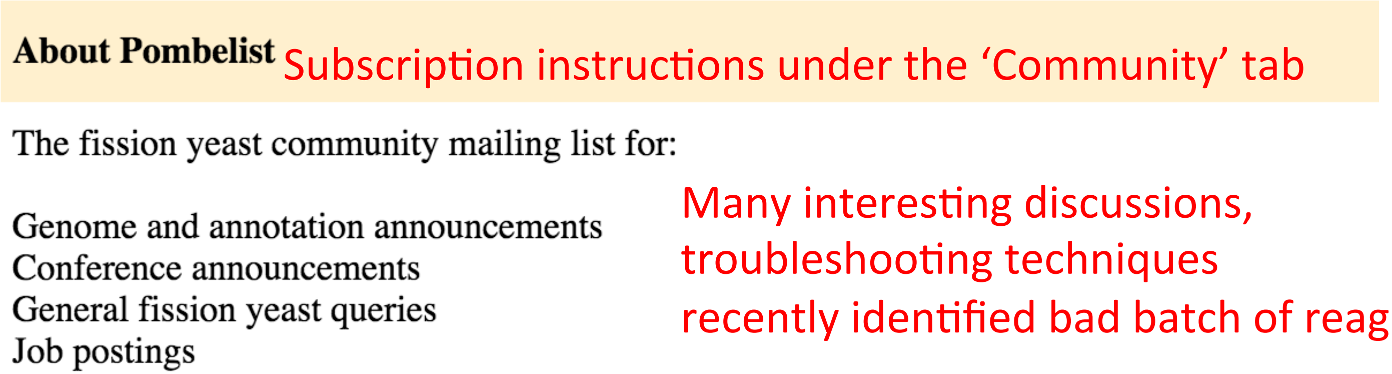
WEDNESDAY, 27TH MAY, 2015

How can I find genes with a specific activity

Can I download all *S. pombe* phenotype data?

How can I find the *S. pombe* ortholog of a human gene?

How can I find genes involved in human disease



About Pombelist

Subscription instructions under the ‘Community’ tab

The fission yeast community mailing list for:

Genome and annotation announcements
Conference announcements
General fission yeast queries
Job postings

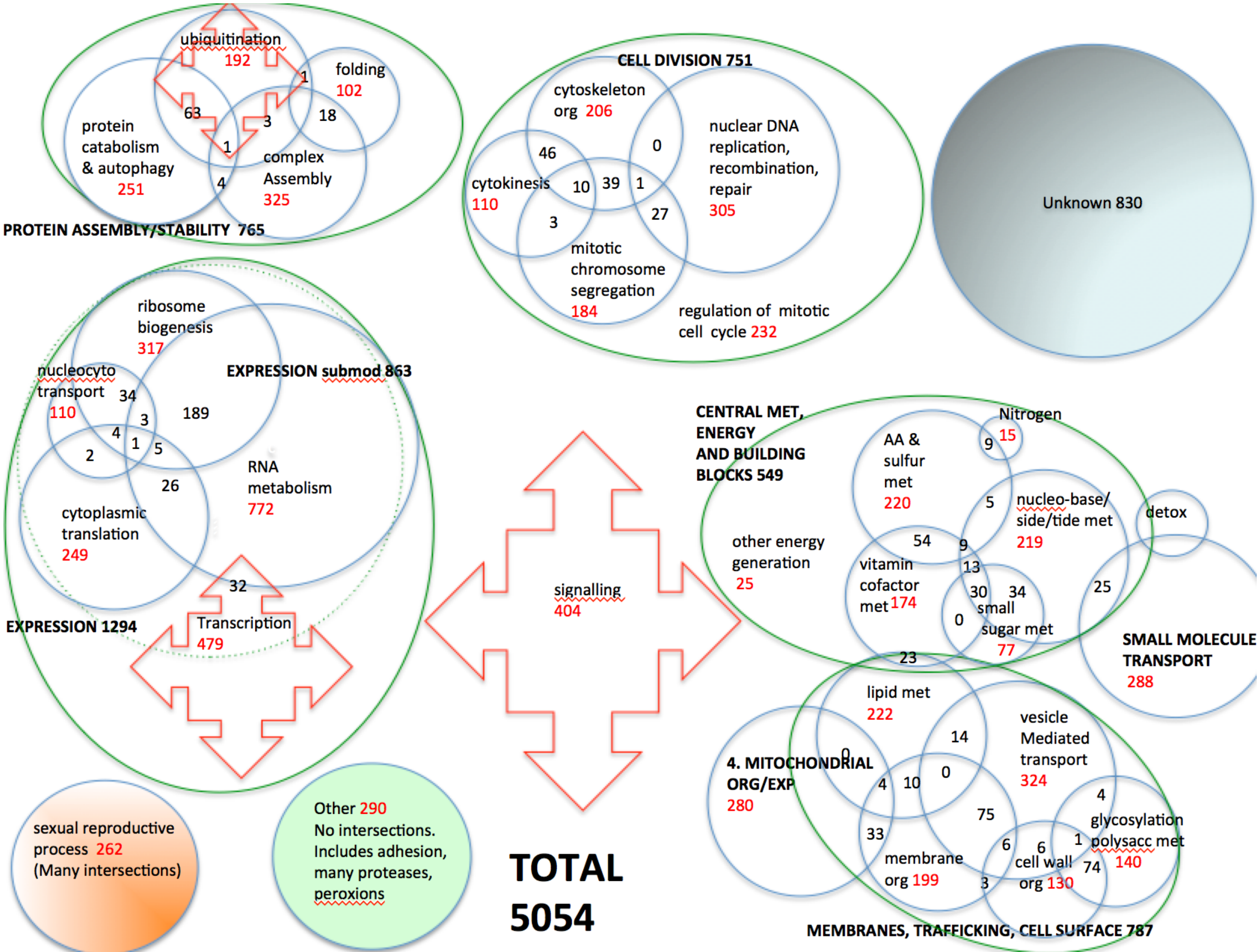
Many interesting discussions,
troubleshooting techniques
recently identified bad batch of reagent…

Final things

- e-mail the helpdesk@pombase.org if you can't find an answer
- Send us feedback corrections, omissions (contact curators link on every gene page)

Spare slides

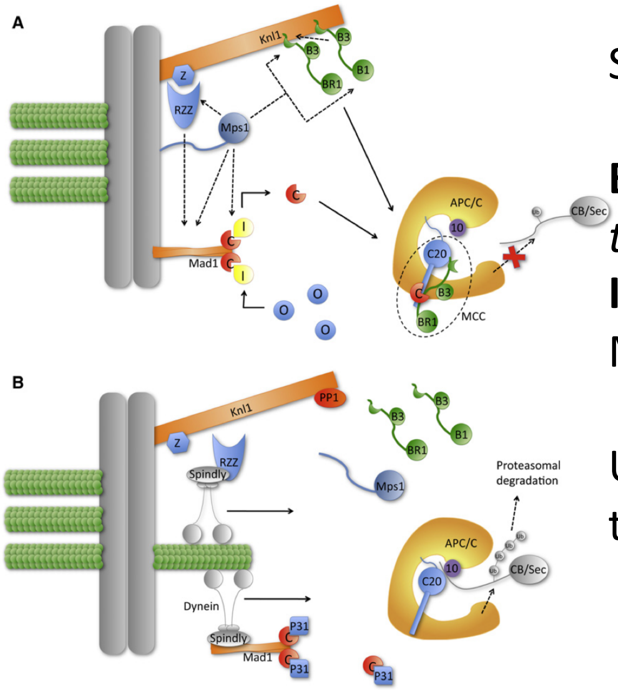
Visual “slim” of all 5054 proteins



Other 290
No intersections.
Includes adhesion, many proteases, peroxions

TOTAL 5054

Biological Processes are ordered assemblies of Molecular Functions…..



From PMID:22898774

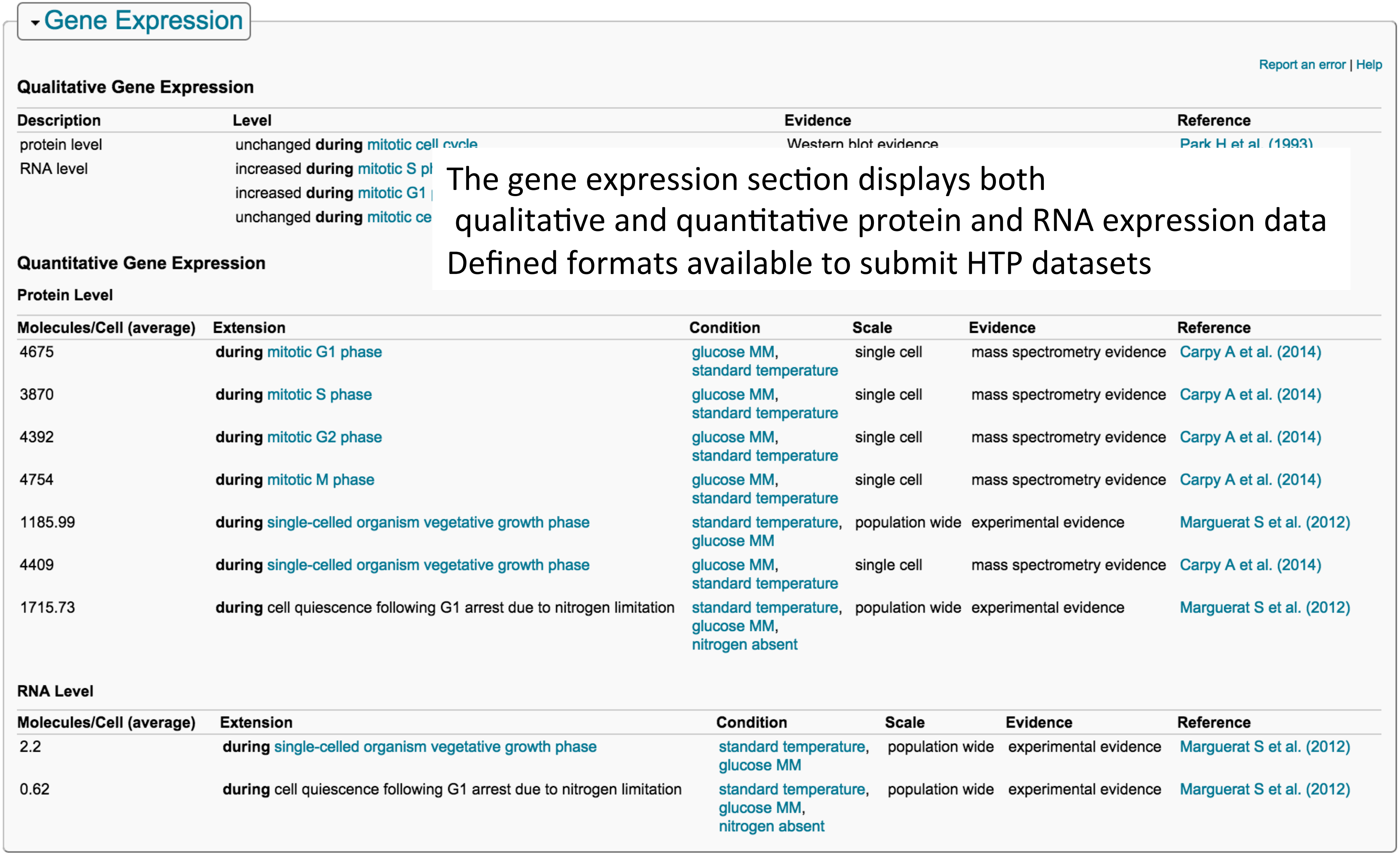
BP *positive regulation of mitotic metaphase/anaphase transition*
Includes Molecular functions:
APC/C MF ubiquitin ligase (substrate securin)
MF ubiquitin conjugation enzyme
Slp1/cdc20 MF Ubiquitin ligase activator activity (substrate APC)

BP *negative regulation of mitotic metaphase/anaphase transition*
Includes Molecular functions:
MCC MF ubiquitin ligase inhibitor activity

Usually we don't know all of the function parts of processes at the molecular level

cdc6

Gene Expression



Gene Expression

Report an error | Help

Qualitative Gene Expression

Description	Level	Evidence	Reference
protein level	unchanged **during** mitotic cell cycle	Western blot evidence	Park H et al. (1993)
RNA level	increased **during** mitotic S pl		
	increased **during** mitotic G1		
	unchanged **during** mitotic ce		

Quantitative Gene Expression

Protein Level

Molecules/Cell (average)	Extension	Condition	Scale	Evidence	Reference
4675	**during** mitotic G1 phase	glucose MM, standard temperature	single cell	mass spectrometry evidence	Carpy A et al. (2014)
3870	**during** mitotic S phase	glucose MM, standard temperature	single cell	mass spectrometry evidence	Carpy A et al. (2014)
4392	**during** mitotic G2 phase	glucose MM, standard temperature	single cell	mass spectrometry evidence	Carpy A et al. (2014)
4754	**during** mitotic M phase	glucose MM, standard temperature	single cell	mass spectrometry evidence	Carpy A et al. (2014)
1185.99	**during** single-celled organism vegetative growth phase	standard temperature, glucose MM	population wide	experimental evidence	Marguerat S et al. (2012)
4409	**during** single-celled organism vegetative growth phase	glucose MM, standard temperature	single cell	mass spectrometry evidence	Carpy A et al. (2014)
1715.73	**during** cell quiescence following G1 arrest due to nitrogen limitation	standard temperature, glucose MM, nitrogen absent	population wide	experimental evidence	Marguerat S et al. (2012)

RNA Level

Molecules/Cell (average)	Extension	Condition	Scale	Evidence	Reference
2.2	**during** single-celled organism vegetative growth phase	standard temperature, glucose MM	population wide	experimental evidence	Marguerat S et al. (2012)
0.62	**during** cell quiescence following G1 arrest due to nitrogen limitation	standard temperature, glucose MM, nitrogen absent	population wide	experimental evidence	Marguerat S et al. (2012)

The gene expression section displays both
qualitative and quantitative protein and RNA expression data
Defined formats available to submit HTP datasets