

Protein Interaction Data Sources

Ian Donaldson
Biotechnology Centre of Oslo
University of Oslo



MBV-INF 4410/9410
Thursday, September 16th, 2010

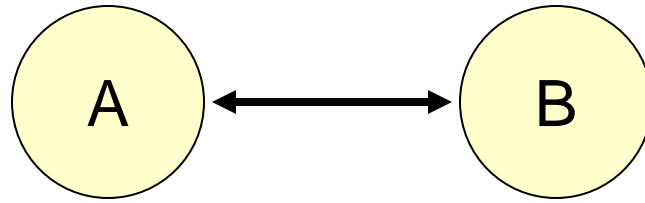
http://donaldson.uio.no/wiki/Protein_Interaction_Resources

<http://irefindex.uio.no>

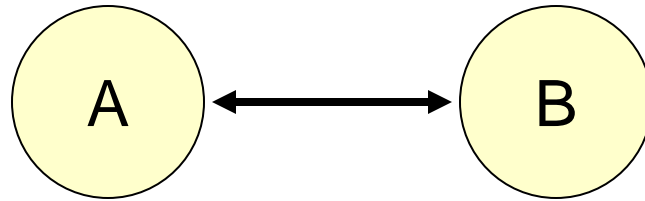
Overview

1. A simple interaction record
2. Experimental methods
3. Representation of interactions
4. Interaction databases
5. Identifiers
6. Text mining sources
7. Predicted interaction sources
8. Consolidated databases
9. Visualization tools
10. Applications

A simple interaction record

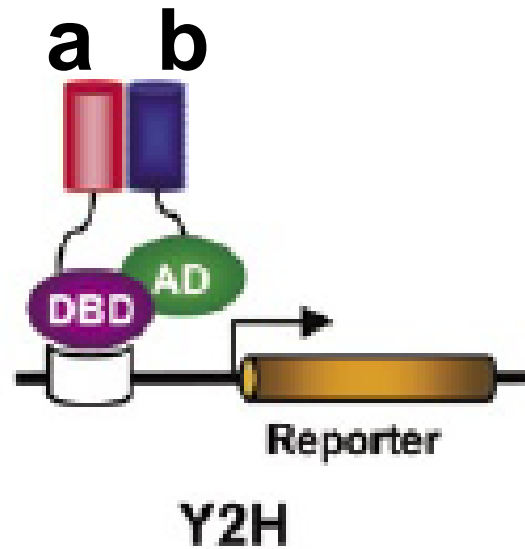


A simple interaction record



- | | |
|--------------------------|------------------------|
| 1. Short label | 5. Short label |
| 2. Type of molecule | 6. Type of molecule |
| 3. Database identifier | 7. Database identifier |
| 4. Origin | 8. Origin |
| 9. Publication reference | |

Experimental methods – Yeast Two Hybrid



Experimental methods

Two Hybrid family

Advantages

- an "in vivo" binding assay
- easy to automate
- detects binary interactions

Experimental methods

Two Hybrid Family

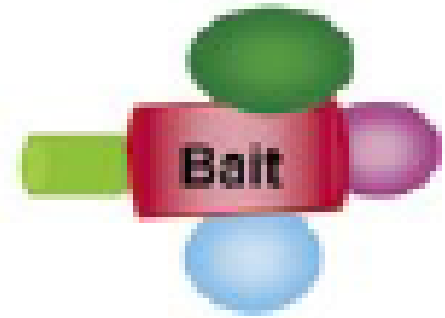
Limitations

- false positives due to
 - sticky or activating bait or prey
 - overexpression
 - mis-localization (contextual specificity)
 - intermediary proteins
- false negatives due to
 - disruption of interaction by fusion construct
 - absence of intermediary proteins
 - absence of native taxon context

Experimental methods

Affinity purification

- interacting proteins are purified by means of an affinity tag on a "bait" protein
- proteins are detected via MS
- readout of this experiment type is a list of proteins that somehow interact with the bait protein
- note this is fundamentally different than readout of the Y2H family of experiments (more later)



Experimental methods

Affinity purification family

Limitations

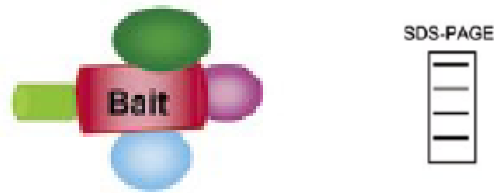
- false positives due to
 - sticky bait or prey
 - mis-identification during MS
 - non-native binding environment
- false negatives due to
 - mis-identification during MS
 - non-native binding environment

Experimental methods

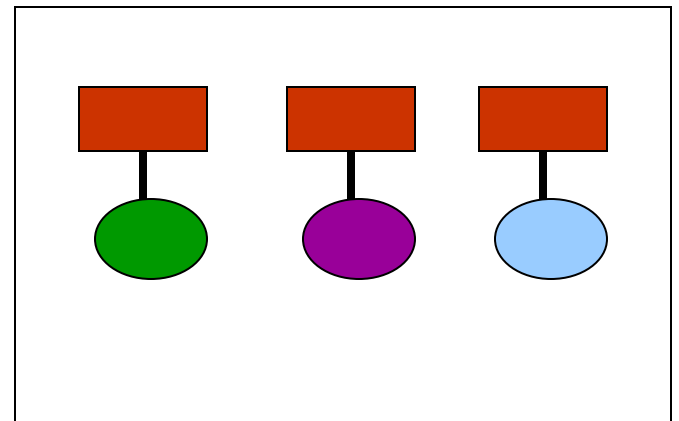
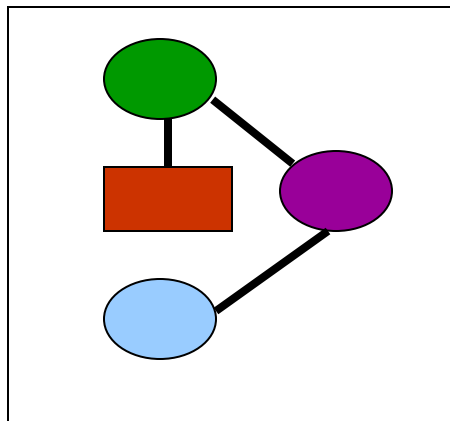
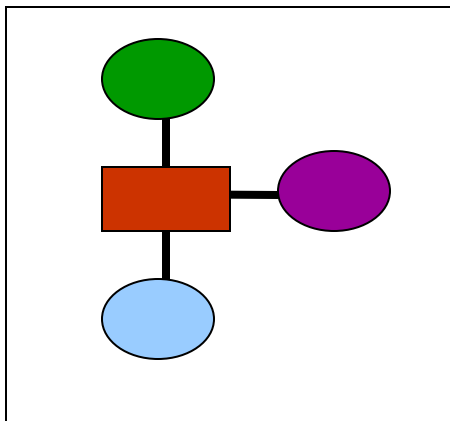
Affinity purification family

Advantages

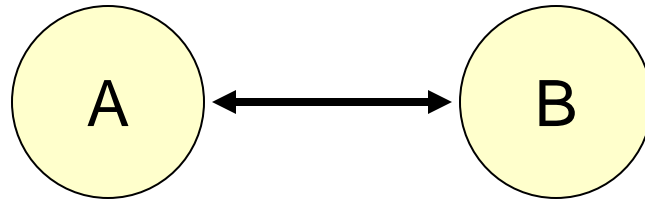
detection of protein complexes but no topology or stoichiometric data



This observation has multiple interpretations:



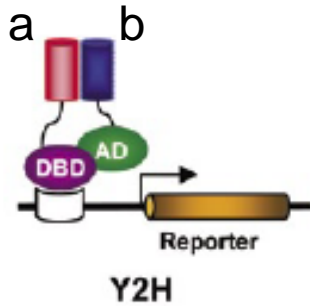
A simple interaction record



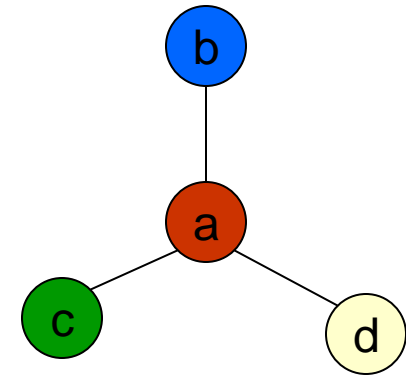
- | | |
|--------------------------|------------------------|
| 1. Short label | 5. Short label |
| 2. Type of molecule | 6. Type of molecule |
| 3. Database identifier | 7. Database identifier |
| 4. Origin | 8. Origin |
| 9. Publication reference | |

Representation of interaction data

binary data versus associative data

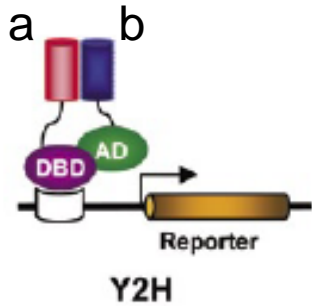


<u>exp</u>	<u>bait</u>	<u>prey</u>
1	a	b
2	a	c
3	a	d

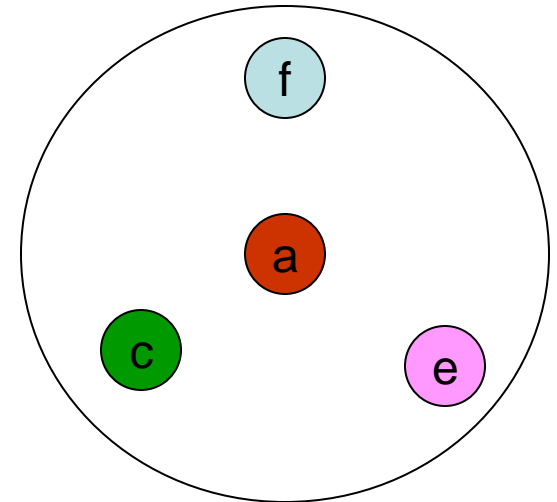
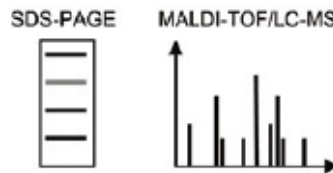
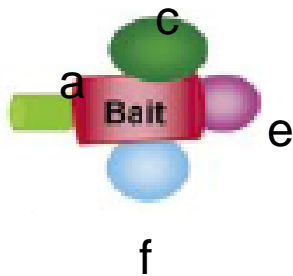
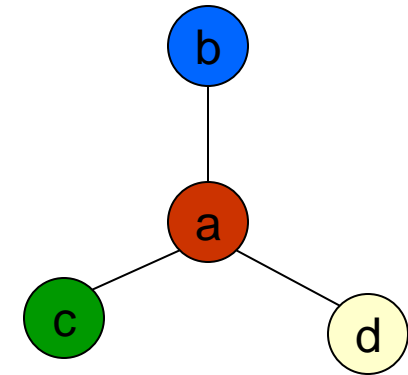


Representation of interaction data

binary data versus associative data

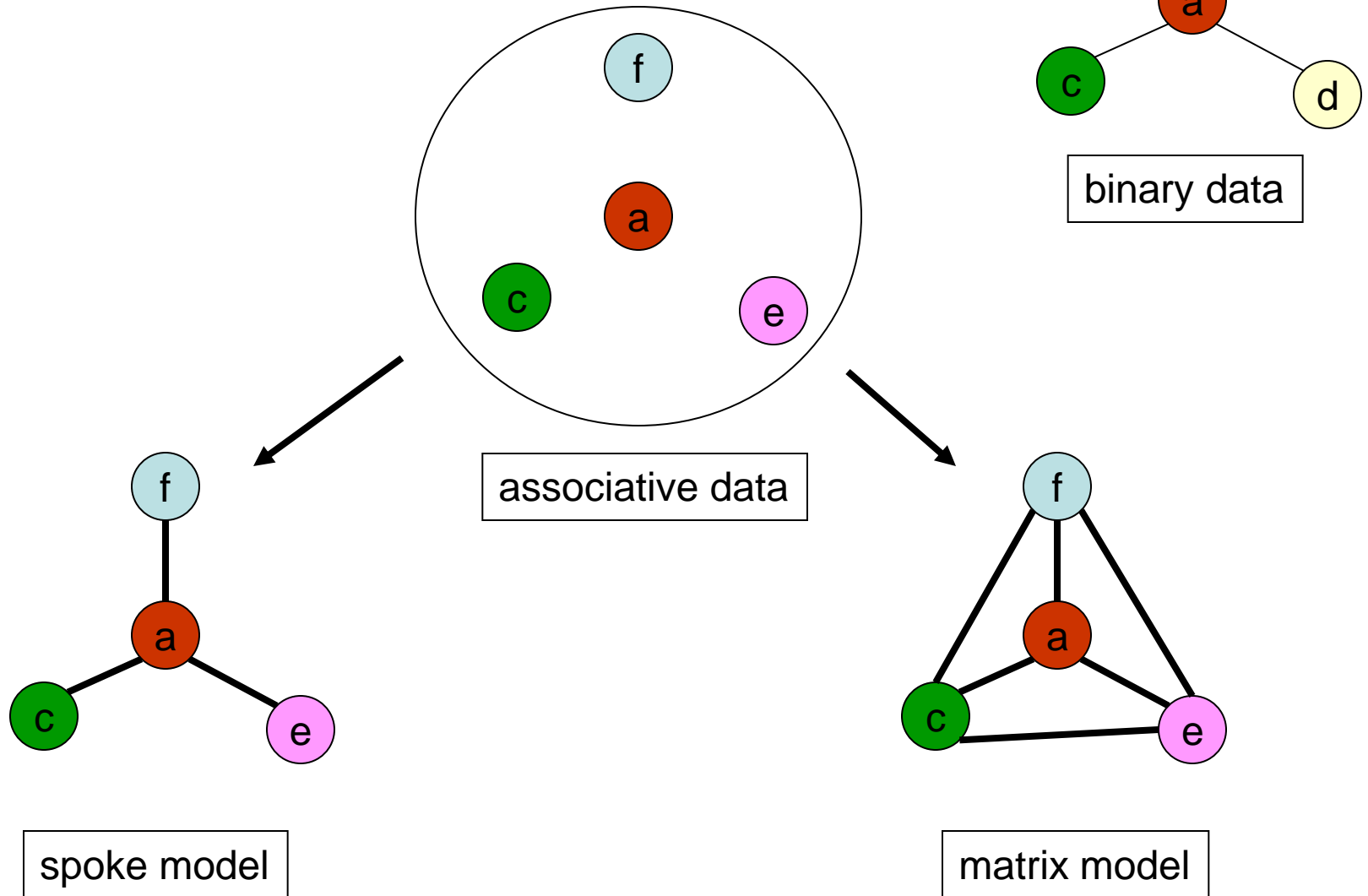


exp	bait	prey
1	a	b
2	a	c
3	a	d

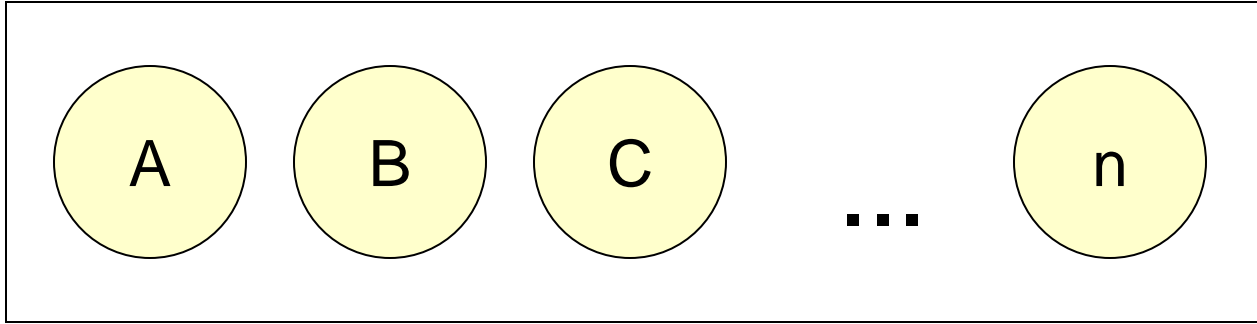


Representation of interaction data

misrepresentations of associative data



A simple interaction record



For each interactor

1. Short label
2. Type of molecule
3. Database identifier
4. Origin

For the interaction

5. Publication reference (method)

More about methods and interaction types that they support

<http://www.ebi.ac.uk/ontology-lookup/browse.do?ontName=MI>

The screenshot displays the Ontology Lookup Service (OLS) website in a Mozilla Firefox browser window. The page title is "Ontology Lookup Service (OLS) - Mozilla Firefox". The main content area is titled "OLS - Ontology Lookup Service" and "MI Ontology Browser". The interface includes a search bar with "All Databases" selected and "Enter Text Here" as the input. A navigation menu at the top includes "Databases", "Tools", "EBI Groups", "Training", "Industry", "About Us", and "Help". A sidebar on the left contains a navigation menu with items like "OLS Home", "Documentation", "Developer Resources", and "Contact Us". The main content area shows a tree view of the "molecular interaction" ontology, with sub-classes such as "interaction detection method", "participant identification method", "feature detection method", "interaction type", "genetic interaction", "colocalization", "association", "alias type", "interactor type", "feature range status", "experimental preparation", "cross-reference type", "database citation", "experimental role", and "biological role". A "Help (hide)" section on the right provides instructions on how to use the browser, including "Double-click a term to see its children" and "Click to highlight a term to see any information associated with it". Below the help section are sections for "Relations", "Term Information" (with fields for ID and Name), "Associated information", and "Term Hierarchy" (with radio buttons for "Paths to Root" and "Child relationships").

Ontology Lookup Service (OLS) - Mozilla Firefox

EMBL-EBI EB-eye Search All Databases Enter Text Here Go Reset Advanced Search Give us feedback

Databases Tools EBI Groups Training Industry About Us Help Site Index

OLS - Ontology Lookup Service

MI Ontology Browser

Help (hide)

Double-click a term to see its children. The ontology browser is populated dynamically. If there are many children for a given term, there may be a small delay while the browser fetches Click to highlight a term to see any information associated with it. Hover over a term to see its relation with its immediate parent. Root terms will not display any relational information.

Relations

Term Information

ID: Zoom

Name:

Associated information

Highlight a term to view its associated information.

Term Hierarchy

Paths to Root: Child relationships:

You can zoom the ontology browser by clicking on a term in the graph.

Done Tor Disabled

Interaction databases

BIND	inactive
BioGRID	IMEX (Observer)
CORUM	complexes
DIP	IMEX (Archival)
HPRD	human resource
IntAct	IMEX (Archival)
MATRIX	extracellular matrix
MINT	IMEX (Topical)
MPACT	IMEX (Topical) yeast
MPIDB	IMEX (Observer) prokaryotic
MPPI	mammalian
OPHID/I2D	predicted
Many others	

See <http://www.pathguide.org/> for URLs

Interaction databases

IMEx

International Molecular Exchange Consortium

IMEx partners exchange data and share curation tasks.

Set standards

They also accept interaction data submissions pre and post publication.

Data consolidation and exchange is an active area of research.

See <http://www.imexconsortium.org/>

Interaction databases: IntAct

The screenshot shows the IntAct website in a Mozilla Firefox browser. The browser's address bar displays the URL <http://www.ebi.ac.uk/intact/site/index.jsf>. The page content includes a search bar, a navigation menu, and a sidebar with statistics. A large text box highlights the URL <http://www.ebi.ac.uk/intact>.

IntAct Home

Search IntAct

To perform a search in the IntAct database use the search box in Examples:

- Gene name: [BRCA2](#)
- UniProtKB Ac: [Q06609](#)
- UniProtKB Id: [dmc1](#)
- Pubmed Id: [10831611](#)

Introduction

IntAct provides a freely available, open source database system and analysis tools for protein interaction data. All interactions are derived from literature curation or direct user submissions and are freely available.

IntAct Basic Statistics

The IntAct Database contains:

- 173,621 binary interactions.
- 54,471 proteins.
- 9,128 experiments.
- 1,463 controlled vocabulary terms.

License

All software, available on this site, is under the [Apache License, version 2](#), and all data under the [Creative Commons Attribution License](#). This means that you are free to copy, distribute, display and make commercial use of all records from the IntAct database provided appropriate credit is given.

Acknowledgements

Identifiers: EntrezGene

The screenshot shows the Entrez Gene website interface. At the top, the browser window title is "Gene Home - Mozilla Firefox". The address bar shows the URL "http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene". The page header includes the NCBI logo and the text "Entrez Gene". Below the header, there are navigation tabs for "All Databases", "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "OMIM", "PMC", "Journals", and "Books". A search bar is present with the text "Search Gene for" and buttons for "Go" and "Clear".

The main content area features a description: "Entrez Gene is a searchable database of genes, from RefSeq genomes, and defined by sequence and/or located in the NCBI Map Viewer". Below this, there are sections for "Limits", "Preview/Index", "History", "Clipboard", and "Details". A "News" section is also visible, with a link to "Limit by Chromosomal Region; Sort Options. News archives...".

A "Sample Searches" section is displayed, listing various search criteria and their corresponding search text:

Find genes by...	Search text
free text	human muscular dystrophy
partial name and multiple species	transporter[title] AND ("Drosophila melanogaster"[organ] OR "Mus musculus"[organ])
chromosome and symbol	(II[chr] OR 2[chr]) AND adh*[sym]
associated sequence accession number	M11313[accn]
gene name (symbol)	BRCA1[sym]
publication (PubMed ID)	11331580[PMID]
Gene Ontology (GO) terms or identifiers	"cell adhesion"[GO] 10030[GO]
Genes with reports of clinical significance (under development)	...[GO]
chromosome	
Enzyme	

A large text box is overlaid on the page, containing the URL: <http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene>

The footer of the page contains several links and information, including "About Entrez", "Entrez Gene: gene-centered information at NCBI", "Nucleic Acids Res. 2005 Jan 1;33:D54-8.", "General help on the Entrez search and retrieval system", "NCBI Handbook Chapter on Entrez Gene (download PDF)", "How to add information about function (GeneRIF)", "Correct or update a Gene record", "Correct or update a reference sequence", and "Report a publication or GeneRIF error".



Entrez Gene

My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search Gene for 7132 [Go] [Clear]

Limits Preview/Index History Clipboard Details

Display Full Report Show 20 Sort by Relevance Send to

All: 1 Current P GeneView: 1

GeneID: 7132

1: TNFRSF1A tumor necrosis factor receptor superfamily, member 1A [Homo sapiens] updated 12-Oct-2008

GeneID: 7132 Summary

Official Symbol: TNFRSF1A aka Gene Name

Official Symbol TNFRSF1A

Official Full Name tumor necrosis factor receptor superfamily, member 1A

Primary source HGNC:11916

See related Ensembl:ENSG00000067182; HPRD:01861; MIM:191190

Gene type protein coding

RefSeq status REVIEWED

Organism Homo sapiens
Lineage Eukaryota; Metazoa; Chordata; Euteleostomi; Mammalia; Eutheria; Primates; Hominidae; Homo

Synonyms: TNF-R-I

Also known as TNFRSF1A; p55; p60; TBP1; TNF-R; TNFAR; TNFR1; p55-R; CD120a; TNFR55; TNFR60; TNF-R-I; TNF-R55; MGC19588

Summary The protein encoded by this gene is a member of the TNF-receptor superfamily. This protein is one of the major receptors for the tumor necrosis factor-alpha. This receptor can activate NF-kappaB, mediate apoptosis, and function as a regulator of inflammation. Antiapoptotic protein BCL2-associated athanogene 4 (BAG4/SODD) and adaptor proteins TRADD and TRAF2

Entrez Gene Home

Table Of Contents

- Summary
- Genomic regions, transcripts...
- Genomic context
- Bibliography
- HIV-1 protein interactions
- Interactions
- General gene information
- General protein information
- Reference Sequences
- Related Sequences
- Additional Links

Links

Explain

- Order cDNA clone
- Conserved Domains
- Genome
- GEO Profiles
- HomoloGene
- Map Viewer
- Nucleotide
- EST
- OMIM
- PubChem Compound
- PubChem Substance
- Full text in PMC
- Probe

protein... have been shown to interact with this receptor, and thus play regulatory roles in the signal transduction mediated by the receptor. Germline mutations of the extracellular domains of this receptor were found to be associated with the autosomal dominant periodic fever syndrome. The impaired receptor clearance is thought to be a mechanism of the disease. [provided by RefSeq]

- Probe
- Protein
- PubMed
- PubMed (OMIM)
- PubMed (GeneRIF)
- SNP
- SNP: Genotype
- SNP: GeneView
- Taxonomy
- UniSTS
- AceView
- CCDS

Genomic regions, transcripts, and products

(minus strand) Go to [reference sequence details](#)

[Try our new Sequence Viewer](#)

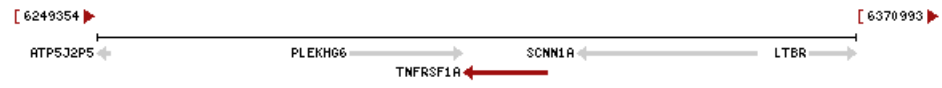


RefSeq Protein Accession:
NP_001056

Genomic context

chromosome: 12; Location: 12p13.2

[See TNFRSF1A in MapViewer](#)



- MGC
- ModelMaker
- PharmGKB
- The registry of TNFRSF1A sequence variants
- UniGene

- Entrez Gene Info
- Feedback
- Subscriptions

Bibliography

Related Articles in PubMed

[PubMed links](#)

GeneRIFs: Gene References Into Function

[What's a GeneRIF?](#)

- oxidative stress promotes TNFR receptor (TNFR1- and TNFR2) self-interaction and ligand-independent and enhanced ligand-dependent TNF signaling
- Data show that Type 2 diabetes is associated with substantially increased TNFR1 and TNF alpha activity

NCBI Sequence Viewer v2.0 - Mozilla Firefox

File Edit View History Bookmarks Digo Tools Help

http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?val=NP_001056.1&dopt=gp

Most Visited Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese...

Stumble! I like it! Send to Channels: Favorites Friends Tools

NCBI Entrez Protein

My NCBI Sign In Register

PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM Books

Search Protein for Go Clear

Display GenPept Show 5 Send to

Range: from begin to end Features: SNP CDD HPRD Refresh

1: NP_001056. Reports tumor necrosis fa...[gi:4507575] Conserved Domains, Links

Comment Features Sequence

LOCUS NP_001056 455 aa

DEFINITION tumor necrosis factor receptor 1 precursor [Homo sapiens]

ACCESSION NP_001056

VERSION NP_001056.1 GI:4507575

DBSOURCE REFSEQ: accession NM_001065.2

KEYWORDS .

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cephalochordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homini

REFERENCE 1 (residues 1 to 455)

AUTHORS Menon,R., Velez,D., Williams,S.M.

TITLE Genetic regulation of tumor necrosis factor receptor concentrations affected by race and preterm birth

JOURNAL Hum. Genet. 124 (3), 243-253 (2008)

PUBMED 18807256

REMARK GeneRIF: Observational study of gene-disease association and gene-environment interaction. (HuGE Navigator)

REFERENCE 2 (residues 1 to 455)

AUTHORS Hosgood,H.D. III, Menashe,I., Shen,M., Yeager,M., Yuenger,J., Rajaraman,P., He,X., Chatterjee,N., Caporaso,N.E., Zhu,Y.,

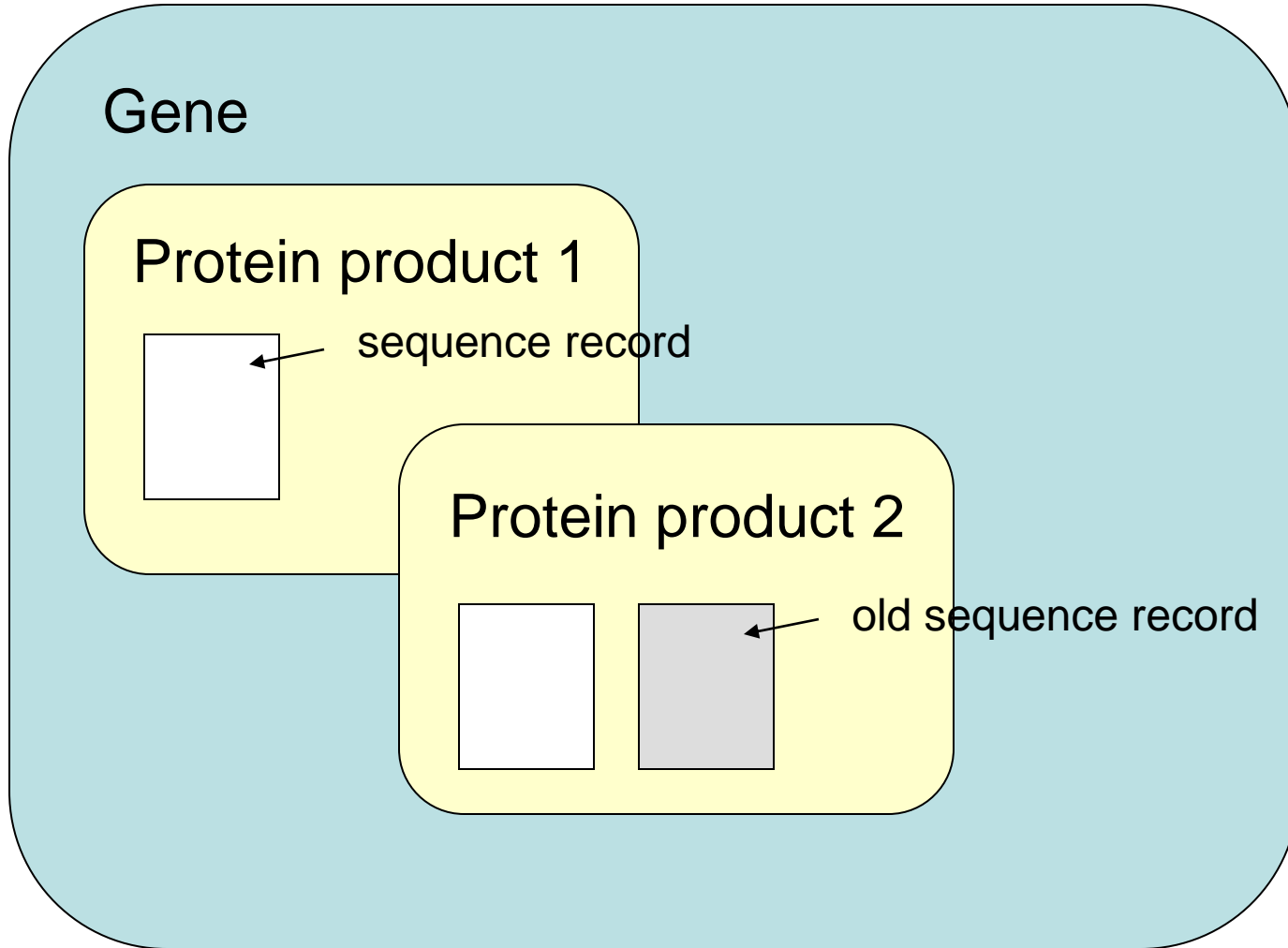
Done

RefSeq Protein Accession: NP_001056

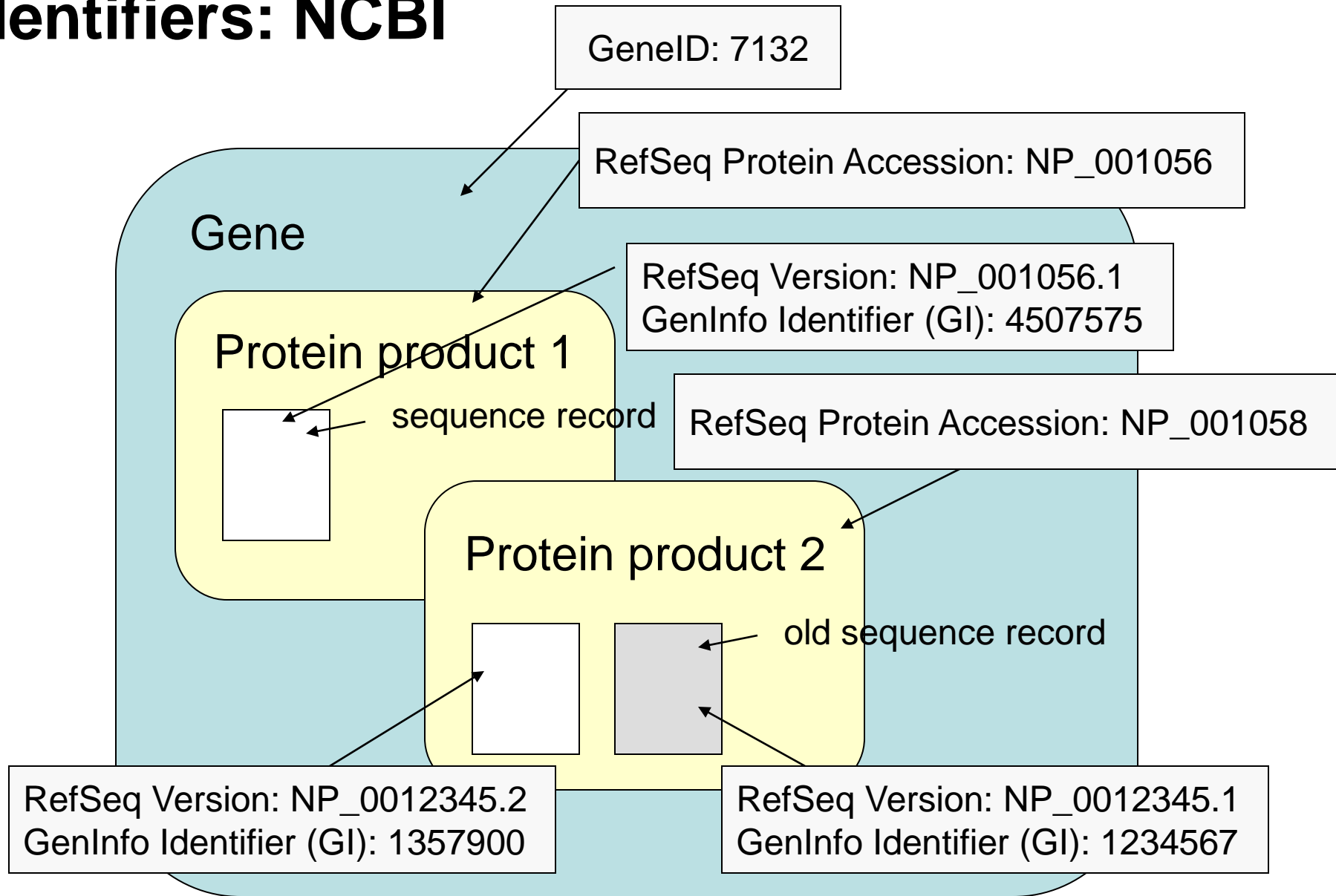
GenInfo Identifier: 4507575

RefSeq Version: NP_001056.1

Identifiers: NCBI



Identifiers: NCBI



Identifiers: UniProtKB

<http://www.uniprot.org/>

UniProt - Mozilla Firefox

File Edit View History Bookmarks Diigo Tools Help

http://www.uniprot.org/

Most Visited Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese...

Stumble! I like it! Send to Channels: All Favorites Friends Tools

UniProt Downloads Contact Documentation/Help

Search in Protein Knowledgebase (UniProtKB) Query Search Clear Fields »

Search Blast Align Retrieve ID Mapping

WELCOME

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

What we provide

UniProtKB	Protein knowledgebase, consists of two sections: <ul style="list-style-type: none">★ Swiss-Prot, which is manually annotated and reviewed.★ TrEMBL, which is automatically annotated and is not reviewed.
UniRef	Sequence clusters, used to speed up similarity searches.
UniParc	Sequence archive, used to keep track of sequences and their identifiers.
Supporting data	Literature citations, taxonomy, keywords and more.

NEWS

Release 14.3 – Oct 14, 2008
10th anniversary of the Swiss Institute of Bioinformatics

- › Statistics for UniProtKB: [Swiss-Prot](#) · [TrEMBL](#)
- › Forthcoming changes
- › News archives

SITE TOUR

Learn how to make best use of the tools and data on this site.

Search in

Query

Protein Knowledgebase (UniProtKB)

UniProtKB Entry: RAD51_HUMAN

★ Reviewed, UniProtKB/Swiss-Prot **Q06609** (RAD51_HUMAN)

Last modified September 23, 2008. Version 100.

UniProtKB Accession: Q06609

Contribute Send feedback

Clusters with 100%, 90%, 50% identity Documents [FASTA] [GFF] [XML]

Names and origin Protein attributes General annotation (Comments) Ontologies Binary interactions Alternative products Sequence annotation (Features) Sequences References Cross-references Entry information Relevant documents

Names and origin

Hide | Top

Protein names

Recommended name:

DNA repair protein

Gene Name: RAD51

Short name=hRAD51

Short name=HsRAD51

Gene names

Name:

RAD51

Synonyms: RAD51A

Synonyms: RAD51A, RECA

Organism

Homo sapiens (Human)

Taxonomic identifier

9606 [NCBI]

Taxon id: 9606 (human)

Taxonomic lineage

Eukaryota › Metazoa › Chordata › Mammalia › Eutheria › Euarchontoglires › Primates › Haplorhini › Catarrhini › Hominoidea › Homo

Binary interactions

Binary interactions from IntAct

With	Entry			
itself		1	EBI-297202,EBI-297202	
BRCA2	P51587	7	EBI-297202,EBI-79792	
RAD51AP1	Q96B01-2	2	EBI-297202,EBI-1178743	
RAD51AP1	Q96B01-3	3	EBI-297202,EBI-1178748	
rhp51	P36601	2	EBI-297202,EBI-926960	From a different organism.
TP53	P04637	1	EBI-297202,EBI-366083	

Alternative products

This entry describes **2** isoforms produced by **alternative splicing**. [\[Align\]](#) [\[Select\]](#)

Isoform 1 (identifier: **Q06609-1**)

This isoform has been chosen as the canonical sequence. This entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.

Isoform identifiers

Isoform 2 (identifier: **Q06609-2**)

The sequence of this isoform differs from the canonical sequence as follows:
77-173: Missing.

Notes: No experimental confirmation available.



Other Resources

SOURCE	Search...
ProtoNet	Search...

Entry information

Hide | Top

Entry name	RAD51_HUMAN	
Accession	Primary (citable) accession number: Q06609 Secondary accession number(s): <u>Q6FHX9, Q9BV60</u>	
Entry history	Integrated into UniProtKB/Swiss-Prot:	June 1, 1994
	Last sequence update:	June 1, 1994
	Last modified:	September 1, 2009
	This is version 100 of the entry and version 1 of the sequence. [Complete history]	
Entry status	Reviewed (UniProtKB/Swiss-Prot)	
Annotation project	HPI (Human Proteome Initiative)	

Secondary (retired) identifiers

Relevant documents

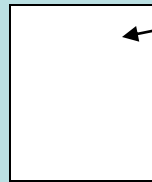
Hide | Top

- [Human chromosome 15](#)
Human chromosome 15: entries, gene names and cross-references to MIM
- [Human entries with polymorphisms or disease mutations](#)
List of human entries with polymorphisms or disease mutations
- [Human polymorphisms and disease mutations](#)
Index of human polymorphisms and disease mutations

Identifiers: UniProt

Q06609/ RAD51_HUMAN

Protein

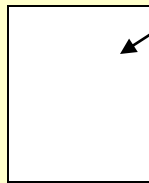


sequence record

Q06609-1

Q06609-2

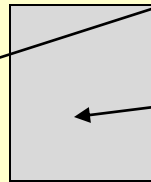
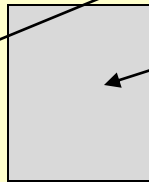
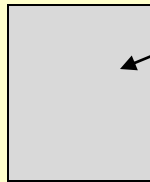
Isoforms



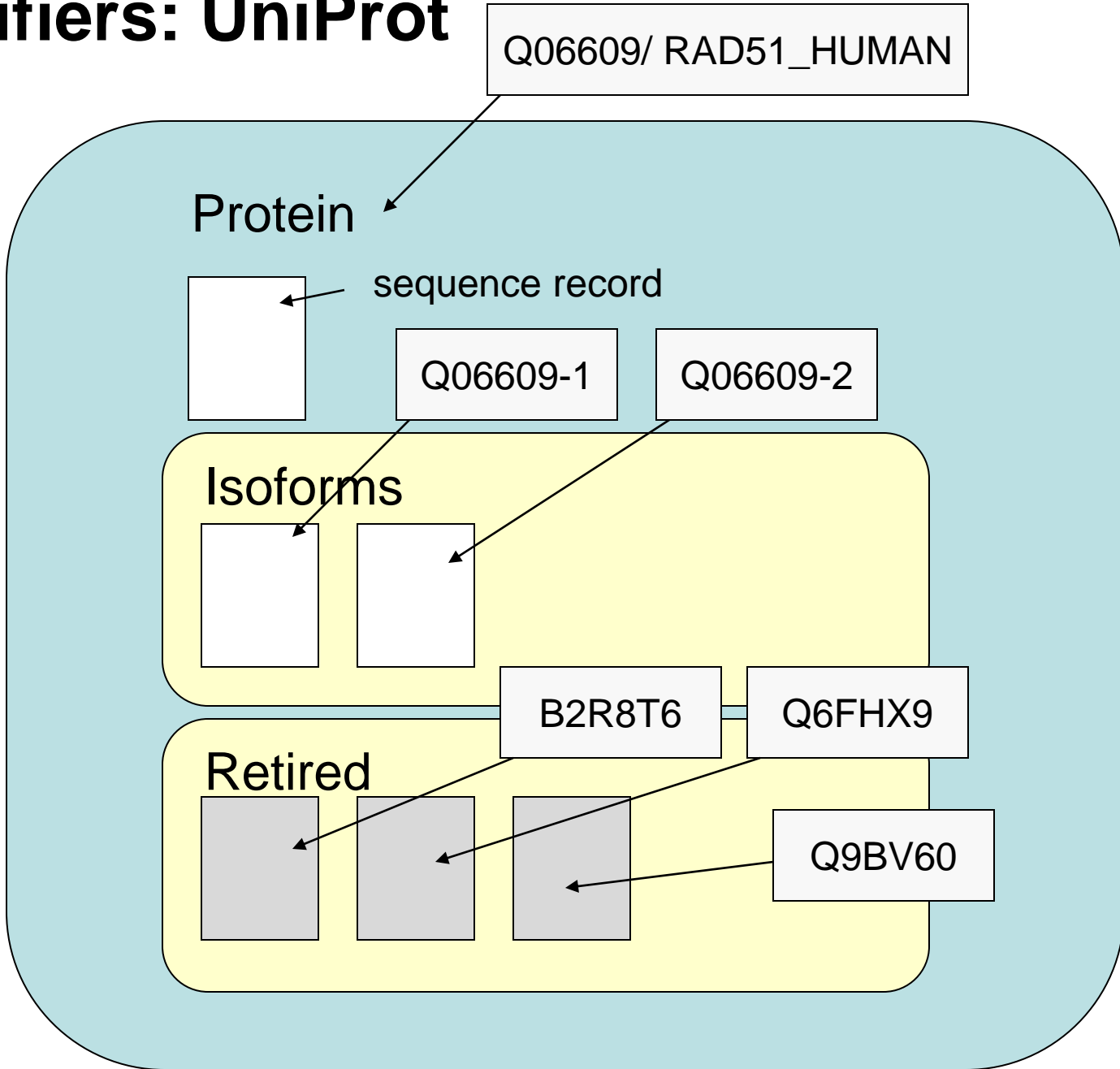
B2R8T6

Q6FHX9

Retired



Q9BV60



Common Identifiers

Gene

Ensembl [ENSG00000139618](#)

Entrez Gene [675](#)

Unigene [Hs.34012](#)

RNA transcript

GenBank [BC026160.1](#)

RefSeq [NM_000059](#)

Ensembl [ENST00000380152](#)

Protein

Ensembl [ENSP00000369497](#)

RefSeq [NP_000050.2](#)

UniProt [BRCA2_HUMAN](#) or

[A1YBP1_HUMAN](#)

IPI [IPI00412408.1](#)

EMBL [AF309413](#)

PDB [1MIU](#)

Species-specific

HUGO HGNC [BRCA2](#)

MGI [MGI:109337](#)

RGD [2219](#)

ZFIN [ZDB-GENE-060510-3](#)

FlyBase [CG9097](#)

WormBase [WBGene00002299](#) or [ZK1067.1](#)

SGD [S000002187](#) or [YDL029W](#)

Annotations

InterPro [IPR015252](#)

OMIM [600185](#)

Pfam [PF09104](#)

Gene Ontology [GO:0000724](#)

SNPs [rs28897757](#)

Experimental Platform

Affymetrix [208368_3p_s_at](#)

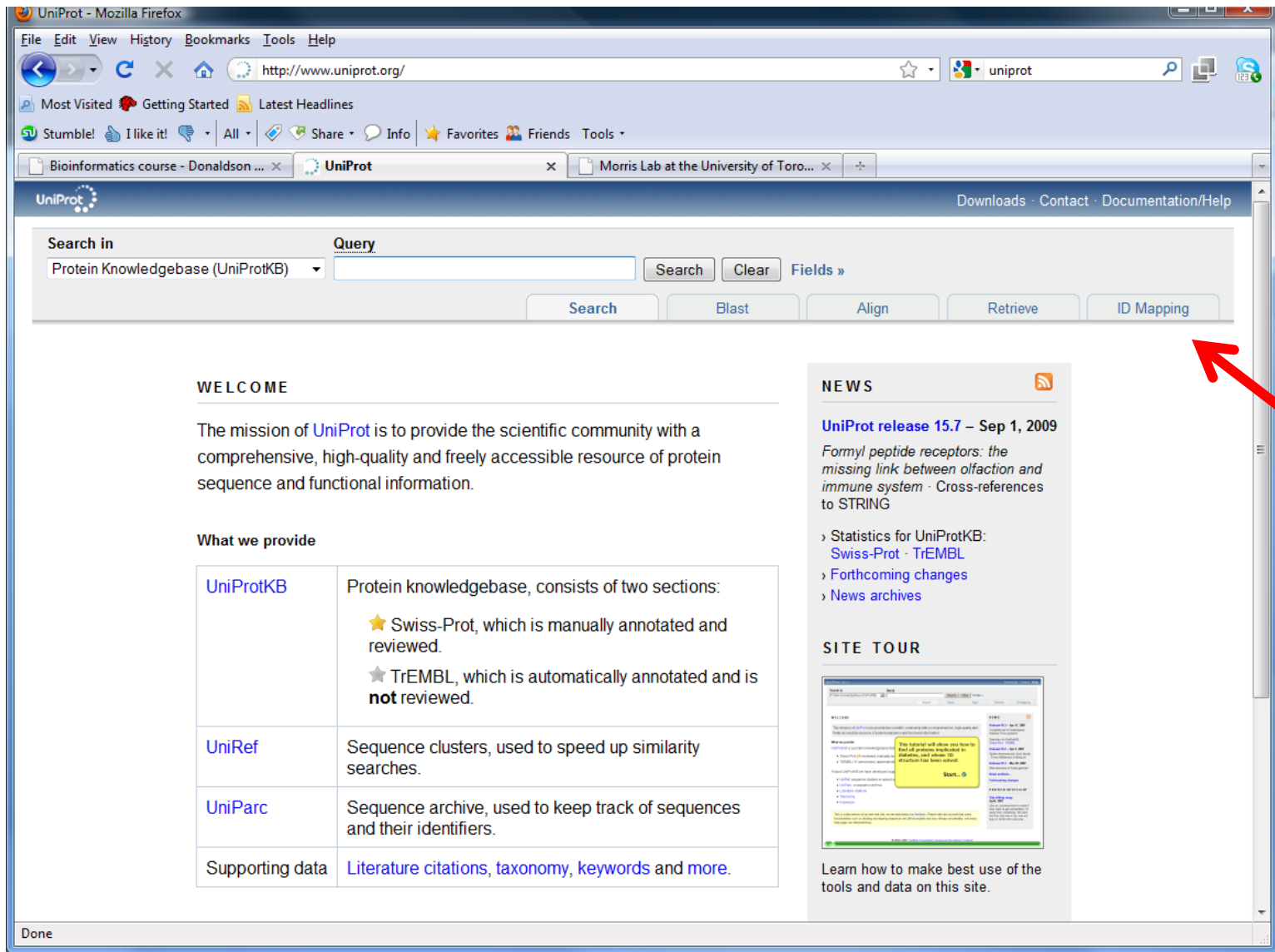
Agilent [A_23_P99452](#)

CodeLink [GE60169](#)

Illumina [GI_4502450-S](#)

Red = Recommended

UniProt ID Mapping Service



The screenshot shows the UniProt website interface. At the top, there is a search bar with a dropdown menu set to 'Protein Knowledgebase (UniProtKB)'. Below the search bar are buttons for 'Search', 'Blast', 'Align', 'Retrieve', and 'ID Mapping'. A red arrow points to the 'ID Mapping' button. The main content area includes a 'WELCOME' section, a 'What we provide' section with a table, a 'NEWS' section, and a 'SITE TOUR' section.

Search in **Query**
Protein Knowledgebase (UniProtKB) [Fields »](#)

WELCOME

The mission of **UniProt** is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

What we provide

UniProtKB	Protein knowledgebase, consists of two sections: <ul style="list-style-type: none">★ Swiss-Prot, which is manually annotated and reviewed.★ TrEMBL, which is automatically annotated and is not reviewed.
UniRef	Sequence clusters, used to speed up similarity searches.
UniParc	Sequence archive, used to keep track of sequences and their identifiers.
Supporting data	Literature citations , taxonomy , keywords and more .

NEWS

UniProt release 15.7 – Sep 1, 2009
Formyl peptide receptors: the missing link between olfaction and immune system · Cross-references to STRING

- › Statistics for UniProtKB:
 - [Swiss-Prot](#) · [TrEMBL](#)
- › [Forthcoming changes](#)
- › [News archives](#)

SITE TOUR

Learn how to make best use of the tools and data on this site.

UniProt ID Mapping Service

The screenshot shows the UniProt website in a Mozilla Firefox browser. The address bar displays <http://www.uniprot.org/>. The browser's address bar includes navigation icons (back, forward, refresh, home), a search icon, and a language dropdown set to 'uniprot'. The browser's tab bar shows the current tab is 'UniProt'.

The UniProt website header features the UniProt logo on the left and navigation links for 'Downloads', 'Contact', and 'Documentation/Help' on the right. Below the header is a search and mapping interface. On the left, under the heading 'Identifiers', there is a text input field containing 'YIL062C'. To the right of this field are two dropdown menus: 'From' (set to 'UniProtKB AC/ID') and 'To' (set to 'Entrez Gene (GeneID)'). Below these are three buttons: 'Map', 'Swap', and 'Clear'. A 'Browse...' button is located below the 'To' dropdown. At the bottom of the mapping section are five buttons: 'Search', 'Blast', 'Align', 'Retrieve', and 'ID Mapping'.

On the right side of the mapping interface, there is a box titled 'Database identifier mapping tips' with the following text: 'To map identifiers to or from UniProtKB: • enter identifiers, e.g.: 1TIA 1FNS • select a source database, e.g.: PDB • or select a target database, e.g.: UniProtKB More...'. Below this box are the 'Search', 'Blast', 'Align', 'Retrieve', and 'ID Mapping' buttons.

The main content area of the website is divided into three columns. The left column has a 'WELCOME' section with the text: 'The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.' Below this is a 'What we provide' section with a sub-section for 'UniProtKB' which states: 'Protein knowledgebase, consists of two sections: ★ Swiss-Prot, which is manually annotated and reviewed. ★ TrEMBL, which is automatically annotated and is'. The right column has a 'NEWS' section with a sub-section for 'UniProt release 15.7 – Sep 1, 2009' which includes the text: 'Formyl peptide receptors: the missing link between olfaction and immune system · Cross-references to STRING'. Below this are links for 'Statistics for UniProtKB: Swiss-Prot · TrEMBL', 'Forthcoming changes', and 'News archives'. At the bottom of the right column is a 'SITE TOUR' section.

Ensembl BioMart

<http://www.biomart.org/biomart>

The screenshot shows the Ensembl BioMart website interface within a Mozilla Firefox browser window. The browser's address bar displays the URL <http://www.biomart.org/biomart/martview/02e2f0a0dabf3992cb35496f164dбаад>. The website header includes the BioMart logo and a navigation menu with links for HOME, MARTVIEW, MARTSERVICE, DOCS, CONTACT, NEWS, and CREDITS. Below the navigation menu, there are buttons for 'New', 'Count', and 'Results', along with utility buttons for 'URL', 'XML', 'Perl', and 'Help'. The main content area is titled 'Please restrict your query using criteria below' and features a sidebar on the left with sections for 'Dataset', 'Features', 'Filters', and 'Attributes'. The 'Filters' section shows '[None selected]'. The 'Attributes' section lists 'Feature Symbol', 'Feature Name', 'GO ID', and 'GO Term'. The main query area is divided into sections: 'Genome Features' (with a sub-section 'Identifiers' containing a dropdown menu for 'MGI ID' and a 'Browse...' button), 'Feature Name' (with an empty text input field), and 'Feature Type' (with a dropdown menu showing options like 'Gene', 'DNA Segment', 'Cytogenetic Marker', and 'QTL'). The browser's status bar at the bottom indicates 'Done'.

ID Mapping Challenges

- Avoid errors: map IDs correctly
- Gene name ambiguity – not a good ID
 - e.g. FLJ92943, LFS1, TRP53, p53
 - Better to use the standard gene symbol: TP53
- Excel error-introduction
 - OCT4 is changed to October-4
- Problems reaching 100% coverage
 - E.g. due to version issues
 - Use multiple sources to increase coverage

Zeeberg BR et al. Mistaken identifiers: gene name errors can be introduced inadvertently when using Excel in bioinformatics BMC Bioinformatics. 2004 Jun 23;5:80

ID Mapping Challenges

- Spot-test any ID mapping service you use.
- Check samples from first, last and middle of your list of identifiers to be converted.
- Ask for help if you are uncertain.

Identifiers: Why can't I find my protein

Are you using a sequence record identifier?

Is the identifier retired?

Are you using the identifier for the correct organism?

Have you tried searching for identifiers for all products of the gene?

Try different identifier systems (Entrez Gene, UniProt)

Remember that only 10% to 50% of all proteins have interaction information.

You always have recourse to text-mining and predicted interactions.

Overview

1. A simple interaction record
2. Experimental methods
3. Representation of interactions
4. Interaction databases
5. Identifiers
6. Text mining sources
7. Predicted interaction sources
8. Consolidated databases
9. Visualization tools
10. Applications

Text mining resources

Most are based on co-occurrence of gene name mentions.

Start at http://zope.bioinfo.cnio.es/bionlp_tools

iHOP	http://www.ihop-net.org/UniPub/iHOP/
EBIMed	http://www.ebi.ac.uk/Rebholz-srv/ebimed/index.jsp
FACTA	http://text0.mib.man.ac.uk/software/facta/
PubGene	http://www.pubgene.org/
InfoPubMed	https://www-tsujii.is.s.u-tokyo.ac.jp/info-pubmed/
Chilibot	http://www.chilibot.net/
MedEvi	http://www.ebi.ac.uk/tc-test/textmining/medevi/
FABLE	http://fable.chop.edu/overview.jsp
BCMS	http://bcms.bioinfo.cnio.es/
BioText	http://biosearch.berkeley.edu/

Text mining resources: iHOP

iHOP - Information Hyperlinked over Proteins - Mozilla Firefox

http://www.ihop-net.org/UniPub/iHOP/index.html?field=all&search=RAD51&organism_id=0

Most Visited Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese... W Wikipedia

Diigo Bookmark Highlight Comment Send Message (0) Options

Stumble! I like it! Send to Channels: Favorites Friends Tools

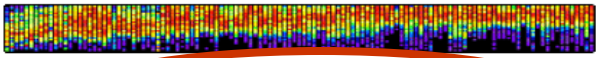




























































iHOP - Inf... EBIMed FACTA PubGene - Find... Chilibot: finding... MedEvi - a per... FABLE - Overvi... BioCreative Met... BioText Search ...

iHOP
Information hyperlinked over proteins

Search Gene
RAD51

Gene Model Developer's Zone
How to cite iHOP

Contact Links Help

Symbol	Name	Synonym/ DB-reference	Organism	Results
	 Life cycles of successful genes			↓
RAD51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)		Homo sapiens	   
Rad51	RAD51 homolog (S. cerevisiae)		Mus musculus	   
RAD51	Strand exchange protein, forms a helical filament with DNA that searches for ...		Saccharomyces cerevisiae	   
rad51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)		Danio rerio	   
RAD51	LeRAD51		Lycopersicon esculentum	   
rad51	rad51 protein		Xenopus laevis	   
RAD51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)		Gallus gallus	   
RAD51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)		Canis familiaris	   
rad51	DNA repair protein Rad51 homolog		Guillardia theta	   
rad51	DNA repair protein		Giardia lamblia ATCC 50803	   
RAD51L1	RAD51-like 1 (S. cerevisiae)		Homo sapiens	   
RAD51C	RAD51 homolog C (S. cerevisiae)		Homo sapiens	   
RAD51L3	RAD51-like 3 (S. cerevisiae)		Homo sapiens	   
RAD51AP1	RAD51 associated protein 1		Homo sapiens	   
Rad51I3	RAD51-like 3 (S. cerevisiae)		Mus musculus	   

Interaction information for this gene.

Find: dmc1 Next Previous Highlight all Match case Reached end of page, continued from top

javascript:if(window.name=="") { window.location.href='./nil'; } else { doaction('IHOP', 91604, 1); }



Information hyperlinked Over Proteins

Search Gene

Show overview

Find in this Page

Filter and options

Gene Model

Developer's Zone

Help

Concept & Implementation by Robert Hoffmann

WARNING: Please keep in mind that gene detection is done automatically and can exhibit a certain error. Read more about synonym ambiguity and the iHOP confidence value

Found in this page: interacts (21)

Find in this Page

(Click here to show sentences in the default order)

Sentences in this view contain interactions of RAD51 - Interaction Information is available whenever you see this symbol - Read more.

Show all Order by relevance

For a summary overview of the information in this page click here. new

The results provide insight into why mutation in just one of the eight BRC repeats would affect the way that BRCA2 protein interacts with the RAD51 filament. [2005]



The RAD51 recombinase interacts directly with the breast cancer-associated tumour suppressor BRCA2, an interaction that is required for normal recombination proficiency, radiation resistance and genome stability. [2003]



It is known that BRCA2 interacts directly with RAD51 through a series of degenerative motifs known as the BRC repeats. [2004]



The breast cancer susceptibility gene BRCA2 encodes a critical accessory protein that interacts with the RAD51 recombinase and this interaction fluctuates during the cell cycle. [2007]



BRCA2 protein interacts directly with the RAD51 recombinase and regulates recombination-mediated DSB repair, accounting for the high levels of spontaneous chromosomal aberrations seen in BRCA2-defective cells. [2007]



Whereas BRCA2 interacts directly with and regulates RAD51, the role of BRCA1 in this process is unclear. [2005]



RAD51 is an important component of double-stranded DNA-repair mechanisms that interacts with both BRCA1 and BRCA2. [2007]



Previous data report that wild-type p53 suppresses homologous recombination and physically interacts with Rad51. [2006]



p53 interacts with hRAD51 and hRAD54, and directly modulates homologous recombination. [2003]



Previously we reported that HsRad51 interacts with XRCC3, and Rad51C interacts with XRCC3, Rad51B, and HsRad51. [2003]



Find: dmc1

Next Previous Highlight all Match case Reached end of page, continued from top

Done

Text mining resources: PubGene

Network based on co-occurrence in article for the "RAD51" Gene and Proteins in Homo sapiens - Mozilla Firefox

File Edit View History Bookmarks Diigo Tools Help

http://www.pubgene.org/ biotext

Most Visited Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese... W Wikipedia Google Translate

Stumble! I like it! Send to Channels: All Favorites Friends Tools

PubGene - Find ... Chilbot: finding ... MedEvi - a perm... FABLE - Overview BioCreative Met... BioText Search ... RAD51 - Google... RAD51 - Google... Network base...

Beta PubGene
FAQ
Users Guide
Demo Tutorials
Feedback

Bio Networks Bio Associations Sequence Homology My Workspace Login

Browse literature or sequence neighbours.

Organism Homo sapiens Select one or "all organisms"

Gene / Protein * RAD51 Terms separated by comma, eg. *BRCA1, LEP*

Biological term and/or One term, eg. *Alzheimer Disease*. Supported categories
Network displays genes/proteins important for the biological term

Submit Advanced Options Network Parameters

Network based on co-occurrence in article for the "RAD51" Gene and Proteins in Homo sapiens

Send selected terms to Choose PubGene Tool Legend ?

Information about 'RAD51'

Key information Related terms

Property	Value
Term	RAD51
Name	RAD51 homolog (Re...
Organism	Homo sapiens
Article	4905
Term T...	PubGene Gene/Prot...
External	EG
Chrom...	
Alt. Sym...	BRCC5
Alt. Sym...	HRAD51
Alt. Sym...	HsRad51
Alt. Sym...	HsT16930
Alt. Sym...	RAD51A
Alt. Sym...	RECA

Network view Table view

The network graph displays RAD51 (red node) at the center, connected to several other genes/proteins (red nodes): POLG2, ATRX, POLA1, BRCA2, DDAH8, and ATP5E. Additionally, there are connections between BRCA2 and TP53, and between TP53 and STK32C. The nodes are arranged in a roughly circular pattern around the central RAD51 node.

Find: dmc1 Next Previous Highlight all Match case Reached end of page, continued from top

Applet pubgenegraph.GraphController started

Overview

1. A simple interaction record
2. Experimental methods
3. Representation of interactions
4. Interaction databases
5. Identifiers
6. Text mining sources
7. Predicted interaction sources
8. Consolidated databases
9. Visualization tools
10. Applications

Predicted interaction resources

STRING

<http://string.embl.de/>

OPHID/I2D

<http://ophid.utoronto.ca/ophidv2.201/>

FunCoup

<http://FunCoup.sbc.su.se>

Finding interaction information

The take home message:

There is no single place to collect interaction information.

Users are placed in a "go-fish" position and are responsible for distinguishing and sorting through multiple data sources.

There are a few resources that are focusing on providing an integrated data resource...

Overview

1. A simple interaction record
2. Experimental methods
3. Representation of interactions
4. Interaction databases
5. Identifiers
6. Text mining sources
7. Predicted interaction sources
8. Consolidated databases
9. Visualization tools
10. Applications

Consolidated databases

APID2NET

<http://bioinfow.dep.usal.es/apid/apid2net.html>

Pathway Commons

<http://www.pathwaycommons.org/pc/>

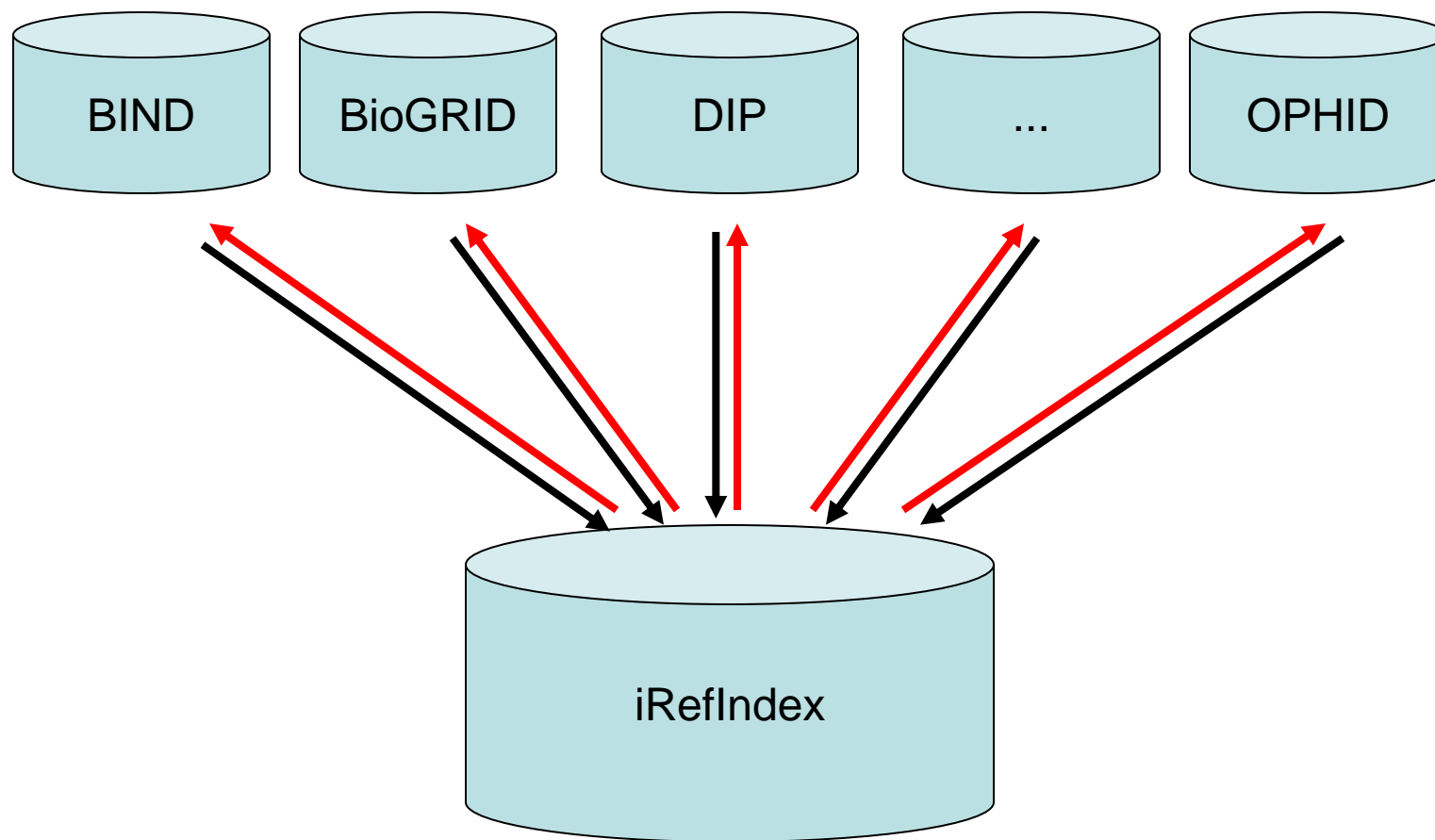
MiMI

<http://mimi.ncibi.org/MimiWeb/main-page.jsp>

iRefIndex

<http://irefindex.uio.no/wiki/iRefIndex>

iRefIndex: Integration and feedback



The problem



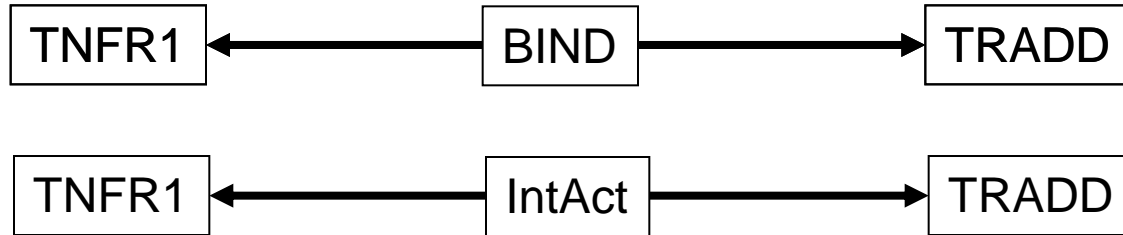
An interaction.

The problem



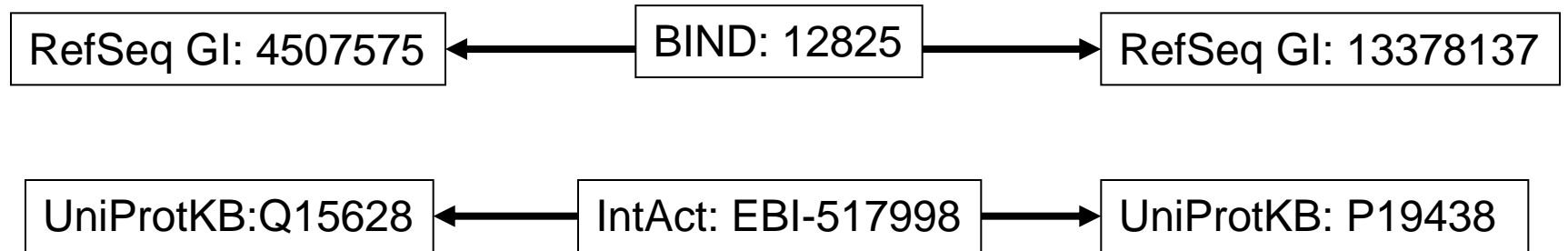
An interaction described in a database.

The problem



An interaction described in two databases.

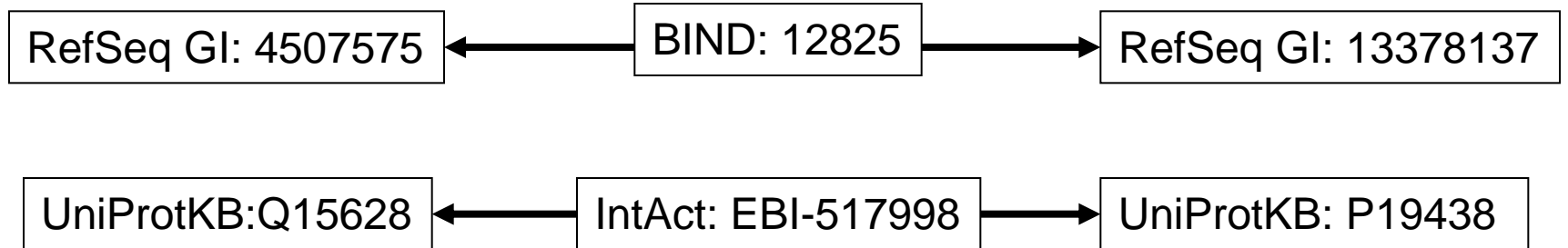
The problem



Different protein accessions are used to describe the same thing.

The solution

MTMDKSELVQ
KAKLAEQAER
YDDMAAAMKA
VTEQGHELNS
EERNLLSVAYK
NVVGARRSSW
RVISS.....



MTMDKSELVQ
KAKLAEQAER
YDDMAAAMKA
VTEQGHELNS
EERNLLSVAYK
NVVGARRSSW
RVISS.....

The solution

MTMDKSELVQ
KAKLAEQAER

SHA-1 digest

EERNLLSVAYK
NVVGARRSSW
RVISS.....

RefSeq GI: 4507575

BIND: 12825

RefSeq GI: 13378137

UniProtKB:Q15628

IntAct: EBI-517998

UniProtKB: P19438

MTMDKSELVQ
KAKLAEQAER
YDDMAAMKA
VTEQGHELNS
EERNLLSVAYK
NVVGARRSSW
RVISS.....

The solution

MTMDKSELVQ
KAKLAEQAER

waTMnn0uXSbf7bracJUscX9ckEc

EERNLLSVAYK
NVVGARRSSW
RVISS.....

RefSeq GI: 4507575

BIND: 12825

RefSeq GI: 13378137

UniProtKB:Q15628

IntAct: EBI-517998

UniProtKB: P19438

MTMDKSELVQ
KAKLAEQAER
YDDMAAMKA
VTEQGHELNS
EERNLLSVAYK
NVVGARRSSW
RVISS.....

The solution

MTMDKSELVQ
KAKLAEQAER

waTMnn0uXSbf7bracJUscX9ckEc

EERNLLSVAYK
NVVGARRSSW
RVISS.....

RefSeq GI: 4507575

BIND: 12825

RefSeq GI: 13378137

UniProtKB:Q15628

IntAct: EBI-517998

UniProtKB: P19438

MTMDKSELVQ
KAKLAEQAER
YDDMAAMKA

SHA-1 digest

EERNLLSVAYK
NVVGARRSSW
RVISS.....

The solution

MTMDKSELVQ
KAKLAEQAER

waTMnn0uXSbf7bracJUscX9ckEc

EERNLLSVAYK
NVVGARRSSW
RVISS.....

RefSeq GI: 4507575

BIND: 12825

RefSeq GI: 13378137

UniProtKB:Q15628

IntAct: EBI-517998

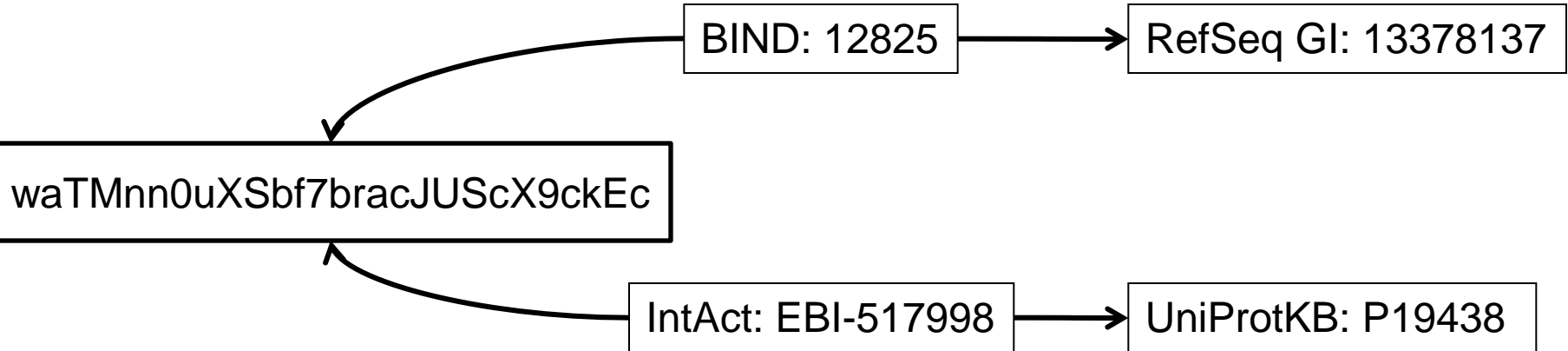
UniProtKB: P19438

MTMDKSELVQ
KAKLAEQAER
YDDMAAMKA

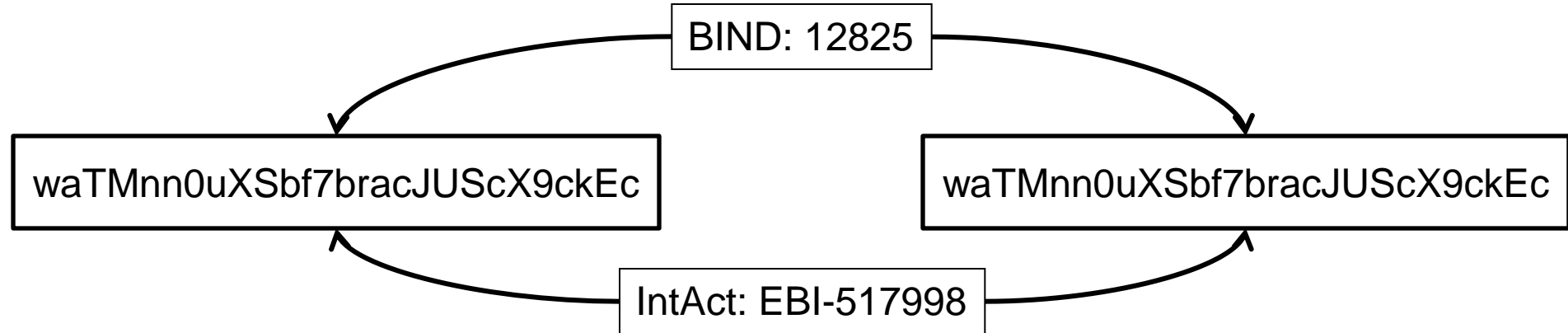
waTMnn0uXSbf7bracJUscX9ckEc

EERNLLSVAYK
NVVGARRSSW
RVISS.....

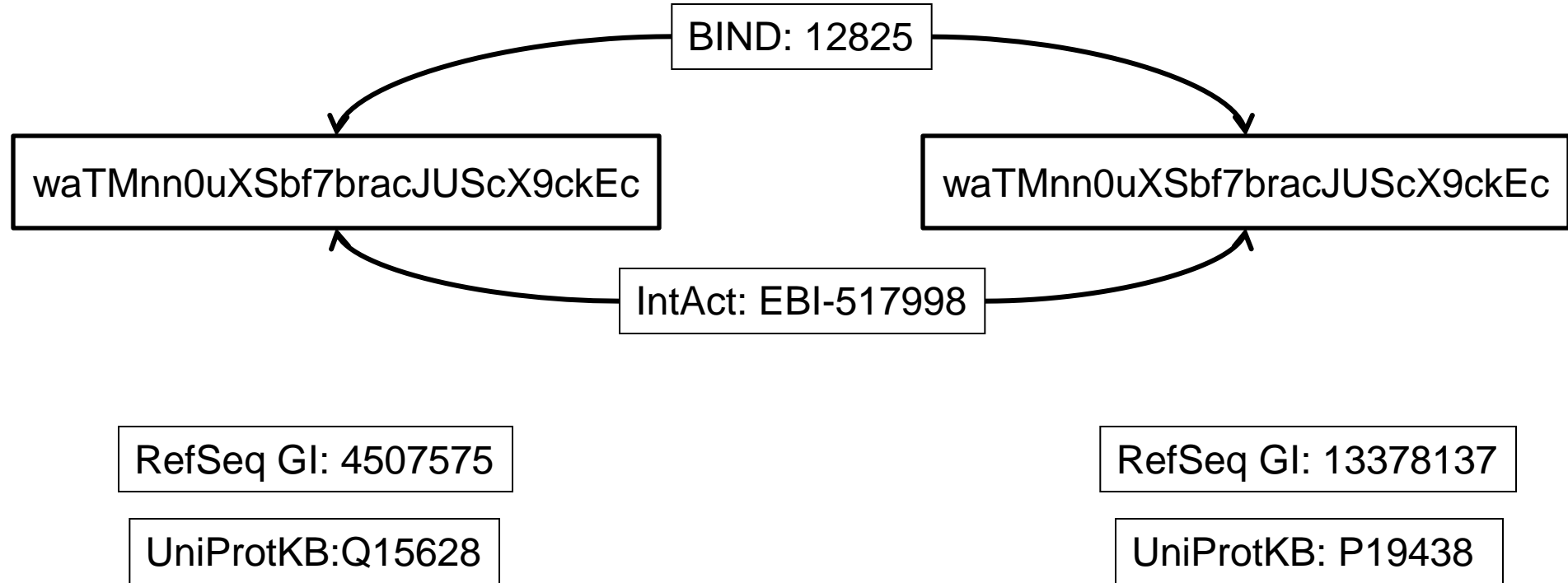
The solution



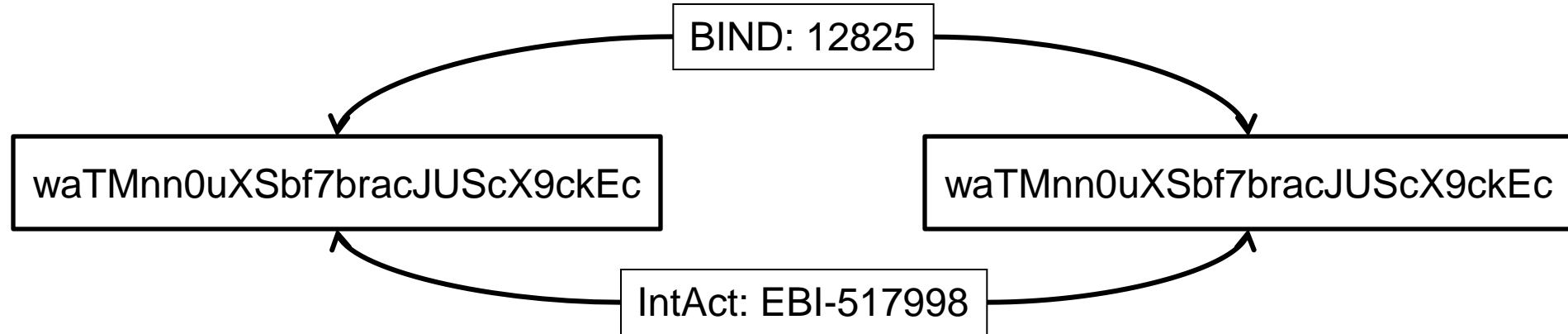
The solution



The solution

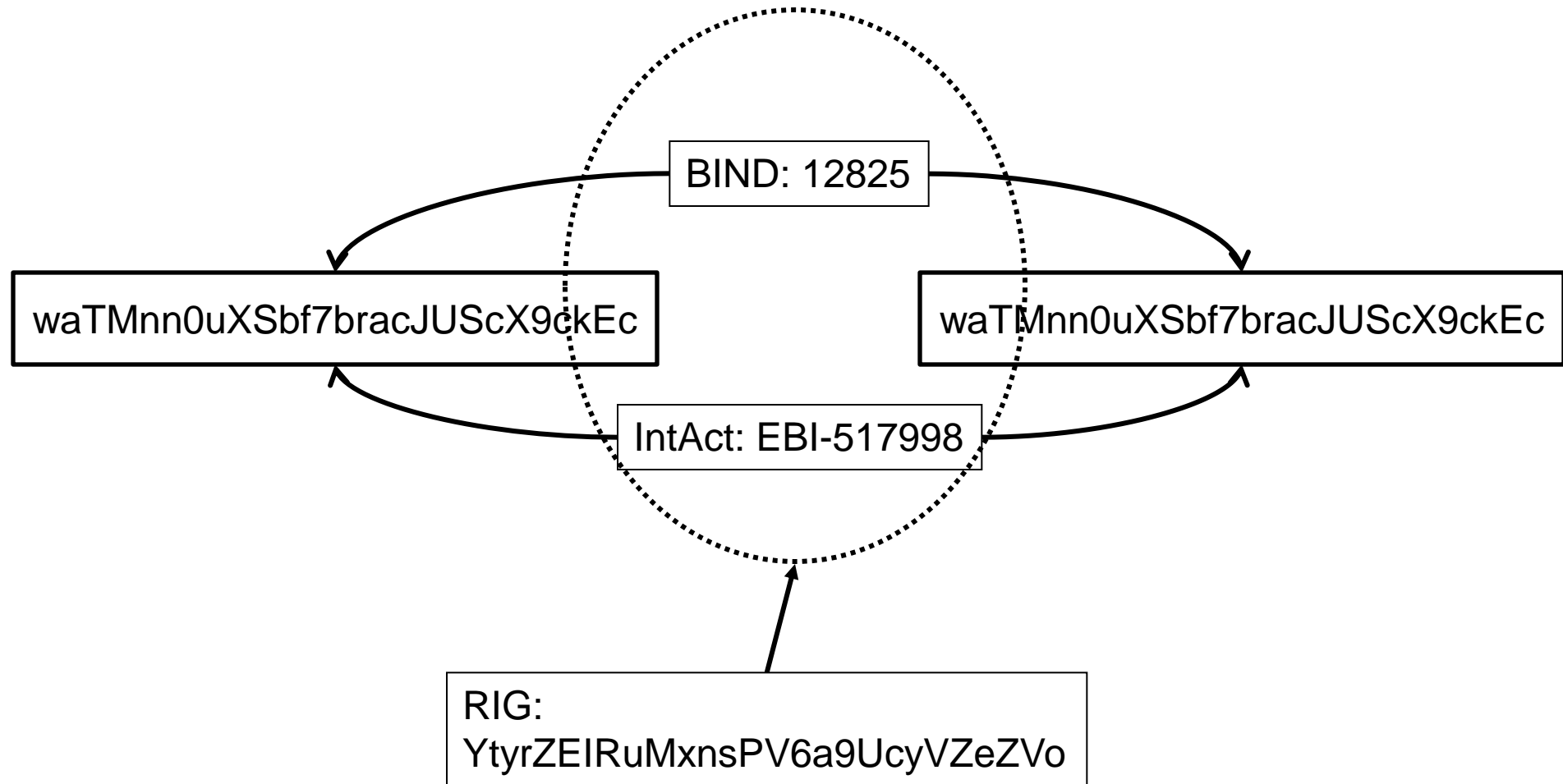


The solution

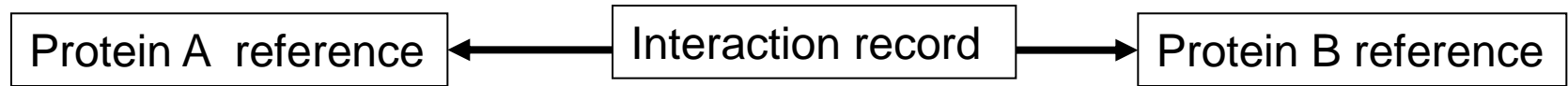


So the solution is based on **exact sequence matches** between interaction records

The solution



Hash keys allow consolidated of interaction data



Results: RIG redundancy between source databases.

Statistics iRefIndex 4.0 - Donaldson Group - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://irefindex.uio.no/wiki/Statistics_iRefIndex_4.0

Most Visited Entrez PubMed UIO Library of Medicine an... UIO The Biotechnology Ce... BBC BBC - bbc.co.uk home... Science/AAAS | Multim... http://bioscape.uio.n...

Start Stumbling... or Sign-in

Page Load Error Statistics iRefIndex 4.0 - Donalds...

Interactions (Corresponds to Table 6 in PMID 18823568) [\[edit\]](#)

BIND	62921										
BioGrid	20497	163891									
DIP	25914	28969	56441								
HPRD	2893	1958	839	37956							
INTACT	24239	25653	24807	8075	111235						
MINT	21991	34654	29988	6270	45260	76607					
MPACT	6904	8489	6777	0	6087	6426	13321				
MPPI	385	26	41	303	89	71	0	829			
OPHID	2210	1333	887	17913	7196	6396	0	183	47297		
CORUM	113	18	29	390	121	66	0	9	158	1919	
	BIND	BioGrid	DIP	HPRD	INTACT	MINT	MPACT	MPPI	OPHID	CORUM	
	(25903)	(111633)	(13594)	(15201)	(55807)	(15712)	(1137)	(238)	(26571)	(1403)	

Interactors [\[edit\]](#)

BIND	40801
BioGrid	14442
DIP	27474

start Skype... Inbox ... Yeast ... analys... C:\Do... irefind... irefind... Statist... Calcul... Search Desktop

Results: RIG redundancy between source databases.

Statistics iRefIndex 4.0 - Donaldson Group - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://irefindex.uio.no/wiki/Statistics_iRefIndex_4.0

Most Visited Entrez PubMed UIO Library of Medicine an... UIO The Biotechnology Ce... BBC BBC - bbc.co.uk home... Science/AAAS | Multim... http://bioscape.uio.n...

Start Stumbling... or Sign-in

Page Load Error

Statistics iRefIndex 4.0 - Donalds...

Interactions (Corresponds to Table 6 in PMID 18823568) [edit]

BIND 62921
BioGrid 20497 163891

All databases contribute some distinct interaction records.

	MPPI	OPHID	CORUM	BIND	BioGrid	DIP	HPRD	INTACT	MINT	MPACT	MPPI	OPHID	CORUM
	20	2210	113	62921	20497	163891	17913	7196	6396	0	183	47297	1919
	41	1333	18	(25903)	(111633)	(13594)	(15201)	(55807)	(15712)	(1137)	(238)	(26571)	(1403)

Interactors [edit]

BIND 40801
BioGrid 14442 27474

start Skype... Inbox ... Yeast ... analys... C:\Do... irefind... irefind... Statist... Calcul... Search Desktop

http://irefindex.uio.no/

iRefIndex - Donaldson Group - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://irefindex.uio.no/wiki/iRefIndex

Most Visited Entrez PubMed iRefIndex Library of Medicine an... iRefIndex The Biotechnology Ce... BBC - bbc.co.uk home... Science/AAAS | Multim... The Research Council ... Wikipedia BBC - Weather Centre...

iGoogle iRefIndex - Donaldson Group

lan.donaldson my talk my preferences my watchlist my contributions log out

page discussion edit history move unwatch

iRefIndex



Contents [hide]

- 1 Please take the biolibrarian survey
- 2 A reference index for protein interaction data
- 3 Long term goals of the iRefIndex project
- 4 Data availability via download
- 5 Data availability via web interface
- 6 Data availability via Cytoscape
- 7 Data availability via Web services
- 8 Feedback files
- 9 License
- 10 Disclaimer
- 11 Statistics
- 12 Credits
- 13 Citation
- 14 References
- 15 Contact
- 16 iRefIndex Google Group

Please take the biolibrarian survey [edit]

See http://irefindex.uio.no/wiki/The_Biolibrarian_Proposal

A reference index for protein interaction data [edit]

iRefIndex provides an index of protein interactions available in a number of primary interaction databases including [BIND](#), [BioGRID](#), [CORUM](#), [DIP](#), [HPRD](#), [IntAct](#), [MINT](#), [MPact](#), [MPPI](#) and [OPHID](#). This index allows the user to search for a protein and retrieve a non-redundant list of interactors for that protein.

iRefIndex uses the Sequence Global Unique Identifier ([SEGUID](#)) to group proteins and interactions into redundant groups. This method allows users to integrate their own data with the iRefIndex in a way that ensures proteins with the exact same sequence will be represented only once.

The iRefIndex paper has been published and is available [here](#).

Done

Finding interaction information

The take home message:

There is no single place to collect interaction information.

Users are placed in a "go-fish" position and are responsible for distinguishing and sorting through multiple data sources.

There are a few resources that are focusing on providing an integrated data resource...

Overview

1. A simple interaction record
2. Experimental methods
3. Representation of interactions
4. Interaction databases
5. Identifiers
6. Text mining sources
7. Predicted interaction sources
8. Consolidated databases
9. Visualization tools
10. Applications

Visualization Tools

Cytoscape

<http://www.cytoscape.org/>

VisAnt

<http://visant.bu.edu/>

There are over 30 visualization tools listed in a recent review

**Suderman M, Hallett M: Tools for visually exploring biological networks.
Bioinformatics 2007, 23(20):2651-2659.**

Cytoscape will be covered in a following talk.

Conclusions

Interaction data is an important bioinformatics data source.

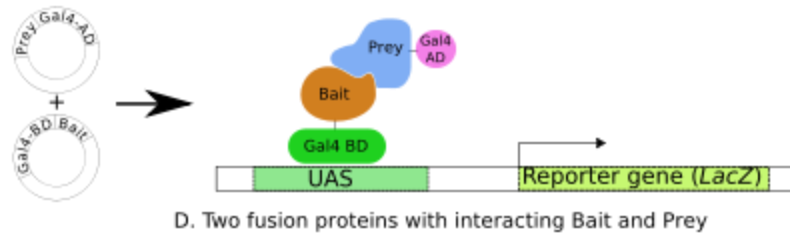
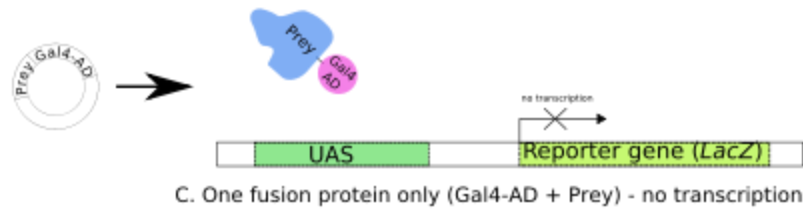
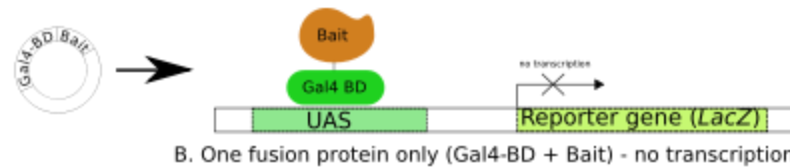
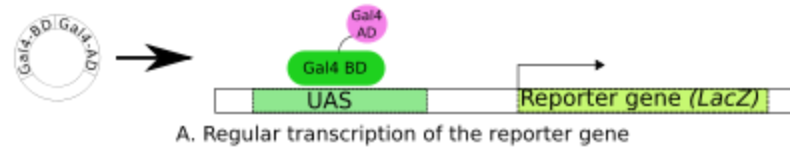
Interaction detection techniques have evolved past the simple Y2H and coimmunoprecipitation methods.

Finding data is still a non-trivial task.

There is no de facto source of data and users are faced with doing their own collection and curation.

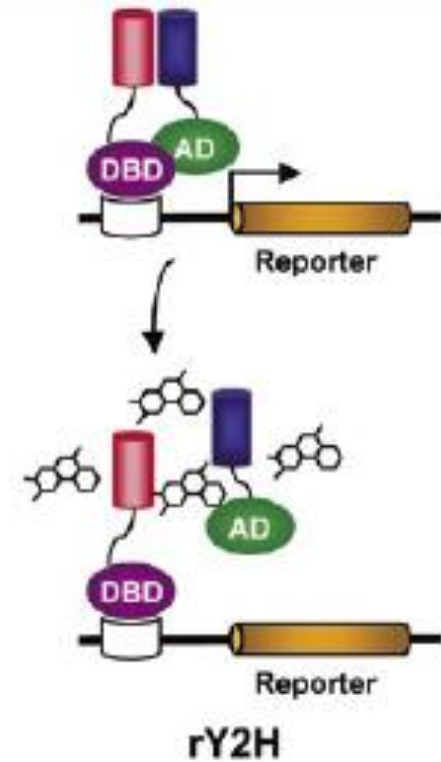
Data integration and accessibility is an ongoing area of research.

Experimental methods – Yeast Two Hybrid



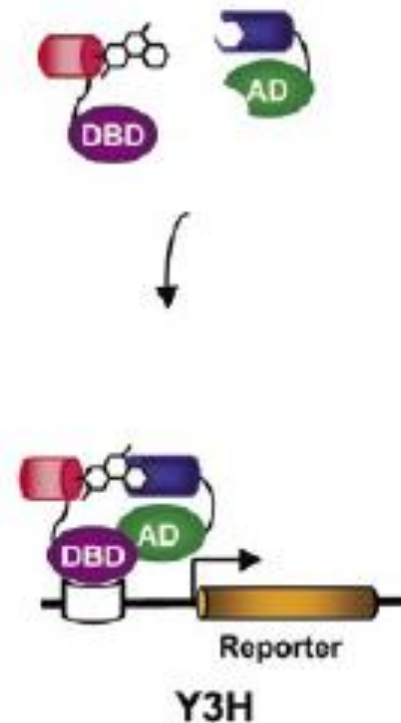
Experimental methods – reverse Two Hybrid

- screens for small molecules that can disrupt an interaction
- uses a counter-selective strategy



Experimental methods – Yeast 3 Hybrid

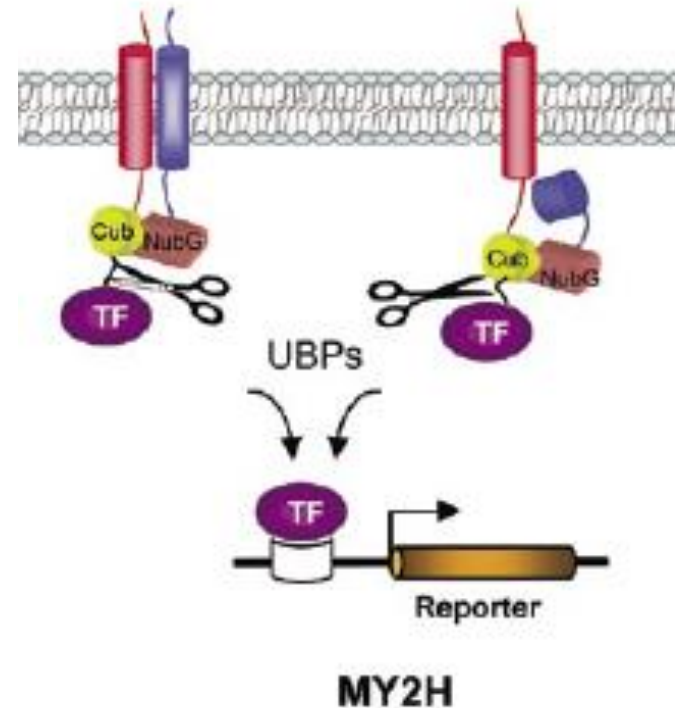
- screens for small molecules or other components that mediate an interaction



Experimental methods – membrane Yeast Two-Hybrid (mY2H)

- screens for proteins that interact within or near the membrane
- nearly one third of all proteins are integral membrane proteins not amenable to Y2H

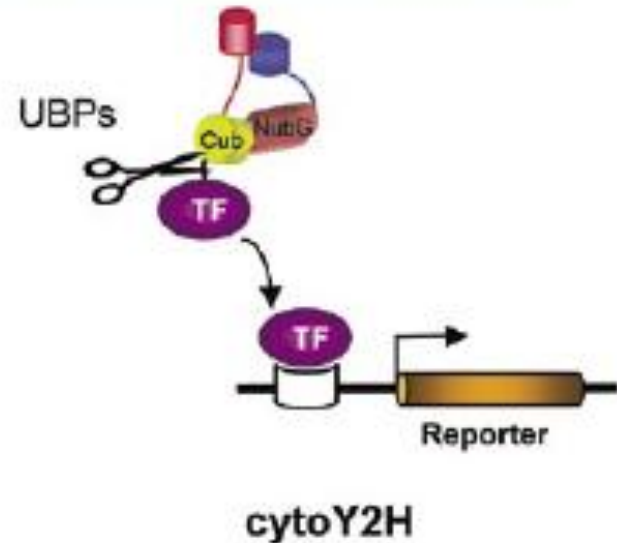
Cub: C-terminus of Ubiquitin
Nub: N-terminus of Ubiquitin



Experimental methods – cytosolic Yeast Two-Hybrid (cY2H)

- screens for proteins that interact within cytosol
- alternative for studying proteins in proper context
- alternative for studying proteins that may self activate in Y2H

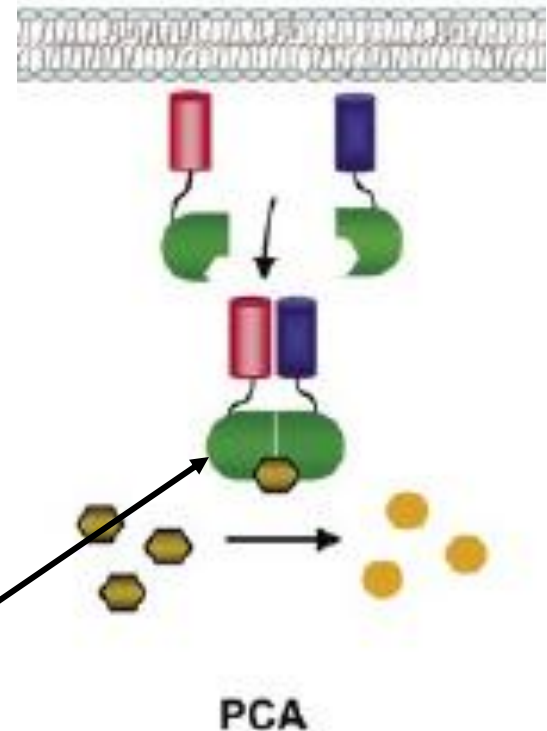
Cub: C-terminus of Ubiquitin
Nub: N-terminus of Ubiquitin



Experimental methods – PCA

Protein Complementation assay

- based on reconstitution of some protein activity
- fusion protein design must prevent spontaneous refolding
- independent of transcription based reporter
- alternatives include assays for luminescence or fluorescence
- useful in a variety of organisms
- other variants of this assay are split TEV protease assay

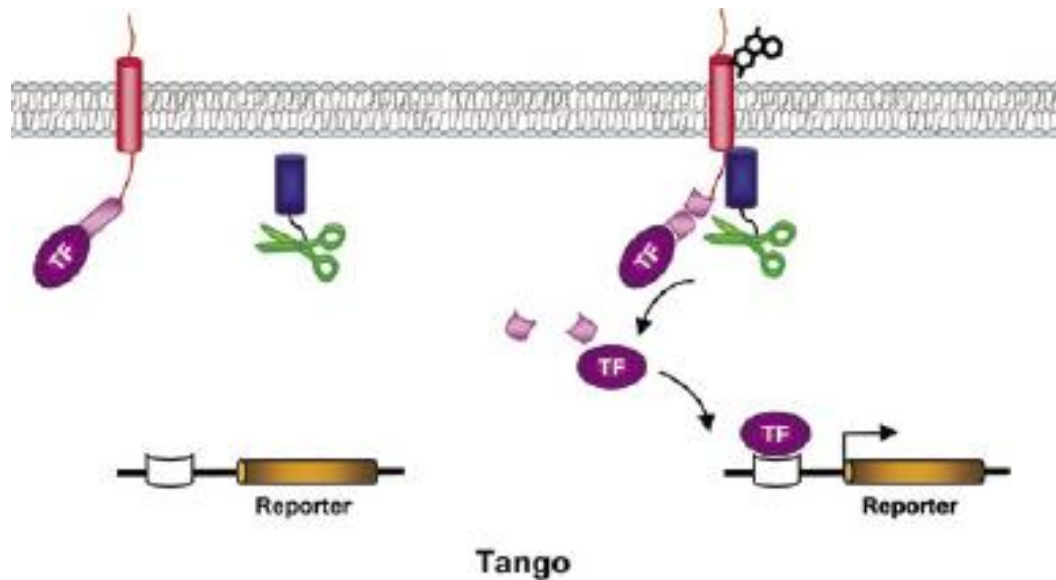


e.g. Reconstituted DHFR activity allows growth on trimethoprim

Experimental methods – Tango

State-specific binding assay

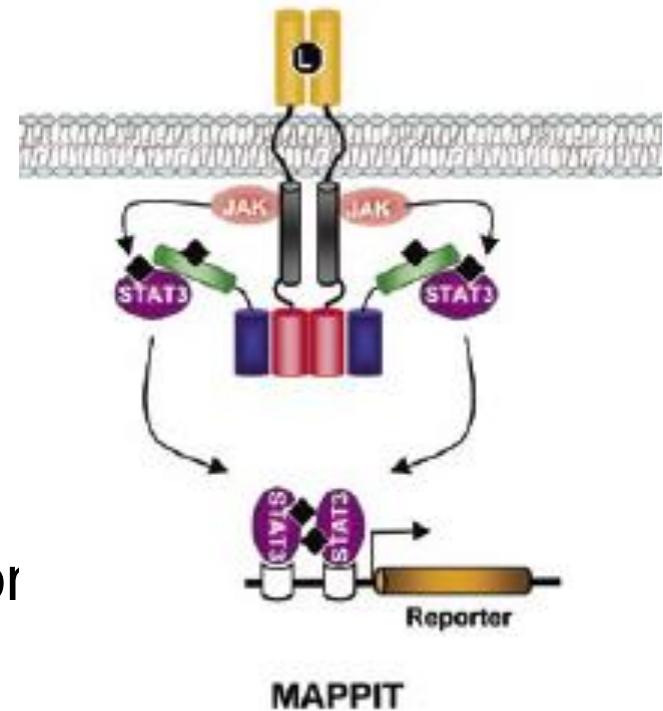
TEV protease is juxtaposed to reporter construct under specific conditions (presence of a ligand).



Experimental methods – MAPPIT

Mammalian protein interaction trap

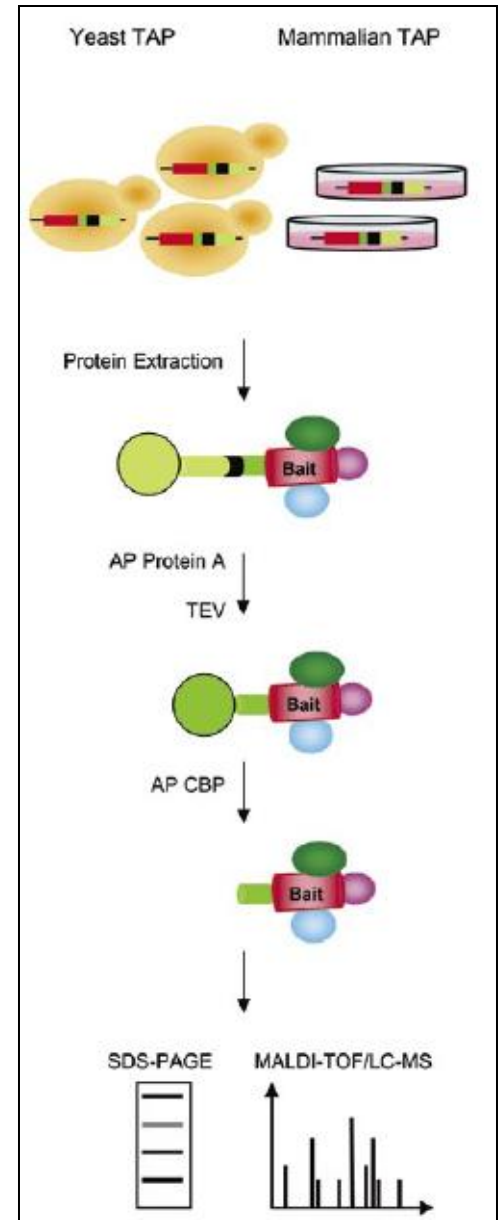
- interaction between bait and prey restores ligand-dependent cytokine receptor signalling
- especially applicable to analysis of cell signalling pathways due to native context of tested bait and prey
- rMAPPIT is used to screen for compounds that disrupt an interaction



Experimental methods

Affinity purification

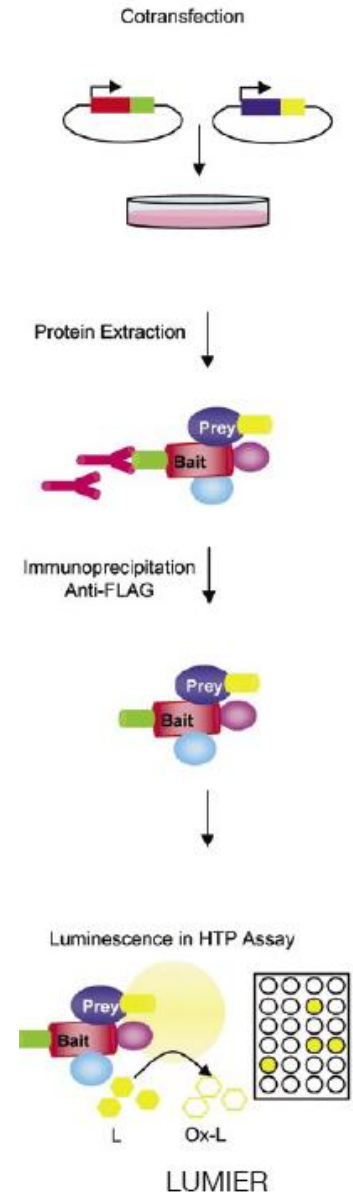
- interacting proteins are purified by means of an affinity tag on a "bait" protein
- proteins are detected via MS
- readout of this experiment type is a list of proteins that somehow interact with the bait protein
- note this is fundamentally different than readout of the Y2H family of experiments (more later)



Experimental methods – LUMIER

Luminescent mammalian interaction assay

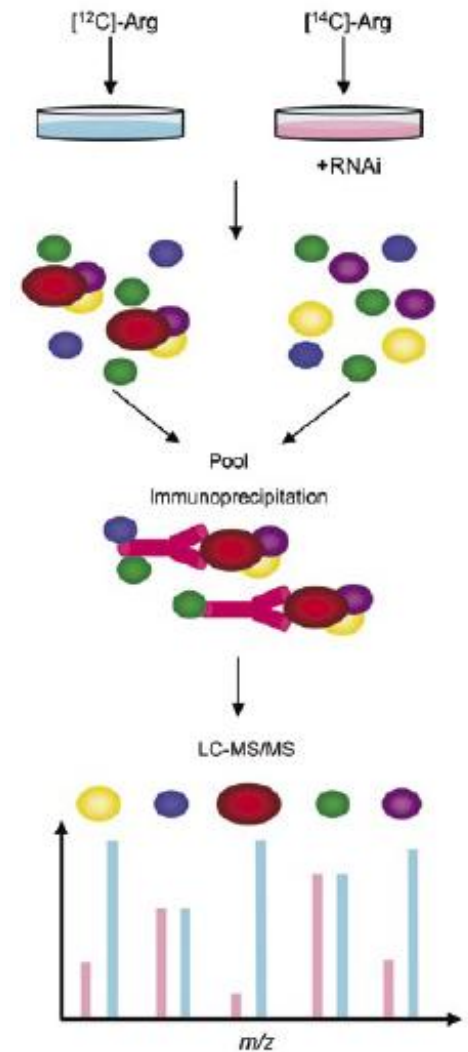
- bait is tagged with FLAG epitope
- prey fused to luciferase
- expression in mammalian cells
- pairwise screening in a high-throughput screen
- pre-tagging of prey aids in non-ambiguous identification



Experimental methods – QUICK

Quantitative IP combined with knockdown

- "bait" is knocked down in heavy-labelled sample
- requires antibodies specific to bait (not high throughput)
- reduces noises from non-specifically binding proteins found in IP sample
- no tags, represents native-like cell



Text mining resources: EBIMed

EBIMed - Mozilla Firefox

File Edit View History Bookmarks Diigo Tools Help

http://www.ebi.ac.uk/Rebholz-srv/ebimed/excerptResults.jsp?queryId=QueryDMYHMSms19102008213500843&key=P3: biotext

Most Visited Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese... Wikipedia

Diigo Bookmark Highlight Comment Send Message (0) Options

Stumble! I like it! Send to Channels: All Favorites Friends Tools

EBIMed FACTA PubGene - Find... Chilbot: finding... MedEvi - a per... FABLE - Overvi... BioCreative Met... BioText Search...

EMBL-EBI EB-eye Search All Databases Enter Text Here Reset Advanced Search Give us feedback

Databases Tools EBI Groups Training Industry About Us Help Site Index

EBIMed
Excerpts table

RAD51 Search
[Advanced Search](#) [Query Syntax](#)

[\[Summary page\]](#)

Abstracts per page
5 Update

first << 1/14 >> last

Rows 1 to 5 (out of 67)

Abstract	Sentences
17991895 <i>Palii Stela S et al. (2007)</i>	Further analysis revealed activation of key strand break repair proteins including ATM, ATR (ATM-Rad3-related), checkpoint kinase 1 (CHK1), BRCA1, NBS1, and RAD51 by Western blotting and immunofluorescence .
17999359 <i>Antoniou Antonis C et al. (2007)</i>	<p>RAD51 is an important component of double-stranded DNA-repair mechanisms that interacts with both BRCA1 and BRCA2 .</p> <p>A single-nucleotide polymorphism (SNP) in the 5' untranslated region (UTR) of RAD51, 135G->C, has been suggested as a possible modifier of breast cancer risk in BRCA1 and BRCA2 mutation carriers .</p> <p>RAD51 is the first gene to be reliably identified as a modifier of risk among BRCA1/2 mutation carriers .</p>
17001309	Several studies have revealed that the recombinational repair genes, RAD51 and BRCA1, and the DNA mismatch repair genes, MLH1 and MSH2, are decreased in

Find: dmc1 Next Previous Highlight all Match case Reached end of page, continued from top

Done

Text mining resources: FACTA

The screenshot shows a Mozilla Firefox browser window titled "RAD51 - FACTA Search - Mozilla Firefox". The address bar contains the URL: <http://text0.mib.man.ac.uk/software/facta/a.cgi?query=RAD51|111111|0|0|8765|0|10&cat=human&cat=disease&cat=sym>. The search results page displays the query "RAD51" and shows 1,329 document(s) hit in 17,702,258 MEDLINE articles (0.00 seconds). Three relevant concepts are listed:

- [Therapeutic exploitation of tumor cell defects in homologous recombination.](#)
In the decade since the **BRCA1** and **BRCA2** genes were cloned, much has been learned about the function of these two major causes of familial breast cancer. **BRCA2** has been shown to play a direct role in the repair of DNA by homologous recombination, by interacting with the **Rad51** protein and facilitating the formation of **Rad51** aggregates at the site of DNA damage. ... **BRCA1** plays a role in sensing DNA damage and replication stress and mediating the signaling responses. ... We have argued that the mechanism of cancer susceptibility from **BRCA1** or **BRCA2** deficiency is mediated via the defect in homologous recombination, since it is the main feature they share in common. ... Core biopsy samples are obtained and immediately irradiated ex vivo, allowing 3-4 hours for the appearance of **Rad51**, **BRCA1** and FancD2 foci. ... We have identified tumors with defects in the ability to form **Rad51** and **BRCA1** foci, where there is no known genetic predisposition, implying that this **BRCA**-dependent repair pathway may be inactivated in sporadic as well as familial breast cancers. ...
PMID:18473729 *Anti-cancer agents in medicinal chemistry* 2008 May
- [Dinucleotide repeat polymorphisms of **RAD51**, **BRCA1**, **BRCA2** gene regions in breast cancer.](#)
... **BRCA1** and **BRCA2**, two breast cancer susceptibility genes, are essential to maintain chromosomal integrity. This is mediated via regulation of **RAD51** during homologous recombination. Dinucleotide polymorphism repeats in the 15q14-21, 17q21 and 13q12-13 regions, where the **RAD51**, **BRCA1** and **BRCA2** genes are located, respectively, have been evaluated. ... Genotypes containing the (CA)(17) or (CA)(19) alleles in the **RAD51** region were found to be associated with a decreased breast cancer risk. ... The results indicate that dinucleotide CA repeat polymorphism at **RAD51** and **BRCA2** gene regions might be associated with genetic susceptibility to breast cancer.
PMID:18429825 *Pathol. Int.* 2008 May
- [DNA methylation inhibitor 5-Aza-2'-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B.](#)
... Further analysis revealed activation of key strand break repair proteins including ATM, ATR (ATM-Rad3-related), checkpoint kinase 1 (CHK1), **BRCA1**, NBS1, and **RAD51** by Western blotting and

The browser window also shows a search bar with "RAD51" entered, buttons for "Find Relevant Concepts" and "Search MEDLINE", and a list of filters: Gene/Protein, Disease, Symptom, Drug, Enzyme, Compound, All, and Clear. The status bar at the bottom shows "Find: dmc1" and "Reached end of page, continued from top".

Text mining resources: Chillibot

Chilibot: finding gene and protein relationships from MEDLINE - a text mining approach - Mozilla Firefox

File Edit View History Bookmarks Digo Tools Help

http://www.chilibot.net/ biotext

Most Visted Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese... W Wikipedia Google Translate

Digo Bookmark Highlight Comment Send Message (0) Options

Stumble! I like it! Send to Channels: All Favorites Friends Tools

Chilibot: finding ... MedEvi - a permut... FABLE - Overview BioCreative MetaSer... BioText Search Eng... RAD51 - Google Se... RAD51 - Google Se... Network based on ...

Chilibot

- New Search
- Saved Results
- Log In/Out
- New Account
- Documentation
- Examples
- Site News
- Download
- Contact Us

Last update: 08/04/07

We recommend FireFox browser.

RAD51 & BRCA2

Retrieving synonyms..

Retrieving abstracts .. [RAD51/BRCA2

Performing linguistic analysis ..

Analyzed 100 *most recent* abstracts out of 207 available.
Found 103 interactive sentences and 70 parallel sentences.

[View all relevant sentences.](#)

Interactive relationship (e.g. stimulation, inhibition, etc)

- These findings indicate that **BRCA2** inhibits **Rad51** cleavage and subsequent apoptosis. [Ref. Radiat Res. 2008](#)
- Similarly, reducing **BRCA2** by 50% reduces **RAD51** and BCCIP foci. [Ref. Mol Cell Biol. 2005](#)
- **BRCA2** expression also modulates **Rad51** protein levels such that **Rad51** protein is relatively decreased in **BRCA2** defective cancer cells. [Ref. Mol Carcinog. 2008](#)
- **BRCA2** is essential for homologous recombination in vivo, but isolated BRC repeat peptides can prevent the assembly of **RAD51** into active nucleoprotein filaments in vitro, suggesting a model in which **BRCA2** sequesters **RAD51** in undamaged cells, and promotes recombinase function after DNA damage. [Ref. Nucleic Acids Res. 2006](#)
- The interactions of **BRCA2** with **RAD51** and DMC1 lead us to suggest that the **BRCA2** tumour suppressor is a universal regulator of recombinase actions. [Ref. Oncogene. 2007](#)
- The **BRCA2** tumour suppressor regulates the RAD 51 [RAD51] recombinase during double strand break (DSB) repair by homologous recombination (HR) but how **BRCA2** executes its functions is not well understood. [Ref. J Mol Biol. 2006](#)
- The crystallographic analysis shows that the BRC motif of the **BRCA2** tumor suppressor is in contact with the subunit subunit interface of **Rad51** and could thus prevent filament formation of **Rad51**. [Ref. Genes Cells. 2008](#)
- From these results, we conclude that (i) **BRCA2** regulates **RAD51** recombination in response to the type of DNA damage and (ii) **BRCA2** suppresses SCRS, suggesting a role for **BRCA2** in sister chromatids cohesion and or alignment. [Ref. Cancer Res. 2005](#)
- Inhibition of filament formation of human **Rad51** protein by a small peptide derived from the BRC motif of the **BRCA2** protein. [Ref. Genes Cells. 2008](#)
- Brh2, the ortholog of the **BRCA2** tumor suppressor in *Ustilago maydis*, works hand in hand with **Rad51** to promote repair of DNA by homologous recombination. [Ref. Proc Natl Acad Sci U S A. 2008](#)
- The **BRCA2** tumor suppressor functions in repair of DNA by homologous recombination through regulating the action of **Rad51**. [Ref. DNA Repair \(Amst\). 2007](#)

Find: dmc1 Next Previous Highlight all Match case Reached end of page, continued from top

Done

Text mining resources: Chillibot

Chilibot: finding gene and protein relationships from MEDLINE - a text mining approach - Mozilla Firefox

File Edit View History Bookmarks Digo Tools Help

http://www.chilibot.net/

Most Visited Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese... W Wikipedia Google Translate

Stumble! I like it! Send to Channels: All Favorites Friends Tools

Chilibot: finding ... MedEvi - a permute... FABLE - Overview BioCreative MetaSer... BioText Search Eng... RAD51 - Google Se... RAD51 - Google Se... Network based on ...

Chilibot

- New Search
- Saved Results
- Log In/Out
- New Account
- Documentation
- Examples
- Site News
- Download
- Contact Us

Last update: 08/04/07

We recommend Firefox browser.

Two lists search, context restriction

Context: substance abuse OR addiction

```
graph TD; P[POLYMORPHISM] --- 8[8] --- TH[TH]; P --- 5[5] --- SCT[SCT]; P --- 6[6] --- GABRG1[GABRG1]; P --- 5[5] --- GABRB1[GABRB1]; P --- 5[5] --- GABRA4[GABRA4]; P --- 29[29] --- GABRA2[GABRA2]; P --- 5[5] --- DRD4[DRD4]; P --- 21[21] --- CCKBR[CCKBR]; P --- BDNF[BDNF]; P --- ADH5[ADH5];
```

Searching for polymorphism of genes associated with substance abuse or addiction. A list of 27 genes was first identified in a data mining study. This Chilibot search found that the polymorphism of 10 of them were known to be associated with substance abuse/addiction.

Hypothesis generation

```
graph TD; CREB[CREB] --- 10[10] --- ACTIN[ACTIN]; CREB --- 23[23] --- CAMKII[CAMKII]; CREB --- 14[14] --- ZIF268[ZIF268]; CREB --- 14[14] --- TAU[TAU]; CREB --- 30[30] --- AMPA[AMPA]; CREB --- 30[30] --- NMDA[NMDA]; CREB --- 13[13] --- PKA[PKA]; CREB --- 5[5] --- PKC[PKC]; CREB --- 13[13] --- TRKB[TRKB]; CREB --- 5[5] --- TRKA[TRKA]; CREB --- 13[13] --- PLC[PLC]; CREB --- 5[5] --- SYNAPSIN_I[SYNAPSIN I]; CREB --- 13[13] --- LTP[LTP]; SYNAPTOTAGMIN[SYNAPTOTAGMIN] --- 13[13] --- ACTIN[ACTIN]; SYNAPTOTAGMIN --- 5[5] --- CAMKII[CAMKII]; SYNAPTOTAGMIN --- 14[14] --- ZIF268[ZIF268]; SYNAPTOTAGMIN --- 14[14] --- TAU[TAU]; SYNAPTOTAGMIN --- 6[6] --- AMPA[AMPA]; SYNAPTOTAGMIN --- 14[14] --- NMDA[NMDA]; SYNAPTOTAGMIN --- 14[14] --- PKA[PKA]; SYNAPTOTAGMIN --- 2[2] --- PKC[PKC]; SYNAPTOTAGMIN --- 14[14] --- TRKB[TRKB]; SYNAPTOTAGMIN --- 14[14] --- TRKA[TRKA]; SYNAPTOTAGMIN --- 2[2] --- PLC[PLC]; SYNAPTOTAGMIN --- 14[14] --- SYNAPSIN_I[SYNAPSIN I]; SYNAPTOTAGMIN --- 2[2] --- LTP[LTP];
```

From a complex of relationship network, Chilibot suggested that CREB may have a relationship (e.g. regulation) with synaptotagmin, based on 1) there is no known relationship between CREB and synaptotagmin; and 2) they are both connected to many gene/protein and are both involved in similar biological process (e.g. LTP). The original search included one keyword ("LTP") and 22 other genes. CREB was first selected by clicking on the graph. Hypothesis was generated by clicking on the "CREB might be related to.." button. See [here](#) for details.

Find: dmc1

Next Previous Highlight all Match case Reached end of page, continued from top

Done

Text mining resources: MedEvi

MedEvi - a permuted concordancer for Medline - Mozilla Firefox

File Edit View History Bookmarks Diigo Tools Help

http://www.ebi.ac.uk/tc-test/textmining/medevi/results.jsp?query=BRCA1+AND+bind*+AND+RAD51&options=side+AND+ef

Most Visited Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese... Wikipedia


Diigo Bookmark Highlight Comment Send Message (0) Options

Stumble! I like it! Send to Channels: All Favorites Friends Tools

MedEvi - a perm... FABLE - Overview BioCreative MetaSer... BioText Search Eng... RAD51 - Google Se... RAD51 - Google Se... Network based on ...

EMBL-EBI EB-eye Search All Databases Enter Text Here Reset Advanced Search Give us feedback

Databases Tools EBI Groups Training Industry About Us Help Site Index



[Show options](#) [HELP?](#)

Query: BRCA1 AND bind* AND RAD51

3 permuted occurrences of 25 filtered documents of 25 hit documents (4.59s). [List keywords](#) [List multiword terms](#)

Jump to result set: [ABC](#) [BCA](#)

Set 1: BRCA1 AND bind* AND RAD51 [Top](#)

[1](#) radiation (IR) and exhibit [defective DNA repair](#) . Both BRCA1 and [BRCA2](#) have been reported to bind Rad51 , [a protein](#) essential for homologous recombination

[2](#) sensitivity to DNA-damaging agents . In human cells , BRCA1 binds to both [Rad50](#) and Rad51 and colocalizes with these proteins at repair foci .

Set 2: bind* AND RAD51 AND BRCA1 [Top](#)

[1](#) the zinc finger domain (exons 2 , 3 and 5) and the binding domain with Rad51 (exon 11) of the BRCA1 protein . An identical nonsense mutation at codon 63

[Terms of Use](#) : [EBI Funding](#) : [Contact EBI](#) : © European Bioinformatics Institute 2006-2008. EBI is an Outstation of the [European Molecular Biology Laboratory](#).

Find: dmc1 Match case

Done

Text mining resources: FABLE

FABLE - Gene Lister query [brca1 AND rad51] - Mozilla Firefox

Fast Automated Biomedical Literature Extraction

Why use FABLE?
FABLE Help

Gene Lister Article Finder

brca1 rad51 Synonyms: Sort by: Frequency Results/page: 25

Next > Last > | 1-25 of 186 unique genes found in articles containing *brca1* AND *rad51* | Download:

Gene	Articles	Synonyms	LitTrack	Links
RAD51	163	▷ RAD51	15q15.1	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
BRCA1	150	▷ BRCA1	17q21.31	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
BRCA2	95	▷ BRCA2	13q13.1	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
TP53	34	▷ TP53	17p13.1	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
NBN	24	▷ NBN	8q21.3	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
ATM	21	▷ ATM	11q22.3	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
NLRP2	20	▷ NLRP2	19q13.42	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
FANCD2	19	▷ FANCD2	3p25.3	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
FANCA	17	▷ FANCA	16q24.3	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
RAD50	17	▷ RAD50	5q31.1	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
FXN	16	▷ FXN	9q21.11	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
MRE11A	16	▷ MRE11A	11q21	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
RAD52	16	▷ RAD52	12p13.33	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
BARD1	15	▷ BARD1	2q35	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
ATRX	12	▷ ATRX	Xq21.1	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
UBE2S	10	▷ UBE2S	19q13.42	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
CHEK2	9	▷ CHEK2	22q12.1	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
ATR	8	▷ ATR	3q23	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx
XRCC3	8	▷ XRCC3	14q32.33	NCBI Gene UCSC SvmAtlas Google GeneCards GeneLynx

Text mining resources: BioText

BioText Search Engine - Mozilla Firefox

File Edit View History Bookmarks Digo Tools Help

http://biosearch.berkeley.edu/index.php?q=RAD51+AND+BRCA2&submit=Search&view=list&sortedby=rel&action=submit_search

Most Visited Entrez PubMed Post to CiteULike Library of Medicine and ... The Biotechnology Cen... BBC - bbc.co.uk home... Home page - The Rese... W Wikipedia

Digo Bookmark Highlight Comment Send Message (0) Options

Stumble! I like it! Send to Channels: All Favorites Friends Tools

BioText SEARCH ENGINE [Home](#) | [About BioText](#) | [Contact Us](#)

RAD51 AND BRCA2

Search Over: Full Text & Abstracts Figure Captions (List) Figure Captions (Grid) Tables Sort By: Relevance

Results 1-1 of 1

[View Fig. & Caption](#)

γ -H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin

Kinner, A., Wu, W., Staudt, C., Iliakis, G. (2008) *Nucleic Acids Research*.

FIGURE CAPTION

Figure 2.. DSB repair pathways. (A) Homologous recombination repair (HRR). After the initial sensing of the DSB by MRN and the activation of ATM, H2AX is phosphorylated, which in turn elicits a sequence of signaling events thought to ultimately cause the activation of nucleases such as Mre11 and CtIP to process the DNA ends and generate ssDNA with 3' overhangs. ssDNA is bound by RPA, which is subsequently exchanged by **Rad51** and **Rad51** paralogs. This exchange is facilitated by Rad52, Rad54 and **BRCA2**. The Rad51-decorated DNA fiber initiates strand invasion into an intact homologous DNA molecule that leads to the formation of a Holiday junction. The DNA sequence around the DSB is copied by DNA synthesis associated with branch migration, and the process is completed by resolution of the Holiday junction. HRR is a templated repair process and is therefore error free (15–17). (B) DNA-PK-dependent nonhomologous end joining (D-NHEJ). DNA ends are recognized by Ku, which recruits, after processing by Tdp1 or PNKP, DNA-PKcs. Upon end-binding, DNA-PKcs is activated and phosphorylates itself and possibly also other proteins (like H2AX on an adjacent nucleosome). Phosphorylated DNA-PKcs is thought to be released from the DNA end, which allows the DNA ligase IV/XRCC4/XLF complex to mediate end-ligation possibly with the help of a DNA polymerase that catalyzes gap filling (19,149,150). (C) Backup pathway of nonhomologous end joining (B-NHEJ). There is evidence that cells of higher eukaryotes with defects in D-NHEJ rejoin the majority of DSBs using an alternative repair pathway that is not utilizing any of the HRR-associated activities (19). This pathway is therefore termed backup NHEJ (B-NHEJ). Although details of this pathway remain to be elucidated, there is evidence that it utilizes the PARP-1/DNA Ligase III/XRCC1 repair module known to be involved in the repair of SSB and base damages (151–155), and that its function is facilitated by the linker histone H1 (156). (D) Single strand annealing (SSA). This repair pathway shares features of HRR and NHEJ, and is best described in yeast (17). After the initial sensing of the DSBs