iRefIndex and iRefWeb: Providing feedback to protein interaction databases.

http://irefindex.uio.no



Ian Donaldson Biotechnology Centre of Oslo University of Oslo

BioPathways SIG ISMB 2009

Acknowledgements

Ian Donaldson Sabry Razick George Magklaras Paul Boddie Katerina Michalickova Antonio Mora

<u>Shoshana Wodak</u> Brian Turner Emerson Cho Kyle Morrison Andrei Turinsky

<u>http://irefindex.uio.no</u> http://donaldson.uio.no

ian.donaldson@biotek.uio.no

http://wodaklab/iRefWeb

brian@wodaklab.org

This talk

1.Overview of what iRefIndex does and is2.Feedback files3.Cytoscape plugin4.iRefWeb5.Comparison of co-sited papers

iRefIndex: Integration and feedback





An interaction.



An interaction described in a database.



An interaction described in two databases.



Different protein accessions are used to describe the same thing.

MTMDKSELVQ KAKLAEQAER YDDMAAAMKA VTEQGHELSN EERNLLSVAYK NVVGARRSSW RVISS.



MTMDKSELVQ KAKLAEQAER YDDMAAAMKA VTEQGHELSN EERNLLSVAYK NVVGARRSSW RVISS.

















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













The solution: integration of data sets



The solution: with provenance and feedback



The solution: with provenance and feedback



BIND **BioGRID** CORUM DIP **HPRD** IntAct MINT **MPact MPPI** OPHID

All of these are distributed in the PSI-MI XML format

The problem and why it's important



2/3 of A interactors are gene-products associated with variations in some disease. But if B and C are the same thing as A then ...

The problem and why it's important



4/5 of A interactors are gene-products associated with variations in some disease



PSI-MI XML record

Interactor A

Primary ref:

Secondary refs:

Taxon identifier:

Sequence:

Interactor A Primary ref:

Secondary refs:

Taxon identifier:

Sequence:

PSI-MI XML record

Interactor A Primary ref: RefSeq:4507575

Secondary refs: UniProtKB:Q15628 database:accession database:accession

Taxon identifier: 9606

Sequence: MAGTRVY....

Intera Prima	actor B ary ref:	
Seco	ndary refs:	
Taxor	identifier:	
Sequ	ence:	

PSI-MI XML record

Interactor A Primary ref: RefSeq:4507575	Db	Acc	Taxid	Seguid
Secondary refs: UniProtKB:Q15628 database:accession	RefSeq	4507575	9606	waTM9606
database:accession				
Taxon identifier: 9606 Sequence: M			I	





Implementation details: assignment score


Implementation details: assignment score

Human readable version : P

• **P** = Primary accession was used

Each binary flag keeps track of some kind of operation used to perform the mapping

Implementation details: assignment score

Binary value : 10000010000100010001

Human readable score : PTUO+

- P = Primary accession was used
- **T** = Taxonomy identifier was not matched
- **U** = Updating performed
- O = Only one matched the SEGUID of original protein
- + = There were multiple possibilities

Implementation details: workflow



Implementation details: workflow



Assignment scores are associated with a path through this workflow



Results: assignments

Protein	720,205
interactors	
Unambiguous	688,190 (95%)
assignments	
Problems	32,515 (5%)
Unique proteins	83,388
(ROGs)	

Release 4.0

Results: assignments by score type

		Examples		
Score type	Total assignments with this score type	ROG Assignment Score	Number of cases	Details for one example
1	598590	Р	512650	UniProt:Q15118 is cited in the interaction record as the primary reference (P).
	(77.43)	SM	2154	Accession NP 191913 is cited in a modified form (M) without the underscore.
	24664	PU	18542	UniProt:O95686 is cited and updated (U) to UniProt/KB:Q9UQK1.
2	(3.19)	PUO+	6	UniProt:P38706 is cited. Two possible updates are possible (+) but only one matches the original (O) sequence in the interaction record (P0C2H6).
	121540	PT	52074	Protein reference cites taxon id as 9534 (African green monkey) but the sequence record cites taxon 9606.
3 (15.	(15.72)	ST	60205	Protein reference cites taxon id as 40674 (mammalia) but the sequence record cites (9606) human.
4	2803 (0.36)	PUTO+	15	UniProt:O04063 is cited with taxon identifier 4530 (rice). More than one updated accession exists (+U). Only one possibility has the same sequence as cited in the interaction record (P0C5B0) with taxon identifier 39947 (a specific strain of rice).
5	9840	SL+	9090	The primary reference cited is not found. 49 secondary references are cited (S). 15 of these were found to map to 8 distinct proteins (+). The protein with the largest (L) SEGUID is arbitrarily chosen.
	(1.27)	PTQ	21	Primary accession P84244 cited as a "see also" (Q) reference with taxon id 9606. The sequence record cites taxon id 10090 (T).
		PN	8909	Q95Q01 is an obsolete accession. The sequence is retrieved from the interaction record. The SEGUID and ROGID are calculated and stored locally as a new entry (N).
6	15649	SEN	5561	RefSeq:NP_010441 is an obsolete accession. The sequence is retrieved using eUtils (E). The SEGUID and ROGID are calculated and stored locally as a new entry (N).
б	(2.02)	STGOEN+	2	EntrezGene 196549 (G) is cited and encodes two proteins (+). The protein accessions cited by EntrezGene are retired. Sequences are retrieved using eUtils (E). One matches the sequence cited in the interaction record (O). The SEGUID and ROGID are calculated and stored locally as a new entry (N).

Results: assignments by score type

		Examples	xamples					
Score type	Total assignments with this score type	ROG Assignment Score	Number of cases	Details for one example				
1	598590 P 512650 UniProt:Q15118 is cited in the interaction record as the primary reference (P).							
	(77.43)	SM	2154	Accession NP 191913 is cited in a modified form (M) without the underscore.				
	24664	PU	18542	UniProt:O95686 is cited and updated (U) to UniProt/KB:Q9UQK1.				
2	2 (3.19) PUO+ 6		6	UniProt:P38706 is cited. Two possible updates are possible (+) but only one matches original (O) sequence in the interaction record (P0C2H6).	the			
	PT 52074 Protein reference cites taxon id as 9534 (African green monkey) but the sequence record cites taxon 9606.							
	IniProt:00	4063 is c	ited witl	n taxon identifier 4530 (rice).	9606)			
►	lore than	one upda	ted acc	ession exists (+U).	ssion			
4 C	Only one p	ossibility	has the	same sequence as cited in the interaction	cord			
re	ecord (P00	C5B0) wit	th taxon	identifier 39947 (a specific strain of rice).	these			
5	9840	SL+	9090	arbitrarily chosen.	S			
	(1.27)	PTQ	21	Primary accession P84244 cited as a "see also" (Q) reference with taxon id 9606. The sequence record cites taxon id 10090 (T).				
		PN	8909	Q95Q01 is an obsolete accession. The sequence is retrieved from the interaction record SEGUID and ROGID are calculated and stored locally as a new entry (N).	ord. The			
6	15649	SEN	5561	RefSeq:NP_010441 is an obsolete accession. The sequence is retrieved using eUtils SEGUID and ROGID are calculated and stored locally as a new entry (N).	(E). The			
	(2.02)	STGOEN+	2	EntrezGene 196549 (G) is cited and encodes two proteins (+). The protein accessions EntrezGene are retired. Sequences are retrieved using eUtils (E). One matches the se cited in the interaction record (O). The SEGUID and ROGID are calculated and stored as a new entry (N).	cited by equence d locally			

Results: interactions (RIGs) assignments

PPI's	761,587
Assigned to RIGs (any score)	754,834 (99.11%)
Distinct interactions	369,457 (48.9%)

Release 4.0

Results: RIG redundancy between source databases.

Statistics iRefIndex 4.0 - Do	🕹 Statistics iRefIndex 4.0 - Donaldson Group - Mozilla Firefox												
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookm	narks <u>T</u> ools	<u>H</u> elp											
🔇 🔊 - C 🗙 🏠	http://ire	efindex.uio.i	no/wiki/Statisti	ics_iRefInde:	×_4.0			<mark>™</mark> €	7 • [Google		ي 🔎	
应 Most Visited 옹 Entrez PubMed 💈	Jio Library of Me	dicine an	Uio The Biote	echnology Ce	e 8888 BBC	- bbc.co.uk	home 🚺	Science/AAA	S Multi	m 📄 http	p://bioscape.uio	.n	»
og Start Stumbling or 🔑 Sign-ir	n												
🔔 Page Load Error	×	Statistic	s iRefIndex	4.0 - Dona	lds🛛								•
	Interac	ctions	(Corre	spond	ls to T	able 6	in Pl	VID 18	823	568 🔊		[edi	it] 🔼
	BIND	62921											
	BioGrid	20497	163891										
	DIP	25914	28969	56441									
	HPRD	2893	1958	839	37956								≡
	INTACT	24239	25653	24807	8075	111235	5						
	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	INTACI	F MINT	MPACT	MPP	I OPHID	CORUM		
		(25903)) (111633) (13594) (15201) (55807) (1571)	2) (1137)	(238) (26571) (1403)		
	Intera	ctors										[edi	it]
													-
	BIND	40801											
Latart Column	Discrid	44440	07474		👝 cup-	6 0	e	The second	1		S. Columb	Correla	Dookton
Staft Skype	Manager Index	Yeas	📲 a	nalys	C:\Do	ire 🐏 ire	arind	irefind		Statist	Calcul	Search	Desktop

Results: RIG redundancy between source databases.

Statistics iRefIndex 4.0 - Dona	aldson Group	- Mozilla Firefox							
<u>F</u> ile <u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookmark	ks <u>T</u> ools <u>H</u> elp	0							***
C ×	http://irefina	dex.uio.no/wiki/Statist	cs_iRefIndex_4.0		<mark>汤</mark> 公	• G• Google		P 💷	
🔎 Most Visited 😂 Entrez PubMed 💴	Library of Medicin	ne an 💴 The Biote	chnology Ce 💵 B	C - bbc.co.uk hom	ne <mark>S</mark> Science/AAAS	5 Multim 📄 http	;//bioscape.uio.n.		>>
到 Start Stumbling or 🔑 Sign-in									
🔔 Page Load Error	🛛 🗋 S	itatistics iRefIndex	4.0 - Donalds 🔀						-
	Interacti	ons (Corre	sponds to	Table 6 ir	n PMID 188	323568 🗗		[edi	t] ^
	BIND 62921 BioGrid 20497 163891								
All da	All databases contribute some								
distin	ct i	nter	actio	on r	eco	rds.			
	OPHID 22 CORUM 11 BI	20 210 1333 13 18 ND BioGrid 5903) (111633	887 1791 29 390 DIP HPRI) (13594) (1520	3 7196 6 121 6 NTACT M 1) (55807) (1	, 0 396 0 6 0 1INT MPACT 15712) (1137)	183 47297 9 158 MPPI OPHID (238) (26571	1919 CORUM) (1403)		
	Interacto	ors						[edi	t]
	BIND 40	0801							
🖌 start 💦 Skype 💽	Jinbox 🛛 🦉	🔹 Yeast 📲 a	nalys 🛅 C:\De	🖭 irefind	l 📑 irefind	🕲 Statist	Calcul	Search	Desktop

Accessing consolidated data

http://irefindex.uio.no

http://irefindex.uio.no/

🕹 iRefIndex - Donaldson	Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	<u>B</u> ookmarks <u>T</u> ools <u>H</u> elp		
🔇 💽 - C 🗙	http://irefindex.uio.no/wiki/iRefIndex	🔊 🗘 🔹 🔽 Google 🖉	
🔎 Most Visited 😂 Entrez Put	bMed 💴 Library of Medicine an 🛄 The Biotechnology	y Ce 🕮 BBC - bbc.co.uk home <mark>S</mark> Science/AAAS Multim 📄 The Research Council W Wikipedia 脑 BBC - Weather Centre	
😽 iGoogle	🖂 📄 iRefIndex - Donaldson Grou	ip 🛛	
BiO	page discussion edit history in iRefIndex	Ian.donaldson my talk my preferences my watchlist my contributions log o	.t _
navigation Donaldson Group Community portal Current events Recent changes Random page Help Donations search Go Search toolbox • What links here	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		
 Related changes Upload file Special pages Printable version 	Please take the biolibrarian surv	'ey [୧ n Pranosal ୟ	dit]
Permanent link	A reference index for protein int	eraction data	dit]
	iRefIndex provides an index of protein interaction MINT &, MPact &, MPPI & and OPHID &. This	ons available in a number of primary interaction databases including BIND &, BioGRID &, CORUM &, DIP &, HPRD &, IntAct &, s 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 Id the iRefIndex in a way that ensures proteins w	dentifier (SEGUID &) to group proteins and interactions into redundant groups. This method allows users to integrate their own data w ith the exact same sequence will be represented only once.	ith
	The iRefIndex paper has been published and is	s available here 🗗	~
Done			

iRefIndex paper

🕹 iRefIndex - Donaldson	Group - Mozilla Firefox			
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	<u>B</u> ookmarks <u>T</u> ools <u>H</u> elp			**** ****
🔇 🖸 - C 🗙			🔊 🗘 🔹 🖸 Google	P
🧖 Most Visited 옹 Entrez Pub	oMed 🔟 Library of Medicine an 🛄 The Biotechnology C	e 🚥 BBC - bbc.co.uk home <mark>S</mark> Science/AAAS Multim 📄 The F	Research Council W Wikipedia 脑 BBC - Weather Ce	entre
😽 iGoogle	🖾 📄 iRefIndex - Donaldson Group			•
BiO	page discussion edit history mo iRefIndex Contents [hide]	ve unwatch	donaldson mytalk mypreferences mywatchlist	my contributions log out
navigation Donaldson Group Community portal Current events Recent changes Random page Help Donations search Go Search toolbox	 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 			2
 Virial links here Related changes Upload file Special pages Printable version Permanent link 	Please take the biolibrarian surve See http://irefindex.uio.no/wiki/The_Biolibrarian_	ሃ Proposal <mark>ይ</mark> ዋ		[edit]
	A reference index for protein inter iRefindex provides an index of protein interaction MINT @, MPact @, MPPI @ and OPHID @. This in iRefindex uses the Sequence Global Unique Ider the iRefindex in a way that ensures proteins with	raction data s available in a number of primary interaction databases includ ndex allows the user to search for a protein and retrieve a non- tifier (SEGUID &) to group proteins and interactions into redur the exact same sequence will be represented only once.	ding BIND &, BioGRID &, CORUM &, DIP &, HP redundant list of interactors for that protein. ndant groups. This method allows users to integ	[edit] RD &, IntAct &, rate their own data with
	The iRefindex paper has been published and is a	vailable here 🗗		~

Data in PSI-MI MITAB format

🕹 iRefIndex - Donaldson	Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	<u>B</u> ookmarks <u>T</u> ools <u>H</u> elp		
🔇 💽 - C 🗙	1 http://irefindex.uio.no/wiki/iRefIndex	🔊 🏠 🔹 🗔 🕞	ogle 🔎 🔒
🔎 Most Visited 옹 Entrez Pub	bMed 🔟 Library of Medicine an 🛄 The Biotechnology (Ce 🚥 BBC - bbc.co.uk home Ѕ Science/AAAS Multim 🗋 The Research Council W Wikipedia 脑 BBC	- Weather Centre
😽 iGoogle	🖂 📄 iRefIndex - Donaldson Group		•
BiO	page discussion edit history mo iRefIndex Contents [hide]	An donaidson my talk my preferences my unwatch	y watchlist my contributions log out
navigation Donaldson Group Community portal Current events Recent changes Random page Help Donations search Go Search toolbox What links here	 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 18 Ependex Google Group 		200
 Related changes Upload file Special pages Printable version Permanent link 	Please take the biolibrarian surve See http://irefindex.uio.no/wiki/The_Biolibrarian_	y Proposal छ	[edit]
	A reference index for protein interaction iRefIndex provides an index of protein interaction MINT @, MPact @, MPPI @ and OPHID @. This i iRefIndex uses the Sequence Global Unique Ide the iRefIndex in a way that ensures proteins with	raction data is available in a number of primary interaction databases including BIND &, BioGRID &, CORUM &, ndex allows the user to search for a protein and retrieve a non-redundant list of interactors for that p ntifier (SEGUID 와) to group proteins and interactions into redundant groups. This method allows use n the exact same sequence will be represented only once.	[edit] DIP &, HPRD &, IntAct &, protein. ers to integrate their own data with
	The iRefIndex paper has been published and is a	available here 🕼.	~

Source databases

🕹 iRefIndex - Donaldson	Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	<u>B</u> ookmarks <u>T</u> ools <u>H</u> elp		
🔇 🖸 - C 🗙	http://irefindex.uio.no/wiki/iRefIndex	S 🗘 🔹 Google	۶۹ 🥄
횓 Most Visited 옹 Entrez Pub	Med 🔟 Library of Medicine an 🔟 The Biotechnology	- Ce 🕮 BBC - bbc.co.uk home <mark>S</mark> Science/AAAS Multim 📄 The Research Council W Wikipedia 脑 BBC - Weather Cen	itre
🔧 iGoogle	📧 📄 iRefIndex - Donaldson Grou		•
navigation Donaldson Group Community portal Current events Recent changes Random page Help Donations search Go Search toolbox	page discussion edit history r 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 Cytoscape 7 Data availability via Web services 8 Feedback files 9 License 10 Disclaimer 11 Statistics 12 Credits 13 Citation 14 References 15 Contact		ry contributions log out
 What links here Related changes Upload file Special pages Printable version Permanent link 	16 iRefIndex Google Group Please take the biolibrarian surve See http://irefindex.uio.no/wiki/The_Biolibrarian	ອ y _Proposal &	[edit]
	A reference index for protein inte	eraction data	[edit]
	iRefIndex provides an index of protein interactic MINT &, MPact &, MPPI & and OPHID &. This iRefIndex uses the Sequence Global Unique Id	ns available in a number of primary interaction databases including BIND &, BioGRID &, CORUM &, DIP &, HPR index allows the user to search for a protein and retrieve a non-redundant list of interactors for that protein. entifier (SEGUID &) to group proteins and interactions into redundant groups. This method allows users to integra	D &, IntAct &, ate their own data with
	the iRefindex in a way that ensures proteins wi	th the exact same sequence will be represented only once.	
	The iRefIndex paper has been published and is	available here 🖉.	~

Most recent statistics

🕹 iRefIndex - Donalds	son Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tor	ry <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp		*** ***
🕢 🖓 - C	A (http://irefindex.uio.no/wiki/iRefIndex	Share Coogle Soogle	P 🔒
🔎 Most Visited 🔗 Entrez	z PubMed 🔟 Library of Medicine an 🔟 The Biotechnology Ce 🕮 BBC - bbc.co.uk home S	Science/AAAS Multim 📄 The Research Council W Wikipedia 脑 BBC - Weather Centre	
🔧 iGoogle	📧 📄 iRefIndex - Donaldson Group 🛛 🔀		-
	page discussion edit history move unwatch	2 Ian.donaldson my talk my preferences my watchlist my contributions I	log out 🔥
	iRefIndex		=
navigation Donaldson Group Community portal Current events Recent changes Random page Help Donations search Go Search toolbox What links here Donations	Contents [hide]1 Please take the biolibrarian survey2 A reference index for protein interaction data3 Long term goals of the iRefindex project4 Data availability via download5 Data availability via web interface6 Data availability via Cytoscape7 Data availability via Web services8 Feedback files9 License10 Disclaimer11 Statistics12 Credits13 Citation14 References15 Contact16 iRefindex Google Group		
 Upload file Special pages 	Please take the biolibrarian survey		[edit]
 Special pages Printable version Permanent link 	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 primar MINT &, MPact &, MPPI & and OPHID &. This index allows the user to search for iRefIndex uses the Sequence Global Unique Identifier (SEGUID &) to group prote the iRefIndex in a way that ensures proteins with the exact same sequence will be	y interaction databases including BIND &, BioGRID &, CORUM &, DIP &, HPRD &, IntAct &, r a protein and retrieve a non-redundant list of interactors for that protein. ins and interactions into redundant groups. This method allows users to integrate their own dat re represented only once.	ta with

The iRefIndex paper has been published and is available here 🗗

License: not all data is distributed....ask

🕹 iRefIndex - Donaldson	Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	<u>B</u> ookmarks <u>T</u> ools <u>H</u> elp		4 ⁴ 4 9 9
< ⊇ · C ×	http://irefindex.uio.no/wiki/iRefIndex	🔊 ☆ 🔹 🗔 Google	P
🔎 Most Visited 옹 Entrez Pub	Med 🔟 Library of Medicine an 🔟 The Biotechnology Ce	💵 BBC - bbc.co.uk home 🗴 Science/AAAS Multim 🗋 The Research Council W Wikipedia 🔝 BBC - Weather Cer	ıtre
🛂 iGoogle	🖂 📄 iRefIndex - Donaldson Group		•
		an.donaldson my talk my preferences my watchlist m	ny contributions log out 🛛 🔥
•	page discussion edit history move	unwatch	Annual Contractor
	iRefIndex		
	Contents [hide] 1 Please take the biolibrarian survey 2 A reference index for protein interaction data		
navigation	3 Long term goals of the iRefIndex project		
 Donaldson Group Community portal Current events Recent changes Random page Help Donations 	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		
search	11 Statistics		
Go Search	12 Credits 13 Citation 14 References		
toolbox	15 Contact		
 What links here Related changes 	16 iRefindex Google Group		
 Upload file 	Please take the biolibrarian survey		[edit]
 Special pages Printable version Permanent link 	See http://irefindex.uio.no/wiki/The_Biolibrarian_Pro	posal ៥	
	A reference index for protein interac	tion data	[edit]
	iRefIndex provides an index of protein interactions av MINT &, MPact &, MPPI & and OPHID &. This inde:	railable in a number of primary interaction databases including BIND &, BioGRID &, CORUM &, DIP &, HPF « allows the user to search for a protein and retrieve a non-redundant list of interactors for that protein.	≀D ଢ₽, IntAct ଢ₽,
	iRefIndex uses the Sequence Global Unique Identifie the iRefIndex in a way that ensures proteins with the	er (SEGUID 🗗 to group proteins and interactions into redundant groups. This method allows users to integra e exact same sequence will be represented only once.	ate their own data with
	The iRefIndex paper has been published and is avail	able here 🗗	~

Feedback to databases

🕹 iRefIndex - Donald	son Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> to	ry <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp		
🕢 🖓 🗸 🗸	A (http://irefindex.uio.no/wiki/iRefIndex	Sight Google	
🔎 Most Visited 옹 Entrea	z PubMed 🔟 Library of Medicine an 🔟 The Biotechnology Ce 💷 BBC - bbc.co.uk home S S	cience/AAAS Multim 📋 The Research Council W Wikipedia 脑 BBC - Weather Centre	
🛂 iGoogle	📧 📄 iRefIndex - Donaldson Group 🛛 🛛		
	page discussion edit history move unwatch	2 lan.donaldson my talk my preferences my watchlist my contributions log out	
	iRefIndex		_
navigation Donaldson Group Community portal Current events Recent changes Random page Help Donations search Go Search	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		
toolbox	14 References 15 Contact		
Vvhat links here Related changes Upload file Special pages Printable version Permanent link	16 iRefIndex Google Group Please take the biolibrarian survey See http://irefindex.uio.no/wiki/The_Biolibrarian_Proposal @	[edit	<u>it]</u>
= Permanent link	A reference index for protein interaction data	[edit	it]
	iRefIndex provides an index of protein interactions available in a number of primary MINT &, MPact &, MPPI & and OPHID &. This index allows the user to search for iRefIndex uses the Sequence Global Unique Identifier (SEGUD Ø) to group protein	interaction databases including BIND &, BioGRID &, CORUM &, DIP &, HPRD &, IntAct &, a protein and retrieve a non-redundant list of interactors for that protein.	
	the iRefindex in a way that ensures proteins with the exact same sequence will be	represented only once.	
	The iDefindex paper has been published and is subjected here fi		

The iRefIndex paper has been published and is available here 🗗

Cytoscape plugin

😉 iRefIndex - Donaldso	on Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	y <u>B</u> ookmarks <u>I</u> ools <u>H</u> elp		
🚱 🗗 🕻 🗙	http://irefindex.uio.no/wiki/iRefIndex	🔊 🏠 🔹 🚺 Google	2 🔒
🔎 Most Visited 😂 Entrez I	PubMed 🔟 Library of Medicine an 🔟 The Biotechnology Ce 💷 BBC - bbc.co.uk home 🗴 Scier	ice/AAAS Multim 📋 The Research Council 👿 Wikipedia 脑 BBC - Weather Centre	
😽 iGoogle	🖂 📑 iRefIndex - Donaldson Group 🛛 🛛		•
	page discussion edit history move unwatch	A lan.donaldson my talk my preferences my watchlist my contributions log	out 🔨
BIO	iRefIndex		=
	Contents [hide] 1 Please take the biolibrarian survey 2 A reference index for protein interaction data 3 Long torm goals of the iEnded variant.		
navigation Donaldson Group Community portal Current events Recent changes Help Donations search Go Search toolbox Vhat links here Related changes	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		
 Upload file Special pages Printable version Permanent link 	Please take the biolibrarian survey See http://irefindex.uio.no/wiki/The_Biolibrarian_Proposal @	[[edit]
	A reference index for protein interaction data iRefIndex provides an index of protein interactions available in a number of primary int MINT @, MPact @, MPPI @ and OPHID @. This index allows the user to search for a iRefIndex uses the Sequence Global Unique Identifier (SEGUID @) to group proteins a the iRefIndex in a way that ensures proteins with the exact same sequence will be re The iRefIndex paper has been published and is available here @.	eraction databases including BIND &, BioGRID &, CORUM &, DIP &, HPRD &, IntAct &, protein and retrieve a non-redundant list of interactors for that protein. and interactions into redundant groups. This method allows users to integrate their own data a presented only once.	(edit) with

PSICQUIC web services

🕹 iRefIndex - Donaldson	n Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	Bookmarks Iools Help		
🔇 🔊 - C 🗙	http://irefindex.uio.no/wiki/iRefIndex	Sociel Google	P 🔒
🔎 Most Visited 😂 Entrez Pu	ubMed 🔟 Library of Medicine an 🔟 The Biotechnology Ce 💷 BBC - bbc.co.uk hom	<mark>S</mark> Science/AAAS Multim 📄 The Research Council W Wikipedia 脑 BBC - Weather Centre	
🔧 iGoogle	📧 📄 iRefIndex - Donaldson Group 🛛 🛛		•
	page discussion edit history move unwatch	2 lan.donaldson my talk my preferences my watchlist my contributions lo	ig out 🔺
	iRefIndex		=
navigation Donaldson Group Community portal Current events Recent changes Random page Help Donations search Go Search toolbox	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		
 What links here Related changes Upload file 	16 iRefIndex Google Group Please take the biolibrarian survey		[edit]
 Special pages Printable version Permanent link 	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 MINT &, MPact &, MPPI & and OPHID &. This index allows the user to se	primary interaction databases including BIND &, BioGRID &, CORUM &, DIP &, HPRD &, IntAct &, arch for a protein and retrieve a non-redundant list of interactors for that protein.	
	iRefIndex uses the Sequence Global Unique Identifier (SEGUID &) to grout the iRefIndex in a way that ensures proteins with the exact same sequence) proteins and interactions into redundant groups. This method allows users to integrate their own data e will be represented only once.	a with
	The iRefIndex paper has been published and is available here 🗗		~

iRefWeb

🕹 iRefIndex - Donaldsor	n Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	Bookmarks Iools Help		
⊘ ≥ • ⊂ ×		Societ Scoole	P 🔒
🔎 Most Visited 옹 Entrez Pu	ubMed 🔟 Library of Medicine an 💴 The Biotechnology Ce 💷 BBC - bbc.co.uk home 🕻	🖞 Science/AAAS Multim 📋 The Research Council W Wikipedia 🔝 BBC - Weather Centre	
🚼 iGoogle	🖂 📋 iRefIndex - Donaldson Group 🛛 🔯		•
BiO	page discussion edit history move unwatch IRefIndex Contents [hide]	Ian.donaldson my talk my preferences my watchlist my contributions log	3 out
navigation Donaldson Group Community portal Current events Recent changes Random page Help Donations search	 Please take the biolibrarian survey A reference index for protein interaction data Long term goals of the iRefindex project Data availability via download Data availability via web interface Data availability via Cytoscape Data availability via Web services Feedback files License Disclaimer Statistics Credits 	200	•
Go Search toolbox • What links here • Related changes • Upload file	13 Citation 14 References 15 Contact 16 iRefindex Google Group Please take the biolibrarian survey		[edit]
 Special pages Printable version Permanent link 	See http://irefindex.uio.no/wiki/The_Biolibrarian_Proposal &		<u> </u>
	A reference index for protein interaction data iRefindex provides an index of protein interactions available in a number of prim MINT &, MPact &, MPPI & and OPHID &. This index allows the user to search iRefindex uses the Sequence Global Unique Identifier (SEGUID) to group pro the iRefindex in a way that ensures proteins with the exact same sequence will The iRefindex paper has been published and is available here &.	ary interaction databases including BIND &, BioGRID &, CORUM &, DIP &, HPRD &, IntAct &, for a protein and retrieve a non-redundant list of interactors for that protein. teins and interactions into redundant groups. This method allows users to integrate their own data I be represented only once.	<u>[eaπ]</u> ↓ with

Google group

🕲 iRefIndex - Donaldsor	n Group - Mozilla Firefox		
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	<u>B</u> ookmarks <u>T</u> ools <u>H</u> elp		*** ***
⊘ ≥ • ⊂ ×	1 http://irefindex.uio.no/wiki/iRefIndex	S 🗘 🔹 Google	P 🔒
🔎 Most Visited 옹 Entrez Pu	ıbMed 🔟 Library of Medicine an 🔟 The Biotechnology	· Ce 🎟 BBC - bbc.co.uk home S Science/AAAS Multim 📄 The Research Council W Wikipedia 脑 BBC - Weather Centre	
😽 iGoogle	🖂 📑 iRefIndex - Donaldson Group	p 🛛 🔀	-
		A lan.donaldson my talk my preferences my watchlist my contri	ibutions log out 🔥
1 A 1	page discussion edit history m	nove unwatch	
A	iRefIndex		
BiO			
	Contents [hide]		
	1 Please take the biolibrarian survey	•	
	2 A reference index for protein interaction data		
navigation	3 Long term goals of the iRefIndex project		
Donaldson Group	4 Data availability via download		• • · /
 Community portal 	5 Data availability via web interface		OTOTO .
 Current events 	6 Data availability via Cytoscape		
 Recent changes 	7 Data availability via Web services		
 Random page Help 	8 Feedback files		
 Donations 	9 License		
	10 Disclaimer		
search	11 Statistics		
	12 Credits		
Go Search	13 Citation		
toolbox	15 Contect		
What links here	16 iRefinder Google Group		
Related changes			
 Upload file 	Please take the biolibrarian surve	ev	[edit]
 Special pages 		-)	
 Printable version Democrat link 	See http://irefindex.uio.no/wiki/The_Biolibrarian	_Proposal 🗗	
Permanent link	A		r
	A reference index for protein inte	eraction data	[edit]
	iRefIndex provides an index of protein interactio MINT & MPact & MPPI를 and OPHID & This	ns available in a number of primary interaction databases including BIND &, BioGRID &, CORUM &, DIP &, HPRD &, In index allows the user to search for a protein and retrieve a non-redundant list of interactors for that protein.	ntAct 🗗
	iRefIndex uses the Sequence Global Unique Ide the iRefIndex in a way that ensures proteins wi	entifier (SEGUID &) to group proteins and interactions into redundant groups. This method allows users to integrate thei th the exact same sequence will be represented only once.	ir own data with

The iRefindex paper has been published and is available here 🗗

File Edit View History Bookmarks Tools Help	0
C X 🟠 🗋 http://irefindex.uio.no/wiki/README_iRefIndex_Feedback_3.0	
🖻 Most Visited 😔 Entrez PubMed 🚾 Library of Medicine an 🚾 The Biotechnology Ce 🚥 BBC - bbc.co.uk home <mark>S</mark> Science/AAAS Multim 🏠 The Research Council W Wikipedia 🍃 biocurator - Twitter S	
외 Start Stumbling or 🔑 Sign-in	
🧟 lan.donaldson my talk my preferences my watchlist my contributions log out	^
page discussion edit history move unwatch README iRefIndex Feedback 3.0 Image: Compare to the second	-
Last edited: , March 26th, 2009 Applies to iRefindex release: 3.0 beta.	
Release date: March 26th, 2009	
Authors: Ian Donaldson and Sabry Razick	
navigation Database: iDefinder (http://irefinder.uio.no.@)	
■ Donaldson Group	
Community portal Organization: Biotechnology Centre of Oslo, University of Oslo (http://www.biotek.uio.no/ &)	
Contents Ibide1	
Recent changes 1 Description	
Random page A Disordary contents	
Postions 2 Changes from last varsion	
Domailoris S Changes nonnaist version	
search 5 License	
B Citation	
Go Search 7 Disclaimer	
B Understanding the Feedback files	
100100X 9 Description of Feedback file	
Vyhat links here 9.1 Column number; 1	
Keiared changes Johoed file 9.2 Column number: 2	
Special pages 9.3 Column number: 3	
Printable version 9.4 Column number: 4	
Permanent link 9.5 Column number: 5	
9.6 Column number: 6	
9.7 Column number: 7	
9.8 Column number: 8	
9.9 Column number: 9	
9.10 Column number: 10	
9.11 Column number: 11	
9.12 Column number: 12	
9.13 Column number: 13	
9.14 Column number: 14	_
9.15 Column number 15	~

C	9	- (°I -) -					hpr	d_feedbad	k_3.0.txt - I	Microsoft Exc	el						x
9	Home	Insert	Pag	ie Layou	it Formula	as Data	Review View	N								0 - 🗖	х
Norma	Page Layout Work	I Page Bro I Custom I Full Scre kbook Views	eak Prev Views en	iew [Ruler Gridlines Message Ba	Formula Headings ar bw/Hide	Bar Zoom 10	00% Zoom Select	to New Windo	Arrange Fr ow All Pa	reeze anes + 🖸 U	olit Dit View ide a‡ Synch nhide de Reset Windov	Side by Side Ironous Scrolling Window Position N	Save Workspace	Switch Windows •	Macros Macros	
	A2	•	0	f;	HPRD												×
	А	В	C	2	D	E	F	G	Н	I.	J	К	L	М	N	0	
1	int_db	int_acc	source	e_file	primary_db	primary_acc	primary_taxid	used_db	used_acc	used_taxid	mapped_c	lb mapped_acc	mapped_taxid	rogid	rogscore	score_typ	e
57311	HPRD	-10	HPRD	SING	HPRD	2995	9606	uniprot	Q13485	9606	uniprot	Q13485	9606	mPu190hr	S	1	
57312	HPRD	-10	HPRD_	SING	HPRD	8915	9606	uniprot	P29375	9606	uniprot	P29375	9606	F6XYSW2/	S	1	
57313	HPRD	-10	HPRD_	SING	HPRD	5397	9606	uniprot	Q96GD4	9606	uniprot	Q96GD4	9606	evvncWJrl	S	1	
57314	HPRD	-10	HPRD	SING	HPRD	11906	40674	entrezge	8352	9606	refseq	NP_003522	9606	ood0Ooh>	SVGO	1	
57315	HPRD	-10	HPRD	SING	HPRD	11915	40674	entrezge	8361	9606	refseq	NP_003531	9606	LZSfhrpw0	SVGO	1	
57316	HPRD	-10	HPRD	SING	HPRD	671	9606	uniprot	P01350	9606	uniprot	P01350	9606	CIVYdfVtl	S	1	
57317	HPRD	-10	HPRD	SING	HPRD	2799	9606	uniprot	P42574	9606	uniprot	P42574	9606	o7/WP52s	S	1	L
57318	HPRD	-10	HPRD	SING	HPRD	3382	9606	uniprot	P17612	9606	uniprot	P17612	9606	7yhAi67qj	S	1	
57319	HPRD	-10	HPRD	SING	HPRD	321	9606	uniprot	P06307	9606	uniprot	P06307	9606	rWXWgq8	S	1	
57320	HPRD	-10	HPRD	SING	HPRD	321	9606	uniprot	P06307	9606	uniprot	P06307	9606	rWXWgq8	S	1	L
57321	HPRD	-10	HPRD	SING	HPRD	321	9606	uniprot	P06307	9606	uniprot	P06307	9606	rWXWgq8	S	1	L
57322	HPRD	-10	HPRD	SING	HPRD	321	9606	uniprot	P06307	9606	uniprot	P06307	9606	rWXWgq8	S	1	
57323	HPRD	-10	HPRD	SING	HPRD	321	9606	uniprot	P06307	9606	uniprot	P06307	9606	rWXWgq8	S	1	L
57324	HPRD	-10	HPRD	SING	HPRD	2926	9606	uniprot	P34947	9606	uniprot	P34947	9606	Wrqm8jds	S	1	L
57325	HPRD	-10	HPRD	SING	HPRD	2925	9606	uniprot	P43250	9606	uniprot	P43250	9606	n7nyE4T60	S	1	L
57326	HPRD	-10	HPRD	SING	HPRD	4113	9606	uniprot	O00585	9606	uniprot	O00585	9606	kCftuxbCJ	S	1	
57327	HPRD	-10	HPRD	SING	HPRD	302	9606	uniprot	P06493	9606	uniprot	P06493	9606	8geyXVp0	S	1	
57328	HPRD	-10	HPRD	SING	HPRD	302	9606	uniprot	P06493	9606	uniprot	P06493	9606	8geyXVp0	S	1	
57329	HPRD	-10	HPRD	SING	HPRD	302	9606	uniprot	P06493	9606	uniprot	P06493	9606	8geyXVp0	S	1	
57330	HPRD	-10	HPRD	SING	HPRD	302	9606	uniprot	P06493	9606	uniprot	P06493	9606	8geyXVp0	s	1	
57331	HPRD	-10	HPRD	SING	HPRD	302	9606	uniprot	P06493	9606	uniprot	P06493	9606	8geyXVp0	S	1	L
57332	HPRD	-10	HPRD	SING	HPRD	302	9606	uniprot	P06493	9606	uniprot	P06493	9606	8geyXVp0	S	1	
57333	HPRD	-10	HPRD	SING	HPRD	302	9606	uniprot	P06493	9606	uniprot	P06493	9606	8geyXVp0	S	1	
57334	HPRD	-10	HPRD	SING	HPRD	302	9606	uniprot	P06493	9606	uniprot	P06493	9606	8geyXVp0	S	1	
57335	HPRD	-10	HPRD	SING	HPRD	1899	9606	uniprot	P08670	9606	uniprot	P08670	9606	GhAhVuXI	S	1	
H + +	→ hpr	d_feedbad	k_3.0	2	/						[∢						
Ready										Avera	ge: 8035.75	Count: 15 Su	m: 64286 🖽 🔲	III 100% (9) .:i

	9	- (° -) =					hp	rd_feedba	ck_3.0.txt - I	Microsoft Ex	cel						x
9	Home	Insert	Page	Layou	t Formula	as Data	Review Vie	ew								💿 💷 💿	x
Norm	al Page Layout Wor	Page Bre I Custom Full Scre kbook Views	eak Previe Views en	:w [Ruler Gridlines Message Ba	Formula I Headings ar bw/Hide	Bar Q Zoom 1	100% Zoon Selec	n to tion	Arrange F ow All P	reeze anes + 🗌 Ur	olit Diew Synch de ⊡‡ Synch nhide 관광 Reset Window	Side by Side ronous Scrolling Window Position v	Save Workspace	Switch Windows *	Macros	
	A2	•	()	f _s	HPRD												×
	А	В	С		D	E	F	G	н	I.	J	K	L	М	N	0	
1	int_db	int_acc	source	file	primary_db	primary_acc	primary_taxid	used_db	used_acc	used_taxid	mapped_d	lb mapped_acc	mapped_taxid	rogid	rogscore	score_typ	e
56094	HPRD	-10	HPRD_S	SINGL	HPRD	18010	40674	4 uniprot	Q9UBE0	40674	uniprot	Q9UBE0	9606	8Az5p8NJ	ST	3	
6095	HPRD	-10	HPRD_S	SING	HPRD	11658	40674	4 uniprot	Q9UBT2	40674	uniprot	Q9UBT2	9606	OwedW+/	ST	3	
6096	HPRD	-10	HPRD_S	SING	HPRD	16283	40674	4 uniprot	Q9NVD7	40674	uniprot	Q9NVD7	9606	KjiRqYX7c	ST	3	
6097	HPRD	-10	HPRD_S	SING	HPRD	3842	40674	4 uniprot	Q13418	40674	uniprot	Q13418	9606	FrK/MsW7	ST	3	
6098	HPRD	-10	HPRD_S	SING	HPRD	3978	40674	4 uniprot	P48059	40674	uniprot	P48059	9606	LzyCXdAH	ST	3	
56099	HPRD	-10	HPRD_9	SINGL	HPRD	5378	40674	4 uniprot	O43639	40674	uniprot	O43639	9606	WV9nJ78I	ST	3	
56100	HPRD	-10	HPRD_9	SINGL	HPRD	3842	40674	4 uniprot	Q13418	40674	uniprot	Q13418	9606	FrK/MsW7	ST	3	
56101	HPRD	-10	HPRD_S	SINGL	HPRD	3978	40674	4 uniprot	P48059	40674	uniprot	P48059	9606	LzyCXdAH	ST	3	
56102	HPRD	-10	HPRD_9	SINGL	HPRD	6260	40674	4 uniprot	014617	40674	uniprot	O14617	9606	vw6bWz0	ST	3	
56103	HPRD	-10	HPRD_S	SINGL	HPRD	16359	40674	4 uniprot	Q9UPQ3	40674	uniprot	Q9UPQ3	9606	iVoZlA+NI	ST	3	
56104	HPRD	-10	HPRD_S	SINGL	HPRD	3880	40674	4 uniprot	P59780	40674	uniprot	P59780	9606	6sX3+WLo	ST	3	
56105	HPRD	-10	HPRD_S	SINGL	HPRD	16106	960	5 uniprot	O15483	9606	uniprot	Q59EK9	9606	Xib1v1Fk/	SU	2	
56106	HPRD	-10	HPRD_S	SINGL	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	
56107	HPRD	-10	HPRD_S	SING	HPRD	16106	960	5 uniprot	O15483	9606	uniprot	Q59EK9	9606	Xib1v1Fk/	SU	2	
56108	HPRD	-10	HPRD_S	SING	HPRD	1624	960	5 uniprot	P10162	9606	uniprot	P10163	9606	t1xOmY9q	SU	2	
56109	HPRD	-10	HPRD_S	SINGL	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	
56110	HPRD	-10	HPRD_S	SING	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	
56111	HPRD	-10	HPRD_S	SING	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	
56112	HPRD	-10	HPRD_S	SING	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	
56113	HPRD	-10	HPRD_S	SINGL	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	
56114	HPRD	-10	HPRD_S	SINGL	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	
6115	HPRD	-10	HPRD_S	SING	HPRD	16106	960	5 uniprot	015483	9606	uniprot	Q59EK9	9606	Xib1v1Fk/	su	2	
6116	HPRD	-10	HPRD_S	SING	HPRD	16106	960	5 uniprot	015483	9606	uniprot	Q59EK9	9606	Xib1v1Fk/	SU	2	
6117	HPRD	-10	HPRD_S	SING	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	
6118	HPRD	-10	HPRD_S	SING	HPRD	9033	960	5 uniprot	O00145	9606	uniprot	Q92835	9606	IM4QLD8E	SU	2	-
H +)	→ hpi	rd_feedbac	.k_3.0	*	7											•	
Ready										Avera	age: 8035.75	Count: 15 Su	m: 64286 🖽 🔲	🛄 100% 🤅) (•) .::

윌 Alpha-parvin - Homo sapiens (Human) -	Mozilla Firefox					- 5	- x
<u>File Edit View History Bookmarks Tools</u>	Help					🥑 🗕	⊐ x
🔇 🖸 🗸 🔂 💭 http://w	www.uniprot.org/uniprot/Q9NVD7 🔝 🏠 🔹 💽 🕻 Google 🔎	e 🤱	View S	ide by Side			
🧖 Most Visited 옹 Entrez PubMed 🛄 Library of Me	edicine an 💴 The Biotechnology Ce 💷 BBC - bbc.co.uk home	>>	de 斗 de Reset \	Window Position	Save Swite	h Macros	
列 Start Stumbling or 🄑 Sign-in			Window		vvorkspace vvindo	Macros	5
UniProt > UniProtKB	Downloads · Contact · Documentation	n/Help					1
•••			К	L	M N	0	
Search in	Query	_	mapped_acc	mapped_taxid	rogid rogsco	re_score_t	ype
Protein Knowledgebase (UniProtKB)	V Search	Clear	Q9UBE0	9606	8Az5p8NJ:ST		3
			Q9UBT2	9606	OwedW+/ST		3
Search Blast *	Align * Retrieve D ID Mapping *	·)	Q9NVD7	9606	KjiRqYX7c ST		3
			Q13418	9606	FrK/MsW7ST		3
★ Reviewed UniProtKB/Swi	iss-Prot O9NVD7 Contribute		P48059	9606	LzyCXdAH ST		3
	Send feedback		O43639	9606	WV9nJ78I ST		3
(PARVA_HUMAN)	Read comments (0) or add your	own	Q13418	9606	FrK/MsW7ST		3
Last modified June 16, 2009. Version 73.		P48059	9606	LzyCXdAH ST		3	
			014617	9606	vw6bWz0`ST		3
🗧 📲 Clusters with 100%, 90%, 50% identi	ity 🗅 Documents (4) 🗐 🛛 🛛 TEXT XML RDF/XML G	FF	Q9UPQ3	9606	iVoZIA+NI ST		3
Third-party data 📾 Customize display	FAS	TA	P59780	9606	6sX3+WLo ST		3
			Q59EK9	9606	Xib1v1Fk/SU		2
Names and origin · Protein attributes · Gene	eral annotation (Comments) - Ontologies - Binary interactions -		Q92835	9606	IM4QLD8E SU		2
Alternative products · Sequence annotation	(Features) · Sequences · References · Cross-references · Entry		Q59EK9	9606	Xib1v1Fk/SU		2
information Relevant documents			P10163	9606	t1xOmY9qSU		2
			Q92835	9606	IM4QLD8E SU		2
Names and origin	Hide	Тор	Q92835	9606	IM4QLD8E SU		2
Names and origin			Q92835	9606	IM4QLD8E SU		2
Protein names	Recommended name:		Q92835	9606	IM4QLD8E SU		2
1 Totelli Hames	Alpha-parvin		Q92835	9606	IM4QLD8E SU		2
			Q92835	9606	IM4QLD8E SU		2
	Coloopin like integrin linked kinese binding protein		Q59EK9	9606	Xib1v1Fk/SU		2
	Calponin-like integrin-linked kinase-binding protein		Q59EK9	9606	Xib1v1Fk/SU		2
	CH-ILKBP		Q92835	9606	IM4QLD8E SU		2
	Matrix-remodeling-associated protein 2		Q92835	9606	IM4QLD8E SU		2
	Actopaxin	~					•
<		>	iount: 15 Sum	n: 64286 🔠 🔲	🛄 100% 😑 —		.
Done		ے او	2				

0) 🖬 🤊	- (°I -) -			h	prd_feedb	ack_3.0.txt	- Microsoft	Excel							x
C	Home	Inse	rt Page La	yout Form	ulas Data	Review V	/iew									 (i) = 1 	∍ x
Nor	rmal Page Layout Wor	I Page I Custo Full S rkbook Vie	Break Preview om Views Screen ews	Gridlines	 ✓ Formula Bas ✓ Headings Bar Show/Hide 	ar Q Zoom	100% Zo Sel	om to ection	New Arrange ndow All	Freeze Panes *	Split Hide Unhide	Vie Vie Syr Res Wind	w Side by Side achronous Scrollin set Window Positi dow	on Worksp	s Switch ace Window	s + Macros Macros	
	A2		• (•	<i>f</i> ∗ HPRD													×
4	А	В	С	D	E	F	G	Н	L.	J		К	L	М	N	0	
1	int_db	int_acc	source_file	primary_db	primary_acc prin	nary_taxid	used_db	used_acc	used_taxid	mappe	d_db_map	ped_acc	mapped_taxic	l rogid	rogscore	score_type	2
77	HPRD	-10	HPRD_SING	I HPRD	11448	9606	uniprot	Q9Y2N3	9606	uniprot	A8C0	534	960	5 zt2YzrYNN	SUXL+	5	
78	HPRD	-10	HPRD_SING	l HPRD	11236	9606	entrezge	138046	9606	refseq	NP_0	0010938	6 960	5 7Vx29NJ3	SVGL+	5	
79	HPRD	-10	HPRD_SING	l HPRD	10516	9606	entrezge	51317	9606	refseq	NP_0	0010952	960	5 pWfOJSil	SVGL+	5	
80	HPRD	-10	HPRD_SING	l HPRD	11448	9606	uniprot	Q9Y2N3	9606	uniprot	A8C0	534	960	5 zt2YzrYNN	SUXL+	5	
81	HPRD	-10	HPRD_SING	I HPRD	11448	9606	uniprot	Q9Y2N3	9606	uniprot	: A8C0	534	960	5 zt2YzrYNM	SUXL+	5	
82	HPRD	-10	HPRD_SING	I HPRD	13868	9606	entrezge	57658	9606	refseq	NP_0	0011371	960	5 C3IJ2QQh	SVGL+	5	
83	HPRD	-10	HPRD_SING	l HPRD	13868	9606	entrezge	57658	9606	refseq	NP_0	0011371	960	5 C3IJ2QQh	SVGL+	5	
84	HPRD	-10	HPRD_SING	I HPRD	10245	9606	entrezge	91544	9606	refseq	NP_0	0010707	960	5 zHLUrNaS	SVGL+	5	
85	HPRD	-10	HPRD_SING	I HPRD	10721	9606	entrezge	284403	9606	refseq	NP_7	775907	960	5 bKUivd80	SVGL+	5	
86	HPRD	-10	HPRD_SING	l HPRD	11448	9606	uniprot	Q9Y2N3	9606	uniprot	A8C0	534	960	5 zt2YzrYNN	SUXL+	5	
87	HPRD	-10	HPRD_SING	I HPRD	11448	9606	uniprot	Q9Y2N3	9606	uniprot	A8C0	534	960	5 zt2YzrYNN	SUXL+	5	
88	HPRD	-10	HPRD_SING	I HPRD	11448	9606	uniprot	Q9Y2N3	9606	uniprot	A8C0	534	960	5 zt2YzrYNN	SUXL+	5	
89	HPRD	-10	HPRD_SING	I HPRD	10245	9606	entrezge	91544	9606	refseq	NP_0	0010707	960	5 zHLUrNaS	SVGL+	5	
90	HPRD	-10	HPRD_SING	I HPRD	11448	9606	uniprot	Q9Y2N3	9606	uniprot	A8C0	534	960	5 zt2YzrYNM	SUXL+	5	
91	HPRD	-10	HPRD_SING	LHPRD	10516	9606	entrezge	51317	9606	refseq	NP_0	0010952	960	5 pWfOJSil	SVGL+	5	
92	HPRD	-10	HPRD_SING	I HPRD	10245	9606	entrezge	91544	9606	refseq	NP_0	0010707	960	5 zHLUrNaS	SVGL+	5	
93	HPRD	-10	HPRD_SING	I HPRD	9571	9606	entrezge	26146	9606	refseq	NP_0	0011329	6 960	5 LJ90KG+0	z SVGL+	5	
94	HPRD	-10	HPRD_SING	IHPRD	10516	9606	entrezge	51317	9606	refseq	NP_0	0010952	960	5 pWfOJSil	SVGL+	5	
95	HPRD	-10	HPRD_SING	I HPRD	10245	9606	entrezge	91544	9606	refseq	NP_0	010707	960	5 zHLUrNaS	SVGL+	5	
96	HPRD	-10	HPRD_SING	IHPRD	10516	9606	entrezge	51317	9606	refseq	NP_0	0010952	960	5 pWfOJSil	SVGL+	5	
97	HPRD	-10	HPRD_SING	L HPRD	9571	9606	entrezge	26146	9606	refseq	NP_0	011329	6 960	5 LJ90KG+0	z SVGL+	5	
98	HPRD	-10	HPRD_SING	I HPRD	2256	9606	entrezge	56159	9606	refseq	NP_0	010038	1 960	5 xzJnHUKk	SVGL+	5	
99	HPRD	-10	HPRD_SING	I HPRD	11131	9606	entrezge	55704	9606	refseq	NP_0	0011290	6 960	5 M/4FOE6	FSVGL+	5	
00	HPRD	-10	HPRD_SING	I HPRD	10516	9606	entrezge	51317	9606	refseq	NP_0	0010952	960	5 pWfOJSil	SVGL+	5	
01	HPRD	-10	HPRD_SING	I HPRD	10516	9606	entrezge	51317	9606	refseq	NP_0	0010952	960	5 pWfOJSil	SVGL+	5	-
4	→ → hp	rd_feedl	back_3.0 🦯							1 4				_			•
Rea	dy								A	/erage: 80	35.75 Cou	int: 15	Sum: 64286 🛛 🏛	1009	% 😑 🗕		÷ .::

hprd_feedback_3.0.txt - Microsoft Excel		- = ×
😌 Q9Y2N3.65 - Mozilla Firefox		🔞 – 📼 🗙
Eile Edit Yiew History Bookmarks Tools Help	de e crolling	
🔇 🕞 🗸 🖒 💭 http://www.uniprot.org/uniprot/Q9Y2N3 📓 🖓 🔹 💽 Google 🔎 💷 🔝	Position Workspace V	Switch Nindows • • Macros
🙍 Most Visited 😒 Entrez PubMed 🔟 Library of Medicine an 🔟 The Biotechnology Ce 🕮 BBC - bbc.co.uk home 😕		*
되 Start Stumbling or 🔑 Sign-in	M	N O
HeiPert Dust I/P	taxid rogid rog	score_score_type
OniProt On I Prot K B Downloads · Contact · Documentation/Help	9606 zt2YzrYNNSUX	(L+ 5
	9606 pWfOIsilssV0	3L+ 5
Search in Query	9606 zt2YzrYNNSU	(L+ 5
Protein Knowledgebase (UniProtKB) 💙 Search Clear Fiel	9606 zt2YzrYNNSUX	<pre>(L+ 5)</pre>
	9606 C3IJ2QQh SV0	GL+ 5
Search Blast Align Retrieve ID Mapping	9606 C3IJ2QQh SV0	GL+ 5
	9606 zHLUrNaS(SVC	GL+ 5
X Obsolete UniProtKB/Swiss-Prot O9Y2N3 (PO121_HUMAN)	9606 bKUivd80l SV0	GL+ 5
	9606 zt2YzrYNNSUX	(L+ 5
Since September 2, 2008. 🔂 History	9606 zt2YzrYNNSUX	(L+ 5
	9606 zt2YzrYNNSUX	(L+ 5
RDF/XML	9606 zHLUrNaS(SVC	GL+ 5
	9606 zt2YzrYNNSUX	(L+ 5
	9606 pWfOJSiI9SVC	GL+ 5
Demerged into A8CG34 and Q96HA1. [List]	9606 zHLUrNaS{SVC	GL+ 5
	9606 LJ90KG+02 SV0	GL+ 5
	9606 pWfOJSiI9SV0	GL+ 5
	9606 zHLUrNaS(SVC	GL+ 5
© 2002, 2009, UniProt Concertium I. 🦛 🧰 🏸 Unicance & Disclaimer I. Contact	9606 pWfOJSIISSVC	3L+ 5
	9606 LJ90KG+02SVC	3L+ 5
	9606 XZJNHUKK SVC	3L+ 5
		3L+ 3
	9606 pWIOJSIISSVC	
	5000 pwr05505	
)
Done		· · · ·

0.) 🖬 🤊	+ (°I +) ∓			hprd_1	eedback_3.0.txt -	Microsoft Excel					x
C	Home	Insert Page Lay	out Formulas	Data	Review View]					i 🕜 🗕 📼	x
Nor	mai Page Layout	Page Break Preview I Custom Views I Full Screen rkbook Views	Ruler Gridlines Message Bar	 Formula B Headings 	Gar Q Door 1009	Zoom to Selection	w Arrange Freeze dow All Panes *	Split Vie Hide ⊒‡ Syn Unhide ≠ Re Wind	ew Side by Side nchronous Scrolling set Window Position dow	Save Switch Workspace Window	h vs • Macros	
	Α2	• (6	A		A							×
	A	вс						к	L	M N	0	
1	int db	int acc source	$\neg \nu$					apped acc	c mapped taxid	rogid rogscore	score type	
2	HPRD	-10 HPRD 5))	11824	9606	0N0mKu4 PN	6	
3	HPRD	-10 HPRD_S						1736	5 40674	TISXIPnRe PN	6	
4	HPRD	-10 HPRD_SING	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
5	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
6	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
7	HPRD	-10 HPRD_SINGI	HPRD	1739	40674 HPR	D 1739	40674 HPRD	1739	40674	H0rBvLrRp PN	6	
8	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
9	HPRD	-10 HPRD_SINGI	HPRD	11824	9606 HPR	D 11824	9606 HPRD	11824	9606	0N0mKu4 PN	6	
10	HPRD	-10 HPRD_SINGI	HPRD	8177	40674 HPR	D 8177	40674 HPRD	8177	7 40674	Xjq2JFMH PN	6	
11	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
12	HPRD	-10 HPRD_SINGI	HPRD	1739	40674 HPR	D 1739	40674 HPRD	1739	40674	H0rBvLrRp PN	6	
13	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
14	HPRD	-10 HPRD_SINGI	HPRD	1739	40674 HPR	D 1739	40674 HPRD	1739	40674	H0rBvLrRp PN	6	
15	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
16	HPRD	-10 HPRD_SINGI	HPRD	1731	40674 HPR	D 1731	40674 HPRD	1731	L 40674	HKrkS34d) PN	6	
17	HPRD	-10 HPRD_SINGI	HPRD	1739	40674 HPR	D 1739	40674 HPRD	1739	40674	H0rBvLrRp PN	6	
18	HPRD	-10 HPRD_SINGI	HPRD	1731	40674 HPR	D 1731	40674 HPRD	1731	L 40674	HKrkS34d) PN	6	
19	HPRD	-10 HPRD_SINGI	HPRD	1731	40674 HPR	D 1731	40674 HPRD	1731	L 40674	HKrkS34d) PN	6	
20	HPRD	-10 HPRD_SINGI	HPRD	1745	40674 HPR	D 1745	40674 HPRD	1745	5 40674	O2b35fDq PN	6	
21	HPRD	-10 HPRD_SINGI	HPRD	1731	40674 HPR	D 1731	40674 HPRD	1731	L 40674	HKrkS34d) PN	6	
22	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
23	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
24	HPRD	-10 HPRD_SINGI	HPRD	1739	40674 HPR	D 1739	40674 HPRD	1739	40674	H0rBvLrRp PN	6	
25	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	
26	HPRD	-10 HPRD_SINGI	HPRD	1736	40674 HPR	D 1736	40674 HPRD	1736	5 40674	TISXIPnRe PN	6	-
• •	→ → hp	rd_feedback_3.0 🥂	J				I ∢ [1				
Rea	dy						Average: 80	35.75 Count: 15	Sum: 64286	100% 🕞 📃) .::

Databases now have these feedback files, along with a set of suggestions on how to handle issues

We will then rebuild iRefIndex (5.0) and compare to build 2.0 starting in early July.

Cytoscape plugin



How can we distinguish High throughput interactions from Low throughput interactions?

Interaction records are supported by Publications (PMIDs)



High PMID reuse indicates htp paper



PMID reuse is 100

Low PMID reuse inidcates Itp paper



PMID reuse is 100

PMID reuse is 2

Low lpr (lowest pmid reuse) indicates Interaction supported by ltp paper.



Lowest PMID reuse is 2
Lowest PMID reuse: lpr score.



Cytoscape plugin



Welcome to Cytoscape 2.6.2

Right-click + drag to ZOOM

Middle-click + drag to PAN

Cytoscape plugin

Cytoscape Desktop (Session Name: ber_short_lpr.cys)						
File Edit View Select Layout Plugins Help						
► 🖬 Q Q Q Q 🔯 🕃 🛱 Search: 🔽 🔽						
Control Panel 🖆 🥵 irefIndex_beta4						
telitor Filters						
Network Nodes Edges						
refIndex_beta4 366(0) 149(149)						
	00/					
		$\bullet \bullet \bullet \bullet \bullet \bullet \bullet$				
Data Panel						6
	i.PMID	i.detection	i.experiment	i.lpr △	i.source	
2920195 (pp) 4303804	12874288	confirmational text mining	HPRD Text Mining Confirmation	10	ophid	
2920195 (pp) 4157895	12874288	confirmational text mining	HPRD Text Mining Confirmation	10	ophid	
844606 (pp) 1613993	-1	coip	Parker-2001a-2	9	ophid	
498749 (pp) 844606	-1	coip	Parker-2001a-2	9 /	ophid	-
498749 (pp) 1013993	-1		Parker-2001a-2	<u> 9</u> /	ophia	-
618017 (pp) 844606	-1	coip	Parker-2001a-2	- <u> </u> 9	ophid	-
844606 (pp) 1613993	-1		Parker-2001a-2	9	ophid	-
618017 (pp) 844606	-1	coip	Parker-2001a-2	9	ophid	~
Node Attribute Browser Edge Attribute Browser Net	work Attribute Browser	· · · · · · · · · · · · · · · · · · ·				

Welcome to Cytoscape 2.6.2

Right-click + drag to ZOOM

Middle-click + drag to PAN

Take home message

One of the concerns in the Cusick paper is the presence of HTP interaction data together with literature curated data.

These data types can be easily distinguished and filtered using either the 1)Cytoscape plugin 2)PSICQUIC web services 3)iRefIndex MITAB files

😻 about iRefWeb: a i	non-redundant inde	x for protein intera	ction data - Mozilla I	Firefox						
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> to	ry <u>B</u> ookmarks <u>T</u> ools	<u>H</u> elp								
🔇 🔊 • C 🔾	< 🏠 🔛 http:/	//wodaklab.org/iRefWeb/					☆ •	G• Google		> 🔒
🔎 Most Visited 옹 Entre:	z PubMed 🛄 Library of	Medicine an 💴 The E	iotechnology Ce 💵 B	BC - bbc.co.uk home	Science/AAAS Mu	ltim 📄 The Research (Council W Wikipedia	🔊 BBC - Weather Cen	tre	
📸 Cytoscape: Analyzing	and Visualizing N 🖂	🚻 about iRefWeb: a	non-redundant 🔀							•
	wodo	ık lab +	Donalds Researc	on h Group			search site	~		^
	home	info +	people	projects +	tools +	resources +	sponsors	papers		
				MERAN	Contraction of the second			A SHE WAS		X
	wodak lab » tools	⊧ » iRefWeb ^{8ETA}	»							
	about	search	statistics	pubmed re	eports	current d	ata source: iRefInde>	_beta2		
		7								

about iRefyreb: a non-redundant index for protein interaction data

What is iRefWeb?

iRerWeb provides a web-based query interface to a relational database containing the latest build of the interaction Reference Index.

The interaction Reference Index (or iRefIndex) integrates protein interaction data from six different interaction databases: **BIND**, **BioGrid**, **IntAct**, **MINT**, **MPPI**, **and OPHID**. This integration is achieved through a rigorously documented procedure for mapping protein IDs across databases, so as to enable systematic backtracking of the links used to establish the identity of the interaction partners. The procedure furthermore assigns unique and global identifiers to protein interactors, interactions and complexes, which may be used by anyone to facilitate exchange and consolidation of data.

The iRefWeb interface thus groups together interaction records from the different databases that all support the same protein interaction or complex. It hence enables a non-redundant view of the data and may be used to locate interaction data for a protein of interest regardless of which interaction database(s) contain the information.

The user may then follow links to the source interaction databases to examine



🕲 search iRefWeb: a n	on-red	undant inde	x for protein inte	raction data - N	lozilla Fire	efox							
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory	<u>B</u> ookr	narks <u>T</u> ools	Help										0
🔇 🖻 • C 🗙	☆	http://v	vodaklab.org/iRefWeb)/searchable?q=QCI	R2_HUMAN&	binaryOnly=	•on			☆ •	G• Google		🔎 🔒
🔎 Most Visited 옹 Entrez P	ubMed	uio Library of M	edicine an 💴 The	Biotechnology Ce	888 BBC -	bbc.co.uk h	ome <mark>S</mark> Sci	ence/AAAS Mul	ltim 📄 The Research 🤇	Council 👿 Wikipedi	a 🔝 BBC - Weather	Centre	
Cytoscape: Analyzing ar	nd Visuali	izing N 🖂 📔	🚻 search iRefWeb	: a non-redundar	🛛 📃								•
	\mathbf{W}	odo	ik lab	+ Done Rese	aldso arch	n Grou	р			search site		×	
	hom	е	info +	people	р	rojects	+ to	ols +	resources +	sponsors	papers		
Martin Street		10 ×		100 DA	Sex 1	6022		e No					
	wodak	lab » tools	» iRefWeb ^{BE}	74 _»									
	al	bout	search	statisti	cs	pub	med report	s	current	data source: iRefIn	dex_beta2		
	QCR2 bina in	_HUMAN ary interaction teractior he blue arrows	is only ns (5 match) s to expand or coll	es)	Searc	h Vie	w help on se	earch terms a	nd usage				
		name				💷 np	🔳 lpr	🔳 hpr		(rigid search score)		
	•	interaction inv QCR2_HUMA	volving RTN4_HUM	IAN and 1AN		1	6	6	m	s3Onuj9UuNgMk4∨	hwRGyEy315U (0.8192345)		
		interaction inv Q6FGD1_HUI B3KSN4_HUI	volving QCR7_HUM MAN and QCR2_H MAN	1an Iuman		0	O	0		8x/wk0azd3r2VJx	:+lqDNrQsTKMI (0.8192345)		
		hla_b-uqcrc2				1	8292	8292	(1xRkOTwF4IemINB	IHZ/W2CVNvq4 (0.6057474)		
		uqoro1-uqoro2	2			1	8292	8292		OEKpetTq1INGUw	tZZb∀Xg5/061k (0.6057474)		
		arf6-uqcrc2				1	8292	8292		xXa5SnqT5VqYK2j	onk9XGtlZxLVU (0.6057474)		

🤓 search iRefWeb: a i	non-re	edundant index	for protein inter	action data - Mozill	a Firefox									
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tor	у <u>В</u> о	okmarks <u>T</u> ools į	<u>H</u> elp											4
< ≥ - C ×	< 👍	http://wo	odaklab.org/iRefWeb,	/searchable?q=QCR2_HL	IMAN&binary	Only=on					☆ •	G• Google		<i>P</i>
🔎 Most Visited 😂 Entrez	PubMe	d 💴 Library of Me	dicine an 💴 The	Biotechnology Ce 💵	BBC - bbc.co	.uk home	S Science	ce/AAAS Multim	📄 The Researc	h Council	W Wikipedi	a <u> </u> BBC - Weath	er Centre	
📸 Cytoscape: Analyzing a	and Visi	ualizing N 🗵 🚦	search iRefWeb:	a non-redundan 🗵										
	W	odak	k lab ·	+ Donalds Researc	son h Gro	oup				seard	∶h site		×	
	hom	ne i	nfo +	people	projec	ts +	tools	+ re	sources +	spon	sors	papers		
						AF6	4	Network			× 11	11 A 14 1 1 1		X e
	wodal	(lab » tools »	iRefWeb ^{BET}	4 »										
	a	bout	search	statistics	F	oubmed r	eports		current	data sour	ce: iRefInde	ex_beta2		
	sea	rch iRefV	Veb: a no	n-redundar	nt inde	ex for	prote	ein intera	ction da	ita				
ſ	0000	1.0.0.4451					1							
l		:_HUMAN		2	earcn	viewneip	on searc	n terms and usa	ige					
		ary interactions t	, in y											
	🕨 ir	nteractions	(5 matche	s)										
	Click t	the blue arrows to	o expand or collap	ose search results.	1	101		I hore				🔲 ri ai d		
		name			in the		ipi 🔹	чпр			(s	earch score)		
	V	interaction invo	olving RTN4_HU	MAN and	1		6	6	ms	:30nuj9Ui	ıNgMk4Vh	wRGyEy315U		
		QCR2_HUMAN	B3KSN4_HUMA	N								(0.8192345)		
			interactors					eneries		typo	cardin	ality		
			QCR2_HUM	IANIB3KSN4_HUMAN				Homo sapiens	3	protein	1	anty		
			RTN4_HUM.	AN				Homo sapiens	3	protein	1			
			evidence											
			interaction	source db (id)		interac	tion type	e name (descrip	ption)		psi-n	ni code		
			ophid (1710))196)		aggreg	ation (NA)				MI:019	91		
	•	interaction involv Q6FGD1_HUMA B3KSN4_HUMA	ing QCR7_HUMA N and QCR2_HU N	N MAN	٥	I	0	0		8x/wk0;	azd3r2∨Jx+	HqDNrQsTKMI (0.8192345)		
		hla_b-uqcrc2			1	8	292	8292		dxRkOTw	F4lemINBIH	HZ/W2C//Nvq4		

🥹 search iRefWeb: a	non-r	edundant index	for	protein interaction data - Mozilla Firefox				
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> to	ry <u>B</u> o	okmarks <u>T</u> ools	<u>H</u> elp					
$\langle \mathbf{D} \cdot \mathbf{C} \rangle$	< /	http://w	odakla	ab.org/iRefWeb/searchable?g=OCR2_HUMAN&binaryOr	ly=on		גֹז י ג י Google	P [S]
🔎 Most Visited 🥳 Entrez	z PubMe	d <u>Uio</u> Library of Me	dicine	e an 🛄 The Biotechnology Ce 💷 BBC - bbc.co.u	k home 📐 Science/AAAS Multim 🔝 The R	esearch Council	W Wikipedia 🔊 BBC - We	ather Centre
Cytoscape: Analyzing	and Vis	ualizing N 🖂 🚦	sea	arch iRefWeb: a non-redundan 🗵 📃				
		namo					(search score	
	•	interaction inv QCR2_HUMAN	olvin B3ł	ng RTN4_HUMAN and 1 KSN4_HUMAN	6 6	ms30nuj9U	uNgMk4VhwRGyEy315U (0.8192345	J)
			int	teractors				
				name	species	type	cardinality	
			•	QCR2_HUMAN B3KSN4_HUMAN	Homo sapiens	protein	1	
				interactor aliases				
				alias	name space			
				QCR2 HUMANIB3KSN4 HUMAN	Display name			
				AAH00484.1 妃	GB			
				AAH03136.1 🚱	GB			
				EAW50592.1 🚱	GB			
				hsa:7385	KEGG			
				P22695 🚱	PROTEIN Accession			
				UQCR2	Short label			
				Ubiquinol cytochrome c reductase core p	rotein II Short label			
				uqcr2_human	Short label			
				12653427 🚱	gi			
				13111931 🚱	gi			
				50592988 🗗	gi			
				NP 003357 🚱	refseq			
				NP_003357 🚱	refseq			
				B3KSN4 🗗	uniprot			
				P22695 🚱	uniprot			
				show/hide interactor mappings and assi	anments			
				RTN4 HIIMAN	Homo saniens	nrotein	1	-
			Ľ.	•••		protoni		
			evi	interaction source db (id)	interaction type name (description)		nsi-mi code	
				ophid (1710196)	aggregation (NA)		MI:0191	
			-					







In either case, multiple interactions are recorded for the same publication where only one interaction is described.



In either case, multiple interactions are recorded for the same publication where only one interaction is described.



How common is this?

🕲 pubmed reports iRefWeb: a non-redundant index for protein interaction data - Mozilla Firefox	
Eile Edit View Hi <u>s</u> tory Bookmarks Tools <u>H</u> elp	
🕢 💽 🗸 🏠 🔛 http://wodaklab.org/iRefWeb/pubReport/list?selectedSourceDbId=0&pubMedId=7535773%0D%0A&overviewSubmit=generate+pubmed+report 🏠 🔹 💽 🕻 Google	P 🔒
🚈 Most Visited 🔗 Entrez PubMed 🔟 Library of Medicine an 🔟 The Biotechnology Ce 🚥 BBC - bbc.co.uk home <mark>S</mark> Science/AAAS Multim 🗋 The Research Council W Wikipedia 脑 BBC - Weather Centre	
🚻 pubmed reports iRefWeb: a non 🔯 🛛 🚻 compare interactions for pubmed id # 7 🔀	•
	* (*) *

wodak lab » tools » iRefWeb^{BETA} »

about	search	statistics	pubmed reports	current data source: iRefIndex_beta2

pubmed reports iRefWeb: a non-redundant index for protein interaction data

PubMed reports provides a way to see which interactions were noted across different source databases for each PubMed ID. This comparison can help to show to what extent the different databases may or may not differ in their interpretation of the same publication. Such differences can have several origins. A frequent origin might be the incensistencies in the choice (often arbitrary) of protein splice forms. Despite the best attempts to find and merge duplicate interaction records across databases, such inconsistencies cannot be reliably resolved at present. Genuine differences and biases in the interpretation of the same reported data may also contribute. We are currently carrying out systematic analyses of the identified discrepancies in order to gain further insight into these issues. Until such insights are obtained, finding that a database failed to record some interactions from a given paper that another database did capture, does not necessarily indicate that there has been an error in either database

Select a target database:

all databases 🔽

or enter specific pubmed id(s)

> 7535773

If no target database is selected—that is the option is left to its default value of "all databases"—then records for every PubMed ID will be shown.

When a target database is selected its number of interactions annotated per PubMed ID are compared against all other databases but results are displayed only if the selected database has possibly missed interactions for a PubMed ID. Thus if the target database has the same number of interactions for a given PubMed ID as found by all other databases, that PubMed ID is skipped in the report, as it is assumed the target noted all possible interactions. If the target database found no interactions, it is assumed that it did not annotate that PubMed ID, and is also skipped.

generate pubmed report

To search on specific PubMed IDs enter the desired IDs in the text input field.

pubmed report — click the blue arrow to expand or collapse the details for each PubMed ID

pubmed id	total	number of	finteractions	s each datab	ase found fo	r the given p	ubmed id
	interactions found	bind	grid	intact	mint	ophid	MPPI
7535773 🔗	5		2		3		

ڬ pubmed reports iRefWeb: a non-redundant index for protein interaction data - Mozilla Firefox	
Eile Edit View History Bookmarks Iools Help	
🕢 💽 - C 🔀 🏠 http://wodaklab.org/iRefWeb/pubReport/list?selectedSourceDbId=0&pubMedId=7535773%0D%0A&overviewSubmit=generate+pubmed+report 🏠 🔹 Google	P 🔒
🚈 Most Visited 😂 Entrez PubMed 🚾 Library of Medicine an 🚾 The Biotechnology Ce 🚥 BBC - bbc.co.uk home S Science/AAAS Multim 🗋 The Research Council W Wikipedia 🔝 BBC - Weather Centre	
🚻 pubmed reports iRefWeb: a non 🔯 🛛 🚻 compare interactions for pubmed id # 7 🖂	•
interaction records across databases, such inconsistencies cannot be reliably resolved at present. Genuine differences and biases in the interpretation of	^

the same reported data may also contribute. We are currently carrying out systematic analyses of the identified discrepancies in order to gain further insight into these issues. Until such insights are obtained, finding that a database failed to record some interactions from a given paper that another database did capture, does not necessarily indicate that there has been an error in either database

Select a target database:

all databases 🗸

or enter specific pubmed id(s)

7535773

If no target database is selected—that is the option is left to its default value of "all databases"—then records for every PubMed ID will be shown.

When a target database is selected its number of interactions annotated per PubMed ID are compared against all other databases but results are displayed only if the selected database has possibly missed interactions for a PubMed ID. Thus if the target database has the same number of interactions for a given PubMed ID as found by all other databases, that PubMed ID is skipped in the report, as it is assumed the target noted all possible interactions. If the target database found no interactions, it is assumed that it did not annotate that PubMed ID, and is also skipped.

generate pubmed report

To search on specific PubMed IDs enter the desired IDs in the text input field.

v pubmed report — click the blue arrow to expand or collapse the details for each PubMed ID

pubmed id	total	numbe	r of interactions	; each database	found fo	r the give	en pubmed id		
	found	bi	nd grid	intact	mint	opl	hid MPPI		
pubmed id		total	bind	grid		intact	mint	ophid	MPPI
7535773 🗗		5		2			3		
	int#10754						1		
	int#34508			1					
	int#155381			1					
	int#271166						1		
	int#330529						1		
	V	liew a si	de-by-side con	parison of the	interacti	ions for	pubmed 75357	73	

🐸 compare int	eractions for pubmed id # 75357	73 - Mozilla Firefox				
<u>File E</u> dit <u>V</u> iew	History <u>B</u> ookmarks <u>T</u> ools <u>H</u> elp					*** ***
<>	C 🗙 🏠 🚻 http://wodakla	b.org/iRefWeb/compareInteractions/list?pubr	ned=7535773&ints=10754%2C34508%2C1	55381%2C271166%2C330529	☆ • Google	P 🔒
Most Visited 🗧	Entrez PubMed 💴 Library of Medicine	an 💴 The Biotechnology Ce 💷 BBC	- bbc.co.uk home <u>S</u> Science/AAAS Mu	tim 📄 The Research Council W Wik	ipedia 🔊 BBC - Weather Centre	
W pubmed repor	rts iRefWeb: a non-redun 🖂 🚻 col	mpare interactions for pubme 🛛				•
WOC	dak lab + R	onaldson esearch Group		search site	<u>×</u>	
home	info + peop	le projects + too	ols + resources +	sponsors papers		
		HAR BUILDING				
wodak lab »	tools » iRefWeb ^{BETA} »					
about	search sta	tistics pubmed report:	s current da	ta source: iRefIndex_beta2		
compare	e interactions for p	ubmed id # 7535773				
Version			an and for the birds for the same			
view pubmed		To expand an interaction or interaction	rr, and [–] to nidercollapse.			_
	[-]	[-]	[-]	[-]	[-]	
Interaction	10754 shc1-jak2	34508 interaction involving Q9HAU2_HUMAN and EG6464	155381 interaction involving EG6464 and JAK2_HUMAN	271166 shc1-jak2	330529 shc1-jak2	
Source DB(s)	mint	grid	grid	mint	mint	
Interactors	[+] Interactor: 1 JAK2_MOUSE	[+] Interactor: 1 EG6464	[+] Interactor: 1 JAK2_HUMAN	[+] Interactor: 1 GRB2_MOUSE Q3U5I5_MOUSE	[+] Interactor: 1 SHC1_HUMAN	
	[+] Interactor: 2 SHC1_HUMAN	[+] Interactor: 2 Q9HAU2_HUMAN	[+] Interactor: 2 EG6464	[+] Interactor: 2 SHC1_MOUSE	[+] Interactor: 2 GHR_MOUSE	
				© 2008 - 2009 Wodak Lat		
/				© 2000, 2009 Woldan Lat		
<						>

🕴 compare in	teractions for pubmed i	id # 75357	73 - Mozilla Firefox				
<u>File E</u> dit <u>V</u> iev	v Hi <u>s</u> tory <u>B</u> ookmarks <u>T</u> i	ools <u>H</u> elp					
$\langle \rangle$	С 🗙 🏡 🎹 ы	ttp://wodaklab	o.org/iRefWeb/compareInteractions/list?pub		155381%2C271166%2C330529	St • G• Google	
					ukin 🗋 The Dessenth Coursel - W/ 118		
	S Entrez Publied <u>de</u> Library	y or medicine a	an 📴 The Biotechnology Ce 🛥 BBC	DDC.CO.UK NOME D SCIENCE/AAAS MU	uidm 🔄 The Research Council 🗰 wi	kipedia 🔝 bbc - weather Centre	
m pubmed repo	orts ikerwed: a non-redun (or Con	npare interactions for pubme 🔯 🗌				
WOC	iak lab -		nalason Search Group		search site	×	
home	info +	neonle		ls + resources +	enoneore nanere		
nome		heobie			sponsors papers		
							
vodak lab » t	tools » iRefWeb ^{BET,}	^A »					
about	search	stati	stics pubmed reports	current dat	a source: iRefIndex_beta2		
compare	e interactions	for pul	bmed id # 7535773				
compare		ioi pu					
View pubmed	id #7535773 🗗 Click	on the [+]	to expand an interaction or interacto	r, and [–] to hide/collapse.			
	[-]		[-]	[-]	[-]	[-]	
Interaction	10754		34508	155381	271166	330529	
	shc1-jak2		interaction involving	interaction involving EG6464 and	shc1-jak2	shc1-jak2	
			Q9HAU2_HUMAN and EG6464	JAK2_HUMAN			
Source DB(s)	mint		grid	grid	mint	mint	
Interactors	[-]		[+]	[+]	[+]	[+]	
	Interactor: 1 .IAK2_MOUSE	-	Interactor: 1 EG6464	Interactor: 1 .IAK2_HUMAN	Interactor: 1 GRB2_MOUSEI03U5I5_MOUSE	Interactor: 1 SHC1_HUMAN	
	aliases	-	200101				
	JAK2_MOUSE D	lisplay	[+]	[+]	[+]	[+]	
	Q62120 샵 P	ROTEIN	Interactor: 2 09HAU2_HUMAN	Interactor: 2 EG6464	Interactor: 2 SHC1_MOUSE	Interactor: 2 GHR_MOUSE	
	162 S.	hort	dellage_lowload	200101		SHIT_NOOSE	
	EG16452 S	hort					
	jak2_mouse S	nort					
	Q62120 G∾ U. [_]	niprot					
	db primary referen	nce ids					
	mint (int id# MINT-27383	(06):					
	Q62120 P	ROTEIN					
	mint /intid# MINT 27202	1601-					



15,813 PMIDs are cited by 2 or more databases



Only 10% of co-cited PMIDs have complete agreement with respect to number of interactors or number of interactions.



Co-cited papers were classified into several overalpping categories.



Average Jaccard coefficient for interactions inside and outside of multiple categories



Percentage of papers with identical coannotations inside and outside of several categories.

HTP, isoforms, complexes and higher eukaryotes are problematic.



Frequency distribution of Jaccard coefficient for interactions, computed for the three largest categories of papers reporting on interactions in human, mouse and yeast.











🕹 iRefWeb: pubmed detail for id # 15640796 - Mozilla Firefox	
Eile Edit View History Bookmarks Tools Help	4 ⁴ 4 9 ₄ 4
C X 🏠 Http://wodaklab.org/iRefWeb/compareInteractions/list?pubmed=15640796 🏠 🔹 💽 Google] 🗗 🔒
🙍 Most Visited 😒 Entrez PubMed 🔟 Library of Medicine an 🔟 The Biotechnology Ce 🕮 BBC - bbc.co.uk home <mark>S</mark> Science/AAAS Multim 🗋 http://bioscape.uio.n	**
🔊 Start Stumbling or 🔑 Sign-in	
Construction 90817 DAD2_YEAST DAD4_YEAST bind interaction 341994 SPC34_YEAST DAM4 YEAST	
Image: Second	

😻 iRef\	Web:	pubm	ed det	ail for	id #	15	5407	96 - N	Mozill	a Fire	efox															
<u>F</u> ile <u>E</u> o	dit <u>V</u> i¢	ew H	i <u>s</u> tory	<u>B</u> ookm	arks	<u>T</u> oo	ols <u>t</u>	<u>t</u> elp																		1.1
$\langle \rangle$	> -	C	×			htt	p://wo	daklab.	.org/iR	efWeb	o/compa	areInter	ractions/l	íist?pubrr	ned=156	40796		☆	•	- Goo	ogle		ρ	d.	8	5
🔎 Most	Visited	S Er	trez Pu	bMed 🛛	Jio Lit	orary	of Me	dicine a	in Ui	o The	Biotec	hnology:	y Ce 💵	BBC -	bbc.co.	uk home.	<mark>S</mark> Scie	ence/AAA	S Multim	🗋	http://bi	ioscape.u	iio.n		>	>
到 Start	Stumbl	ng	or 🌽	Sign-ir	٦																					
				je je na se je																					Į	
Done																									3	~



Current hypothesis

Disagreements between databases may be largely accounted for by different choices of

1)Splice isoform2)Taxon

Current hypothesis

But other causes of disagreement include

Minor sequence variations (e.g. UniProt vs. RefSeq)
 Different update mechanisms by int db's.
 Different curation policies on what to curate
 Spoke vs. Matrix representation.
 Simple curator error.

Take home message

This is a multi-factorial problem.

We want to clean up as many of the mechanical causes using changes implemented in the feedback study.

Then we will re-examine remaining causes of disagreement through a second feedback paper.

In the mean time, iRefWeb is there to look at.

Acknowledgements

Ian Donaldson Sabry Razick George Magklaras Paul Boddie Katerina Michalickova Antonio Mora <u>Shoshana Wodak</u> Brian Turner Emerson Cho Kyle Morrison Andrei Turinsky

http://irefindex.uio.no http://www.cytoscape.org/

ian.donaldson@biotek.uio.no

http://wodaklab/iRefWeb

brian@wodaklab.org