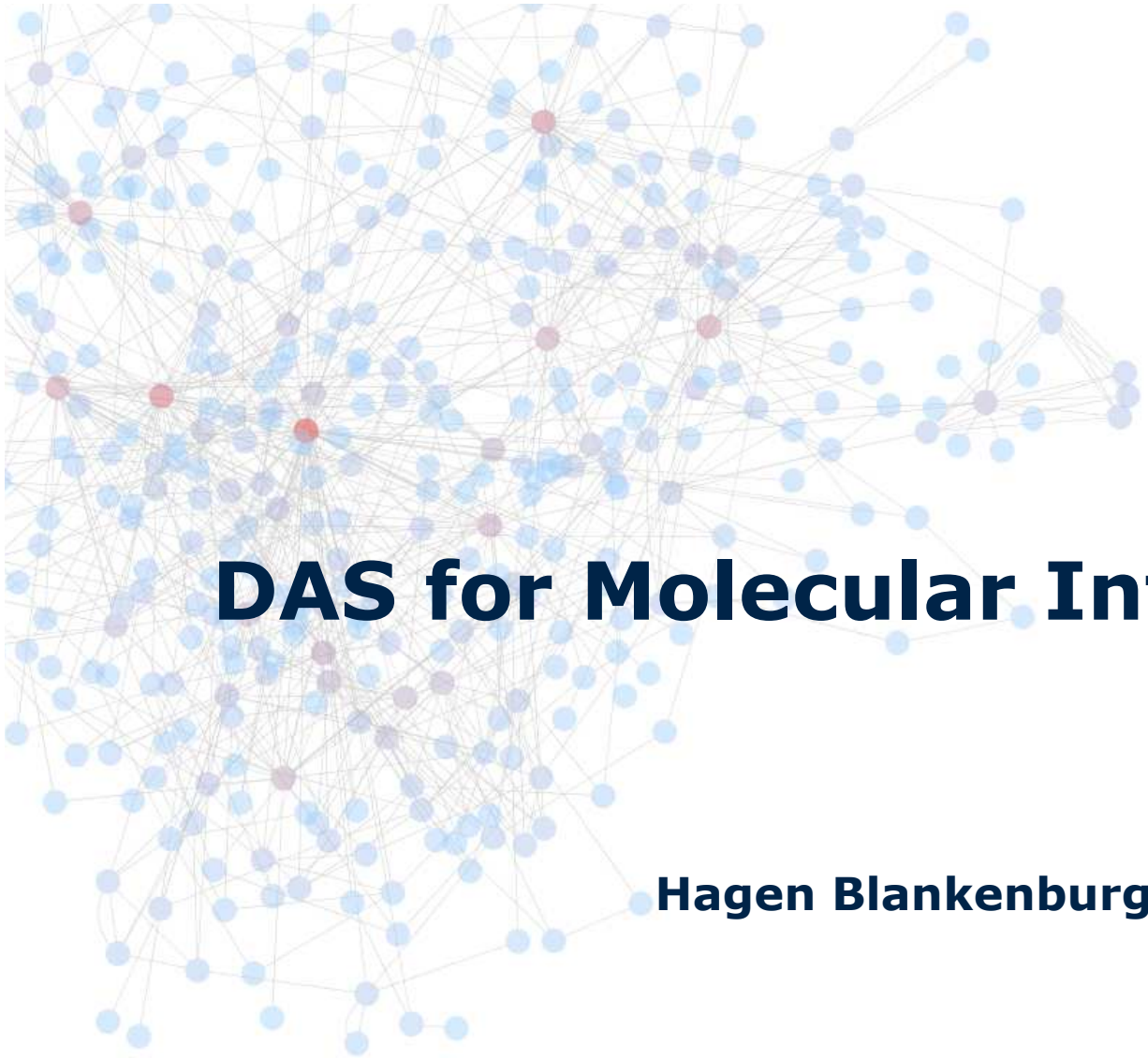


DAS for Molecular Interactions

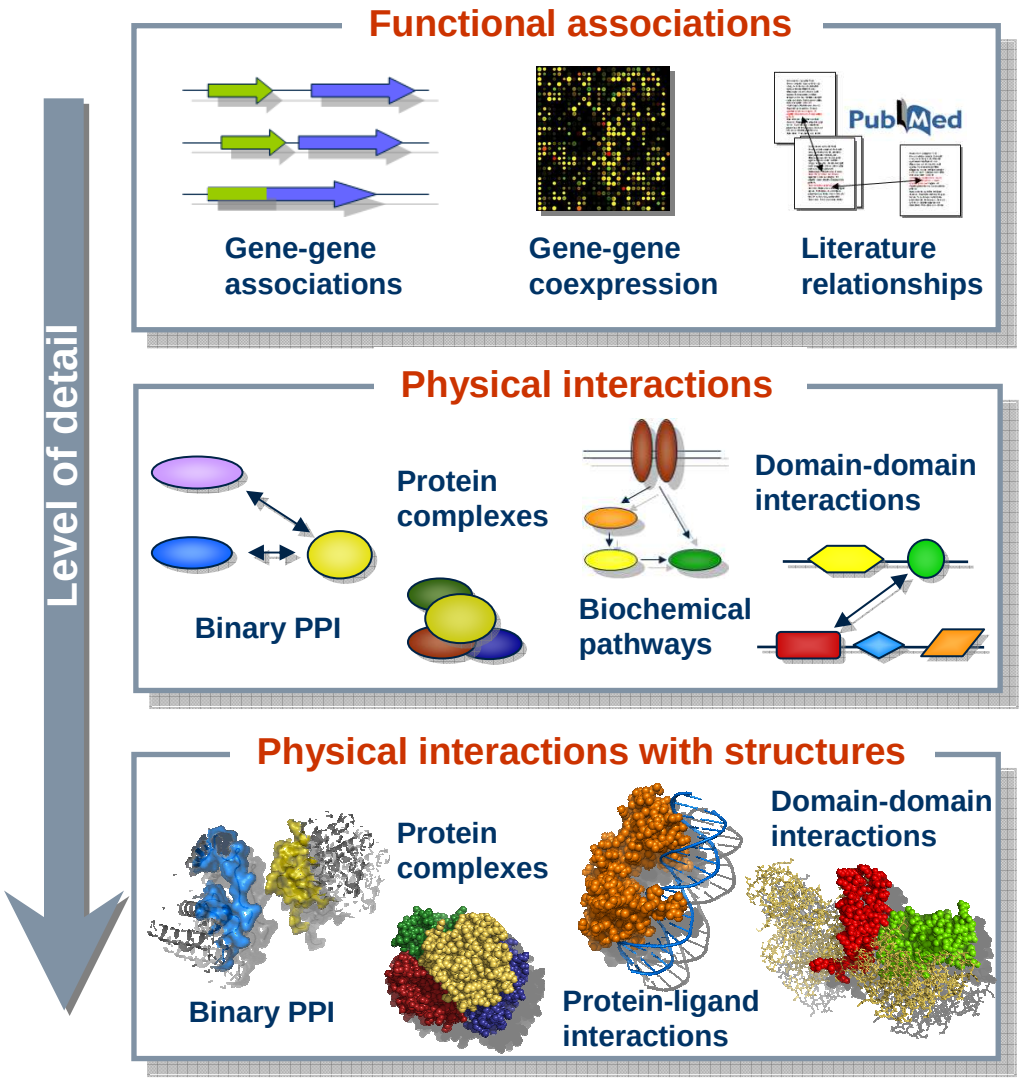
Hagen Blankenburg

max planck institut
informatik

MAX-PLANCK-GESELLSCHAFT

Importance of molecular interactions

- Fundamental for understanding of cellular processes
  - Prediction of protein function
  - Importance in certain diseases
  - Essentiality of hub proteins?
- The next big thing?!
  - „Interactomics“, „network medicine“, …



Problem I: Data abundance and distribution

Proteins Interacting in the Nucleus

3DID - 3D interacting domains

max planck institut informatik

PIN Database

PPID Protein-Protein Interaction Database

PSIBASE

PIPs Human Protein-Protein Interaction Prediction

predictome

BOND Biomolecular Object Network Databank

BIND Inside

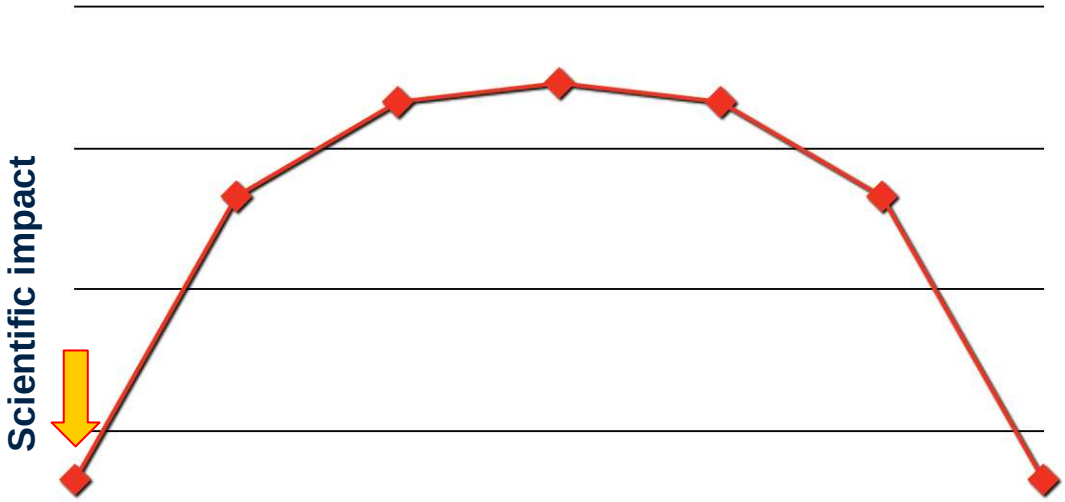
IntAct

HiMAP HUMAN INTERACTOME MAP

www.HPID.org

PSIMAP

Domino
domain peptide interactions database



STRING 8.0

MPact Protein Interaction & Complex DB

(Credit: Tim Hubbard)

Bioverse

OPHID

Too little bioinformatics

Too many databases

Too diverse interfaces

Human Protein Reference Database

PDZBase

MINT a Molecular INTeractions database

DIMA domain interaction map

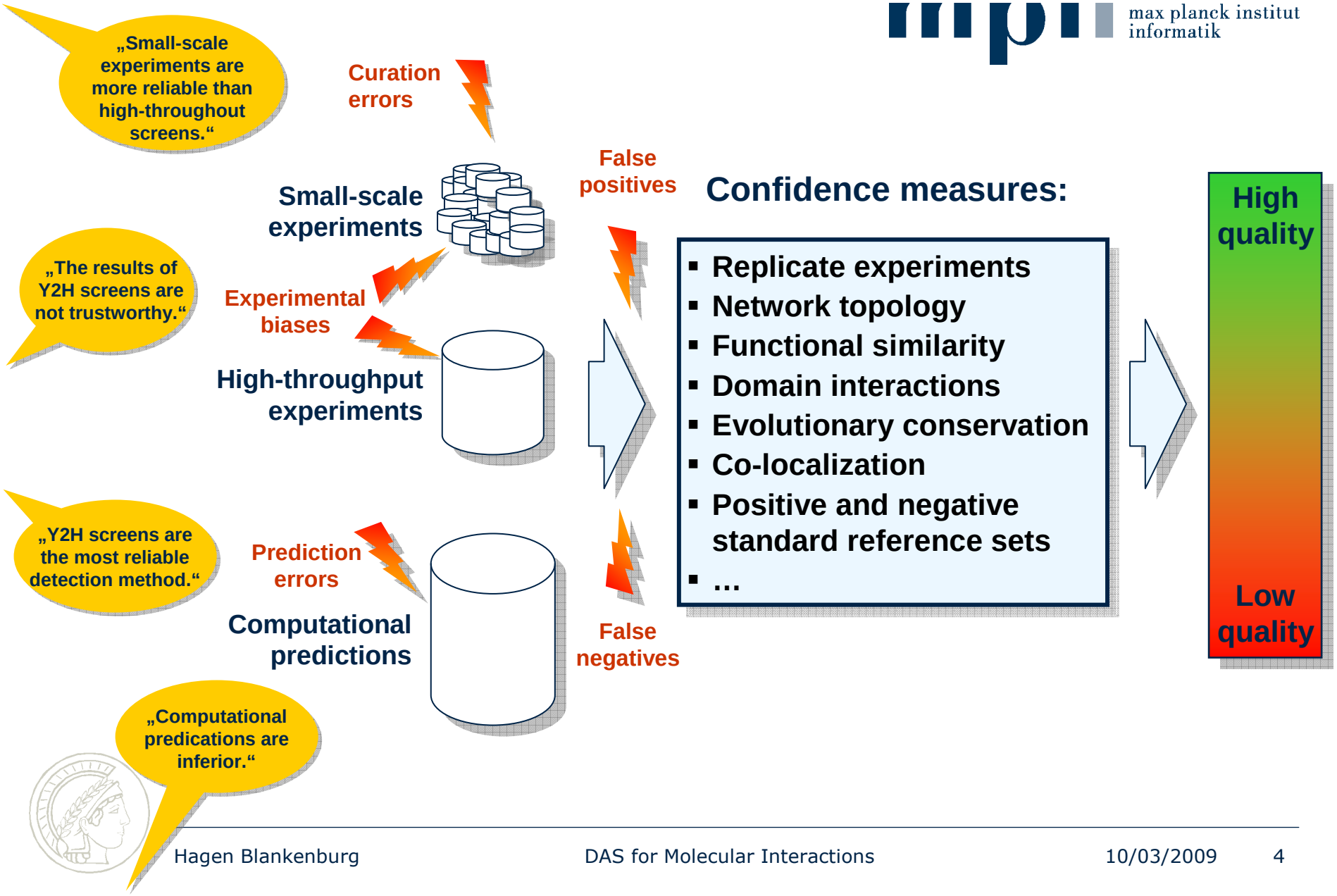
HomoMINT: an inferred human network

Human Annotated and Predicted Protein Interaction (HAPPI)

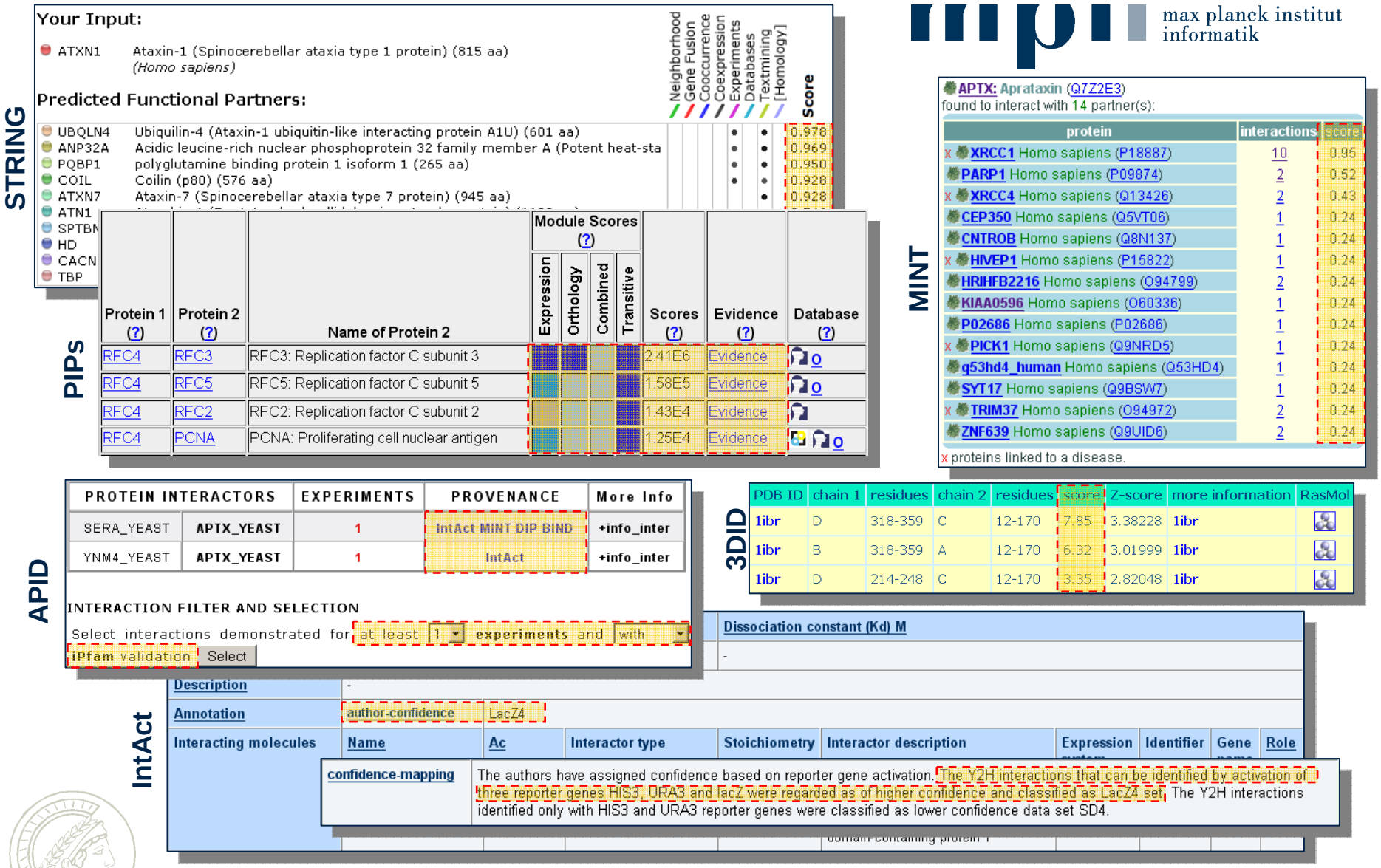
InterDom Database of Interacting Domains

Human protein-protein interaction network database search

Problem II: Data quality



Problem II: Data quality



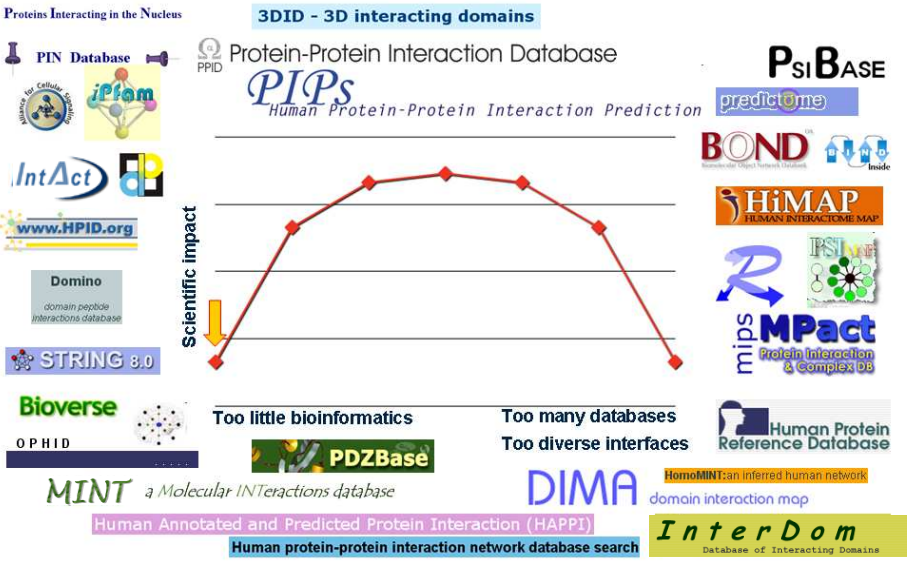
Solution: Distributed System

max planck institut informatik

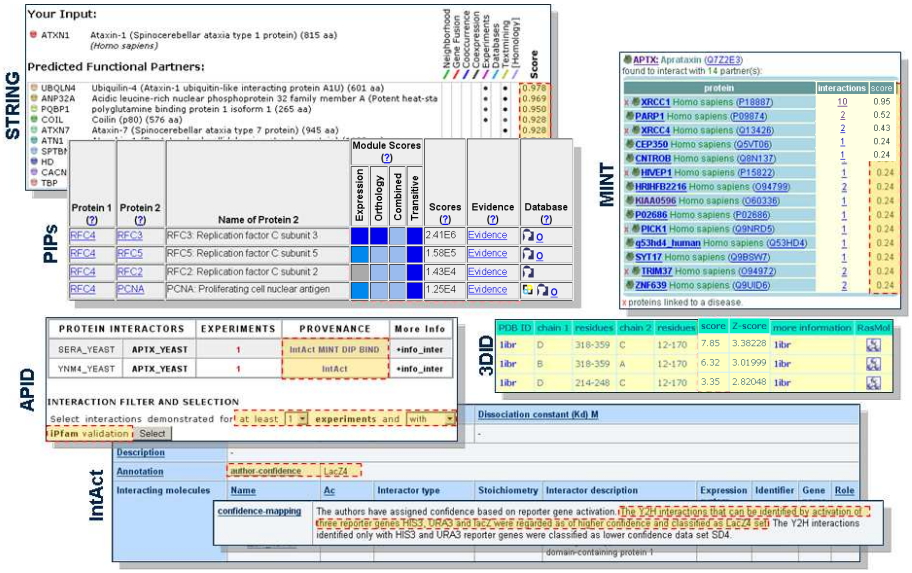
Interaction data servers

Interaction confidence scoring servers

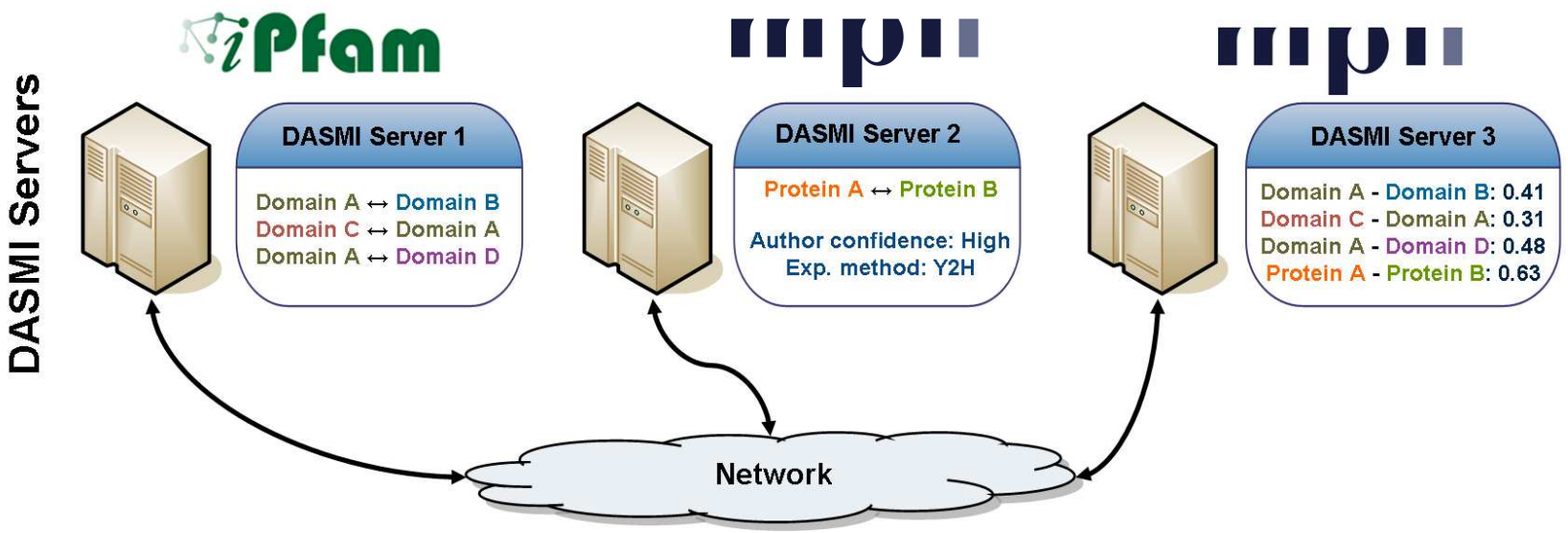
Problem I: Data abundance and distribution



Problem II: Data quality



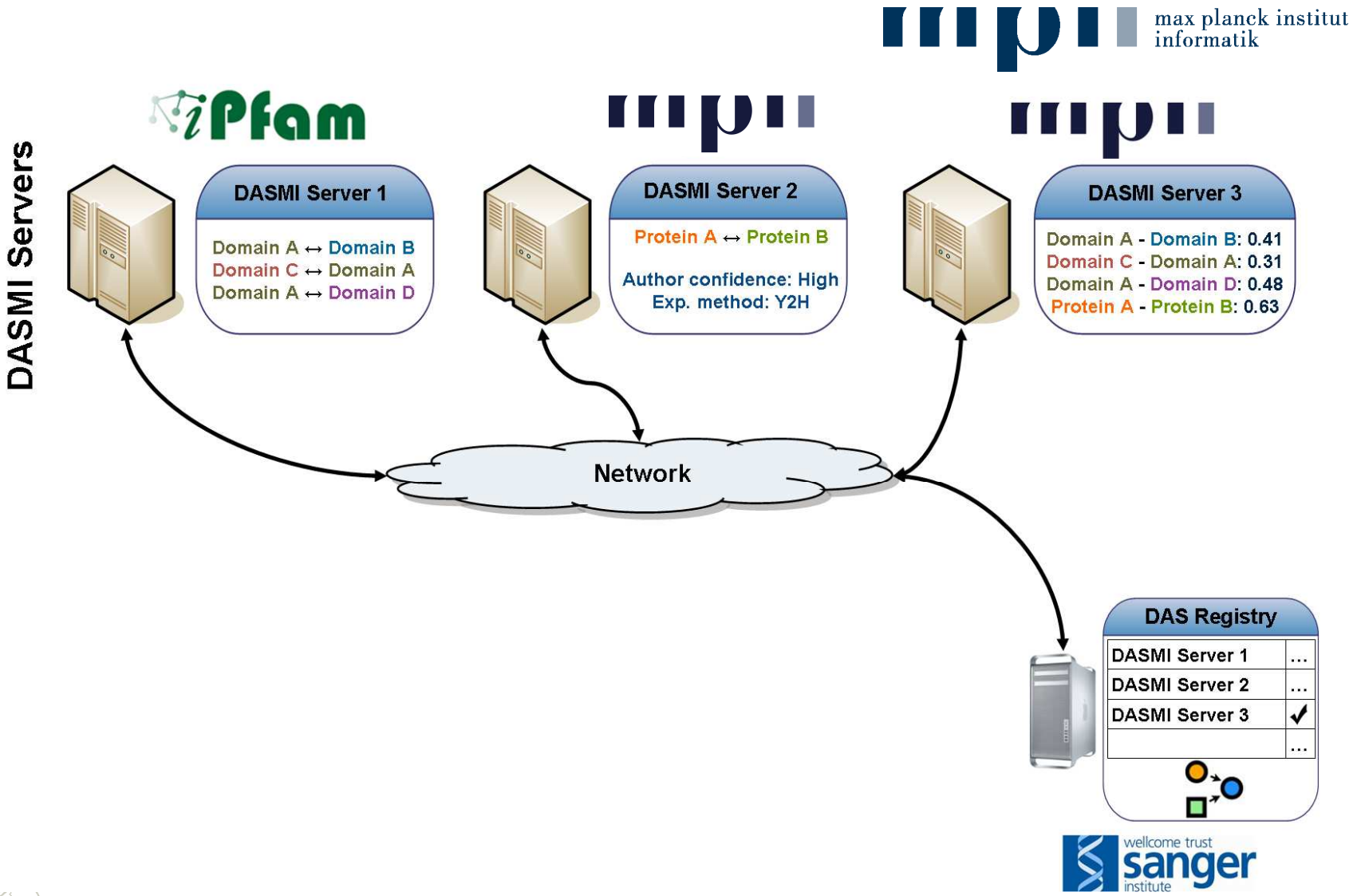
DAS for Molecular Interactions (DASMI) - Servers



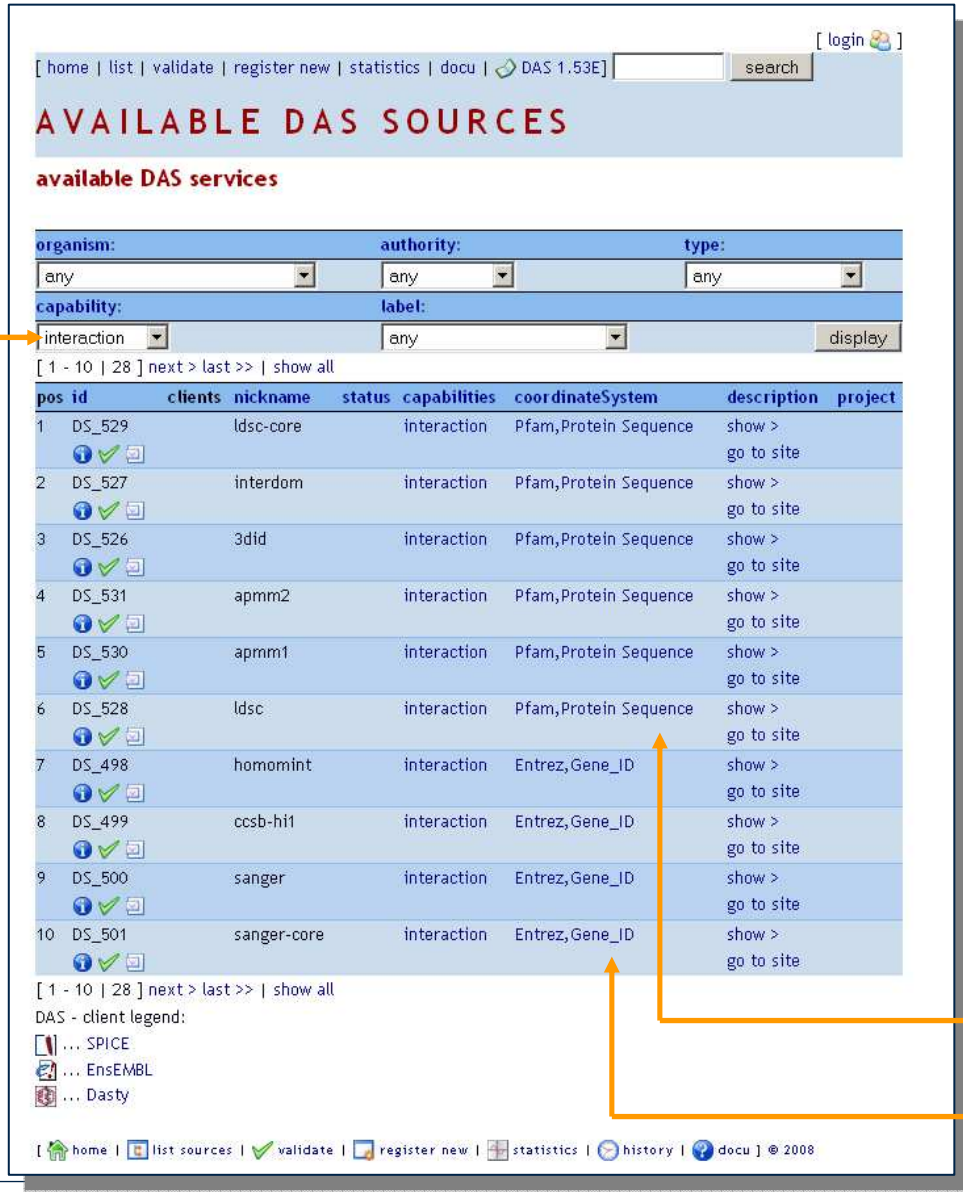
- Servers have coordinate / identifier systems

UniProtKB:	**P51587, BRCA2_HUMAN**
Entrez Gene:	**675**
GeneInfo:	**119395734, 28400649, 1177438, 14424438, 2315186, 27065822, 37675289, 1161384, 16116616,**
RefSeq:	**NP_000050.2, NM_000059**
IPI:	**IPI00412408**
Ensembl:	**ENSG00000139618**

DAS for Molecular Interactions (DASMI) - Registry



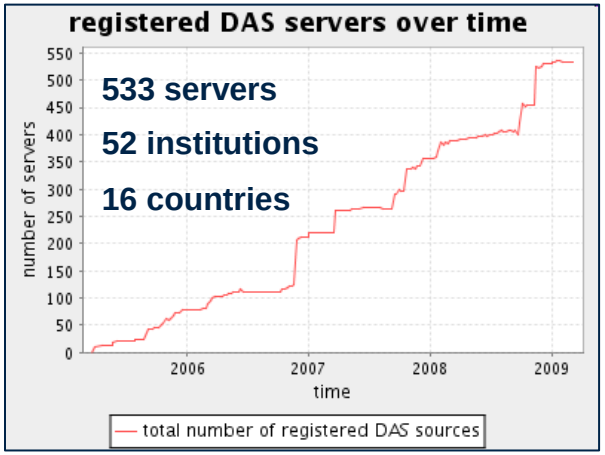
DAS registry – http://www.dasregistry.org



All DAS servers with interaction capability

max planck institut informatik

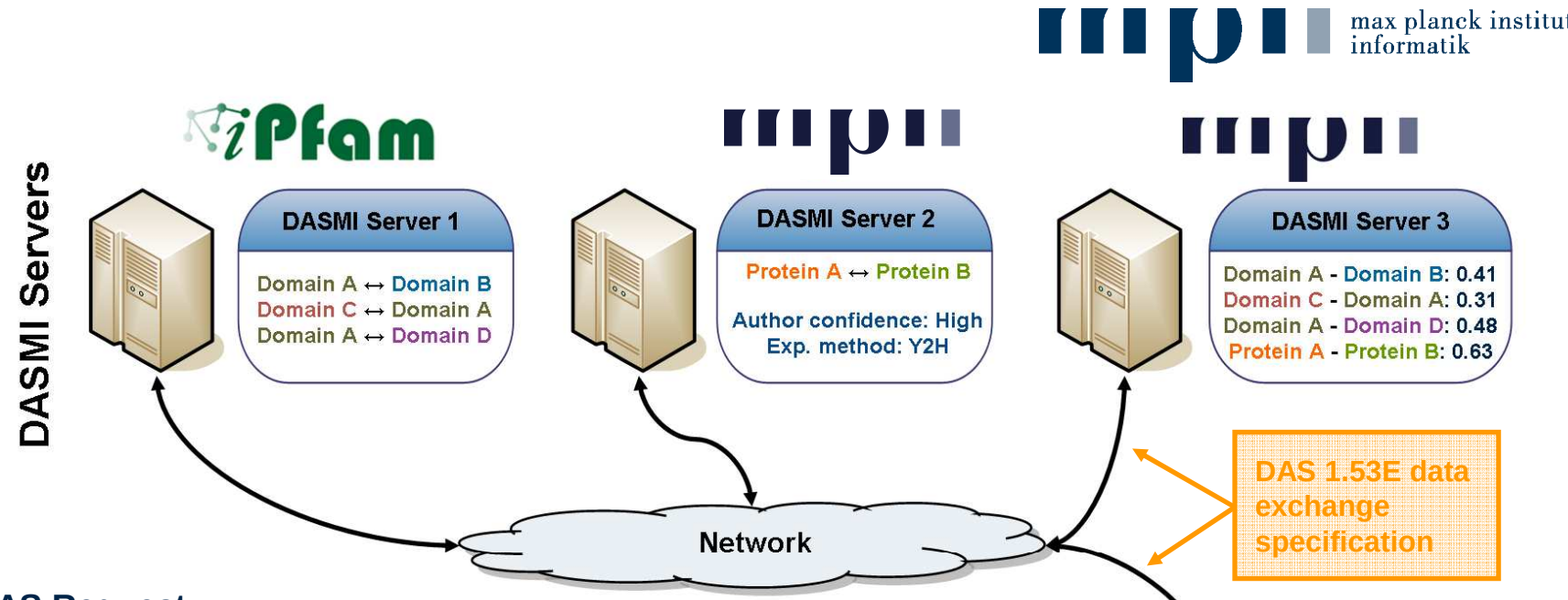
- Maintained at Sanger Institute



Domain interactions

Protein interactions

DAS for Molecular Interactions (DASMI) - Data exchange



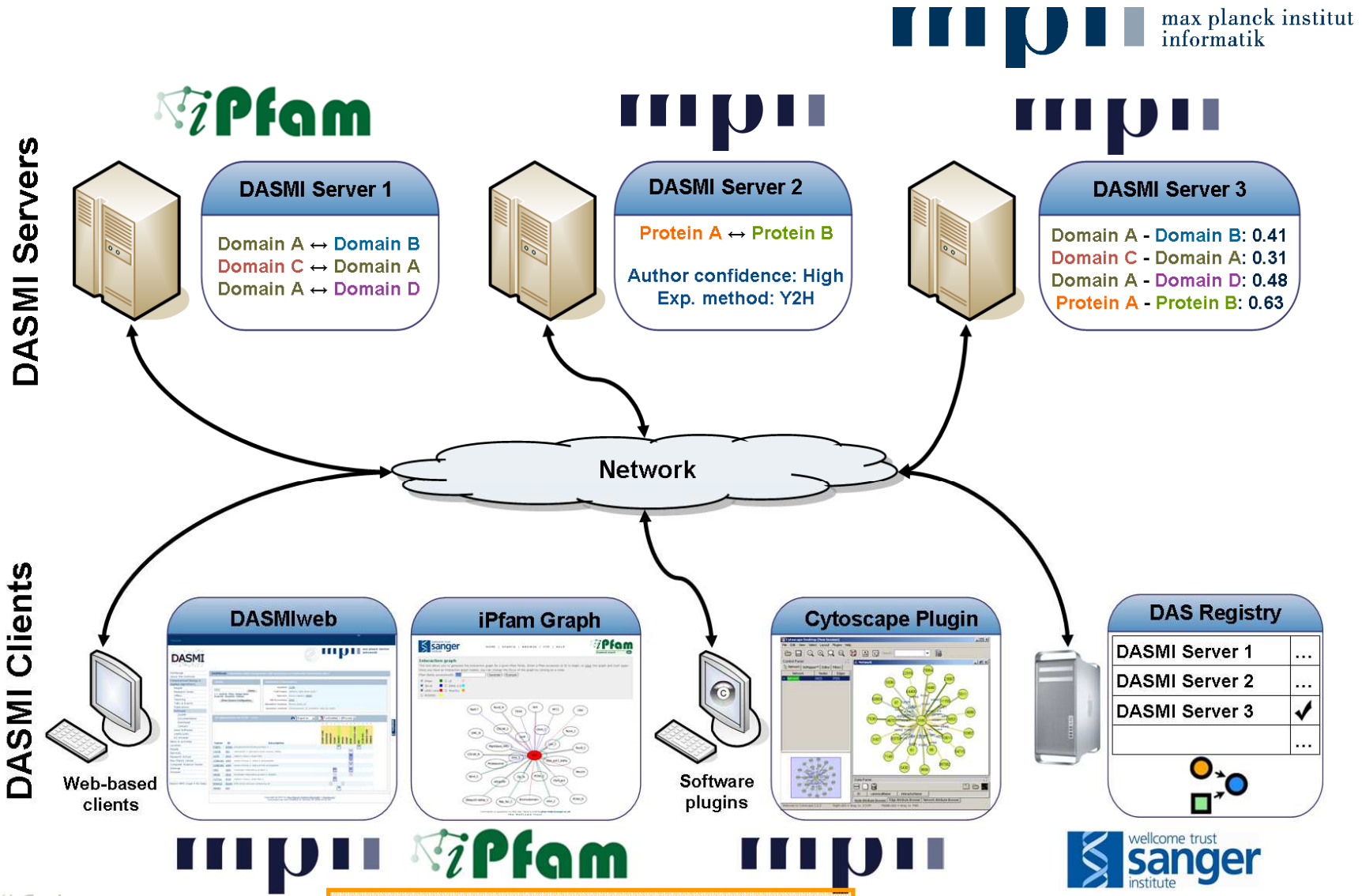
DAS Request

```
http://www.dasmi.de/das/funsimmat/interaction?interactor=P09497&interactor=O60828
                                              &detail=property:bpscore
```

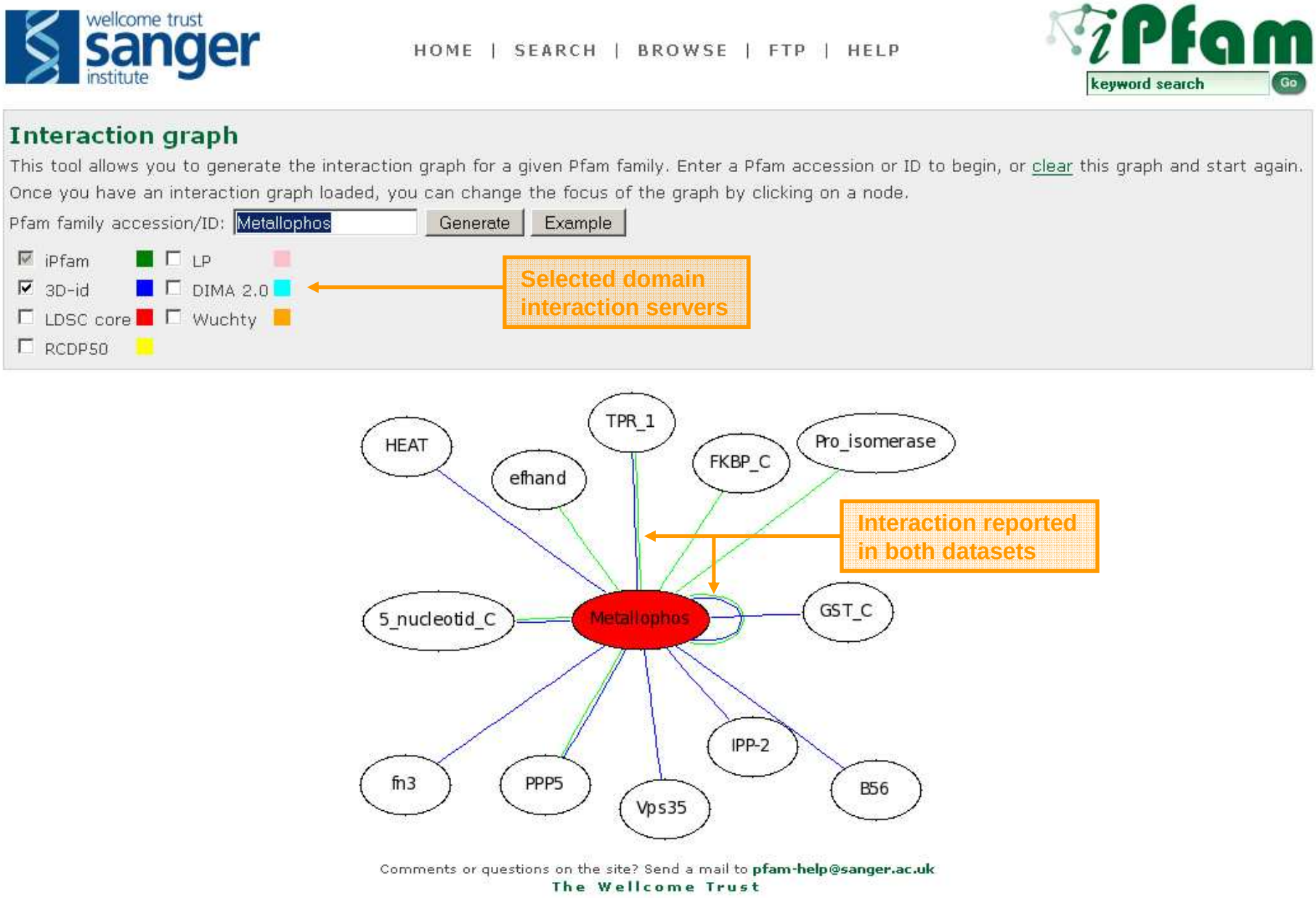
DASINT XML Response

```
<?xml version="1.0" encoding="UTF-8"?>
<DASINT xmlns="http://www.dasmi.de/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.dasmi.de/ http://www.dasmi.de/dasint.xsd">
   <INTERACTOR intId="1" shortLabel="CLCB_HUMAN" dbSource="uniprotkb"dbSourceCvId="MI:0486"
       dbVersion="12.1" dbAccessionId="P09497" dbCoordSys="UniProt,Protein Sequence"/>
   <INTERACTOR intId="2" shortLabel="PQBP1_HUMAN" dbSource="uniprotkb" dbSourceCvId="MI:0486"
       dbVersion="12.1" dbAccessionId="O60828" dbCoordSys="UniProt,Protein Sequence"/>
   <INTERACTION name="P09497-O60828" dbSource="" dbAccessionId="">
       <DETAIL property="bpscore" value="0.154429056459711"/>
       <PARTICIPANT intId="1"/>
       <PARTICIPANT intId="2"/>
   </INTERACTION>
</DASINT>
```

DAS for Molecular Interactions (DASMI) - Clients

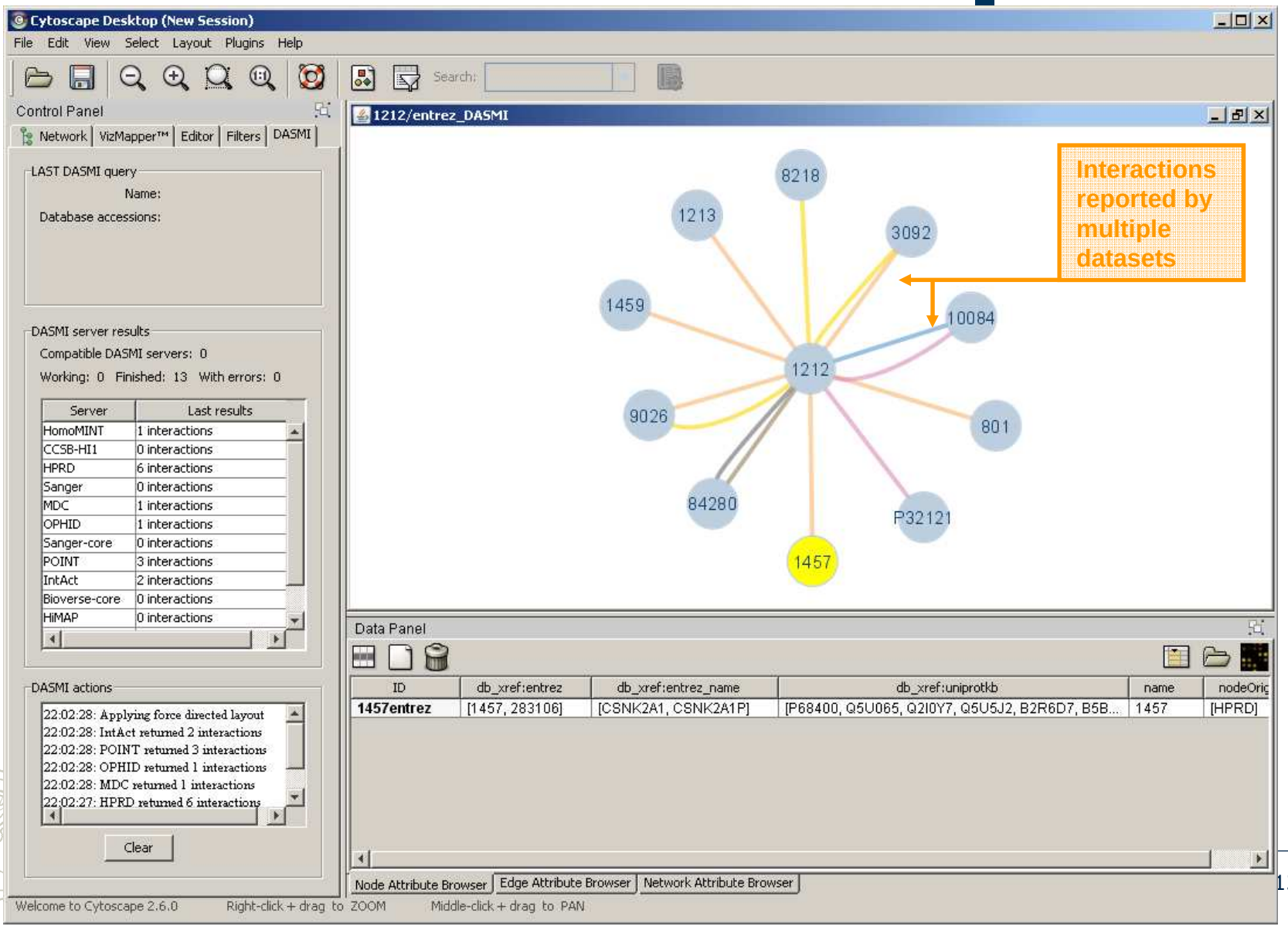


iPfam graphical domain interaction browser



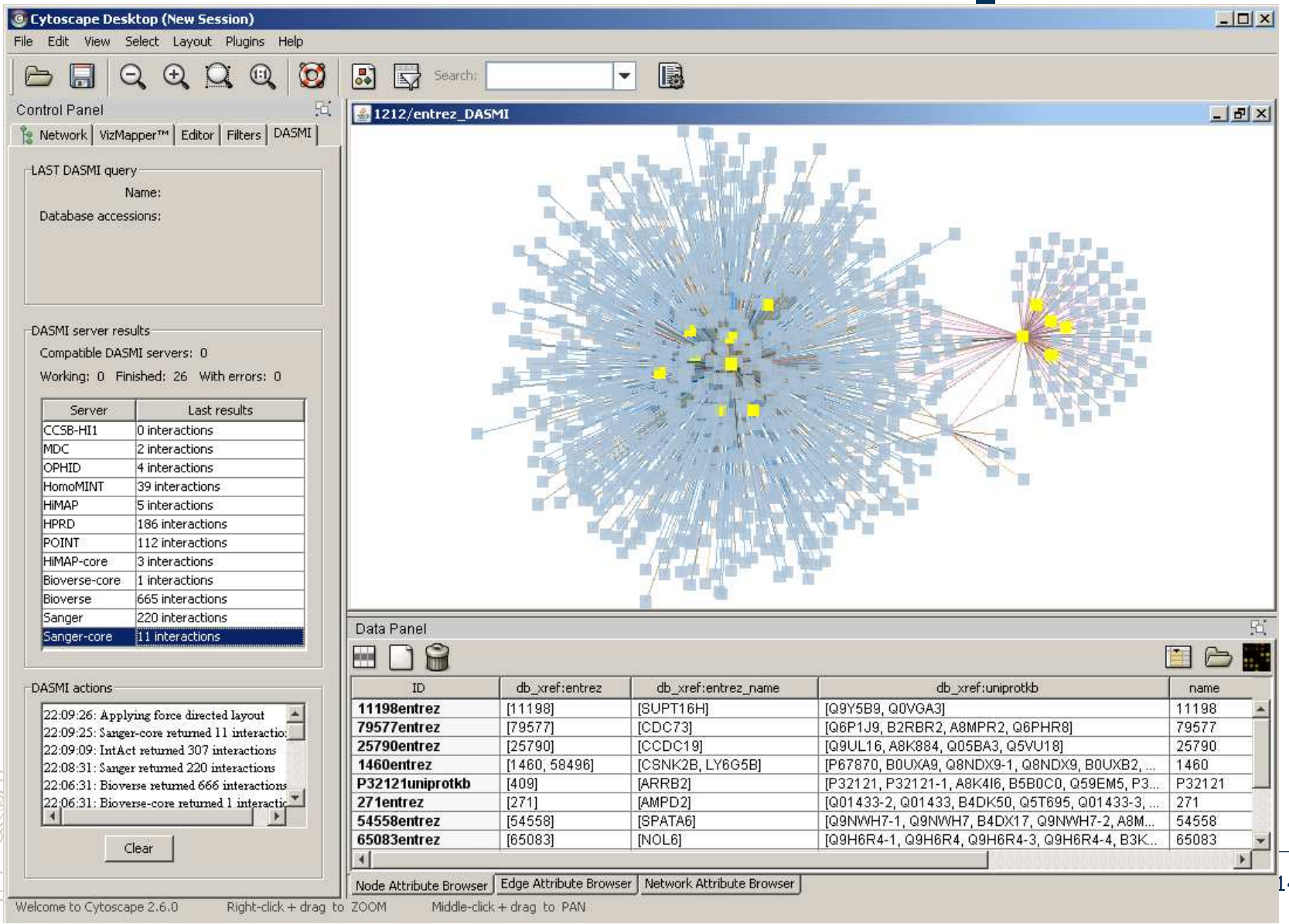
DASMI Cytoscape Client

max planck institut informatik



DASMI Cytoscape Client

max planck institut informatik



Deutsch

http://www.dasmi.de/web

DASMI

max planck institut informatik

- Homepage
- About the Institute
- Computational Biology & Applied Algorithmics
  - People
  - Research Areas
  - Offers
  - Teaching
  - Talks & Events
  - Publications
  - Software
    - DASMI
    - DASMIweb
    - Documentation
    - Download
    - Contact
  - more Software
  - Useful Links
- News & Activities
- Location
- People
- Services
- Research School
- Max Planck Center
- Computer Science Cluster
- Sitemap

Search MPII (type ? for help)

DASMIweb - dynamic online integration and annotation of molecular interaction data

Query

961 Query

e.g. Ensembl, Entrez Gene, GeneInfo, Pfam, RefSeq UniProtKB

Show Source Configuration

Interactor Information

Identifier: 961 (Entrez Gene) (Leukocyte surface antigen CD47 precursor (Integrin-associated protein)(IAP) (Antigenic surface determinant protein OA3) (Protein MER6).) (CD47 molecule)

UniProtKB	CD47_HUMAN, A8K198_HUMAN, Q71A41_HUMAN,
UniGene	Hs.446414,
Entrez Gene	961,
UniParc	UPI0000049C6C, UPI00001BDF27,

22 interactions (showing 11 to 20) previous next

Export as ... Select confidence measure ... Query Confidence Sources

Predictions

Literature curation / experiments

Name	ID	Description	SANGER-CORE (0)	INTACT (1)	CCSB-HI1 (0)	HPRD (11)	BIOVERSE (5)	SANGER (0)	MINT (0)	HOMOMINT (1)	HIMAP (1)	HIMAP-CORE (0)	POINT (6)	OPHID (2)	MPIDB (0)	MDC (0)	DIP (0)	BIOVERSE-CORE (1)
P2RY2	5029	P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)(Purinergic receptor).				■												
EPB42	2038	Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)(P4.2).				■												
THBS1	7057	Hypothetical protein (Fragment).				■	■							■				■
UBQLN1	29979	Ubiquilin-1 (Protein linking IAP with cytoskeleton 1) (PLIC-1) (hPLIC-1).				■												
RHAG	6005	Rhesus blood group-associated glycoprotein (Rhesus blood group-associated ammonia channel) (Erythrocyte plasma membrane 50 kDaglycoprotein) (Rh50A) (CD241 antigen).				■												
PTK2	5747	Focal adhesion kinase 1 (EC 2.7.10.2) (FADK 1) (pp125FAK) (Protein-tyrosine kinase 2).				■												
ITGA6	3655	Integrin alpha-6 variant (Fragment).									■							
COMP	1311	Cartilage oligomeric matrix protein variant (Fragment).					■											
THBS2	7058	Hypothetical protein DKFZp686G02190 (Fragment).					■											
THBS3	7059	Thrombospondin 3 variant (Fragment).					■											
THBS4	7060	Thrombospondin-4 precursor.					■											

Interaction reported by multiple datasets

myDASMI

Copyright © 2008 by Max-Planck-Institut Informatik | Impressum
Document was last modified on March 08 2009 12:41:59.

Deutsch

http://www.dasmi.de/web

DASMI

max planck institut informatik

- Homepage
- About the Institute
- Computational Biology & Applied Algorithmics
  - People
  - Research Areas
  - Offers
  - Teaching
  - Talks & Events
  - Publications
  - Software
    - DASMI
    - DASMIweb
    - Documentation
    - Download
    - Contact
  - more Software
  - Useful Links
- News & Activities
- Location
- People
- Services
- Research School
- Max Planck Center
- Computer Science Cluster
- Sitemap

Search MPII (type ? for help)

DASMIweb - dynamic online integration and annotation of molecular interaction data

Query

961 Query

e.g. Ensembl, Entrez Gene, GeneInfo, Pfam, RefSeq UniProtKB

Show Source Configuration

Interactor Information

Identifier: 961 (Entrez Gene) (Leukocyte surface antigen CD47 precursor (Integrin-associated protein)(IAP) (Antigenic surface determinant protein OA3) (Protein MER6).) (CD47 molecule)

UniProtKB	CD47_HUMAN, A8K198_HUMAN, Q71A41_HUMAN,
UniGene	Hs.446414,
Entrez Gene	961,
UniParc	UPI0000049C6C, UPI00001BDF27,

22 interactions (showing 11 to 20) previous next

Export as ...

Select confidence measure ...

Query Confidence Sources

All confidence scoring methods that returned results for the current interactions

Select confidence measure ...
Bioverse (originalConfidenceScore)
Bioverse-core (originalConfidenceScore)
Domain support (predicted)
FunSimMat2.1 (BPscore)
FunSimMat2.1 (CCscore)
FunSimMat2.1 (MFscore)
HiMAP (originalConfidenceScore)

Name	ID	Description	POINT (6)	OPHID (2)	MPIDB (0)	MDC (0)	DIP (0)	BIOVERSE-CORE (1)
P2RY2	5029	P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)(Purinergic receptor).	■					
EPB42	2038	Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)(P4.2).	■					
THBS1	7057	Hypothetical protein (Fragment).	■	■	■			■
UBQLN1	29979	Ubiquilin-1 (Protein linking IAP with cytoskeleton 1) (PLIC-1) (hPLIC-1).	■					
RHAG	6005	Rhesus blood group-associated glycoprotein (Rhesus blood group-associated ammonia channel) (Erythrocyte plasma membrane 50 kDaglycoprotein) (Rh50A) (CD241 antigen).	■					
PTK2	5747	Focal adhesion kinase 1 (EC 2.7.10.2) (FADK 1) (pp125FAK) (Protein-tyrosine kinase 2).	■					
ITGA6	3655	Integrin alpha-6 variant (Fragment).	■					
COMP	1311	Cartilage oligomeric matrix protein variant (Fragment).	■					
THBS2	7058	Hypothetical protein DKFZp686G02190 (Fragment).	■					
THBS3	7059	Thrombospondin 3 variant (Fragment).	■					
THBS4	7060	Thrombospondin-4 precursor.	■					

myDASMI

Copyright © 2008 by Max-Planck-Institut Informatik | Impressum
Document was last modified on March 08 2009 12:41:59.

Deutsch

http://www.dasmi.de/web

DASMI

max planck institut informatik

- Homepage
- About the Institute
- Computational Biology & Applied Algorithmics
  - People
  - Research Areas
  - Offers
  - Teaching
  - Talks & Events
  - Publications
  - Software
    - DASMI
    - DASMIweb
    - Documentation
    - Download
    - Contact
  - more Software
  - Useful Links
- News & Activities
- Location
- People
- Services
- Research School
- Max Planck Center
- Computer Science Cluster
- Sitemap

Search MPII (type ? for help)

DASMIweb - dynamic online integration and annotation of molecular interaction data

Query

961 Query

e.g. Ensembl, Entrez Gene, GeneInfo, Pfam, RefSeq UniProtKB

Show Source Configuration

Interactor Information

Identifier: 961 (Entrez Gene) (Leukocyte surface antigen CD47 precursor (Integrin-associated protein)(IAP) (Antigenic surface determinant protein OA3) (Protein MER6).) (CD47 molecule)

UniProtKB	CD47_HUMAN, A8K198_HUMAN, Q71A41_HUMAN,
UniGene	Hs.446414,
Entrez Gene	961,
UniParc	UPI0000049C6C, UPI00001BDF27,

22 interactions (showing 11 to 20) previous next

Export as ... | Bioverse (originalConfidenceScore) | Query Confidence Sources

Original confidence score provided by the authors

Name	ID	Description	SANGER-CORE	INTACT	CCSB-HI1	HPRD	BIOVERSE	SANGER	MINT	HOMOMINT	HIMAP	HIMAP-CORE	POINT	OPHID	MPIDB	MDC	DIP	BIOVERSE-CORE
			0	1	0	11	5	0	0	1	1	0	6	2	0	0	0	1
P2RY2	5029	P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)(Purinergic receptor).																
EPB42	2038	Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)(P4.2).																
THBS1	7057	Hypothetical protein (Fragment).				0.92	0.92							0.92				0.92
UBQLN1	29979	Ubiquilin-1 (Protein linking IAP with cytoskeleton 1) (PLIC-1) (hPLIC-1).																
RHAG	6005	Rhesus blood group-associated glycoprotein (Rhesus blood group-associated ammonia channel) (Erythrocyte plasma membrane 50 kDaglycoprotein) (Rh50A) (CD241 antigen).																
PTK2	5747	Focal adhesion kinase 1 (EC 2.7.10.2) (FADK 1) (pp125FAK) (Protein-tyrosine kinase 2).																
ITGA6	3655	Integrin alpha-6 variant (Fragment).																
COMP	1311	Cartilage oligomeric matrix protein variant (Fragment).					0.26											
THBS2	7058	Hypothetical protein DKFZp686G02190 (Fragment).					0.57											
THBS3	7059	Thrombospondin 3 variant (Fragment).					0.27											
THBS4	7060	Thrombospondin-4 precursor.					0.27											

myDASMI

Details for Bioverse (originalConfidenceScore)

The confidence score as reported in the original publication.
Range: 0-1
http://nar.oxfordjournals.org/cgi/content/abstract/31/13/3736

Copyright © 2008 by Max-Planck-Institut Informatik | Impressum
Document was last modified on March 08 2009 12:41:59.

Deutsch

http://www.dasmi.de/web

DASMI

max planck institut informatik

- Homepage
- About the Institute
- Computational Biology & Applied Algorithmics
  - People
  - Research Areas
  - Offers
  - Teaching
  - Talks & Events
  - Publications
  - Software
    - DASMI
    - DASMIweb
    - Documentation
    - Download
    - Contact
  - more Software
  - Useful Links
- News & Activities
- Location
- People
- Services
- Research School
- Max Planck Center
- Computer Science Cluster
- Sitemap

Search MPII (type ? for help)

DASMIweb - dynamic online integration and annotation of molecular interaction data

Query

961 Query

e.g. Ensembl, Entrez Gene, GeneInfo, Pfam, RefSeq UniProtKB

Show Source Configuration

Interactor Information

Identifier: 961 (Entrez Gene) (Leukocyte surface antigen CD47 precursor (Integrin-associated protein)(IAP) (Antigenic surface determinant protein OA3) (Protein MER6).) (CD47 molecule)

UniProtKB	CD47_HUMAN, A8K198_HUMAN, Q71A41_HUMAN,
UniGene	Hs.446414,
Entrez Gene	961,
UniParc	UPI0000049C6C, UPI00001BDF27,

22 interactions (showing 11 to 20) previous next

Export as ...

Domain support (predicted)

Query Confidence Sources

Protein interactions supported by underlying domain interactions

Name	ID	Description
P2RY2	5029	P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)(Purinergic receptor).
EPB42	2038	Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)(P4.2).
THBS1	7057	Hypothetical protein (Fragment).
UBQLN1	29979	Ubiquilin-1 (Protein linking IAP with cytoskeleton 1) (PLIC-1) (hPLIC-1).
RHAG	6005	Rhesus blood group-associated glycoprotein (Rhesus blood group-associated ammonia channel) (Erythrocyte plasma membrane 50 kDaglycoprotein) (Rh50A) (CD241 antigen).
PTK2	5747	Focal adhesion kinase 1 (EC 2.7.10.2) (FADK 1) (pp125FAK) (Protein-tyrosine kinase 2).
ITGA6	3655	Integrin alpha-6 variant (Fragment).
COMP	1311	Cartilage oligomeric matrix protein variant (Fragment).
THBS2	7058	Hypothetical protein DKFZp686G02190 (Fragment).
THBS3	7059	Thrombospondin 3 variant (Fragment).
THBS4	7060	Thrombospondin-4 precursor.

0	1	0	11	5	0	0	1	1	0	6	2	0	0	0	1
SANGER-CORE	INTACT	CCSB-HI1	HPRD	BIOVERSE	SANGER	MINT	HOMOMINT	HIMAP	HIMAP-CORE	POINT	OPHID	MPIDB	MDC	DIP	BIOVERSE-CORE

myDASMI

Details for Domain support (predicted)

This measure indicates protein-protein interactions that are supported by underlying domain-domain interactions. The domain-domain interaction datasets used in the predicted subset are based on diverse computational predication algorithms. Please see the respective interaction details for more information on the datasets.
Range: 1
http://www.dasmi.de

Copyright © 2008 by Max-Planck-Institut Informatik | Impressum
Document was last modified on March 08 2009 12:41:59.

Deutsch

http://www.dasmi.de/web

DASMI

max planck institut informatik

- Homepage
- About the Institute
- Computational Biology & Applied Algorithmics
  - People
  - Research Areas
  - Offers
  - Teaching
  - Talks & Events
  - Publications
  - Software
    - DASMI
    - DASMIweb
    - Documentation
    - Download
    - Contact
  - more Software
  - Useful Links
- News & Activities
- Location
- People
- Services
- Research School
- Max Planck Center
- Computer Science Cluster
- Sitemap

Search MPII (type ? for help)

DASMIweb - dynamic online integration and annotation of molecular interaction data

Query

961 Query

e.g. Ensembl, Entrez Gene, GeneInfo, Pfam, RefSeq UniProtKB

Show Source Configuration

Interactor Information

Identifier: 961 (Entrez Gene) (Leukocyte surface antigen CD47 precursor (Integrin-associated protein)(IAP) (Antigenic surface determinant protein OA3) (Protein MER6).) (CD47 molecule)

UniProtKB	CD47_HUMAN, A8K198_HUMAN, Q71A41_HUMAN,
UniGene	Hs.446414,
Entrez Gene	961,
UniParc	UPI0000049C6C, UPI00001BDF27,

22 interactions (showing 11 to 20) previous next | Export as ... | Domain support (predicted) | Query Confidence Sources

Name	ID	Description	SANGER-CORE (0)	INTACT (1)	CCSB-HI1 (0)	HPRD (11)	BIOVERSE (5)	SANGER (0)	MINT (0)	HOMOMINT (1)	HIMAP (1)	HIMAP-CORE (0)	POINT (8)	OPHID (2)	MPIDB (0)	MDC (0)	DIP (0)	BIOVERSE-CORE (1)
P2RY2	5029	P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)(Purinergic receptor).																
EPB42	2038	Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)(P4.2).																

Details from source HPRD

link	http://www.hprd.org/interactions?protein=01517&isoform_id=01517_1&isoform_name=
link	http://www.hprd.org/interactions?protein=03017&isoform_id=03017_1&isoform_name=
DAS query string	http://dasmi.bioinf.mpi-inf.mpg.de/das/hprd/interaction?interactor=Q08722&interactor=A8K198&interactor=Q71A41&operation=union

Details from source Domain support

predicted	PF00927-PF04549: IPPRI core (0.03977)
predicted	PF00927-PF08204: IPPRI core (0.03977)
predicted	PF00868-PF04549: IPPRI core (0.061411)
predicted	PF00868-PF08204: IPPRI core (0.061411)
predicted	PF01841-PF04549: IPPRI core (0.061411)
predicted	PF01841-PF08204: IPPRI core (0.061411)

Details from source FunSimMat2.1

MFscore	0.976656009021968
CCscore	0.846404924988201

Details on domain interactions

Details for Domain support (predicted)

This measure indicates protein-protein interactions that are supported by underlying domain-domain interactions. The domain-domain interaction datasets used in the predicted subset are based on diverse computational predication algorithms. Please see the respective interaction details for more information on the datasets.
Range: 1
http://www.dasmi.de

myDASMI

Deutsch

http://www.dasmi.de/web

DASMI

max planck institut informatik

- Homepage
- About the Institute
- Computational Biology & Applied Algorithmics
  - People
  - Research Areas
  - Offers
  - Teaching
  - Talks & Events
  - Publications
  - Software
    - DASMI
    - DASMIweb
    - Documentation
    - Download
    - Contact
  - more Software
  - Useful Links
- News & Activities
- Location
- People
- Services
- Research School
- Max Planck Center
- Computer Science Cluster
- Sitemap

Search MPII (type ? for help)

DASMIweb - dynamic online integration and annotation of molecular interaction data

Query

961 Query

e.g. Ensembl, Entrez Gene, GeneInfo, Pfam, RefSeq UniProtKB

Show Source Configuration

Interactor Information

Identifier: 961 (Entrez Gene) (Leukocyte surface antigen CD47 precursor (Integrin-associated protein)(IAP) (Antigenic surface determinant protein OA3) (Protein MER6).) (CD47 molecule)

UniProtKB	CD47_HUMAN, A8K198_HUMAN, Q71A41_HUMAN,
UniGene	Hs.446414,
Entrez Gene	961,
UniParc	UPI0000049C6C, UPI00001BDF27,

22 interactions (showing 11 to 20) previous next

Export as ...

FunSimMat2.1 (BPscore)

Query Confidence Sources

Functional similarity based on GO annotation of interactors

Name	ID	Description
P2RY2	5029	P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)(Purinergic receptor).
EPB42	2038	Erythrocyte membrane protein band 4.2 (Erythrocyte protein 4.2)(P4.2).
THBS1	7057	Hypothetical protein (Fragment).
UBQLN1	29979	Ubiquilin-1 (Protein linking IAP with cytoskeleton 1) (PLIC-1) (hPLIC-1).
RHAG	6005	Rhesus blood group-associated glycoprotein (Rhesus blood group-associated ammonia channel) (Erythrocyte plasma membrane 50 kDaglycoprotein) (Rh50A) (CD241 antigen).
PTK2	5747	Focal adhesion kinase 1 (EC 2.7.10.2) (FADK 1) (pp125FAK) (Protein-tyrosine kinase 2).
ITGA6	3655	Integrin alpha-6 variant (Fragment).
COMP	1311	Cartilage oligomeric matrix protein variant (Fragment).
THBS2	7058	Hypothetical protein DKFZp686G02190 (Fragment).
THBS3	7059	Thrombospondin 3 variant (Fragment).
THBS4	7060	Thrombospondin-4 precursor.

myDASMI

Details for FunSimMat2.1 (BPscore)

The BPscore is based on biological process annotation of the Gene Ontology.
Range: 0-1
http://funsimmat.bioinf.mpi-inf.mpg.de/help.php

Copyright © 2008 by Max-Planck-Institut Informatik | Impressum
Document was last modified on March 08 2009 12:41:59.

Conclusions & Outlook

max planck institut informatik

- Usage of DASMI servers and clients surprisingly good, but more external DASMI servers are desirable
- DAS client and server libraries (Dazzle, ProServer, (MyDAS), Dasobert, Bio-DAS-lite) support DASMI

- Considerable overlap with HUPO-PSI initiatives for distributed interaction data retrieval (PSICQUIC) and confidence scoring (PSISCORE)
- Develop methods for combining different interaction confidence scoring schemas in DASMI clients

Acknowledgements

- Joachim Büch
- Fidel Ramírez
- Dorothea Emig
- Sven-Eric Schelhorn
- Andreas Schlicker
- Thomas Lengauer
- Mario Albrecht
- Robert Finn
- Andreas Prlić
- Jonathan Warren
- Andrew Jenkinson
- Henning Hermjakob