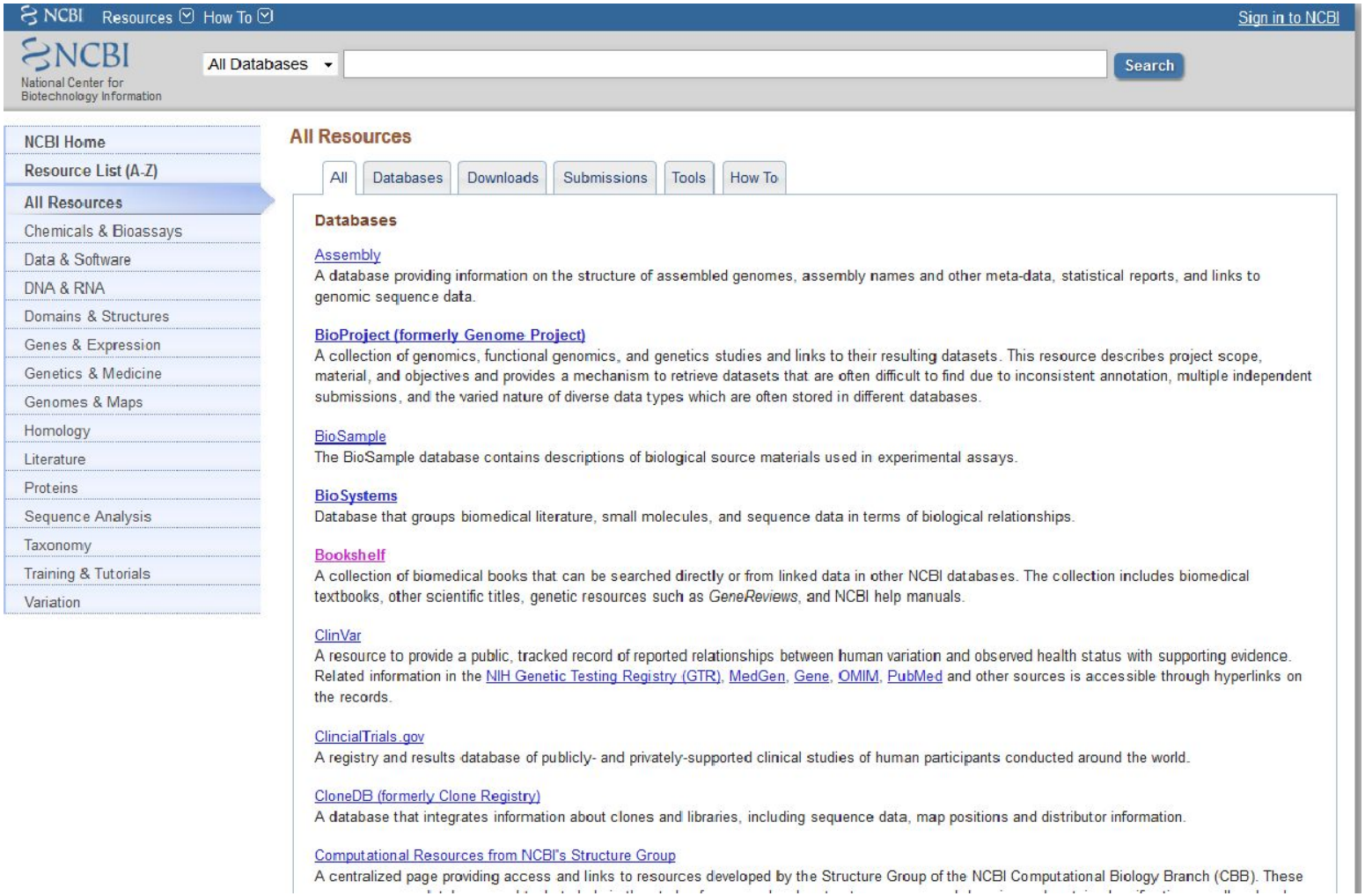
Задание по базам данных

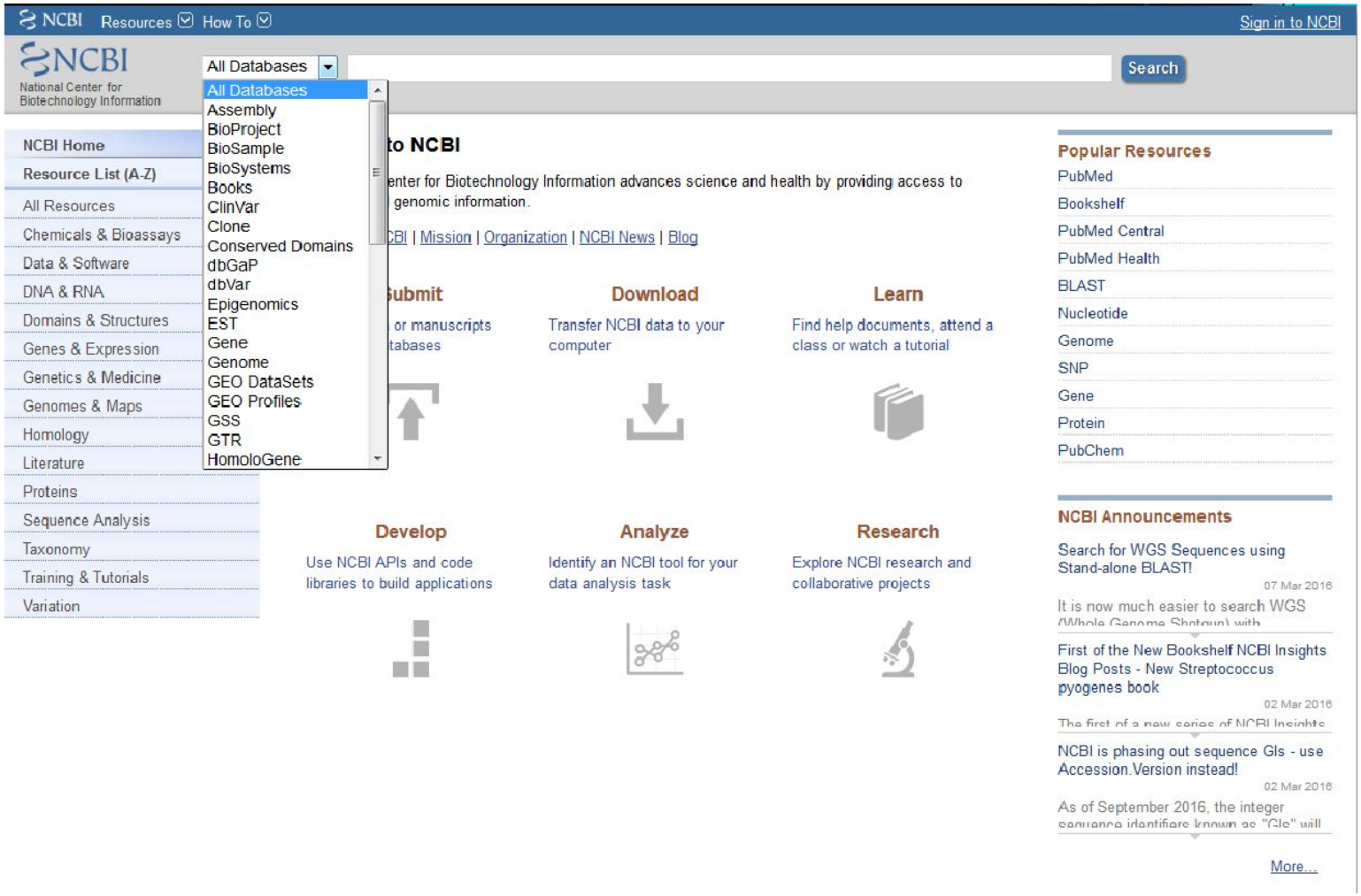
Список баз данных которые можно использовать

База данных национального института рака NCBI (http://www.ncbi.nlm.nih.gov)



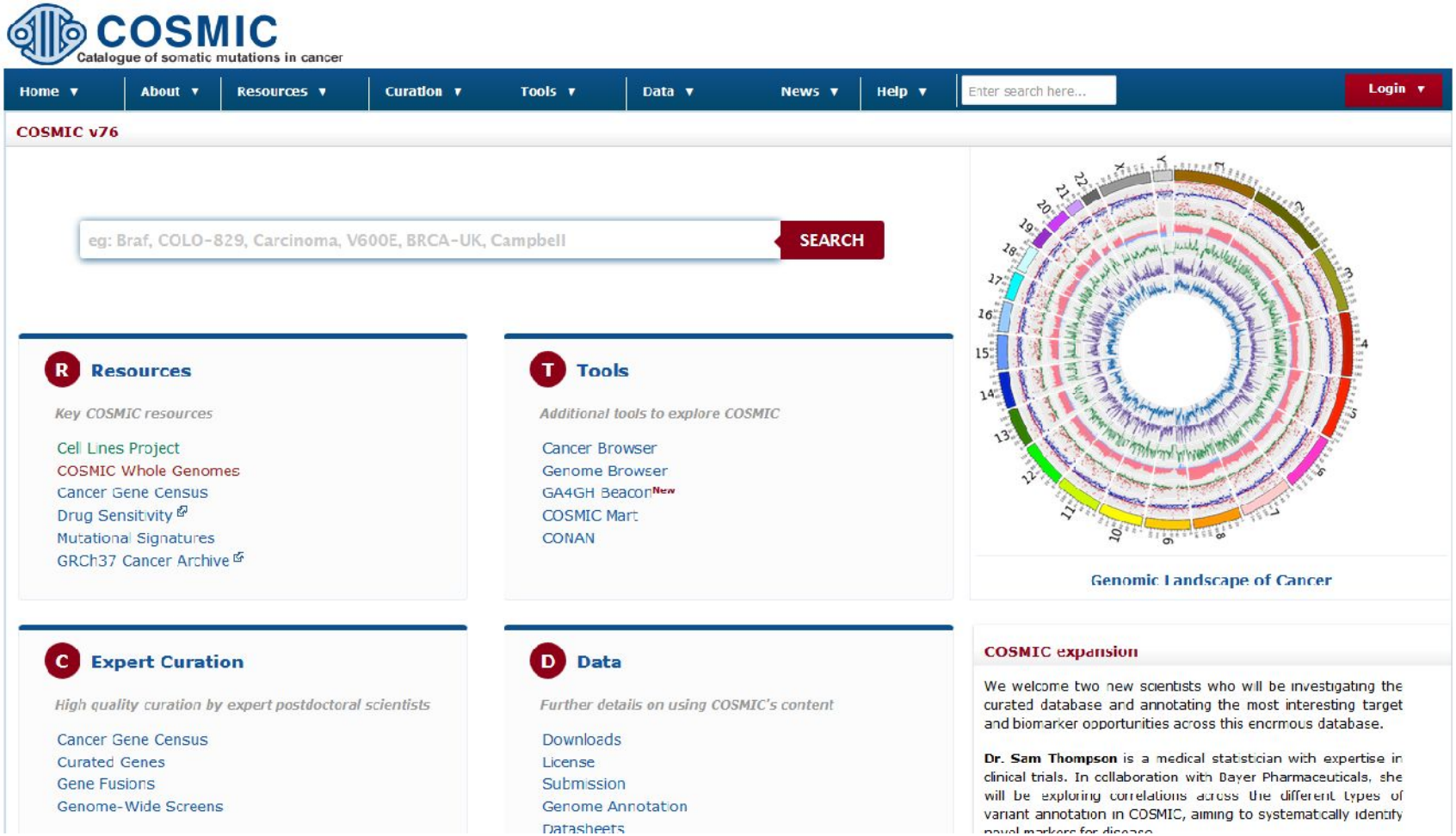
Необходимо описать базу данных, структуру , основные вкладки, какую информацию они несут.

Описание вкладок базы данных NSBI



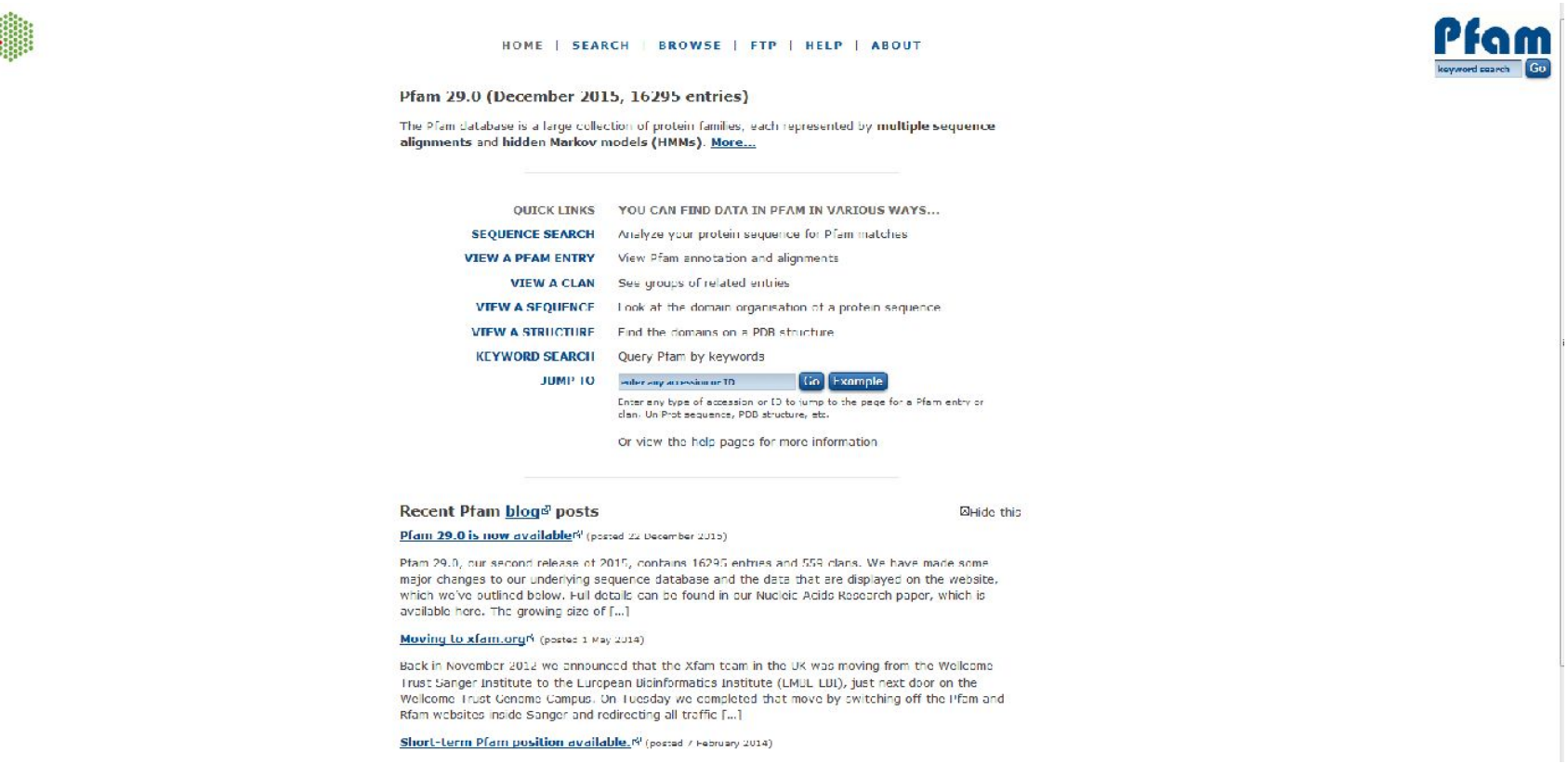
Всего 43 вкладки, одному человеку достаточно описать одну вкладку с указанием информации этой вкладки и примером работы во вкладке (примеры в виде скриншотов).

COSMIC **C**atalogue **O**f **S**omatic **M**utations **I**n **C**ancer
баз данных соматических мутаций, свойственных различным типам опухолевых заболеваний человека.
http://cancer.sanger.ac.uk/cosmic



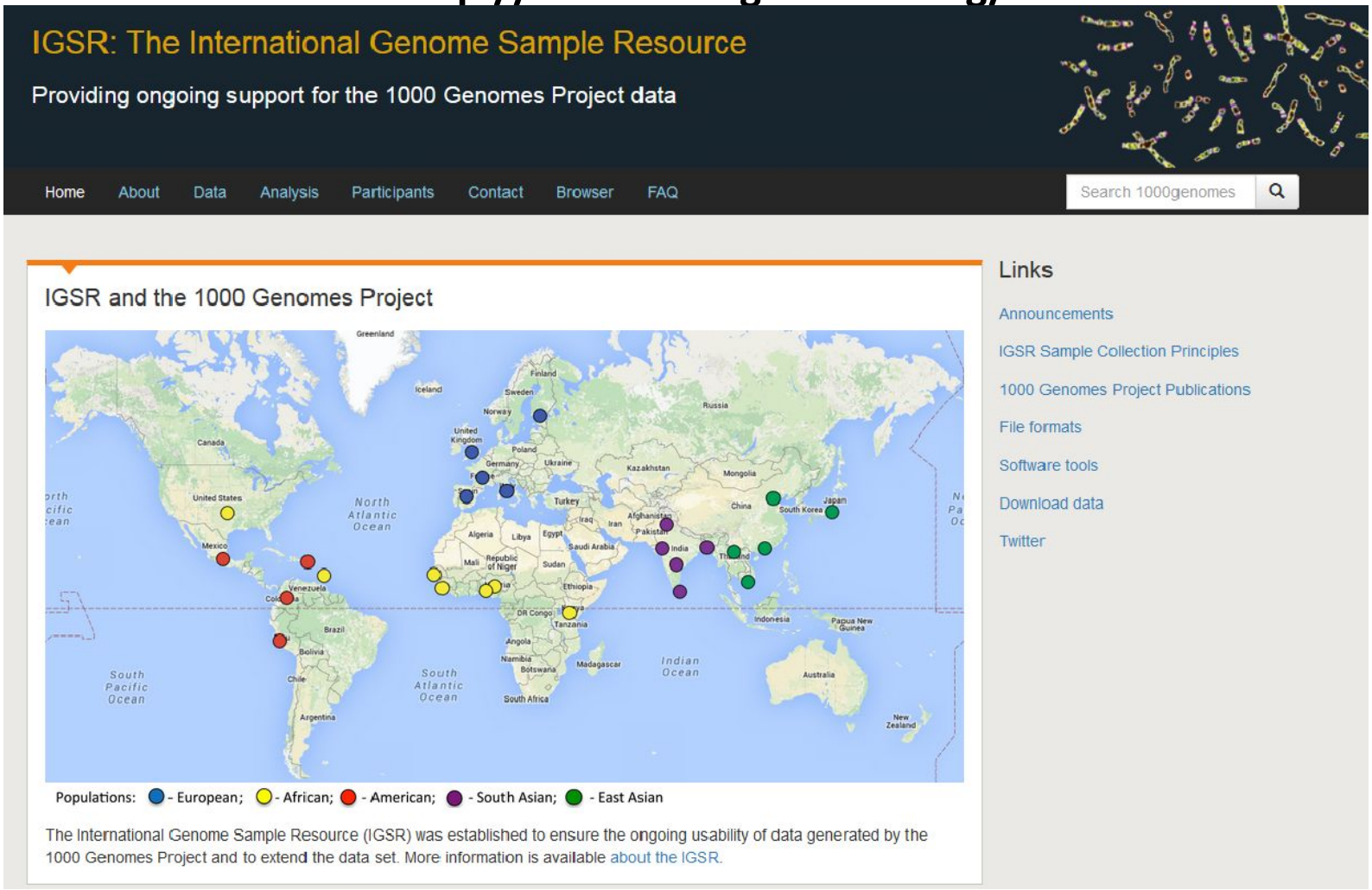
Описание базы данных, описание вкладок, примеры работы с базой данных. (2 человека)

PFAM - **P**rotein **fam**ilies database of alignments and HMMs
коллекция белковых семейств
http://pfam.xfam.org/



Описание базы данных, описание вкладок, примеры работы с базой данных.

1000 Genomes
http://www.1000genomes.org/

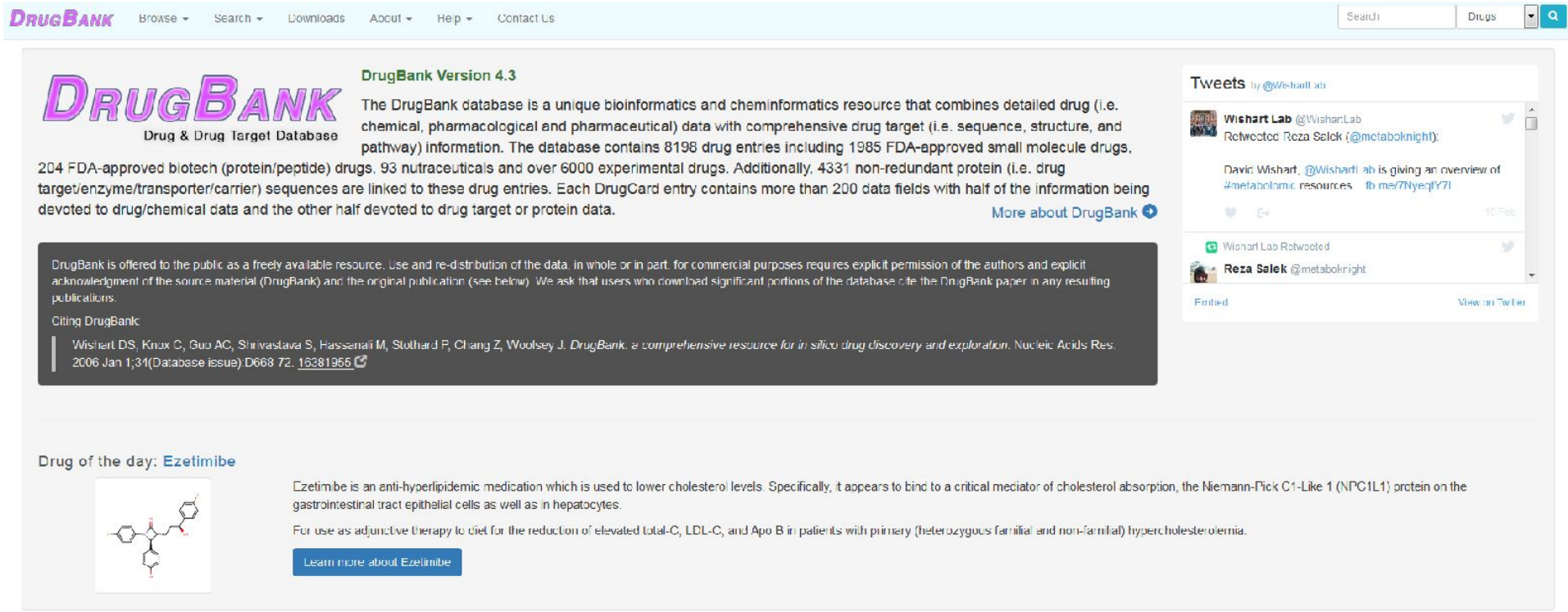


Описание базы данных, описание вкладок, примеры работы с базой данных. (2 человека)

DrugBank

база данных лекарственных веществ с химической, фармакологической и фармацевтической информацией.

http://www.drugbank.ca/



Описание базы данных, описание вкладок, примеры работы с базой данных.

Reactome

База данных о биологических путях у человека

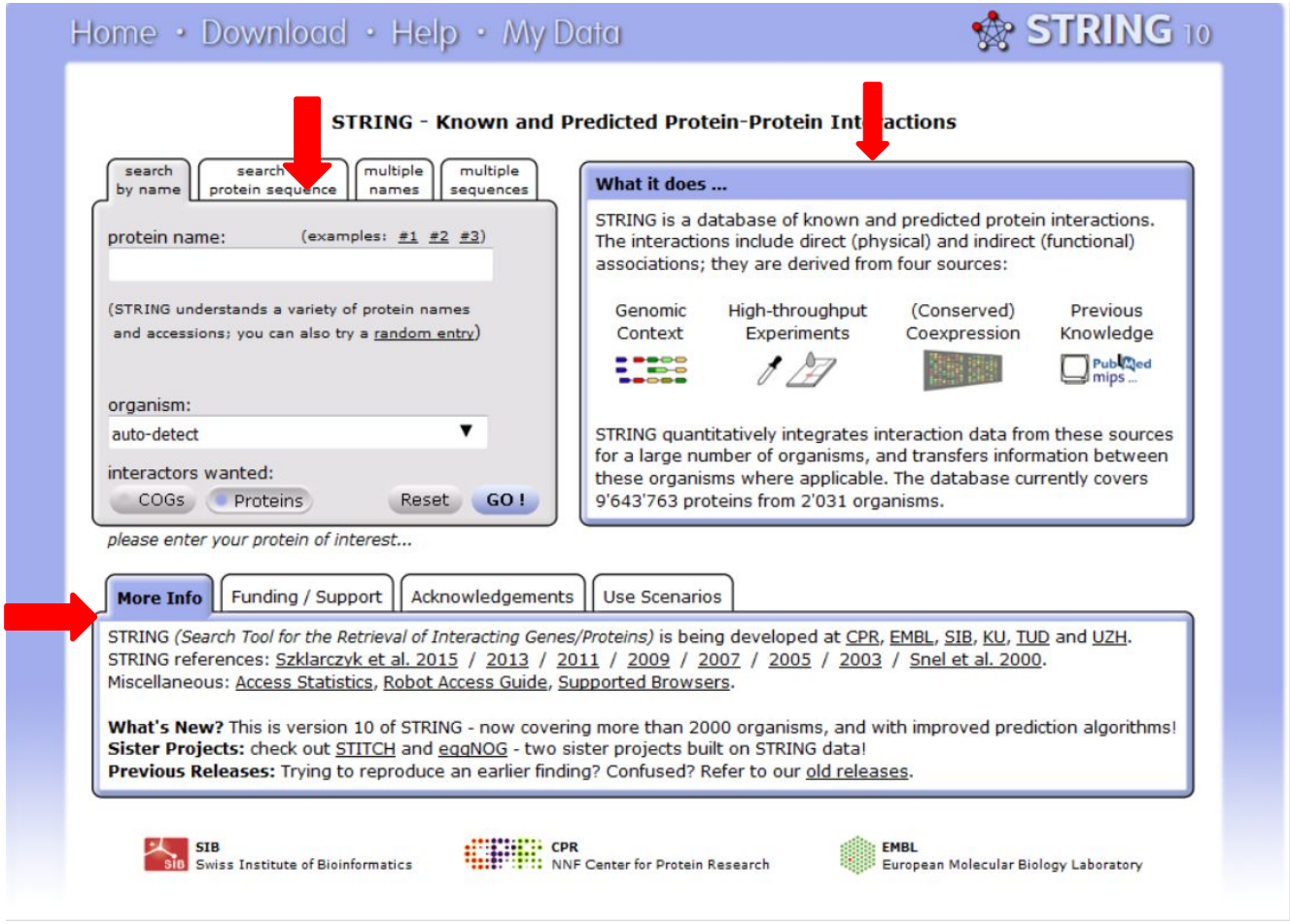
http://www.reactome.org/

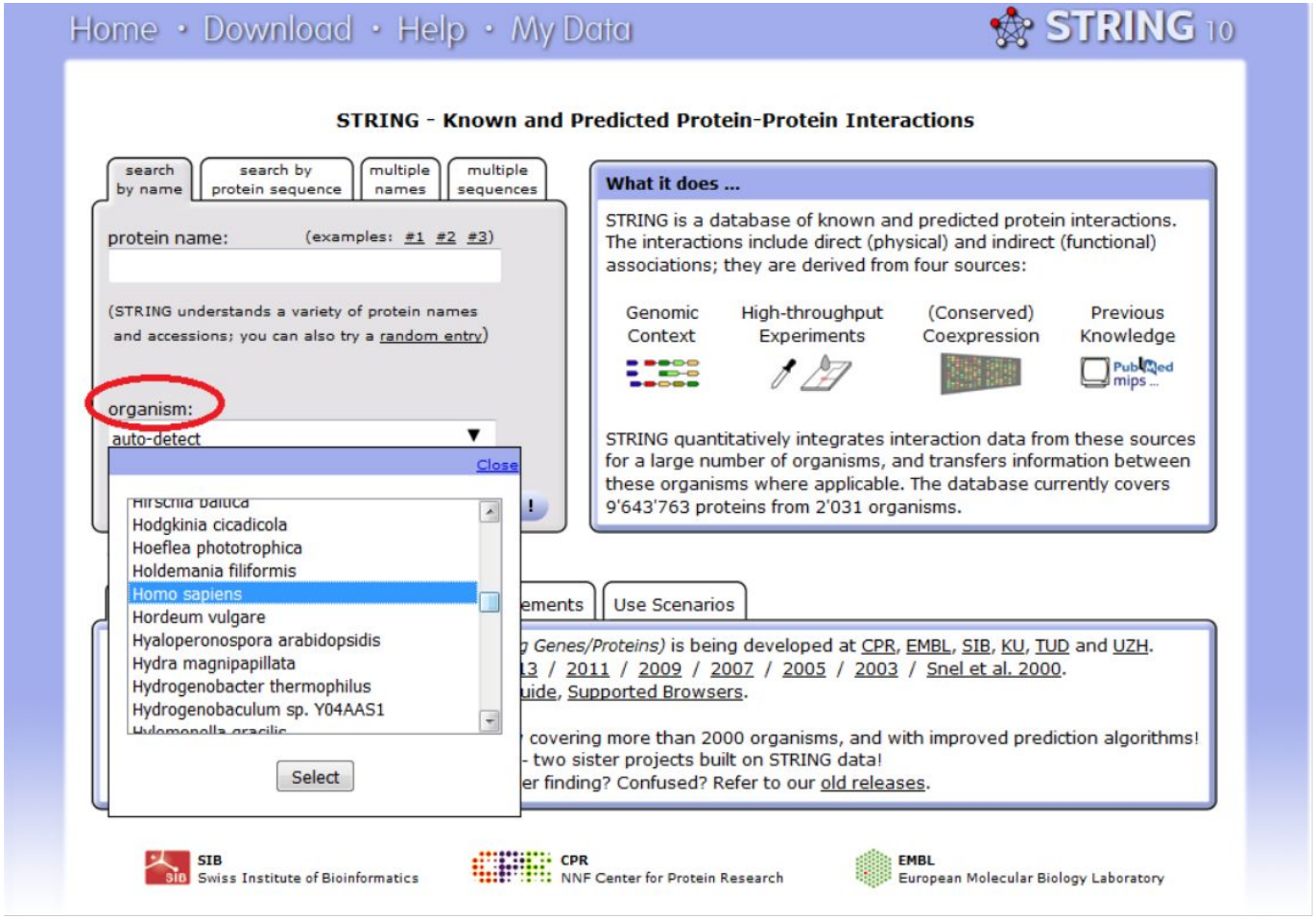
Описание базы данных, описание вкладок, примеры работы с базой данных. (2 человека)

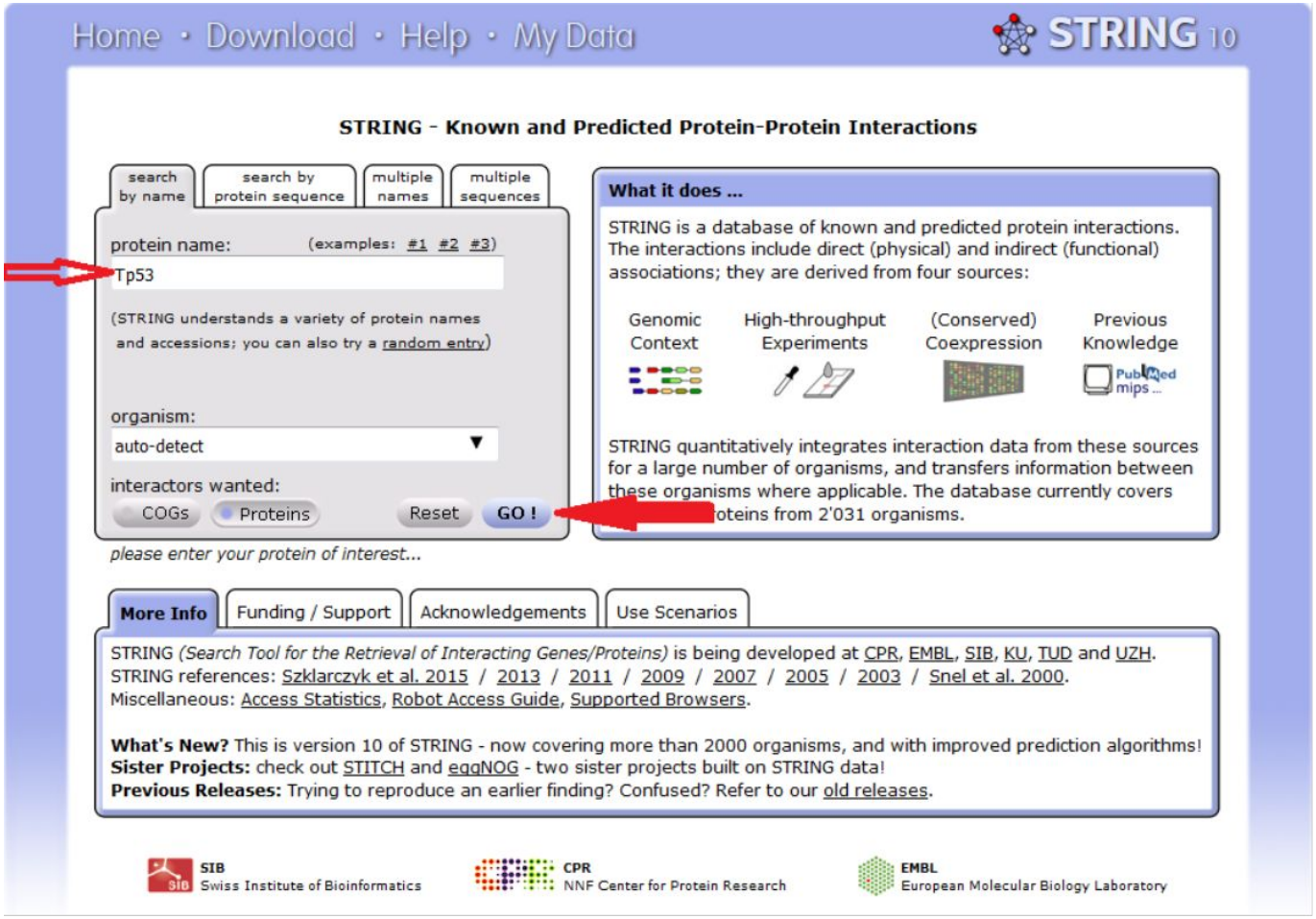
Пример описания базы данных

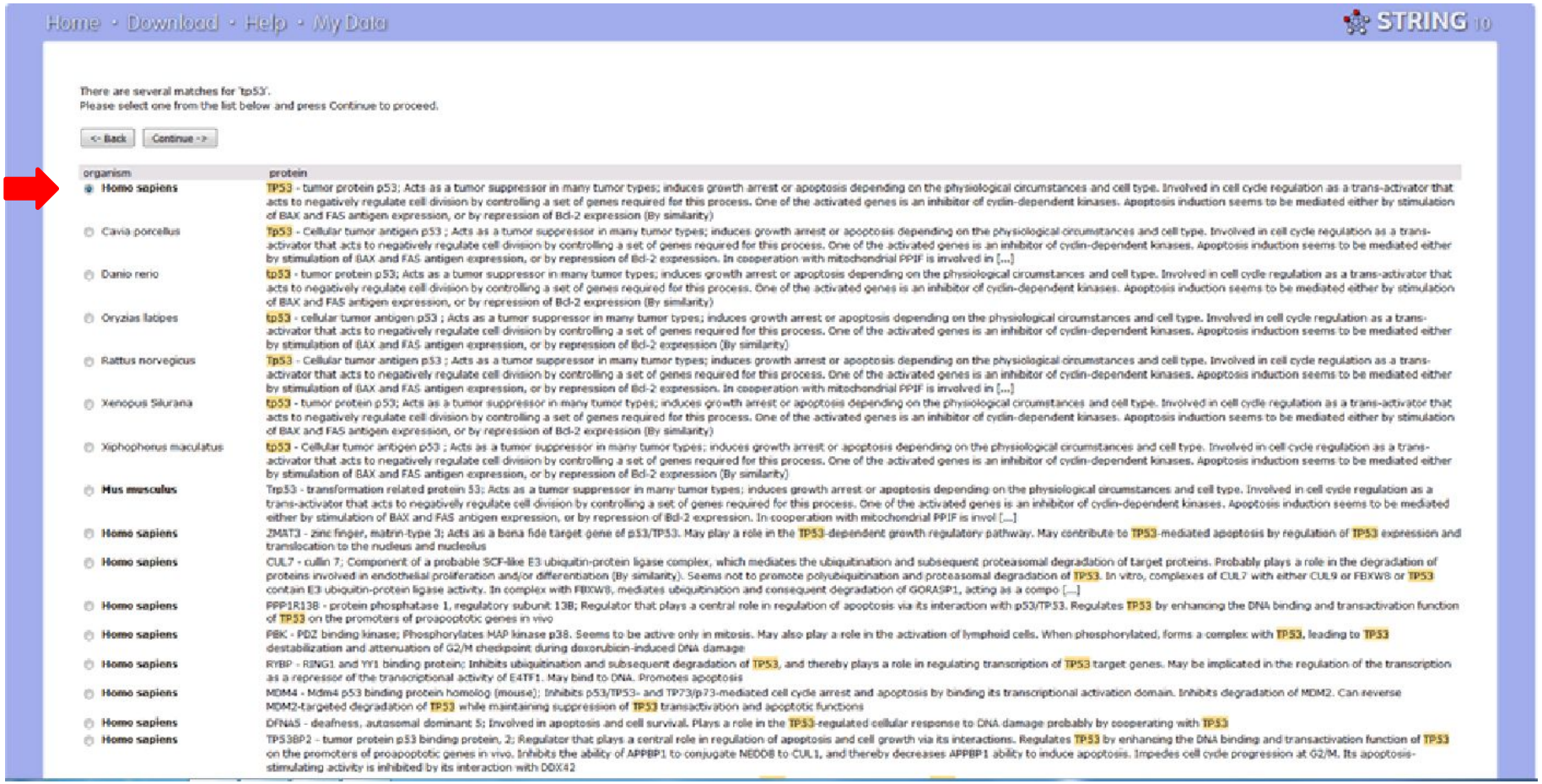
База данных STRING - *Search Tool for the Retrieval of Interacting Genes/Proteins (*http://string-db.org/)

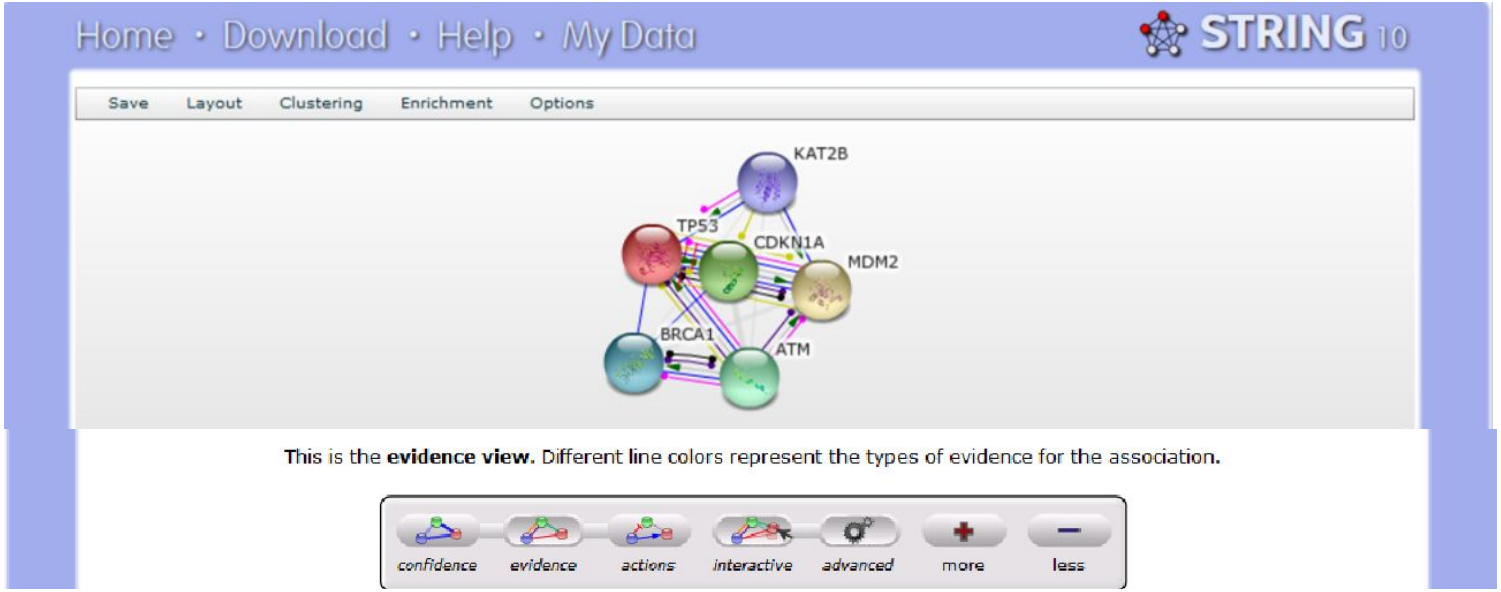
Это биологическая база данных и web-ресурс описывает и предсказывает белок-белковые взаимодействия. База данных аккумулирует в себя информацию из большого числа источников включающих экспериментальные данные. База содержит в себе информацию о 9,6 млн белков более 2000 организмов.




Home • Download • Help • My Data
STRING 10
STRING - Known and Predicted Protein-Protein Interactions
search by name
search by protein sequence
multiple names
multiple sequences
protein name:
(examples: #1 #2 #3)
(STRING understands a variety of protein names and accessions; you can also try a random entry)
organism:
auto-detect
Close
Hodgkinia cicadicola
Hoeflea phototrophica
Holdemania filiformis
Homo sapiens
Hordeum vulgare
Hyaloperonospora arabidopsidis
Hydra magnipapillata
Hydrogenobacter thermophilus
Hydrogenobaculum sp. Y04AAS1
Select
What it does ...
STRING is a database of known and predicted protein interactions. The interactions include direct (physical) and indirect (functional) associations; they are derived from four sources:
Genomic Context
High-throughput Experiments
(Conserved) Coexpression
Previous Knowledge
STRING quantitatively integrates interaction data from these sources for a large number of organisms, and transfers information between these organisms where applicable. The database currently covers 9'643'763 proteins from 2'031 organisms.
Use Scenarios
SIB Swiss Institute of Bioinformatics
CPR NNF Center for Protein Research
EMBL European Molecular Biology Laboratory


Home • Download • Help • My Data
STRING 10
STRING - Known and Predicted Protein-Protein Interactions
search by name
search by protein sequence
multiple names
multiple sequences
protein name:
(examples: #1 #2 #3)
Tp53
(STRING understands a variety of protein names and accessions; you can also try a random entry)
organism:
auto-detect
interactors wanted:
COGs
Proteins
Reset
GO !
please enter your protein of interest...
What it does ...
STRING is a database of known and predicted protein interactions. The interactions include direct (physical) and indirect (functional) associations; they are derived from four sources:
Genomic Context
High-throughput Experiments
(Conserved) Coexpression
Previous Knowledge
STRING quantitatively integrates interaction data from these sources for a large number of organisms, and transfers information between these organisms where applicable. The database currently covers oteins from 2'031 organisms.
More Info
Funding / Support
Acknowledgements
Use Scenarios
STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) is being developed at CPR, EMBL, SIB, KU, TUD and UZH.
STRING references: Szklarczyk et al. 2015 / 2013 / 2011 / 2009 / 2007 / 2005 / 2003 / Snel et al. 2000.
Miscellaneous: Access Statistics, Robot Access Guide, Supported Browsers.
What's New? This is version 10 of STRING - now covering more than 2000 organisms, and with improved prediction algorithms!
Sister Projects: check out STITCH and eggNOG - two sister projects built on STRING data!
Previous Releases: Trying to reproduce an earlier finding? Confused? Refer to our old releases.
SIB
Swiss Institute of Bioinformatics
CPR
NNF Center for Protein Research
EMBL
European Molecular Biology Laboratory


Home • Download • Help • My Data
STRING 10
There are several matches for 'tp53'.
Please select one from the list below and press Continue to proceed.
<- Back
Continue ->
organism
protein
Homo sapiens
TP53 - tumor protein p53; Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression (By similarity)
Cavia porcellus
Tp53 - Cellular tumor antigen p53 ; Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression. In cooperation with mitochondrial PPIF is involved in [...]
Danio rerio
tp53 - tumor protein p53; Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression (By similarity)
Oryzias latipes
tp53 - cellular tumor antigen p53 ; Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression (By similarity)
Rattus norvegicus
Tp53 - Cellular tumor antigen p53 ; Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression. In cooperation with mitochondrial PPIF is involved in [...]
Xenopus Silurana
tp53 - tumor protein p53; Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression (By similarity)
Xiphophorus maculatus
tp53 - Cellular tumor antigen p53 ; Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression (By similarity)
Mus musculus
Trp53 - transformation related protein 53; Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression. In cooperation with mitochondrial PPIF is invol [...]
Homo sapiens
ZMAT3 - zinc finger, matrin-type 3; Acts as a bona fide target gene of p53/TP53. May play a role in the TP53-dependent growth regulatory pathway. May contribute to TP53-mediated apoptosis by regulation of TP53 expression and translocation to the nucleus and nucleolus
Homo sapiens
CUL7 - cullin 7; Component of a probable SCF-like E3 ubiquitin-protein ligase complex, which mediates the ubiquitination and subsequent proteasomal degradation of target proteins. Probably plays a role in the degradation of proteins involved in endothelial proliferation and/or differentiation (By similarity). Seems not to promote polyubiquitination and proteasomal degradation of TP53. In vitro, complexes of CUL7 with either CUL9 or FBXW8 or TP53 contain E3 ubiquitin-protein ligase activity. In complex with FBXW8, mediates ubiquitination and consequent degradation of GORASP1, acting as a compo [...]
Homo sapiens
PPP1R13B - protein phosphatase 1, regulatory subunit 13B; Regulator that plays a central role in regulation of apoptosis via its interaction with p53/TP53. Regulates TP53 by enhancing the DNA binding and transactivation function of TP53 on the promoters of proapoptotic genes in vivo
Homo sapiens
PBK - PDZ binding kinase; Phosphorylates MAP kinase p38. Seems to be active only in mitosis. May also play a role in the activation of lymphoid cells. When phosphorylated, forms a complex with TP53, leading to TP53 destabilization and attenuation of G2/M checkpoint during doxorubicin-induced DNA damage
Homo sapiens
RYBP - RING1 and YY1 binding protein; Inhibits ubiquitination and subsequent degradation of TP53, and thereby plays a role in regulating transcription of TP53 target genes. May be implicated in the regulation of the transcription as a repressor of the transcriptional activity of E4TF1. May bind to DNA. Promotes apoptosis
Homo sapiens
MDM4 - Mdm4 p53 binding protein homolog (mouse); Inhibits p53/TP53- and TP73/p73-mediated cell cycle arrest and apoptosis by binding its transcriptional activation domain. Inhibits degradation of MDM2. Can reverse MDM2-targeted degradation of TP53 while maintaining suppression of TP53 transactivation and apoptotic functions
Homo sapiens
DFNA5 - deafness, autosomal dominant 5; Involved in apoptosis and cell survival. Plays a role in the TP53-regulated cellular response to DNA damage probably by cooperating with TP53
Homo sapiens
TP53BP2 - tumor protein p53 binding protein, 2; Regulator that plays a central role in regulation of apoptosis and cell growth via its interactions. Regulates TP53 by enhancing the DNA binding and transactivation function of TP53 on the promoters of proapoptotic genes in vivo. Inhibits the ability of APPBP1 to conjugate NEDD8 to CUL1, and thereby decreases APPBP1 ability to induce apoptosis. Impedes cell cycle progression at G2/M. Its apoptosis-stimulating activity is inhibited by its interaction with DDX42



This is the **evidence view.** Different line colors represent the types of evidence for the association.

Info & Parameters ...

Network Display - Nodes are either colored (if they are directly linked to the input - as in the table) or white (nodes of a higher iteration/depth). Edges, i.e. predicted functional links, consist of up to eight lines: one color for each type of evidence. Hover or click to reveal more information about the node/edge.

Active Prediction Methods:

☑ Neighborhood ☑ Gene Fusion ☑ Co-occurrence
☑ Co-expression ☑ Experiments ☑ Databases ☑ Textmining

required confidence (score): medium confidence (0.400)

- highest confidence (0.900)
- high confidence (0.700)
- medium confidence (0.400)
- low confidence (0.150)

interactors shown: no more than 10 interactors

or custom limit:

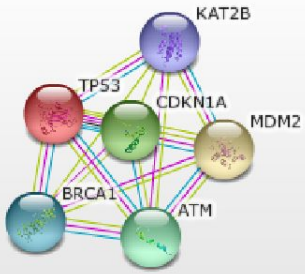
0

☐ Disable Structure Previews inside Network Bubbles

Home · Download · Help · My Data

STRING 10

Save Layout Clustering Enrichment Options



This is the **evidence view.** Different line colors represent the types of evidence for the association.

ATM Close

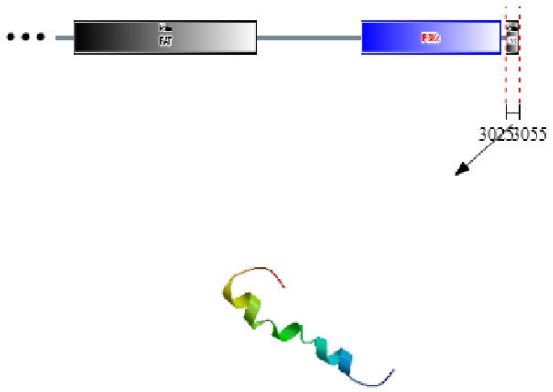
Actions
- re-center network on this node
- add this node to input nodes

Information

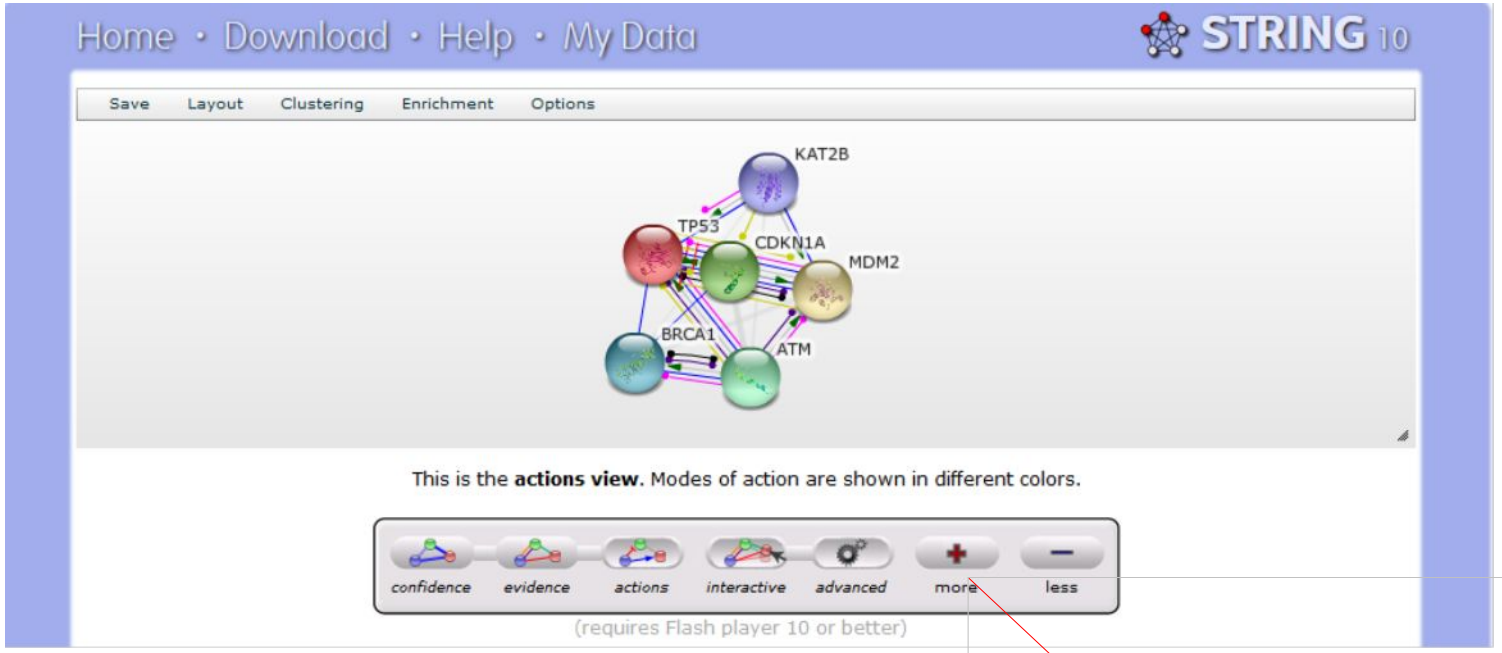
ataxia telangiectasia mutated; Serine/threonine protein kinase which activates checkpoint signaling upon double strand breaks (DSBs), apoptosis and genotoxic stresses such as ionizing ultraviolet A light (UVA), thereby acting as a DNA damage sensor. Recognizes the substrate consensus sequence [ST]-Q. Phosphorylates 'Ser-139' of histone variant H2AX/H2AFX at double strand breaks (DSBs), thereby regulating DNA damage response mechanism. Also plays a role in pre-B cell allelic exclusion, a process leading to expression of a single immunoglobulin heavy chain allele to enforce clonality and [...]

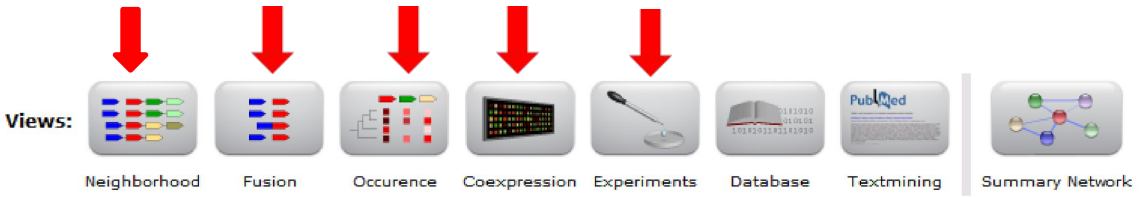
Identifier: ENSP00000278616

- show protein sequence
- homologs among STRING organisms



homology model (**1w1nA**
identity: 38.7%


Home • Download • Help • My Data
STRING 10
Save
Layout
Clustering
Enrichment
Options
KAT2B
TP53
CDKN1A
MDM2
BRCA1
ATM
This is the actions view. Modes of action are shown in different colors.
confidence
evidence
actions
interactive
advanced
more
less
(requires Flash player 10 or better)



Nighborhood: Подобный геномный контекст у разных видов предполагают аналогичную функцию белков.

Fusion: белки, которые связаны внутри геномов, весьма вероятно, могут быть функционально связаны между.

Occurence: Белки, которые имеют аналогичную функцию или находятся в том же пути метаболизма, должны быть отображены вместе и имеют схожие филогенетическое профиль.

Coexpression: Прогнозируемая связь между генами на основе наблюдаемых закономерностей одновременной экспрессии генов.

Info & Parameters ...

Coexpression Display.

Active Prediction Methods:

☑ Neighborhood ☑ Gene Fusion ☑ Co-occurrence
☑ Co-expression ☑ Experiments ☑ Databases ☑ Textmining

required confidence (score): medium confidence (0.400)

or custom value:

interactors shown: no more than 10 interactors

or custom limit:

Update Parameters

Info & Parameters ...

Network Display - Nodes are either colored (if they are directly linked to the input - as in the table) or white (nodes of a higher iteration/depth). Edges, i.e. predicted functional links, consist of up to eight lines: one color for each type of evidence. Hover or click to reveal more information about the node/edge.

Active Prediction Methods:

☑ Neighborhood ☑ Gene Fusion ☑ Co-occurrence
☑ Co-expression ☑ Experiments ☑ Databases ☑ Textmining

required confidence (score): medium confidence (0.400)

highest confidence (0.900)
high confidence (0.700)
medium confidence (0.400)
low confidence (0.150)

interactors shown: no more than 5 interactors

or custom limit:

0

☐ Disable Structure Previews inside Network Bubbles

Update Parameters

Info & Parameters ...

Network Display - Nodes are either colored (if they are directly linked to the input - as in the table) or white (nodes of a higher iteration/depth). Edges, i.e. predicted functional links, consist of up to eight lines: one color for each type of evidence. Hover or click to reveal more information about the node/edge.

Active Prediction Methods:

☑ Neighborhood ☑ Gene Fusion ☑ Co-occurrence
☑ Co-expression ☑ Experiments ☑ Databases ☑ Textmining

required confidence (score): medium confidence (0.400)

or custom value:

interactors shown: no more than 5 interactors

no more than 5 interactors
no more than 10 interactors
no more than 20 interactors
no more than 50 interactors

additional (white) nodes

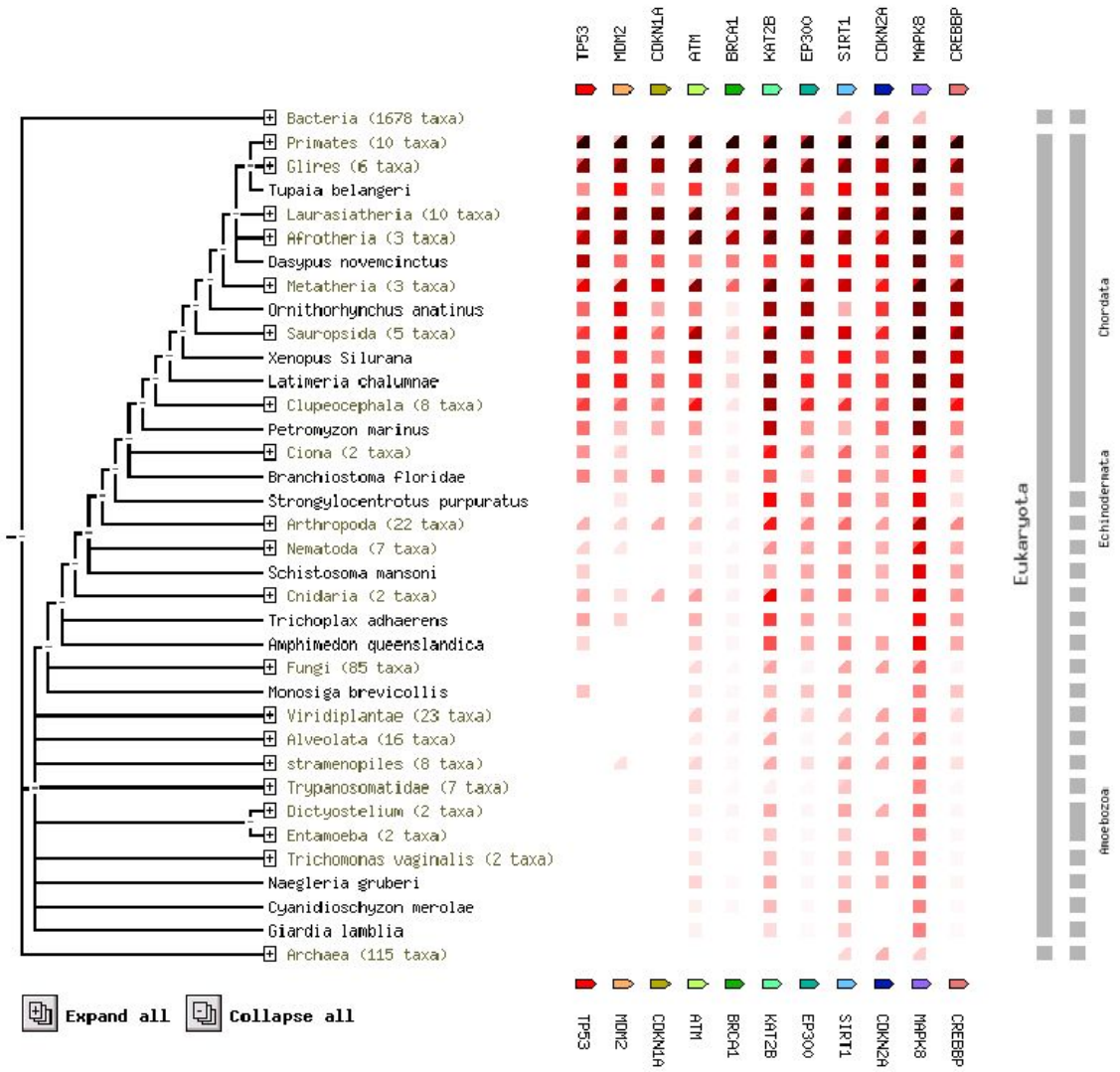
0

☐ Disable Structure Previews inside Network Bubbles

Update Parameters

- Occurrence View -

In which organisms is my protein conserved ?
Functional partners often have similar occurrence patterns.

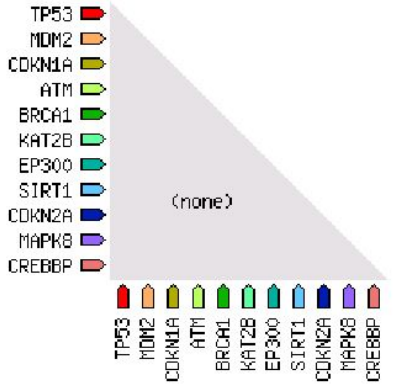


STRING 10

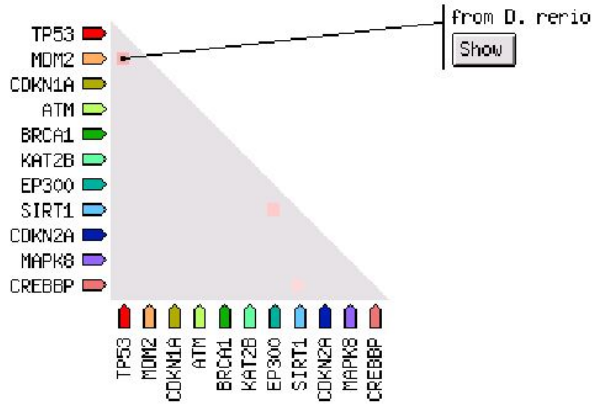
- Coexpression -

association score 0.0 1.0

... from Coexpression in Homo sapiens:



... from Coexpression in other species (transferred):



Relevant datasets in Homo sapiens:

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0676"(tandem affinity purification) assay
CDKN1A EP300 CREBBP [... and 24 other proteins]

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0676"(tandem affinity purification) assay
CDKN1A EP300 CREBBP [... and 24 other proteins]

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0413"(electrophoretic mobility shift assay) assay
TP53 MDM2 CDKN1A

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0412"(electrophoretic mobility supershift assay) assay
TP53 MDM2 CDKN1A

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0402"(chromatin immunoprecipitation assay) assay
TP53 MDM2 CDKN1A

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0402"(chromatin immunoprecipitation assay) assay
TP53 MDM2 CDKN1A

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0402"(chromatin immunoprecipitation assay) assay
TP53 MDM2 CDKN1A

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0413"(electrophoretic mobility shift assay) assay
TP53 MDM2 CDKN1A

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0412"(electrophoretic mobility supershift assay) assay
TP53 MDM2 CDKN1A

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0402"(chromatin immunoprecipitation assay) assay
TP53 MDM2 CDKN1A

Showing 10 out of 956 sets. Show All

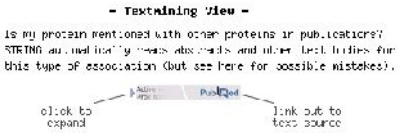
Relevant information transferred from other species:

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0676"(tandem affinity purification) assay
Saccharomyces cerevisiae: SIR2 GCN5 HST1 RSC4 BDF1 RSC1 RSC2 [... and 181 other proteins]

protein-protein interaction (intact) info
Detected by psi-mi:"MI:0676"(tandem affinity purification) assay
Saccharomyces cerevisiae: HST1 GCN5 SIR2 RSC4 BDF1 RSC2 RSC1 [... and 181 other proteins]

protein-protein interaction (intact) info

protein-protein interaction (intact) info



VarWalker: personalized mutation network analysis of putative cancer genes from next-generation sequencing data. PubMed

Ye P, Zhao Z

PLoS Comput Biol. 10(2):e1003460 (2014).

Abstract:

A major challenge in interpreting the large volume of mutation data identified by next-generation sequencing (NGS) is to distinguish driver mutations from neutral passenger mutations to facilitate the identification of targetable genes and new drugs. Current approaches are primarily based on mutation frequencies of single-genes, which lack the power to detect infrequently mutated driver genes and ignore functional interconnection and regulation among cancer genes. We propose a novel mutation network method, VarWalker, to prioritize driver genes in large scale cancer mutation data. VarWalker fits generalized additive models for each sample based on sample specific mutation profiles and builds on the joint frequency of both mutation genes and their close interactors. These interactors are selected and optimized using the Random Walk with Restart algorithm in a protein-protein interaction network. We applied the method in >300 tumor genomes in two large-scale NGS benchmark datasets: 183 lung adenocarcinoma samples and 121 melanoma samples. In each cancer, we derived a consensus mutation subnetwork containing significantly enriched consensus cancer genes and cancer related functional pathways. These cancer specific mutation networks were then validated using independent datasets for each cancer. Importantly, VarWalker prioritizes well known, infrequently mutated genes, which are shown to interact with highly recurrently mutated genes yet have been ignored by conventional single gene based approaches. Utilizing VarWalker, we demonstrated that network-assisted approaches can be effectively adapted to facilitate the detection of cancer driver genes in NGS data.

Excerpts from full text:

... important interactor. For example, TP53 is inhibited by the protein MDM2, but it is activated by **ATM** (●), both of which have a direct interaction with **TP53** (●) [35]. In such cases, consideration of only the most accessible interactor would [...] that function in the regulation of nuclear SMAD2/3 signaling pathways (SMAD2, SMAD4, MYOD1, **CREBBP** (●), JUN, SNIP1, NCOA1, NCOR1, CDK2, AKT1, CDK4, and **KAT2B** (●), pBonferroni−1.17×10−6); and (ii) the proteins that play key roles in the [...] (12), PIK3CG (11), and NTRK3 (11). These genes could also be rarely frequently mutated (e.g., in **CDKN2A** (●) (9), SMAD4 (6), NTRK1 (5), RB1 (4), AKT1 (1), HRAS (1), and MDM2 (1). Functional [...] mutated gene (in Table S7). These interactions were among 28 infrequently mutated genes (in **BRCA1** (●) (= CGC gene) interacts with TP53 (= high-frequency gene, known LIAD gene, and CGC ...

Mechanisms of radiation toxicity in transformed and non-transformed cells. PubMed

Panganiban RA, Snow AL, Day RM

Int J Mol Sci. 14(8):15931-58 (2013).

● SIRT1, sirtuin-1 ● CIP1, Cip1, Sdi1, WAF1, Waf1, waf1 ● p300 ● PCAF, pcaf ● P53, TP53, Tp53, glioblastoma, p53 ● ATM ● JNK ● INK4a, ink4a ● Hdm2, Mdm2 ● BRCA1 ...

Mitochondrial dysfunction in cancer. PubMed

Boland ML, Chourasia AH, Macleod KF

Front Oncol. 3:292 (2013).

● SIRT1, SirT1, Sirt1, Sirtuin, sirtuin ● GCN5 ● glioblastoma, p53 ● ATM, Atm ● JNK ● ARF, CDKN2A, INK4A, p16 ● BRCA1 ...

p53 (●)-Based cytotherapy: exploiting the 'guardian of the genome' to protect normal cells from cytotoxic therapy. PubMed

Rao B, Lain S, Thompson AM

Br J Cancer. 109(12):2954-8 (2013).

● Cip1, Waf1, p21, p21Cip1 ● p53 ● ATM ● Arf, p14ARF, p19ARF ● MDM2, Mdm2, mdm2 ● BRCA1 ...

Suppressed expression of T-box transcription factors is involved in senescence in chronic obstructive pulmonary disease. PubMed

Acquaah-Mensah GK, Malhotra D, Vulimiri M, McDermott JE, Biswal S

PLoS Comput Biol. 8(7):e1002597 (2012).

● SIRT1 ● p21 ● EP300 ● TP53, p53 ● MAPK8 ● ARF, CDKN2A, p16INK4A ● MDM2 ...