

Functional Protein Association  
Networks

# The STRING Database

Presented by Molly Weglarz





# Road Map





# Our Goal

How to Navigate the STRING  
Database



# 1. Live Demo!

Let's start with a tour of the interface





# Critique

## Cons

- No definition on homepage

## Pros

- User friendly!
  - User guide integrated
- Organized
- Manageable amount of info

## Other

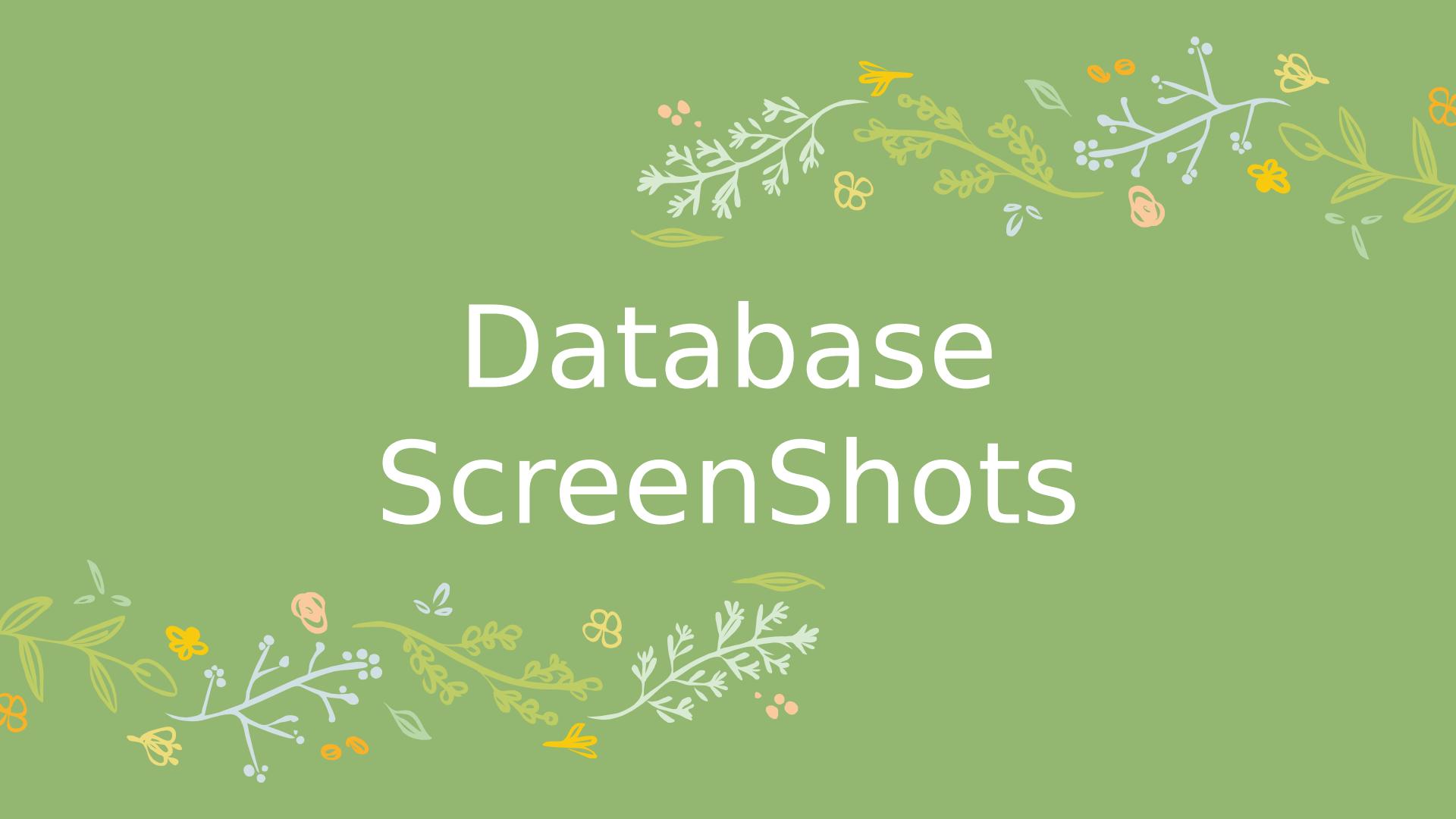
- So much better than BioGrid

# Thank You!

**Any questions?**

You can find me at  
**mweglarz@udallas.edu**





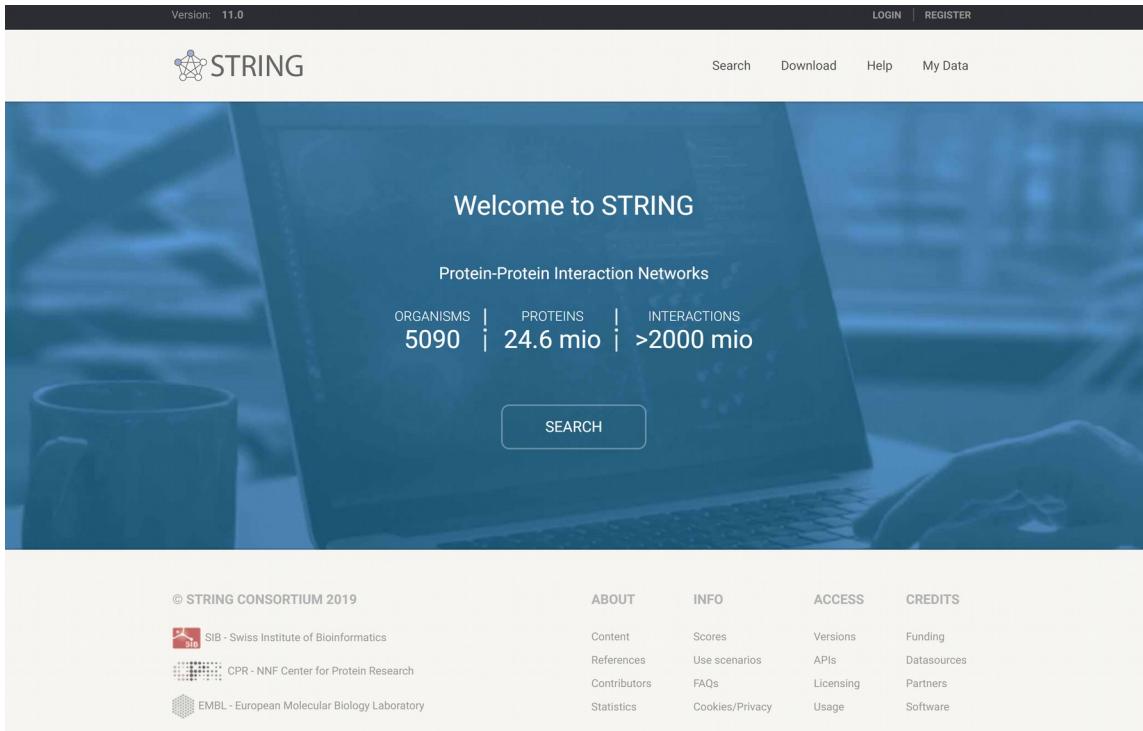
# Database ScreenShots

# What is STRING? &Relevance

- + Protein-Protein Interaction Networks
- + Known and predicted protein-protein interactions.



# 1. First Interface.



The screenshot shows the STRING 11.0 interface. At the top, it displays "Version: 11.0" and navigation links for "LOGIN" and "REGISTER". The main header "STRING" is accompanied by a logo of a molecular structure. Below the header, there are links for "Search", "Download", "Help", and "My Data". The central image is a blue-toned photograph of a person's hands on a keyboard with a computer monitor displaying the STRING logo. The text "Welcome to STRING" and "Protein-Protein Interaction Networks" is overlaid. Key statistics are shown: "ORGANISMS 5090", "PROTEINS 24.6 mio", and "INTERACTIONS >2000 mio". A large "SEARCH" button is centered below these stats. At the bottom, there are links for "© STRING CONSORTIUM 2019", "ABOUT", "INFO", "ACCESS", and "CREDITS", along with logos for SIB, CPR, and EMBL.



# 2. Search Interface.



Version: 11.0

LOGIN | REGISTER

 STRING

Search Download Help My Data

Protein by name >

Protein by sequence >

Multiple proteins >

Multiple sequences >

Proteins with Values/Ranks New >

Organisms >

Protein families ("COGs") >

Examples >

Random entry >

**SEARCH**

Single Protein by Name / Identifier

Protein Name: (examples: #1 #2 #3)

Organism: auto-detect

**SEARCH**

© STRING CONSORTIUM 2019

 SIB - Swiss Institute of Bioinformatics

 CPR - NNF Center for Protein Research

 EMBL - European Molecular Biology Laboratory

ABOUT INFO ACCESS CREDITS

Content Scores Versions Funding

References Use scenarios APIs Datasources

Contributors FAQs Licensing Partners

Statistics Cookies/Privacy Usage Software

# 3. Select the correct protein.

Version: 11.0

LOGIN | REGISTER

## STRING

Search Download Help My Data

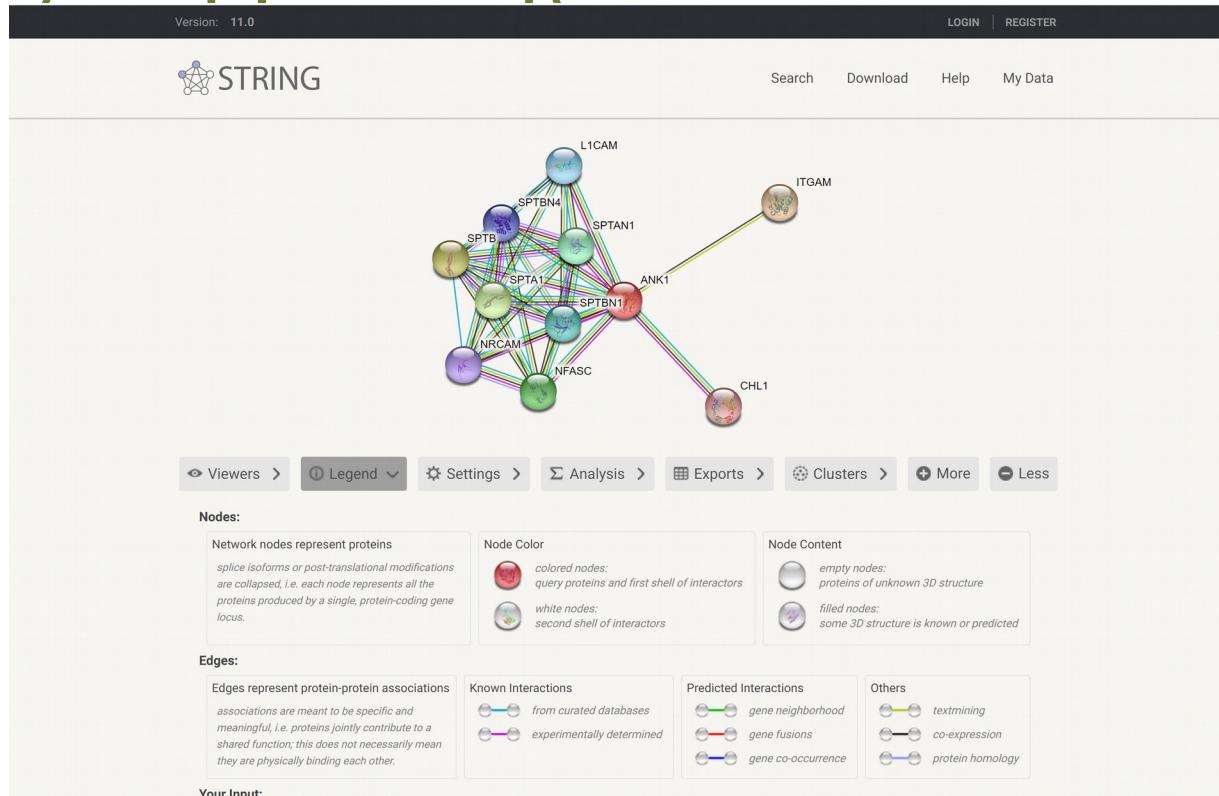
There are several matches for 'Ankyrin 1'. Please select one from the list below and press Continue to proceed.

<- BACK CONTINUE ->

organism	protein
<input checked="" type="checkbox"/> Homo sapiens	ANK1 - Ankyrin-1; Attaches integral membrane proteins to cytoskeletal elements; binds to the erythrocyte membrane protein band 4.2, to Na-K ATPase, to the lymphocyte membrane protein GP85, and to the cytoskeletal proteins fodrin, tubulin, vimentin and desmin. Erythrocyte ankyrins also link spectrin (beta chain) to the cytoplasmic domain of the erythrocytes anion exchange protein; they retain most or all of these binding functions [a.k.a. ANK, NP_000028.3, BC030957, <a href="#">ankyrin 1</a> ]
<input type="checkbox"/> Homo sapiens	ANKRD42 - Ankyrin repeat domain-containing protein 42; Protein phosphatase 1 regulatory subunits; Ankyrin repeat domain containing [a.k.a. ENSP00000377049, OTTHUMT00000392933, FLJ37874]
<input type="checkbox"/> Homo sapiens	PPP1R27 - Protein phosphatase 1 regulatory subunit 27; Inhibits phosphatase activity of protein phosphatase 1 (PP1) complexes; Ankyrin repeat domain containing [a.k.a. DYSFIP1, dysferlin interacting protein 1, NM_001007533.3]
<input type="checkbox"/> Homo sapiens	FANK1 - Fibronectin type 3 and ankyrin repeat domains protein 1; Through the activation of JUN and AP-1-mediated transcription, may regulate apoptosis; Ankyrin repeat domain containing [a.k.a. UNQ6504/PRO21382, HSD13, XM_006718058.1]
<input type="checkbox"/> Homo sapiens	GPANK1 - G-patch domain and ankyrin repeats 1 [a.k.a. BAT4, G5, GPATCH10]
<input type="checkbox"/> Homo sapiens	BANK1 - B-cell scaffold protein with ankyrin repeats; Involved in B-cell receptor (BCR)-induced Ca(2+) mobilization from intracellular stores. Promotes Lyn-mediated phosphorylation of IP3 receptors 1 and 2; Ankyrin repeat domain containing [a.k.a. OTTHUMT00000363163, uc003hv2, Q8NDB2]
<input type="checkbox"/> Homo sapiens	DZANK1 - Double zinc ribbon and ankyrin repeat domains 1 [a.k.a. C20orf12, C20orf84, UP10000456F6B]
<input type="checkbox"/> Homo sapiens	TRANK1 - Tetrastricopeptide repeat and ankyrin repeat containing 1 [a.k.a. KIAA0342, LBA1, XM_006713432]
<input type="checkbox"/> Homo sapiens	SHANK1 - SH3 and multiple ankyrin repeat domains protein 1; Seems to be an adapter protein in the postsynaptic density (PSD) of excitatory synapses that interconnects receptors of the postsynaptic membrane including NMDA-type and metabotropic glutamate receptors via complexes with GKAP/PSD-95 and Homer, respectively, and the actin-based cytoskeleton. Plays a role in the structural and functional organization of the dendritic spine and synaptic junction; Ankyrin repeat domain containing [a.k.a. NX_Q9Y566, OTTHUMT00000268071, CCDS12799.1]
<input type="checkbox"/> Homo sapiens	ANKRD1 - Ankyrin repeat domain-containing protein 1; May play an important role in endothelial cell activation. May act as a nuclear transcription factor that negatively regulates the expression of cardiac genes. Induction seems to be correlated with apoptotic cell death in hepatoma cells; Ankyrin repeat domain containing [a.k.a. CARP, C193, HATA2]



# 4. The STRING Network!!!



# 5. Click on the node...Protein

The screenshot shows a protein interaction network centered on the ANK1 protein. The network includes nodes for L1CAM, SPTBN4, SPTAN1, ITGAM, SPTB, and others. A detailed pop-up window for ANK1 provides the following information:

**Information**  
Ankyrin-1; Attaches integral membrane proteins to cytoskeletal elements; binds to the erythrocyte membrane protein band 4.2, to Na-K ATPase, to the lymphocyte membrane protein GP85, and to the cytoskeletal proteins fodrin, tubulin, vimentin and desmin. Erythrocyte ankyrins also link spectrin (beta chain) to the cytoplasmic domain of the erythrocytes anion exchange protein; they retain most or all of these binding functions  
Identifier: ENSP0000265709, ANK1  
Organism: Homo sapiens

**Viewers** (links to various databases)

**active interaction sources:**

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

**minimum required interaction score:**

The pop-up window also displays a protein sequence alignment from position 10 to 324, a 3D homology model of the protein structure, and navigation controls for the homology model.



# 6. Click on the “Edge”-Interaction Box!

The screenshot shows a protein interaction viewer interface. At the top, there is a network graph with various proteins represented as nodes and interactions as edges. One specific edge between the proteins ANK1 and CHL1 is highlighted with a yellow box, indicating it is the selected interaction.

**Interaction**

**ANK1 [ENSP00000265709]**  
Ankyrin-1; Attaches integral membrane proteins to cytoskeletal elements; binds to the erythrocyte membrane protein band 4.2, to Na-K ATPase, to the lymphocyte membrane protein GP85, and to the cytoskeletal proteins fodrin, tubulin, vimentin and desmin. Erythrocyte ankyrins also link spectrin (beta chain) to the cytoplasmic domain of the erythrocytes anion exchange protein; they retain most or all of these binding functions

**CHL1 [ENSP00000256509]**  
Neural cell adhesion molecule L1-like protein; Extracellular matrix and cell adhesion protein that plays a role in nervous system development and in synaptic plasticity. Both soluble and membranous forms promote neurite outgrowth of cerebellar and hippocampal neurons and suppress neuronal cell death. Plays a role in neuronal positioning of pyramidal neurons and in regulation of both the number of interneurons and the efficacy of GABAergic synapses. May play a role in regulating cell migration in nerve regeneration and cortical development. Potentiates integrin-dependent cell migration [...]

**Less**

**Evidence suggesting a functional link:**

Neighborhood in the Genome:	none / insignificant.
Gene Fusions:	none / insignificant.
Cooccurrence Across Genomes:	none / insignificant.
Co-Expression:	none, but putative homologs are coexpressed in other species (score 0.062). <a href="#">Show</a>
Experimental/Biochemical Data:	yes (score 0.190). In addition, putative homologs were found interacting in other species (score 0.265). <a href="#">Show</a>
Association in Curated Databases:	yes (score 0.900).
Co-Mentioned in PubMed Abstracts:	yes (score 0.041). In addition, putative homologs are mentioned together in other species (score 0.127). <a href="#">Show</a>
Combined Score:	0.942

**Predictions for specific actions:**

Binding: yes (score: 0.923) [Show](#)

1st shelf: no more than 10 interactors [Show](#)

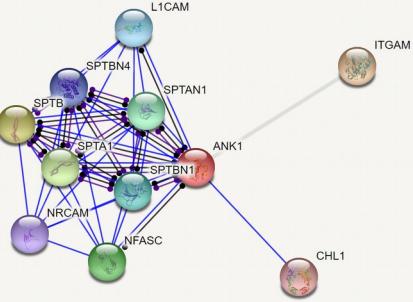


# 7. Settings

Version: 11.0 [LOGIN](#) [REGISTER](#)

 STRING

Search Download Help My Data



Viewers > Legend > Settings > Analysis > Exports > Clusters > More > Less

Basic Settings

meaning of network edges:

- evidence (   line color indicates the type of interaction evidence )
- confidence (   line thickness indicates the strength of data support )
- molecular action (   line shape indicates the predicted mode of action )

active interaction sources:

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

minimum required interaction score:

medium confidence (0.400)

[UPDATE](#)



# 8. Analysis

Viewers > Legend > Settings > Analysis > Exports > Clusters > More Less

**Network Stats**

number of nodes: 11	expected number of edges: 11
number of edges: 36	PPI enrichment p-value: 2.74e-09
average node degree: 6.55	your network has significantly more interactions than expected ( <a href="#">what does that mean?</a> )
avg. local clustering coefficient: 0.929	

**Functional enrichments in your network** Note: some enrichments may be expected here ([why?](#))

**Biological Process (GO)**

GO-term	description	count in gene set	false discovery rate
GO:0007411	axon guidance	9 of 220	9.25e-14
GO:0040011	locomotion	10 of 1144	1.66e-10
GO:0051693	actin filament capping	5 of 32	2.40e-10
GO:0006928	movement of cell or subcellular component	10 of 1355	6.65e-10
GO:0032535	regulation of cellular component size	7 of 333	2.63e-09

(more ...)

**Molecular Function (GO)**

GO-term	description	count in gene set	false discovery rate
GO:0005200	structural constituent of cytoskeleton	6 of 106	5.67e-10
GO:0030506	ankyrin binding	4 of 19	9.74e-09
GO:0008092	cytoskeletal protein binding	7 of 882	1.49e-06
GO:0005088	Ras guanyl-nucleotide exchange factor activity	5 of 243	1.49e-06
GO:0003779	actin binding	5 of 413	1.05e-05

(more ...)

**Cellular Component (GO)**

GO-term	description	count in gene set	false discovery rate
GO:0008091	spectrin	5 of 8	4.01e-12
GO:0014731	spectrin-associated cytoskeleton	4 of 8	2.16e-09
GO:0044304	main axon	5 of 67	1.07e-08
GO:0033267	axon part	6 of 341	2.95e-07
GO:0043194	axon initial segment	3 of 14	1.81e-06

(more ...)

**Reference publications**

publication	(year) title	count in gene set	false discovery rate
PMID:21877111	(2012) Spectrin-based skeleton as an actor in cell signalling	6 of 49	6.98e-09



# 9. Exports

CHL1

Viewers > Legend > Settings > Analysis > Exports > Clusters > More > Less

Export your current network:

- ... as a bitmap image: [download](#) file format is 'PNG'; portable network graphic
- ... as a high-resolution bitmap: [download](#) same PNG format, but resolution at 400 dpi
- ... as a vector graphic: [download](#) SVG: scalable vector graphic - can be opened and edited in Illustrator, CorelDraw, Dia, etc
- ... as simple tabular text output: [download](#) TSV: tab separated values - can be opened in Excel
- ... as an XML summary: [download](#) structured XML interaction data, according to the PSI-MI data standard
- ... network coordinates: [download](#) a flat-file format describing the coordinates and colors of nodes in the network
- ... protein sequences: [download](#) MFA: multi-fasta format - containing the aminoacid sequences in the network
- ... protein annotations: [download](#) a tab-delimited file describing the names, domains and annotated functions of the network proteins

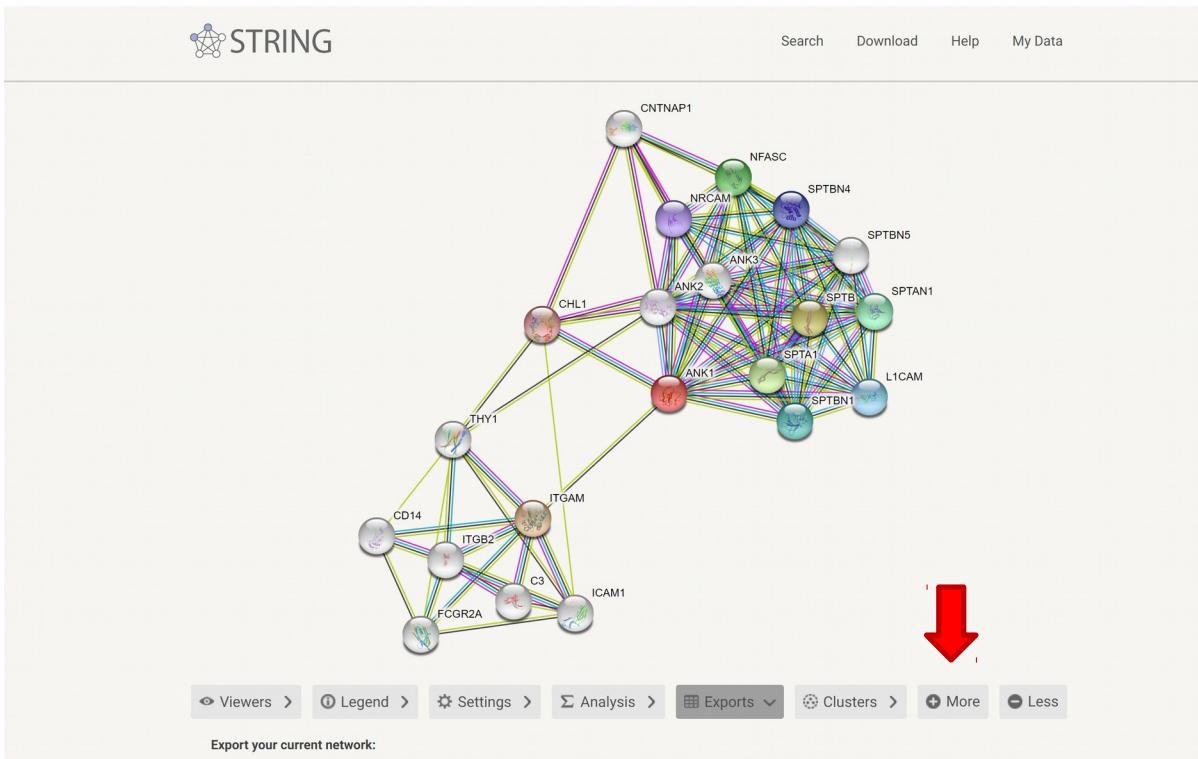
Browse interactions in tabular form:

node1	node2	node1 accession	node2 accession	node1 annotation	node2 annotation	score
ANK1	CHL1	ENSP00000265709	ENSP00000256509	Ankyrin-1; Attaches integral membrane protein...	Neural cell adhesion molecule L1-like pro...	0.942
ANK1	ITGAM	ENSP00000265709	ENSP00000441691	Ankyrin-1; Attaches integral membrane protein...	Integrin alpha-M; Integrin ITGAM/ITGB2 L...	0.999
ANK1	L1CAM	ENSP00000265709	ENSP00000359077	Ankyrin-1; Attaches integral membrane protein...	Neural cell adhesion molecule L1; Neural...	0.957
ANK1	NFASC	ENSP00000265709	ENSP00000344786	Ankyrin-1; Attaches integral membrane protein...	Neurofascin; Cell adhesion, ankyrin-bind...	0.974
ANK1	NRCAM	ENSP00000265709	ENSP00000368314	Ankyrin-1; Attaches integral membrane protein...	Neuronal cell adhesion molecule; Cell ad...	0.943
ANK1	SPTA1	ENSP00000265709	ENSP00000357129	Ankyrin-1; Attaches integral membrane protein...	Spectrin alpha chain, erythrocytic 1; Spec...	0.977
ANK1	SPTAN1	ENSP00000265709	ENSP00000361824	Ankyrin-1; Attaches integral membrane protein...	Spectrin alpha chain, non-erythrocytic 1; ...	0.964
ANK1	SPTB	ENSP00000265709	ENSP00000451752	Ankyrin-1; Attaches integral membrane protein...	Spectrin beta chain, erythrocytic; Spectrin...	0.996
ANK1	SPTBN1	ENSP00000265709	ENSP00000349259	Ankyrin-1; Attaches integral membrane protein...	Spectrin beta chain, non-erythrocytic 1; F...	0.958
ANK1	SPTBN1	ENSP00000265709	ENSP00000263373	Ankyrin-1; Attaches integral membrane protein...	Spectrin beta chain, non-erythrocytic 4; Pl...	0.945
CHL1	ANK1	ENSP00000256509	ENSP00000265709	Neural cell adhesion molecule L1-like pro...	Ankyrin-1; Attaches integral membrane protein...	0.942
ITGAM	ANK1	ENSP00000441691	ENSP00000265709	Integrin alpha-M; Integrin ITGAM/ITGB2 L...	Ankyrin-1; Attaches integral membrane protein...	0.999
L1CAM	ANK1	ENSP00000359077	ENSP00000265709	Neural cell adhesion molecule L1; Neural...	Ankyrin-1; Attaches integral membrane protein...	0.957
L1CAM	SPTA1	ENSP00000359077	ENSP00000357129	Neural cell adhesion molecule L1; Neural...	Spectrin alpha chain, erythrocytic 1; Spec...	0.902
L1CAM	SPTAN1	ENSP00000359077	ENSP00000361824	Neural cell adhesion molecule L1; Neural...	Spectrin alpha chain, non-erythrocytic 1; ...	0.910
L1CAM	SPTB	ENSP00000359077	ENSP00000451752	Neural cell adhesion molecule L1; Neural...	Spectrin beta chain, erythrocytic; Spectrin...	0.902
L1CAM	SPTBN1	ENSP00000359077	ENSP00000349259	Neural cell adhesion molecule L1; Neural...	Spectrin beta chain, non-erythrocytic 1; F...	0.907
L1CAM	SPTBN4	ENSP00000359077	ENSP00000263373	Neural cell adhesion molecule L1; Neural...	Spectrin beta chain, non-erythrocytic 4; Pl...	0.926
NFASC	ANK1	ENSP00000344786	ENSP00000265709	Neurofascin; Cell adhesion, ankyrin-bind...	Ankyrin-1; Attaches integral membrane protein...	0.974
NFASC	NRCAM	ENSP00000344786	ENSP00000368314	Neurofascin; Cell adhesion, ankyrin-bind...	Neuronal cell adhesion molecule; Cell ad...	0.944

◀◀ page 1 of 4 ▶▶



# 10. The “More” View



# 11. The Download Tab



STRING

Search Download Help My Data

## DOWNLOAD

**FILES TOO LARGE?**

Some of the files below can be made smaller (prior to download), by restricting the data to one organism of interest. Choose an organism here:

Homo sapiens ▾ UPDATE

**INTERACTION DATA**

File	Description	Access
<a href="#">9606.protein.links.v11.0.txt.gz (71.2 Mb)</a>	protein network data (scored links between proteins)	
<a href="#">9606.protein.links.detailed.v11.0.txt.gz (110.1 Mb)</a>	protein network data (incl. subscores per channel)	
<a href="#">9606.protein.links.full.v11.0.txt.gz (127.6 Mb)</a>	protein network data (incl. distinction: direct vs. interologs)	
<a href="#">9606.protein.actions.v11.0.txt.gz (14.4 Mb)</a>	interaction types for protein links	
<a href="#">COG.links.v11.0.txt.gz (288.3 Mb)</a>	association scores between orthologous groups	
<a href="#">COG.links.detailed.v11.0.txt.gz (366.7 Mb)</a>	association scores (incl. subscores per channel)	

**ACCESSORY DATA**

File	Description	Access
<a href="#">9606.protein.info.v11.0.txt.gz (1.9 Mb)</a>	list of STRING proteins incl. their display names and descriptions	
<a href="#">9606.protein.sequences.v11.0.fa.gz (6.7 Mb)</a>	sequences of the proteins in STRING (can be used as a blast db)	
<a href="#">COG.mappings.v11.0.txt.gz (381.4 Mb)</a>	orthologous groups (COGs,NOGs,KOGs,...) and their proteins	

# 12. References



 STRING

Search Download Help My Data

## ABOUT

- Content >
- References > **STRING Database – References**
- Contributors >
- Statistics >

When using STRING, please consult (and cite) any of the following references:

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, von Mering C. **STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.** Nucleic Acids Res. 2019 Jan; 47:D607-613. [PubMed](#)

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. **The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible.** Nucleic Acids Res. 2017 Jan; 45:D362-68. [PubMed](#)

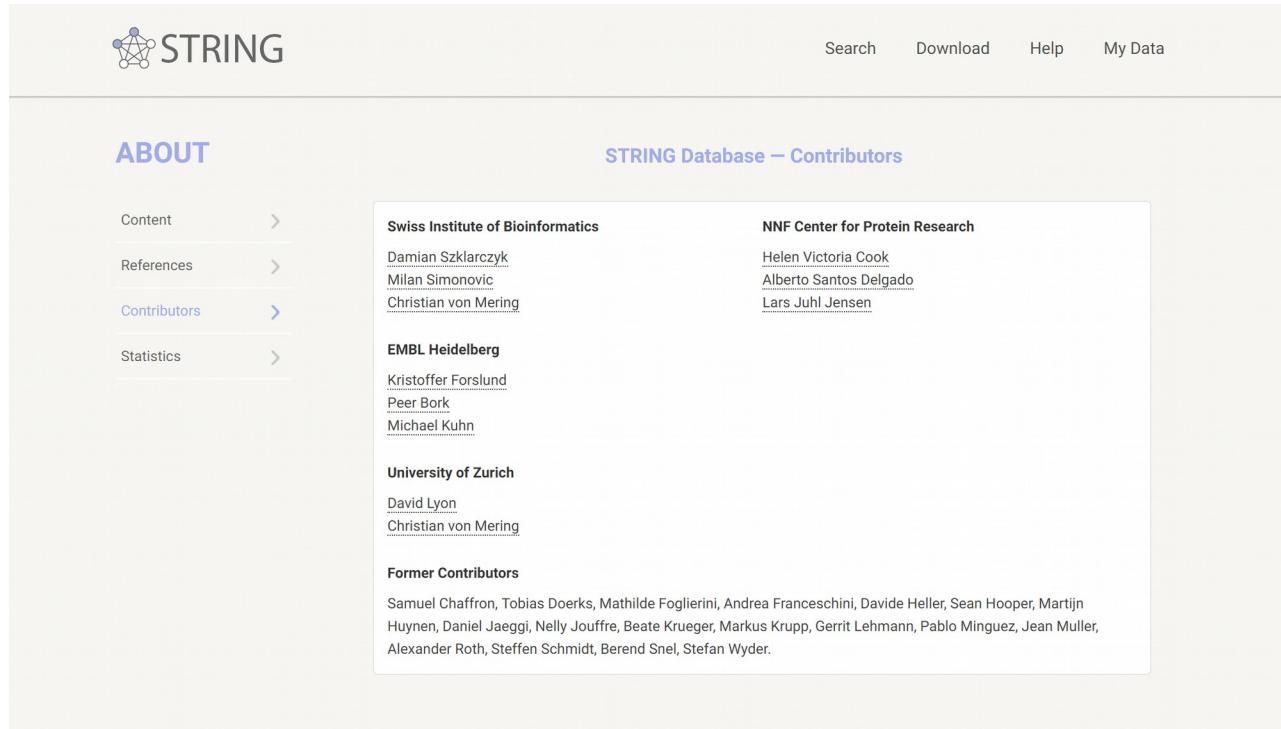
Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C. **STRING v10: protein-protein interaction networks, integrated over the tree of life.** Nucleic Acids Res. 2015 Jan; 43:D447-52. [PubMed](#)

Franceschini A, Lin J, von Mering C, Jensen LJ. **SVD-phy: improved prediction of protein functional associations through singular value decomposition of phylogenetic profiles.** Bioinformatics. 2015 Nov; btv696. [PubMed](#)

Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ. **STRING v9.1: protein-protein interaction networks, with increased coverage and integration.** Nucleic Acids Res. 2013 Jan; 41:D808-15. [PubMed](#)

Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, Jensen LJ, von Mering C.

# 13. Contributors



The screenshot shows the 'ABOUT' section of the STRING Database. The top navigation bar includes links for Search, Download, Help, and My Data. Below this, the 'ABOUT' section has a sidebar with links for Content, References, Contributors (which is currently selected), and Statistics. The main content area is titled 'STRING Database – Contributors'. It lists contributions from several institutions:

- Swiss Institute of Bioinformatics**: Damian Szkłarczyk, Milan Simonovic, Christian von Mering
- NNF Center for Protein Research**: Helen Victoria Cook, Alberto Santos Delgado, Lars Juhl Jensen
- EMBL Heidelberg**: Kristoffer Forslund, Peer Bork, Michael Kuhn
- University of Zurich**: David Lyon, Christian von Mering
- Former Contributors**: Samuel Chaffron, Tobias Doerks, Mathilde Foglierini, Andrea Franceschini, Davide Heller, Sean Hooper, Martijn Huynen, Daniel Jaeggi, Nelly Jouffre, Beate Krueger, Markus Krupp, Gerrit Lehmann, Pablo Minguez, Jean Muller, Alexander Roth, Steffen Schmidt, Berend Snel, Stefan Wyder.



# 14. Statistics



 STRING

Search Download Help My Data

## ABOUT

### STRING Database – Statistics

**Number of organisms**

4'445 Bacteria  
477 Eukaryotes  
168 Archaea  
*5'090 total organisms*

**Number of proteins**

24'584'628 proteins

**Number of interactions (by confidence level)**

52'857'362 interactions at highest confidence (score  $\geq 0.900$ )  
153'506'478 interactions at high confidence or better (score  $\geq 0.700$ )  
530'027'879 interactions at medium confidence or better (score  $\geq 0.400$ )  
3'099'673'831 interactions at low confidence or better (score  $\geq 0.150$ )  
*3'123'056'667 total interactions*

**Top organisms (by query volume)**

1. Homo sapiens
2. Mus musculus
3. Arabidopsis thaliana
4. Saccharomyces cerevisiae
5. Escherichia coli
6. Caenorhabditis elegans
7. Rattus norvegicus
8. Drosophila melanogaster
9. Bacillus subtilis
10. Pseudomonas aeruginosa PAO1



# 15. Content



**STRING**

Search Download Help My Data

## ABOUT

Content >

References >

Contributors >

Statistics >

### STRING Database – Content

STRING is a database of known and predicted protein-protein interactions. The interactions include direct (physical) and indirect (functional) associations; they stem from computational prediction, from knowledge transfer between organisms, and from interactions aggregated from other (primary) databases.

#### Data Sources

Interactions in STRING are derived from five main sources:



Genomic Context  
Predictions



High-throughput Lab  
Experiments



(Conserved) Co-  
Expression



Automated  
Textmining



Previous Knowledge in  
Databases

#### Coverage

The STRING database currently covers 9'643'763 proteins from 2'031 organisms.

# 16. Methods- Interaction



## INFO

- Scores >
- Use scenarios >
- FAQs >
- Cookies/Privacy >

### Interaction Scores

#### The basic principle

In STRING, each protein-protein interaction is annotated with one or more 'scores'. Importantly, these scores do **not** indicate the strength or the specificity of the interaction. Instead, they are indicators of **confidence**, i.e. how likely STRING judges an interaction to be true, given the available evidence. All scores rank from 0 to 1, with 1 being the highest possible confidence. A score of 0.5 would indicate that roughly every second interaction might be erroneous (i.e., a false positive).

#### Transfer scores

For most types of evidence, there are two types of scores: the 'normal' score, and the 'transferred' score. The latter is computed from data that is not originally observed in the organism of interest, but instead in some other organism and then transferred via homology/orthology. All potential source organisms are searched for evidence, but the actual transfers to the receiving organism are made non-redundant (according to 'clades' of closely related organisms in the tree of life).

#### A typical organism

As an example, the model organism 'Escherichia coli K12 MG1655' is shown below – indicating the number of interactions per score type, at a confidence of 'medium' or better (score  $\geq 0.400$ ):

gene neighborhood, normal:	7851 interactions
gene neighborhood, transferred:	11177 interactions
gene fusion:	514 interactions
gene cooccurrence:	35497 interactions
gene coexpression, normal:	12376 interactions
gene coexpression, transferred:	3154 interactions
experiments/biochemistry, normal:	5301 interactions
experiments/biochemistry, transferred:	4113 interactions
annotated pathways, normal:	6726 interactions
annotated pathways, transferred:	1727 interactions
textmining, normal:	27445 interactions
textmining, transferred:	7119 interactions
combined-score, total:	210914 interactions



# 17. FAQ'S



 STRING

[Search](#) [Download](#) [Help](#) [My Data](#)

## INFO

- [Scores](#)
- [Use scenarios](#)
- [FAQs](#)
- [Cookies/Privacy](#)

### Frequently Asked Questions

**Q: How can I obtain the complete data set?**  
STRING has recently changed its licensing model, at the request of the [ELIXIR initiative](#). This means that all its data is now freely available, from the [Download](#) section. The licensing model is [CC BY 4.0](#), requesting proper attribution as the only condition for usage.

**Q: How are scores computed?**  
Q: I am interested in downloading a limited set of interactions, for one or a few proteins only. How can I do that?  
Q: How can I save a certain network?  
Q: For my latest manuscript, I would like to use a network image produced by STRING. Must I ask for permission?  
Q: How can I trace the origin of the different evidences for a given interaction?  
Q: How can I cite STRING?  
Q: Which databases does STRING extract experimental/biochemical data from?  
Q: From which databases does STRING extract curated data?  
Q: How do I extract purely experimental data?  
Q: I need PPIs for a given species, but only from experimental data and not transferred from other species.  
Q: I want to differentiate physical interactions from functional ones within STRING.  
Q: STRING is said to be 'locus-based' and only a single translated protein per locus is stored. What does this mean?  
Q: Does STRING contain any pathway or Gene Ontology information? I see that there is a table called 'funcats' ... ?  
Q: Is there any phenotype or disease information contained in STRING?  
Q: Does the database provide a PubMed Reference ID for each interaction?  
Q: Regarding the 'sets', pathways and complexes ... what is the difference between a "set" and a "collection"?

# 18. Versions.



STRING

Search Download Help My Data

## ACCESS

Version history >

APIs >

Licensing >

Usage statistics >

### STRING Database Version History

Version	Date	Address	Content
11.0	current: since January 19, 2019	<a href="https://string-db.org/">https://string-db.org/</a>	24'584'628 proteins from 5090 organisms; 3'123'056'667 interactions.
10.5	May 14, 2017 to Jan 19, 2019	<a href="https://version-10-5.string-db.org/">https://version-10-5.string-db.org/</a>	9'643'763 proteins from 2031 organisms; 1'380'838'440 interactions.
10.0	Apr 16, 2016 to May 14, 2017	<a href="http://version10.string-db.org/">http://version10.string-db.org/</a>	9'643'763 proteins from 2031 organisms; 932'553'897 interactions.
10.0a	Apr 12, 2105 to Apr 16, 2016	<a href="http://version10a.string-db.org/">http://version10a.string-db.org/</a>	9'643'763 proteins from 2031 organisms; 932'553'897 interactions.
9.1	Dec 27, 2013 to Apr 12, 2015	<a href="http://string91.embl.de/">http://string91.embl.de/</a>	5'214'234 proteins from 1133 organisms; 332'235'675 interactions.
9.05	Mar 3, 2013 to Dec 27, 2013	<a href="http://string905.embl.de/">http://string905.embl.de/</a>	5'214'234 proteins from 1133 organisms; 336'561'678 interactions.
9.0	May 29, 2011 to Mar 3, 2013	<a href="http://string90.embl.de/">http://string90.embl.de/</a>	5'214'234 proteins from 1133 organisms; 224'346'017 interactions.
8.3	May 26, 2010 to May 29, 2011	<a href="http://string83.embl.de/">http://string83.embl.de/</a>	2'590'259 proteins from 630 organisms; 89'236'924 interactions.
8.2	Oct 18, 2009 to May 26, 2010	<a href="http://string-db.org/version_8_2">http://string-db.org/version_8_2</a>	2'590'259 proteins from 630 organisms; 88'633'860 interactions.
8.1	June 14, 2009 to Oct 18, 2009	<a href="http://string-db.org/version_8_1">http://string-db.org/version_8_1</a>	2'590'259 proteins from 630 organisms; 84'076'989 interactions.
8.0	Nov 9, 2008 to June 14, 2009	<a href="http://string.embl.de/version_8_0/">http://string.embl.de/version_8_0/</a>	2'483'276 proteins from 630 organisms; 78'001'069 interactions.
7.1	Oct 19, 2007 to Nov 9, 2008	<a href="http://string.embl.de/version_7_1/">http://string.embl.de/version_7_1/</a>	1'513'782 proteins from 373 organisms; 38'573'579 interactions.
7.0	Jan 15, 2007 to Oct 19, 2007	- no longer available online -	1'513'782 proteins from 373 organisms; 36'252'552 interactions.
6.3	Dec 12, 2005 to Jan 15, 2007	- no longer available online -	736'429 proteins from 179 organisms; 12'014'052 interactions.
6.2	June 26, 2005 to Dec 12, 2005	- no longer available online -	736'429 proteins from 179 organisms; 15'052'353 interactions.