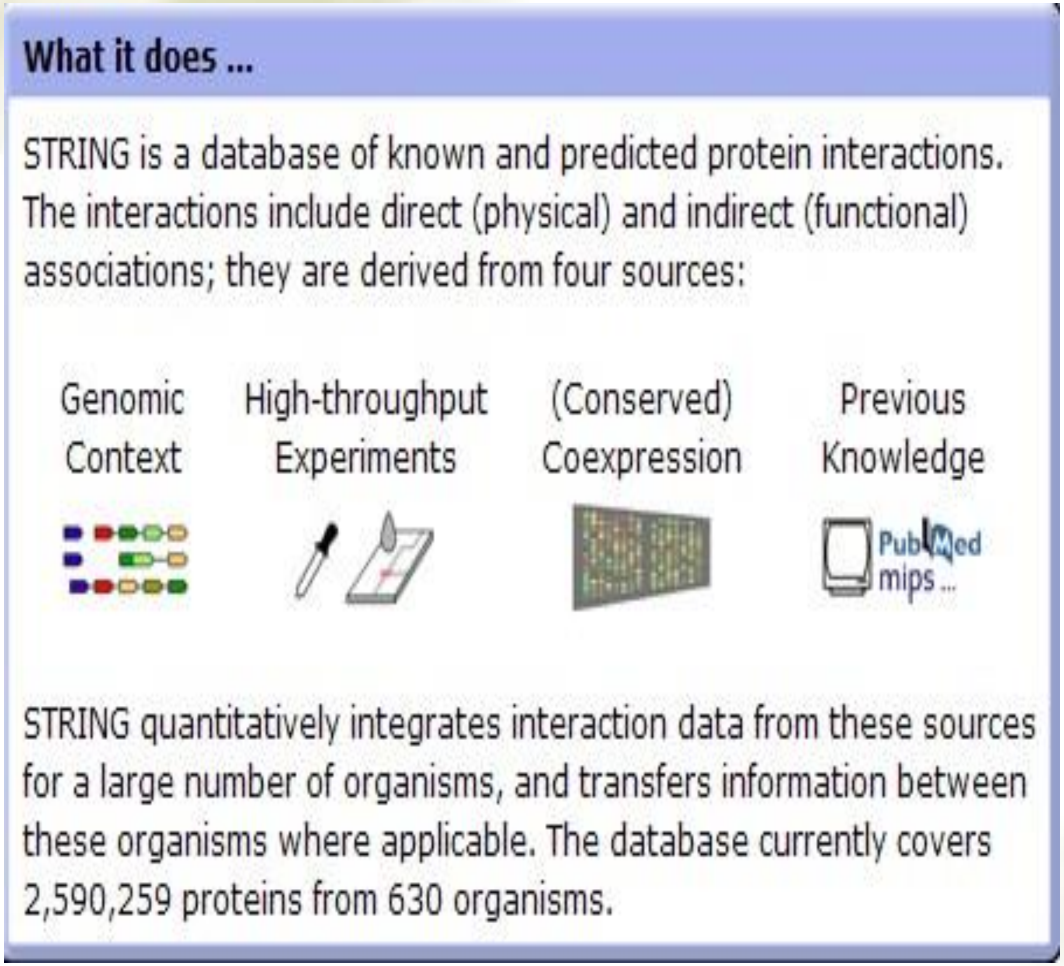
STRING

Introduction

胡奔
2010.10.15

- Background introduction and supporters information
- Network access adresses
- Type of data storage
- Database capacity and update frequency (way)
- Query interface and query method
- Usage

Background introduction and supporters information



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:

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 2,590,259 proteins from 630 organisms.

Supporters:

CPR
NNF Center for Protein Research

EMBL
European Molecular Biology Laboratory

SIB
Swiss Institute of Bioinformatics

SUND-KU
University of Copenhagen

TUD
Technical University Dresden, Biotec

UZH
University of Zurich

CPR--- Center for Protein Research

http://www.cpr.ku.dk/

- The Novo Nordisk Foundation has been established at the Faculty of Health Sciences, University of Copenhagen, to promote basic and applied discovery research on human proteins of medical relevance. The establishment of the center announced in April 2007 has been made possible by a donation of 600 million DKK (~113 MUSD) from the Novo Nordisk Foundation and through significant contributions from the University of Copenhagen for the renovation of the Center laboratories.
- The Center comprises a wide range of expertise and skills within its research departments, with activities in the areas of disease systems biology, proteomics, high throughput protein production and characterisation, chemical biology, disease biology, and protein therapeutics. The Center will also contribute to the progress of translational research within medicine and provide fundamental insights which can be used to promote drug discovery and development.

KU---University of Copenhagen

http://healthsciences.ku.dk/

EMBL--European Molecular Biology Laboratory .

http://www.embl.org/

The Laboratory has five units: the main Laboratory in Heidelberg, and outstations in Hinxton near Cambridge (the European Bioinformatics Institute), Grenoble, Hamburg, and Monterotondo near Rome.

EMBL is international, innovative and interdisciplinary.

- **Brief History**
- The European Molecular Biology Laboratory was the idea of prominent scientists such as the American physicist and molecular biologist Leo Szilárd and Nobel Prize winners James D. Watson and John C. Kendrew. Their goal was to create a CERN-like supranational research centre to redress the balance in the strongly US-dominated field of molecular biology.

- The founding contract of this centre of excellence was signed in July 1974 on a basis of an intergovernmental treaty of nine European countries plus Israel. Since then, the number of member states has increased progressively, until Luxembourg became the twentieth member in 2007, and Australia joined as an associate member in 2008.

- **Mission**
- EMBL was set up in order to promote molecular biology across Europe, and to provide an attractive alternative to the United States as a workplace for Europe's leading young molecular biologists. To accomplish this, EMBL has pursued five major missions:

- **Basic Research in Molecular Biology**

- **Technology and Instrumentation**

- **Facilities and Services**

- **Teaching and Training**

- **Technology Transfer**

SIB—Swiss Institute of Bioinformatics

http://www.isb-sib.ch/

- The SIB Swiss Institute of Bioinformatics is an academic, non-profit foundation established in 1998. SIB coordinates research and education in bioinformatics throughout Switzerland and provides high quality bioinformatics services to the national and international research community.

TUD-- Technische Universität Dresden

http://www.biotec.tu-dresden.de/

- The Biotechnology Center (BIOTEC) of the **Technische Universität Dresden** is a unique interdisciplinary center focusing on research and teaching in molecular bio-engineering. The BIOTEC hosts top international research groups working on genomics, proteomics, biophysics, cellular machines, molecular genetics, tissue engineering, and bioinformatics.

UZH—University of Zurich

http://www.uzh.ch/

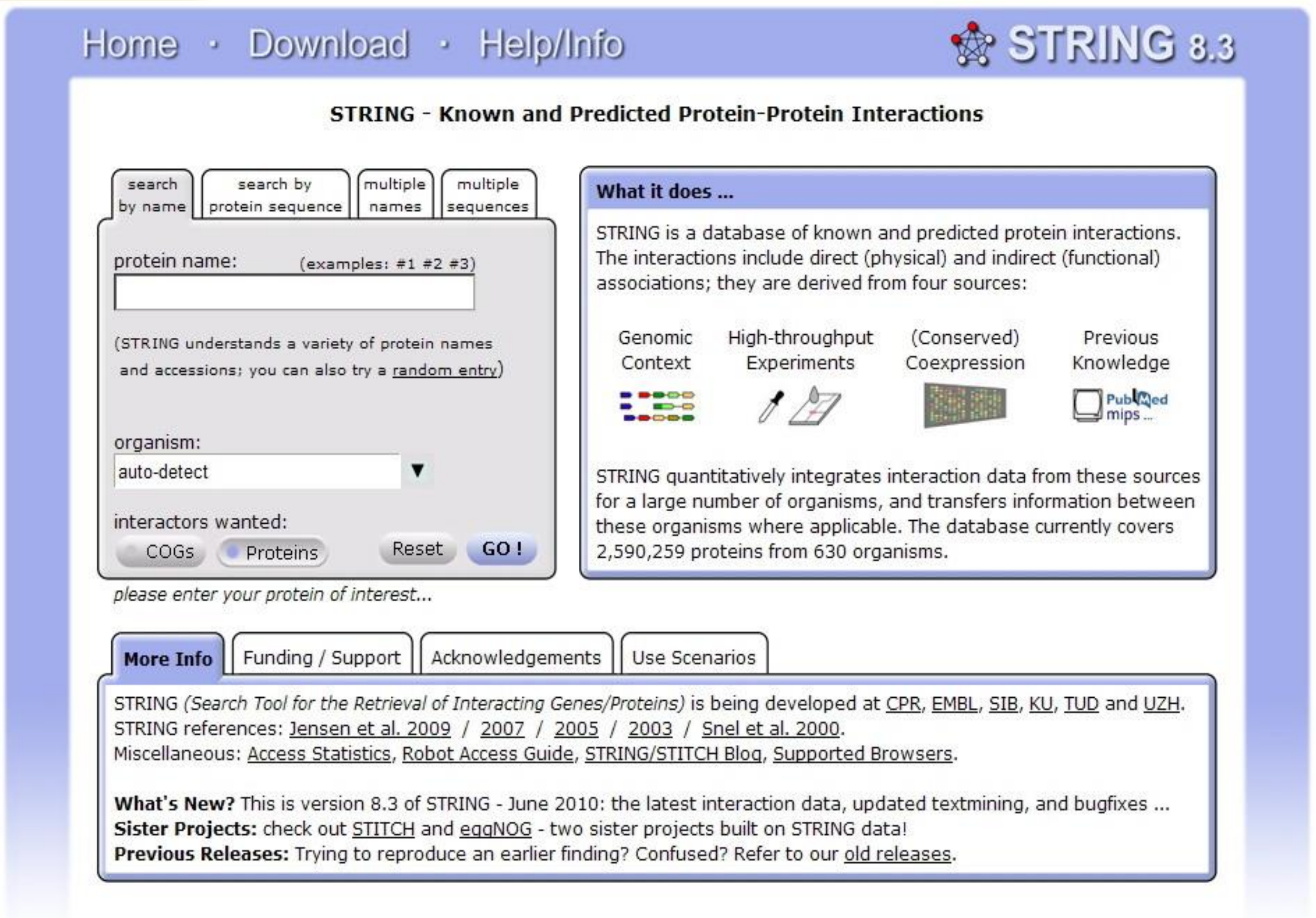
- The University of Zurich – the largest university and the broadest range of courses in Switzerland

- The University of Zurich was founded in 1833, though its origins stretch back as far as 1525 and the days of protestant reformer Ulrich Zwingli. Today the University enjoys international renown as a place of education and research. Two thousand lecturers in 140 special institutes provide the broadest range of subjects and courses available from any Swiss seat of higher education. With 24,000 students and 1,900 graduates every year, Zurich is also Switzerland’s largest university. The University provides academic services, works with the private sector and considers itself part of a national and global network for the acquisition and dissemination of knowledge.

- Background introduction and supporters information
- <u>Network access adresses</u>
- Type of data storage
- Database capacity and update frequency (way)
- Query interface and query method
- Usage

Network access adresses

http://string-db.org/
http://string.embl.de/



STRING - Known and Predicted Protein-Protein Interactions

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 ▼

interactors wanted: ○ COGs ● Proteins

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 2,590,259 proteins from 630 organisms.

STRING (*Search Tool for the Retrieval of Interacting Genes/Proteins*) is being developed at CPR, EMBL, SIB, KU, TUD and UZH.
STRING references: Jensen et al. 2009 / 2007 / 2005 / 2003 / Snel et al. 2000.
Miscellaneous: Access Statistics, Robot Access Guide, STRING/STITCH Blog, Supported Browsers.

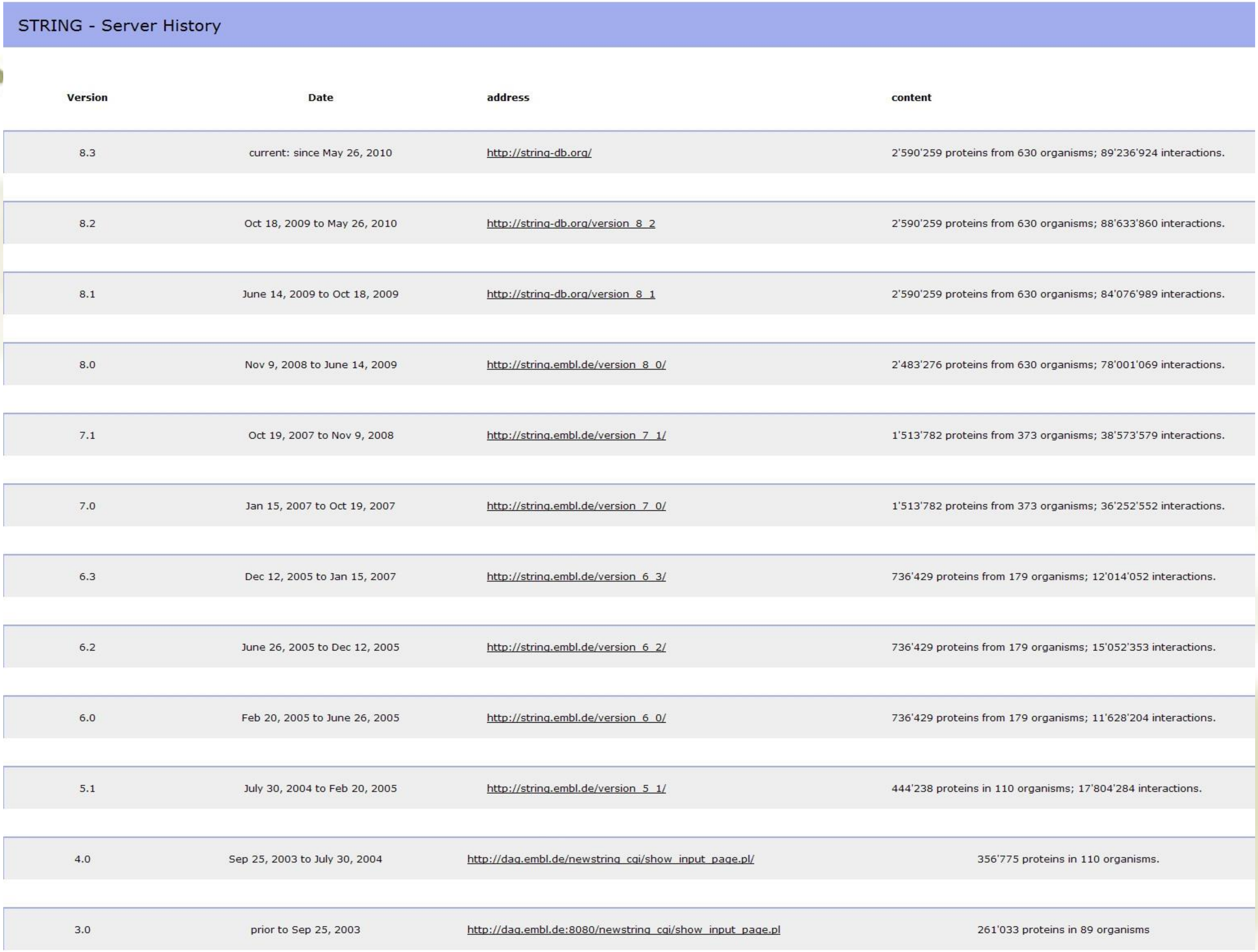
What's New? This is version 8.3 of STRING - June 2010: the latest interaction data, updated textmining, and bugfixes ...
Sister Projects: check out STITCH and eggNOG - two sister projects built on STRING data!
Previous Releases: Trying to reproduce an earlier finding? Confused? Refer to our old releases.

- Background introduction and supporters information
- Network access adresses
- <u>Type of data storage</u>
- Database capacity and update frequency (way)
- Query interface and query method
- Usage

- Background introduction and supporters information
- Network access adresses
- Type of data storage
- Database capacity and update frequency (way)
- Query interface and query method
- Usage

Database capacity and update frequency (way)

- The database currently covers 2,590,259 proteins from 630 organisms and 89,236,924 interactions.
- The database has 12 versions now,which was updated from September 2003 and the details as follow:

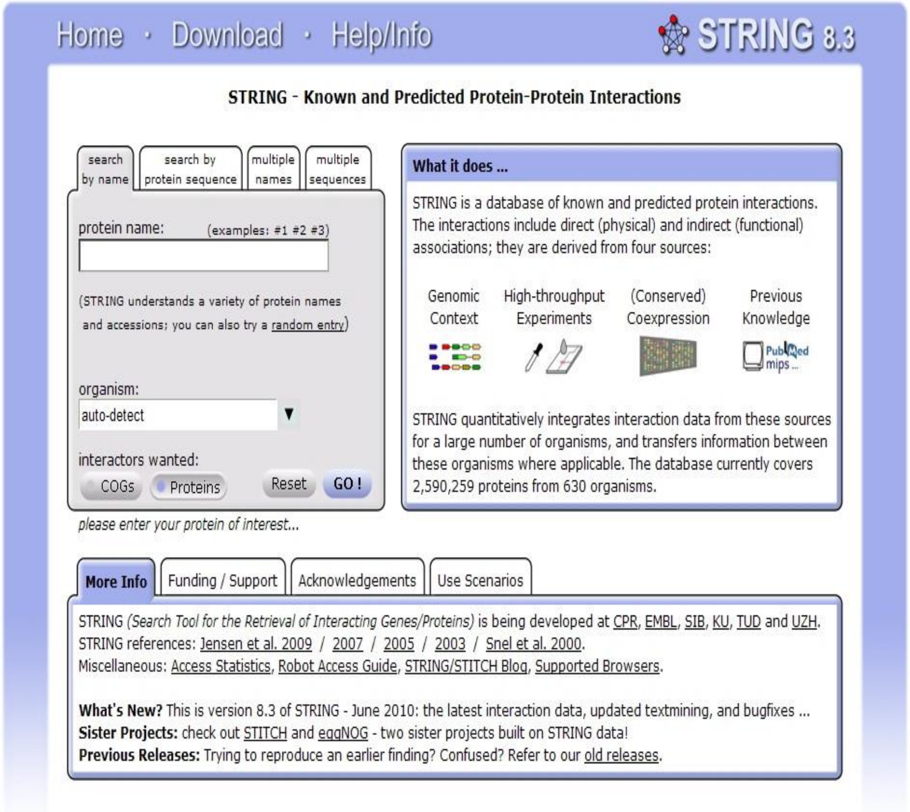


STRING - Server History

Version	Date	address	content
8.3	current: since May 26, 2010	http://string-db.org/	2'590'259 proteins from 630 organisms; 89'236'924 interactions.
8.2	Oct 18, 2009 to May 26, 2010	http://string-db.org/version_8_2	2'590'259 proteins from 630 organisms; 88'633'860 interactions.
8.1	June 14, 2009 to Oct 18, 2009	http://string-db.org/version_8_1	2'590'259 proteins from 630 organisms; 84'076'989 interactions.
8.0	Nov 9, 2008 to June 14, 2009	http://string.embl.de/version_8_0/	2'483'276 proteins from 630 organisms; 78'001'069 interactions.
7.1	Oct 19, 2007 to Nov 9, 2008	http://string.embl.de/version_7_1/	1'513'782 proteins from 373 organisms; 38'573'579 interactions.
7.0	Jan 15, 2007 to Oct 19, 2007	http://string.embl.de/version_7_0/	1'513'782 proteins from 373 organisms; 36'252'552 interactions.
6.3	Dec 12, 2005 to Jan 15, 2007	http://string.embl.de/version_6_3/	736'429 proteins from 179 organisms; 12'014'052 interactions.
6.2	June 26, 2005 to Dec 12, 2005	http://string.embl.de/version_6_2/	736'429 proteins from 179 organisms; 15'052'353 interactions.
6.0	Feb 20, 2005 to June 26, 2005	http://string.embl.de/version_6_0/	736'429 proteins from 179 organisms; 11'628'204 interactions.
5.1	July 30, 2004 to Feb 20, 2005	http://string.embl.de/version_5_1/	444'238 proteins in 110 organisms; 17'804'284 interactions.
4.0	Sep 25, 2003 to July 30, 2004	http://dag.embl.de/newstring_cgi/show_input_page.pl/	356'775 proteins in 110 organisms.
3.0	prior to Sep 25, 2003	http://dag.embl.de:8080/newstring_cgi/show_input_page.pl	261'033 proteins in 89 organisms

- Background introduction and supporters information
- Network access adresses
- Type of data storage
- Database capacity and update frequency (way)
- <u>Query interface and query method</u>
- Usage

Query interface and query method



STRING - Known and Predicted Protein-Protein Interactions

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

interactors wanted:

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:

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 2,590,259 proteins from 630 organisms.

STRING *(Search Tool for the Retrieval of Interacting Genes/Proteins)* is being developed at CPR, EMBL, SIB, KU, TUD and UZH.
STRING references: Jensen et al. 2009 / 2007 / 2005 / 2003 / Snel et al. 2000.
Miscellaneous: Access Statistics, Robot Access Guide, STRING/STITCH Blog, Supported Browsers.

What's New? This is version 8.3 of STRING - June 2010: the latest interaction data, updated textmining, and bugfixes ...
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.

- Search by name
- Search by protein sequence
- Multiple names
- Multiple sequences

- a red line - fusion evidence;
- a green line - neighborhood evidence;
- a blue line - coocurrence evidence;
- a black line - coexpression evidence
- a purple line - experimental evidence;
- a yellow line - textmining evidence;
- a light blue line - database evidence;

- Background introduction and supporters information
- Network access adresses
- Type of data storage
- Database capacity and update frequency (way)
- Query interface and query method
- Usage

Usage

- To search the information of known protein interactions
- To predicted protein interactions

END
THANKS!