

CBI Resource Review

Chen Lei, Feng Shengfu, Pan Runyu, Wang Fei, Wang Han,
Wang Xin, Yang Jin, Zhao Mo

Presenter: Yang Jin 杨津

School of Life Sciences, Peking University

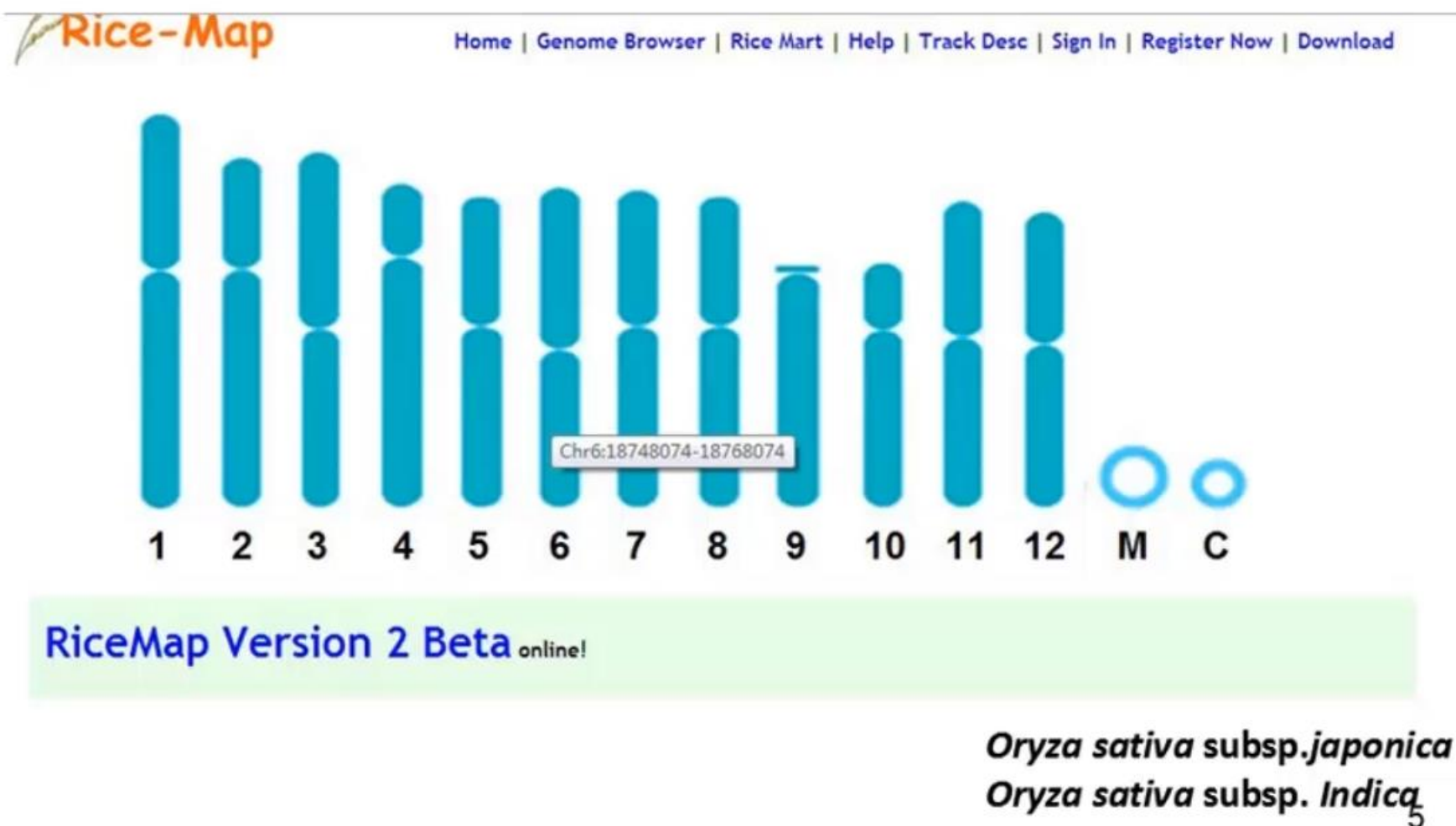
Catalogue

- Genome Analysis and Comparison
- Protein Families of Biomedical Signification
- Gene Expression Regulation
- Bioinformatics Infrastructure

Genome Analysis and Comparison

- 07-Rice-Map -- A Map Like Rice Genome Browser and Related Resource
- 07-CVTree -- Composition Vector Tree constructs phylogeny trees
- 06-ColinearScan -- Dynamic programming for colinearity analysis of genomes
- 96-PGAAS -- prokaryotic genome assembly assistant system
- 96-PSMY -- Potential Synergy mapping of Yeast Genes

Rice-Map -- A Map Like Rice Genome Browser and Related Resource



Go to

in

GO!

[Advanced Search](#)
[Tracks](#)
[Landmarks](#)
[Ensembl Plants](#)
[GRAMENE](#)


Copyright © 2007-2012, Center for Bioinformatics, Dalian University. All rights reserved!

Current Tracks	Entry Details	My Landmarks
rice:chr8:User Annotation	☆☆☆☆☆ (0 evaluations)	
rice:chr8:MSU_Os1_Rice_Gene_Models	☆☆☆☆☆ (2 evaluations)	
rice:chr8:RGP_Predicted_Geneset	☆☆☆☆☆ (1 evaluations)	
rice:chr8:Epigenetics_modification (shoots_of_four-leaf_stage_seedlings_of_japonica)	☆☆☆☆☆ (1 evaluations)	
rice:chr8:PhastCons	☆☆☆☆☆ (0 evaluations)	

Composition Vector Tree

- CVTree constructs whole-genome based phylogenetic trees without sequence alignment by using a Composition Vector (CV) approach.
- It was first developed to infer evolutionary relatedness of microbial organisms and then successfully applied to viruses, chloroplasts, and fungi.
- It circumvents the ambiguity of choosing the genes for phylogenetic reconstruction and avoids the necessity of aligning sequences of essentially different length and gene content.

Composition Vector Tree

First Use:

If this is your first time to run CVTree or you wish to start another project, just press [Create a new project] button. You will get a project number and the CVTree Project Page appears.

Create a new project

Example project

Modify parameters and rerun:

If you want to revise some parameters of a previous project and run it again, you can enter your Project Number here and load the project. **NOTE:** project will be kept for 2 days after the last run.

Project Number:

Reload project

Server load average in 1 min: 50 %

Composition Vector Tree

First Use:

If this is your first time to run CVTree press [Create a new project] button. Your Project Page appears.

Create a new project

Modify parameters and rerun:

If you want to revise some parameters of can enter your Project Number here and for 2 days after the last run.

Project Number:

Server load average in 1 min: 50 %

Parameters:

K-tuple length :

Sequence type : ☐ DNA ☒ Protein

Email me result:

All parameters are fine, run project!

User's files:

File used for Protein CVTree (*.faa):

out-group	Date	File Size	File Name
<input checked="" type="radio"/>			Random (auto-chosen by Phylip package)

File used for DNA CVTree (*.ffn):

out-group	Date	File Size	File Name
-----------	------	-----------	-----------

Delete selected files or Upload this file: 未选择文件 (100M)

[upload Protein sequences \(can be compressed\)](#)

Inbuilt genomic data:

Distribution of selected genomes:
None

ColinearScan -- Dynamic programming for colinearity analysis of genomes

Overview

Dynamic programming which is widely used in sequence alignment can also be used to detect colinearity between/within chromosome(s).

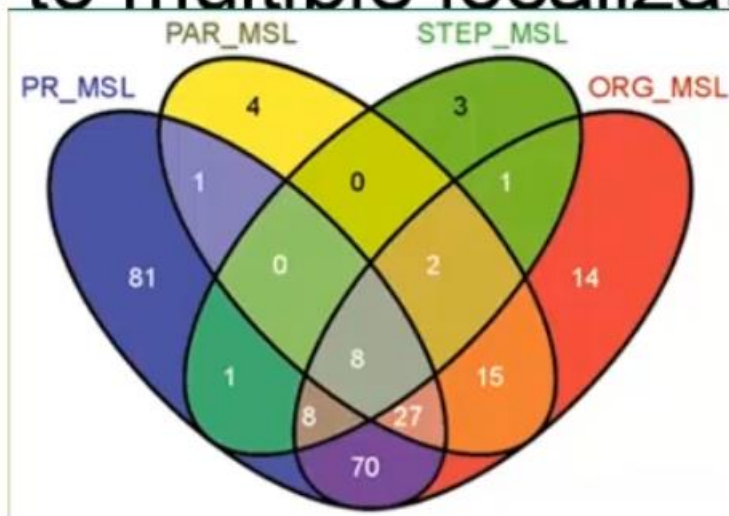
ColinearScan is a free software package distributed under GNU Public Licence v2 on a statistical model for colinearity prediction.

Protein Families of Biomedical Signification

- 12-TSdb -- A database of transporter substrates linking metabolic pathways and transporter systems on a genome scale via their shared substrates
- 10-PathLocdb -- A comprehensive database for the subcellular localization of metabolic pathways and its application to multiple localization analysis

PathlocDB

- A comprehensive database for
- subcellular localization of metabolic pathways
- application to multiple localization analysis




PathlocDB

- Our database allows you to:
- Search & browse the metabolic pathways by their subcellular localizations and organisms
- Systematic comparing the localization profiles of metabolic pathways between different organisms
- Discover the potential regulatory mechanisms and suspicious localization of metabolic pathways
- Clarify the pathway boundary from the view of subcellular localization
- Discover the mechanism of intermediates communication between different subcellular localizations

PathlocDB

- Superpathways:
 - 337 (SwissProt dataset)
 - 215 (KEGG dataset)
 - 337 (UniProt dataset)
- Pathways with localization annotation:
 - 43014
- Proteins in database:
 - 80676
- Pathways with multiple localization annotation:
 - 4477
- Superpathways with multiple localization annotation:
 - 682

PathlocDB

 **PathLoc** Pathway Localization

[Home](#) [Browser](#) [Query](#) [Discussion](#) [Blast](#) [Unsubscribed](#)

[fungi](#) | [plants](#) | [invertebrates](#) | [vertebrates](#) | [mammals](#) | [rodents](#) | [human](#)

Browser PathLoc by different features

[+ Expand all](#) - [Collapse all](#)

☒ Subcellular localization

- ☒ cell membrane
 - [PLSP63](#) - Amino-acid degradation ; L-phenylalanine degradation ; acetoacetic acid and fumarate from L-phenylalanine (mitochondrion, cell membrane, golgi apparatus, endoplasmic reticulum, cytoplasm,)
 - [PLSP85](#) - Carbohydrate degradation ; glycolysis ; pyruvate from D-glyceraldehyde 3-phosphate (cytoplasm, cell membrane, extracellular, myofibril, cell wall, chloroplast, nucleus, plastid, membrane,)
 - [PLSP183](#) - Lipid metabolism ; malonyl-CoA biosynthesis ; malonyl-CoA from acetyl-CoA (mitochondrion, cell membrane, chloroplast, plastid, membrane, cytoplasm,)
 - [PLSP244](#) - Photosynthesis ; C4 acid pathway (mitochondrion, cell membrane, chloroplast, plastid, cytoplasm,)
 - [PLSP262](#) - Porphyrin biosynthesis ; chlorophyll biosynthesis (cyanelle, mitochondrion, cell membrane, etioplast prolamellar body, chloroplast, plastid,)
 - [PLSP287](#) - Protein modification ; protein ubiquitination (extracellular matrix, mitochondrion, vesicle, cell membrane, midbody, lipid raft, golgi apparatus, myofibril, nucleus, lysosome, endosome, endoplasmic reticulum, peroxisome, vacuole, cytoplasm, extracellular, synapse, cytoskeleton, membrane,)
 - [PLKE6](#) - 00051_Fructose and mannose metabolism (endosome, mitochondrion, endoplasmic reticulum, peroxisome, cytoplasm, cell membrane, golgi apparatus, chloroplast, nucleus, plastid, ER-Golgi intermediate compartment, membrane, lysosome,)
 - [PLKE17](#) - 00150_Androgen and estrogen metabolism (mitochondrion, endoplasmic reticulum, cytoplasm, cell membrane, extracellular, golgi apparatus, nucleus, microsome, membrane, lysosome,)
 - [PLKE22](#) - 00230_Purine metabolism (extracellular matrix, mitochondrion, endoplasmic reticulum, peroxisome, vacuole, cytoplasm, extracellular, cell membrane, golgi apparatus, chloroplast, nucleus, cytoskeleton, plastid, cell projection, lysosome,)
 - [PLKE24](#) - 00240_Pyrimidine metabolism (mitochondrion, endoplasmic reticulum, vacuole, cytoplasm, cell membrane, extracellular, golgi apparatus, chloroplast, cytoskeleton, nucleus, plastid, membrane, lysosome,)
 - [PLKE36](#) - 00350_Tyrosine metabolism (extracellular matrix, mitochondrion, chromaffin granule, cell membrane, golgi apparatus, nucleus, plastid, endosome, endoplasmic reticulum, peroxisome, vacuole, cytoplasm, extracellular, synapse, cytoskeleton, chloroplast, membrane,)
 - [PLKE38](#) - 00360_Phenylalanine metabolism (extracellular matrix, mitochondrion, cell membrane, golgi apparatus, nucleus, plastid, lysosome, endosome, endoplasmic reticulum, peroxisome, vacuole, cytoplasm, extracellular, synapse, cytoskeleton, chloroplast, membrane,)
 - [PLKE39](#) - 00361_gamma-Hexachlorocyclohexane degradation (mitochondrion, endoplasmic reticulum, vacuole, cytoplasm, extracellular, cell membrane, cell wall, nucleus, microsome, membrane, lysosome,)
 - [PLKE49](#) - 00460_Cyanoamino acid metabolism (mitochondrion, cell membrane, extracellular, cell wall, peroxisome, cytoplasm, lysosome,)
 - [PLKE53](#) - 00500_Starch and sucrose metabolism (amyloplast, mitochondrion, extracellular matrix, cell membrane, golgi apparatus, cell wall, nucleus, plastid, microsome, lysosome, endoplasmic reticulum, apoplast, cytoplasm, extracellular, chloroplast, membrane,)
 - [PLKE60](#) - 00530_Aminosugars metabolism (endosome, mitochondrion, extracellular matrix, endoplasmic reticulum, vacuole, cytoplasm, extracellular, cell membrane, golgi apparatus, cell wall, chloroplast, nucleus, plastid, membrane, lysosome,)
 - [PLKE61](#) - 00531_Glycosaminoglycan degradation (endosome, cell membrane, extracellular, apoplast, nucleus, cytoplasm, lysosome,)
 - [PLKE64](#) - 00534_Heparan sulfate biosynthesis (extracellular matrix, extracellular, cell membrane, golgi apparatus, endoplasmic reticulum, membrane,)
 - [PLKE66](#) - 00561_Glycerolipid metabolism (endosome, mitochondrion, apoplast, endoplasmic reticulum, vacuole, cytoplasm, cell membrane, extracellular, golgi apparatus, chloroplast, nucleus, plastid, ER-Golgi intermediate compartment, lysosome,)

Gene Expression Regulation

- 11-**AutismKB** -- AutismKB is an evidence-based knowledgebase of autism
- 11-LSD -- Leaf Senescence Database
- 10-**PlantTFDB 3.0** -- Plant Transcription Factor Database
- 10-AHD 2.0 -- Arabidopsis Hormone Database
- 09-RLEDB -- Rating-Limiting Enzyme regulation database
- 09-HomeoDB -- A Database of Homeobox Gene Diversity (Mirror at Oxford)
- 08-SeedGeneDB -- Seed Specific Gene DataBase
- 08-KARG -- Knowledgebase for Addiction Related Genes
- 07-hESlincRNABrowser-- Browser for transcriptional profiling and transcriptional regulation in human embryonic stem cells
- 07-GSDS -- Gene Structure Display Server
- 06-CPC -- Coding Potential Calculator
- 06-CEAS -- Cis-regulatory Element Annotation System
- 05-NATsDB -- Natural antisense transcripts database
- 96-DEPD -- The Differentially Expressed Protein Database
- ? RStGEP -- Organ-specific expression profiling of early development of rice stamen
- ? -NTAP -- NimbleGen tiling array data analysis package

Autism KB

- AutismKB is an evidence-based knowledgebase of ASD genetics.
- URL : <http://autismkb.cbi.pku.edu.cn/index.php>

The screenshot shows the AutismKB website. The header is blue with the AutismKB logo and navigation links: Home, Datasets, Analysis, Tools, Search, Document, and Contact. On the right of the header is the Center for Bioinformatics, PKU logo. The main content area has a blue sidebar on the left with a deprecation notice and a definition of ASD. The main text area contains information about the database's scope (3075 genes, 99 syndromic, 3022 non-syndromic), experimental methods (6), and a list of main datasets (Syndromic, Non-syndromic, Core Dataset, View CNV Regions, View Linkage Regions, View SNPs/VNTRs). The right sidebar features a quick search bar, lists of syndromic and non-syndromic genes, and a statistics section. The footer includes a citation and a copyright notice.

AutismKB Home Datasets Analysis Tools Search Document Contact Center for Bioinformatics, PKU

Deprecated: mysql_connect(): The mysql extension is deprecated and will be removed in the future: use mysqli or PDO instead in /rd1/www/autismkb_www/mysql_connect.php on line 14

Autism spectrum disorder (ASD) is a heterogeneous **neurodevelopmental disorder** with a prevalence of **1.0-2.6%**. The three core symptoms of ASD are:

1. impairments in reciprocal social interaction;
2. communication impairments;
3. presence of restricted, repetitive and stereotyped patterns of behavior, interests and activities.

Genetic factors contribute significantly to ASD. AutismKB is an evidence-based knowledgebase of ASD genetics. The current version contains **3075 genes** (**99 syndromic autism related genes** and **3022 non-syndromic autism related genes**), **4964 Copy Number Variations (CNVs)** and **158 linkage regions** associated with ASD by one or more of the following **six experimental methods**:

1. Genome-Wide Association Studies (GWAS);
2. Genome-wide CNV studies;
3. Linkage analysis;
4. Low-scale genetic association studies;
5. Expression profiling;
6. Other low-scale gene studies.

Based on a scoring and ranking system ([More](#)), **99 syndromic autism related genes** and **109 non-syndromic autism related genes** (171 genes in total) were designated as having high confidence ([Core Dataset](#)).

Click to view main datasets:

[Syndromic](#) [Non-syndromic](#) [Core Dataset](#) [View CNV Regions](#) [View Linkage Regions](#) [View SNPs/VNTRs](#)

Quick Search:
(e.g. NLGN3)

Syndromic Genes

MECP2
FMR1
TSC1
TSC2
UBE3A
[More](#)

Non-syndromic Genes

NLGN4X
RELN
SLC6A4
GABRB3
NLGN3
[More](#)

AutismKB Statistics

Studies: 616
Genes: 3,075
SNPs/VNTRs: 3,386
CNVs: 4,917
Linkage Regions: 158
Last Update: 05/25/2012
[Click to Download Data](#)

Please Cite: Xu, L.M., Li, J.R., Huang, Y., Zhao, M., Tang, X. and Wei, L. (2012) AutismKB: an evidence-based knowledgebase of autism genetics. *Nucleic Acids Res*, **40**, D1016-1022.

Registered with NIF

Copyright © 2011, Center for Bioinformatics, Peking University | All Rights Reserved

Autism KB-Datasets

- Browse by gene categories:
 - Syndromic
 - Non-syndromic
 - Core datasets
 - CNVs
 - linkage regions
 - SNPs & VNTRs
- Browse by methods:
 - Genome-Wide Association Studies (GWAS)
 - Genome-wide CNV studies
 - Linkage analysis
 - Low-scale genetic association studies
 - Expression profiling
 - Other low-scale gene studies.
- Browse by chromosome

Autism KB-Gene List

Syndromic Autism Gene List ([Chromosome View](#))

Some autism cases are observed in the context of a recognized syndrome (e.g., the fragile X syndrome) with known causal genes. These **99 genes** were classified as syndromic autism genes in AutismKB.

This dataset comes from Catalina Betancur's review published in *Brain Research* in 2011([Betancur, 2011](#))

”” Reference:

Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res* 1380, 42-77.

Abbreviations: AD, autosomal dominant; AR, autosomal recessive; ASD, autism spectrum disorder; ID, intellectual disability; XL, X linked.

Gene_ID	Gene_Symbol	Band	Gene Name	Inheritance	Disorder	OMIM	Detail
AutG775	CACNA1C	12p13.3	calcium channel, voltage-dependent, L type, alpha 1C subunit	AD	Timothy syndrome	Timothy syndrome (601005)	▼
AutG7337	UBE3A	15q11.2	ubiquitin protein ligase E3A	AD	Angelman syndrome	Angelman syndrome (105830)	▼
AutG7249	TSC2	16p13.3	tuberous sclerosis 2	AD	Tuberous sclerosis	Tuberous sclerosis-2 (191100)	▼
AutG7248	TSC1	9q34	tuberous sclerosis 1	AD	Tuberous sclerosis	Tuberous sclerosis-1 (191100)	▼

Autism KB-Gene Details

Evidence Detail for CACNA1C

[View Evidences](#)[View Annotations](#)

Basic Information

[Top](#)

Gene Symbol:	CACNA1C (CACH2,CACN2,CACNL1A1,CCHL1A1,CaV1.2,MGC120730,TS)
Gene Full Name:	calcium channel, voltage-dependent, L type, alpha 1C subunit
Band:	12p13.3
Quick Links	Entrez ID: 775 ; OMIM: 114205 ; Uniprot ID: CAC1C_HUMAN ; ENSEMBL ID: ENSG00000151067

Sequences

[Top](#)

Evidence Statistic

[Top](#)

Syndromic Autism Gene

[Top](#)

Low Scale Association Studies (by Ethnic Group)

[Top](#)

Genome-Wide Association Studies(By Ethnic Group)

[Top](#)

Other Low Scale Gene Studies

[Top](#)

Large Scale Expression Studies

[Top](#)

CNV Studies

[Top](#)

Linkage Studies

[Top](#)

Autism KB-analysis

- Using 171 confident genes:
 - Enriched GO by DAVID
 - Enriched Pathway by KOBAS
 - Enriched GO Map by WebGestalt
- Tools
 - Rank the genes by yourself
 - BLAST
 - Search tool

PlantTFDB 3.0 -- Plant Transcription Factor Database



Plant Transcription Factor Database

v3.0

Center for Bioinformatics, Peking University, China Previous versions: v1.0 v2.0

[Home](#) | [Blast](#) | [Search](#) | [Download](#) | [Prediction](#) | [Help](#) | [About](#) | [Links](#)

(eg: LFY)

Jin JP, Zhang H, Kong L, Gao G and Luo JC. [PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors](#). Nucleic Acids Research 2013; doi: 10.1093/nar/gkt1016.

Or: Zhang H, Jin JP, Tang L, Zhao Y, Gu XC, Gao G, Luo JC. [PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database](#). Nucleic Acids Research 2011. 39: D1114-D1117.

©2010-2013, Center for Bioinformatics, Peking University

Last Modified: 2013-8-23

Questions or Comments, please contact

planttfdb@mail.cbi.pku.edu.cn



Supported By

Other tools

- SynDB -- Synapse DataBase
 - An online resource of proteins known or predicted to be related to the synapse or synaptic activity.
- SPD -- Secreted Protein database
 - A collection of secreted proteins from Human, Mouse and Rat proteomes
 - Includes sequences from SwissProt, Trembl, Ensembl and Refseq.
- KOBAS -- a web server for annotation and identification of enriched pathways and diseases

Bioinformatics Infrastructure

- 06-ABCGrid -- Application of Bioinformatics Computing Grid
- 04-WebLab -- Your lab on the web

ABCGrid

- Application of Bioinformatics Computing Grid.
- For small biology laboratories
- Use large amount of heterogeneous computing resources and access many bioinformatics applications
- From one master node without being aware of where the data and computational resources are located.
- Source code is released under the GNU General Public License(GPL).

WebLab -- Your lab on the web

WebLab is a multifunctional bioinformatics analysis platform integrating diversified tools with unified, user-friendly web interface. However, WebLab is not a mere bioinformatics toolbox, it also offers powerful data management function, group strategy and knowledge sharing mechanism, which will bring considerable advance of efficiency for both wet bench and in silicon scientists working in biomedicine community.

- Thank you for your kind attention!
- <http://www.cbi.pku.edu.cn/research/>