SoyNet

Functional gene network for Glycine Max

INDEX

- About SoyNet (2016)

SoyNet is a functional gene network and co-expression network database for a legume crop, **soybean (*Glycine max*)**. Soybean is a major sources of oil and protein for human and livestock, and accounts for more than 50% of globally consumed edible oil. To facilitate genetic dissection of economically important traits of soybean via network biology approaches, we developed a network wiring **40,812 soybean genes** (~73% of the coding genome) with **~2 million functional links**, which were inferred by Bayesian statistics framework. Details about network construction methods previously developed are described in *Nature Biotechnology 2010*.

SoyNet web database can generate functional hypotheses for soybean genes using two network algorithms: 'Find new members of a pathway', 'Find context associated genes', and 'Find functional modules'.

Licensing

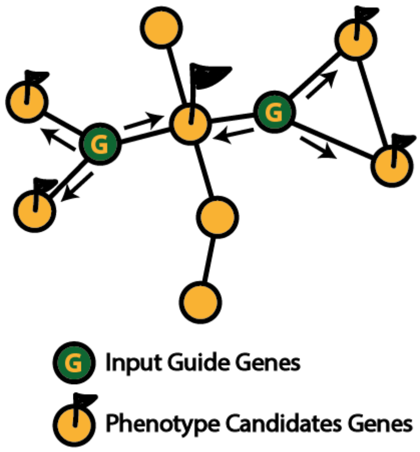
Contents of this database are **freely available for academia and public entities**. If you are a commercial entity, please send an email for further information about licensing.

- Network-search tools

Network search allows users to identify predict new candidate genes for a phenotype or biological pathways by prioritizing soybean genes. Functionally associated genes are wired to another in the network. Thus, genes that are highly connected to genes of known function (user submitted guide genes: genes that are known to participate in certain biological pathways or that lead to certain phenotypic outcome) are candidate genes of the pathway.

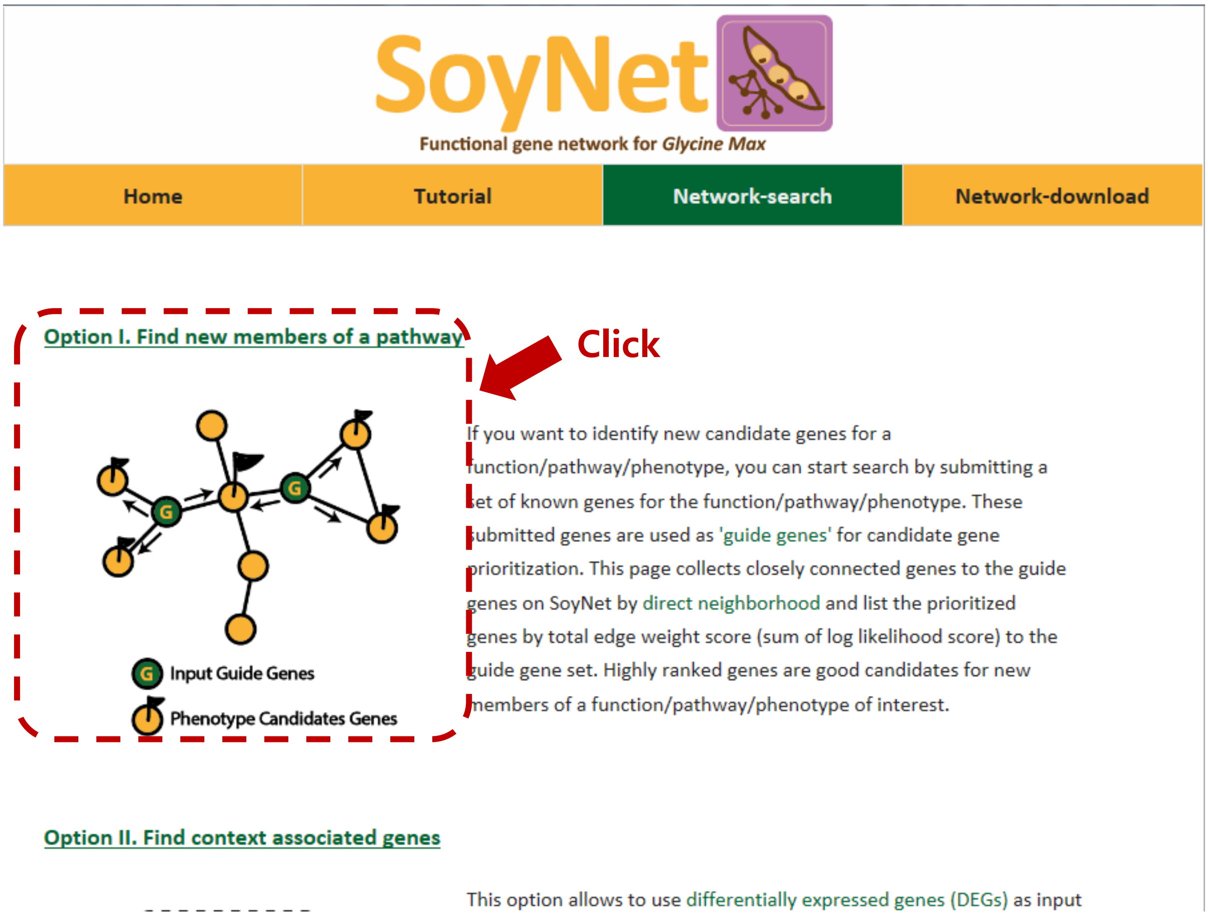
Option 1. Find new members of a pathway

Overview



If you want to identify new candidate genes for a function/pathway/phenotype, you can start search by submitting a set of known genes for the function/pathway/phenotype. These submitted genes are used as 'guide genes' for candidate gene prioritization. This page collects closely connected genes to the guide genes on SoyNet by direct neighborhood and list the prioritized genes by total edge weight score (sum of log likelihood score) to the guide gene set. Highly ranked genes are good candidates for new members of a function/pathway/phenotype of interest.

A. Submit your guide genes



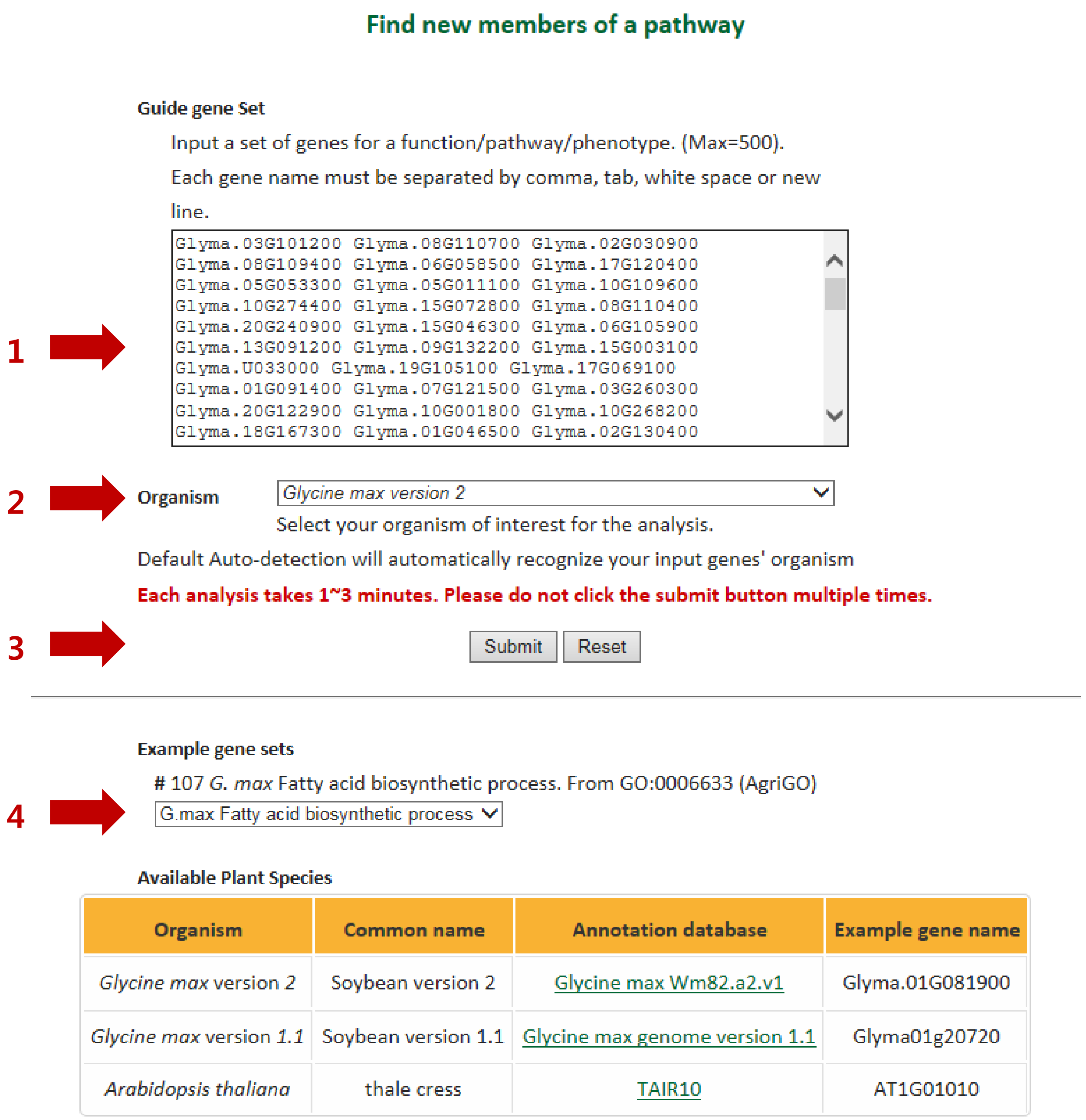
Option I. Find new members of a pathway

If you want to identify new candidate genes for a function/pathway/phenotype, you can start search by submitting a set of known genes for the function/pathway/phenotype. These submitted genes are used as 'guide genes' for candidate gene prioritization. This page collects closely connected genes to the guide genes on SoyNet by direct neighborhood and list the prioritized genes by total edge weight score (sum of log likelihood score) to the guide gene set. Highly ranked genes are good candidates for new members of a function/pathway/phenotype of interest.

Option II. Find context associated genes

This option allows to use differentially expressed genes (DEGs) as input

Click ‘Find new members of a pathway’ on Network-search page.



Find new members of a pathway

Guide gene Set

Input a set of genes for a function/pathway/phenotype. (Max=500).

Each gene name must be separated by comma, tab, white space or new line.

```
Glyma.03G101200 Glyma.08G110700 Glyma.02G030900
Glyma.08G109400 Glyma.06G058500 Glyma.17G120400
Glyma.05G053300 Glyma.05G011100 Glyma.10G109600
Glyma.10G274400 Glyma.15G072800 Glyma.08G110400
Glyma.20G240900 Glyma.15G046300 Glyma.06G105900
Glyma.13G091200 Glyma.09G132200 Glyma.15G003100
Glyma.U033000 Glyma.19G105100 Glyma.17G069100
Glyma.01G091400 Glyma.07G121500 Glyma.03G260300
Glyma.20G122900 Glyma.10G001800 Glyma.10G268200
Glyma.18G167300 Glyma.01G046500 Glyma.02G130400
```

Organism *Glycine max version 2*

Select your organism of interest for the analysis.

Default Auto-detection will automatically recognize your input genes' organism

Each analysis takes 1~3 minutes. Please do not click the submit button multiple times.

Example gene sets

107 *G. max* Fatty acid biosynthetic process. From GO:0006633 (AgriGO)

G.max Fatty acid biosynthetic process

Available Plant Species

Organism	Common name	Annotation database	Example gene name
Glycine max version *2*	Soybean version 2	Glycine max Wm82.a2.v1	Glyma.01G081900
Glycine max version *1.1*	Soybean version 1.1	Glycine max genome version 1.1	Glyma01g20720
Arabidopsis thaliana	thale cress	TAIR10	AT1G01010

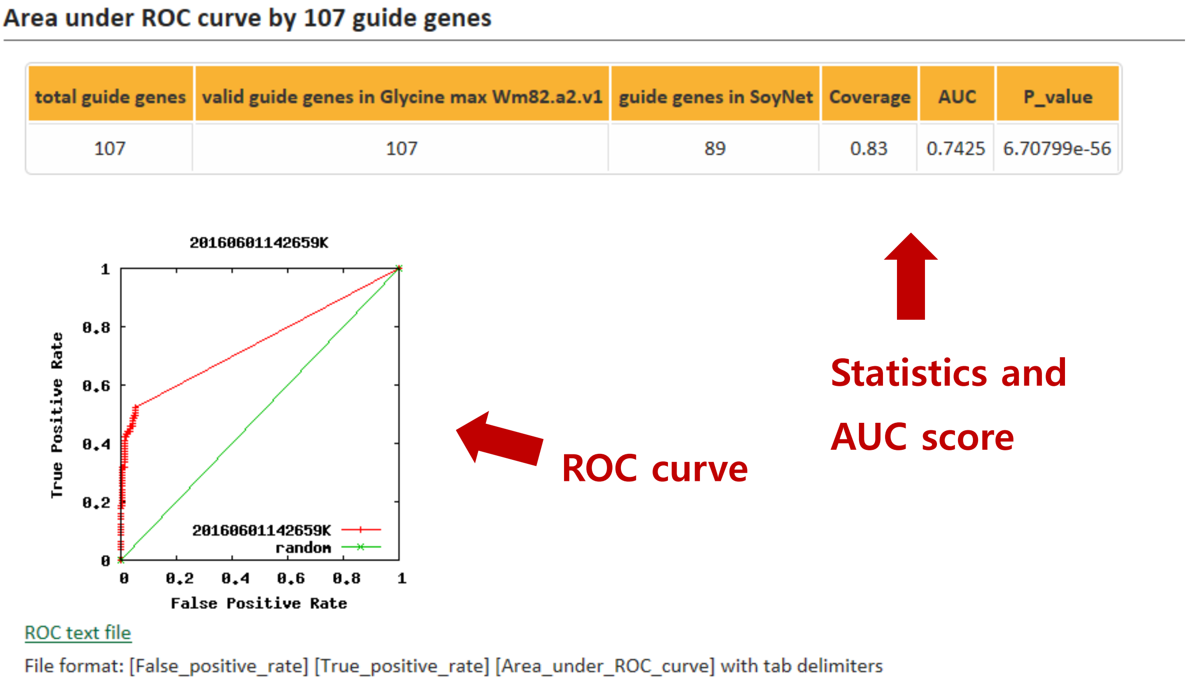
1. Insert your guide genes in the text box. The gene names must be separated by tab, comma, white space or new line. A maximum of 500 genes are available. User can input not only *Glycine max* genes and *Arabidopsis thaliana* genes for porting well studied pathways to Soybean.

2. Organism selection box. User can select an organism among *Glycine max* version 2 genome, *Glycine max* version 1.1 genome, *Arabidopsis thaliana*, and auto-detection, which option will automatically recognize your species of 'Guide gene set'.

3. Click 'submit' button to start the analysis. This may take about 1~3 minutes.

4. You could select our example guide genes for a test run. In this tutorial, *Glycine max* Fatty acid biosynthetic process genes were selected.

B. Result page - Area Under ROC curve



Area under ROC curve by 107 guide genes

total guide genes	valid guide genes in Glycine max Wm82.a2.v1	guide genes in SoyNet	Coverage	AUC	P_value
107	107	89	0.83	0.7425	6.70799e-56

ROC text file

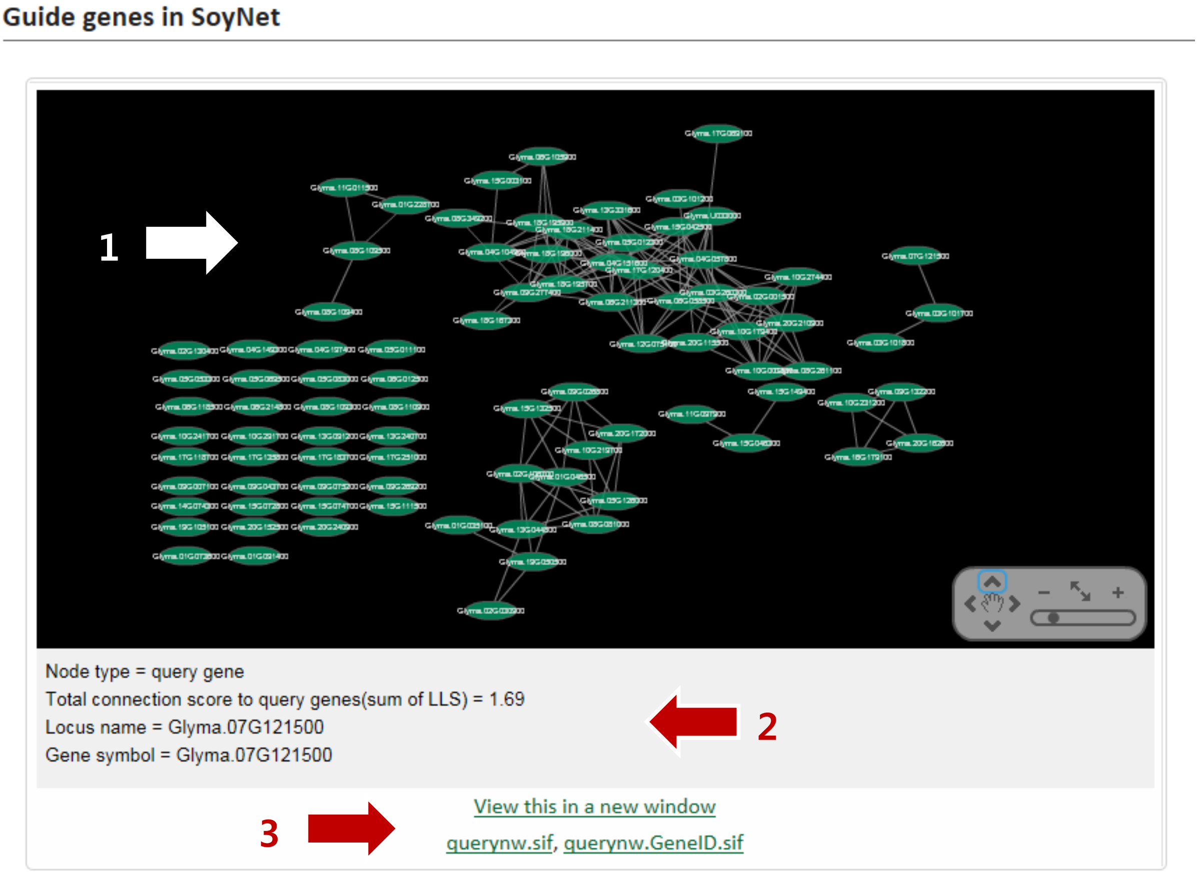
File format: [False_positive_rate] [True_positive_rate] [Area_under_ROC_curve] with tab delimiters

ROC curve

Statistics and AUC score

Statistics of query genes. This table shows the number of guide genes input, valid guide genes in *Glycine max*, guide genes in SoyNet, AUC score, and P-value. AUC is 'Area Under ROC Curve' which is shown below and represents closeness of guide genes. Generally, AUC above 0.7 means those genes form closed in SoyNet.

C. Result page - Network visualization of guide genes



Guide genes in SoyNet

Node type = query gene
Total connection score to query genes(sum of LLS) = 1.69
Locus name = Glyma.07G121500
Gene symbol = Glyma.07G121500

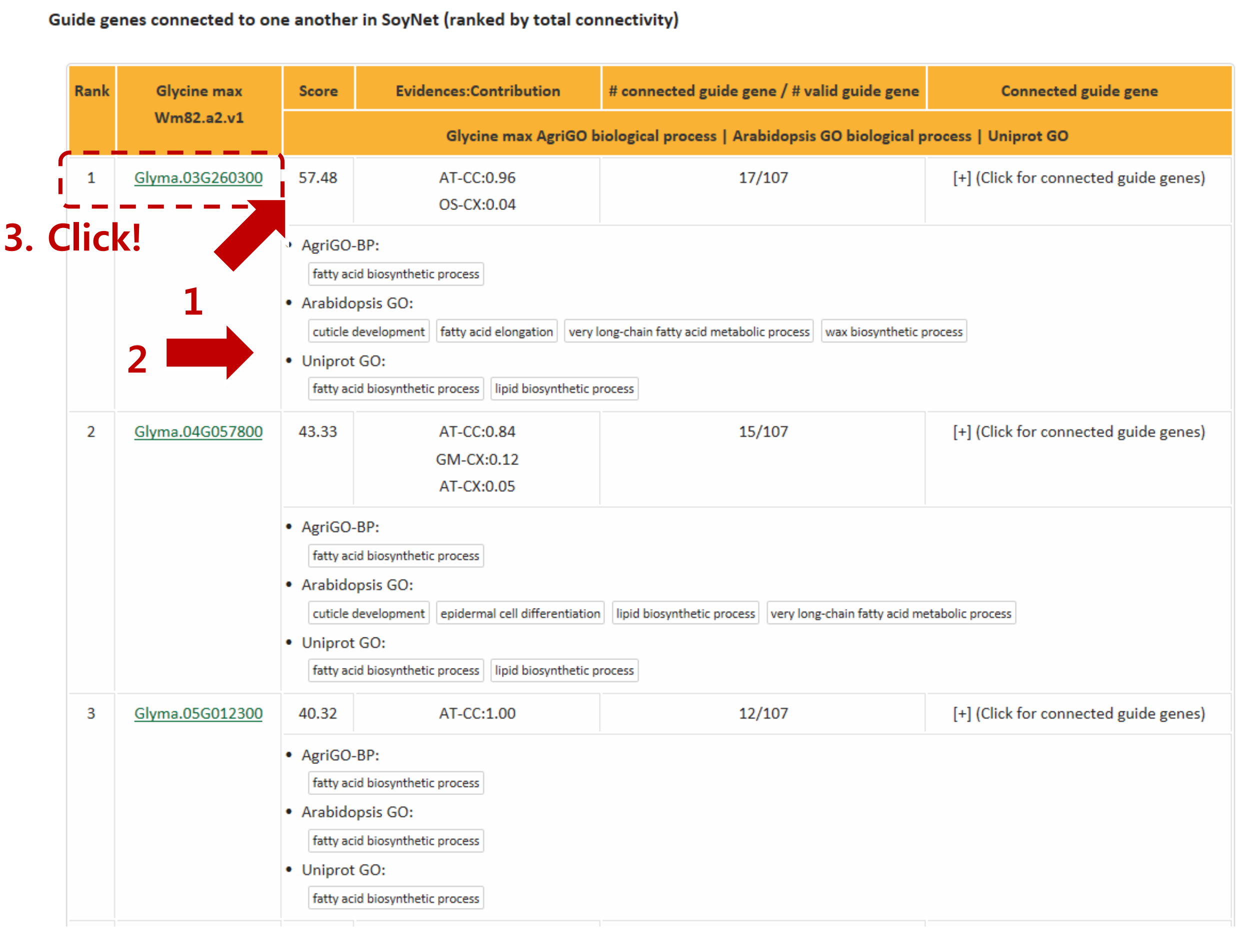
View this in a new window
querynw.sif, querynw.GeneID.sif

1. This plot shows network of guide genes using Cytoscape Web. This requires installation of adobe flash player. Users can click ‘Node’ or ‘Edge’ to show additional information of edges or nodes.

2. Additional information of edges or nodes is shown here.

3. Users can view this graph in a new window, or download Cytoscape Network Files to see the network using Cytoscape standard alone program.

D. Result page - Prioritization of guide genes



Guide genes connected to one another in SoyNet (ranked by total connectivity)

Rank	Glycine max Wm82.a2.v1	Score	Evidences:Contribution	# connected guide gene / # valid guide gene	Connected guide gene
		Glycine max AgriGO biological process | Arabidopsis GO biological process | Uniprot GO			
1	Glyma.03G260300	57.48	AT-CC:0.96 OS-CX:0.04	17/107	[+] (Click for connected guide genes)
		• AgriGO-BP: fatty acid biosynthetic process • Arabidopsis GO: cuticle development fatty acid elongation very long-chain fatty acid metabolic process wax biosynthetic process • Uniprot GO: fatty acid biosynthetic process lipid biosynthetic process			
2	Glyma.04G057800	43.33	AT-CC:0.84 GM-CX:0.12 AT-CX:0.05	15/107	[+] (Click for connected guide genes)
		• AgriGO-BP: fatty acid biosynthetic process • Arabidopsis GO: cuticle development epidermal cell differentiation lipid biosynthetic process very long-chain fatty acid metabolic process • Uniprot GO: fatty acid biosynthetic process lipid biosynthetic process			
3	Glyma.05G012300	40.32	AT-CC:1.00	12/107	[+] (Click for connected guide genes)
		• AgriGO-BP: fatty acid biosynthetic process • Arabidopsis GO: fatty acid biosynthetic process • Uniprot GO: fatty acid biosynthetic process			

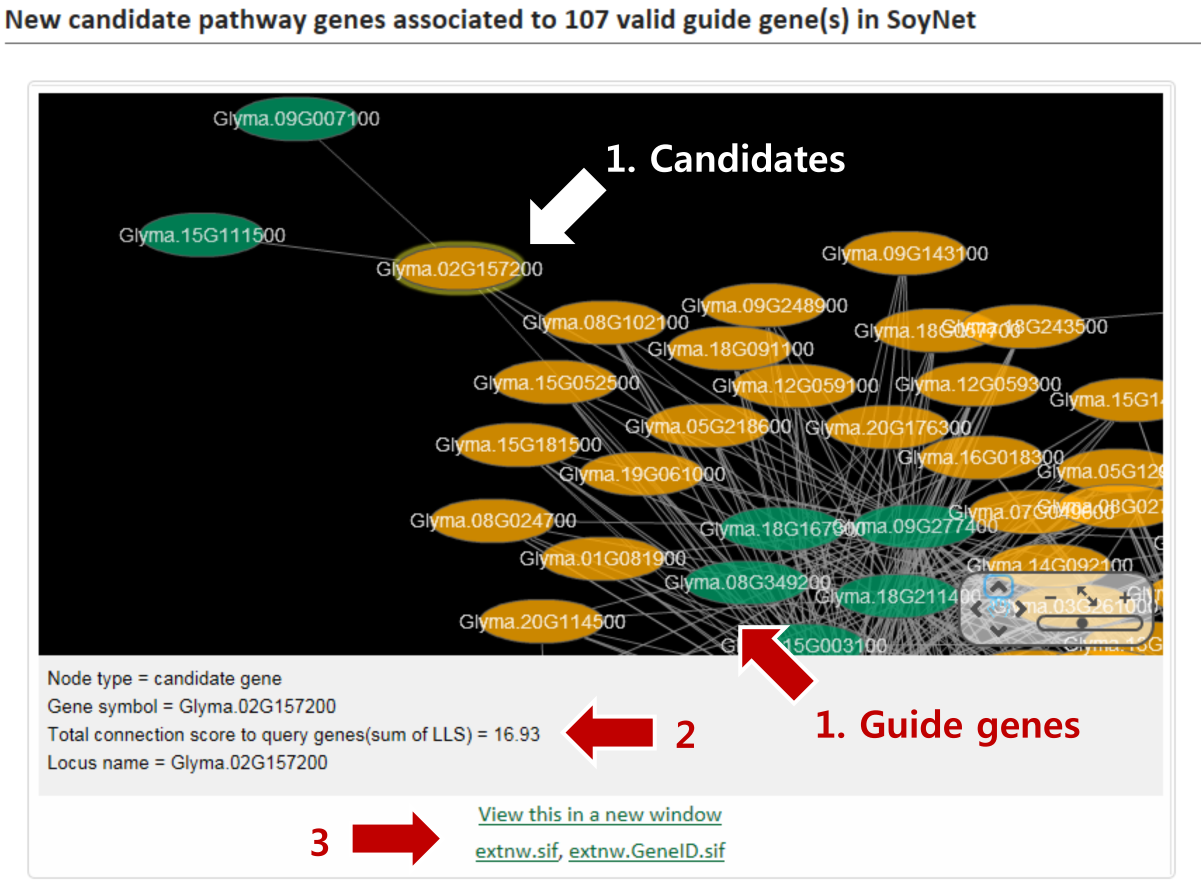
This table shows a list of guide genes. Those genes are prioritized using network connectivity scores. Therefore higher rank genes could be considered more important for pathways.

1. Score is calculated by sum of LLS (Log Likelihood Scores) of network links to all other guide genes. Evidence is contribution source of the link.

2. SoyNet also presents various pathway terms of genes to provide comprehensive analysis of pathways. The annotation of each gene are provided using AgiGO, UniprotGOA and Arabidopsis Gene ontology biological process term.

3. Clicking the locus names lead to corresponding annotation databases for detailed descriptions.

E. Result page - Network visualization of new candidates of pathway



New candidate pathway genes associated to 107 valid guide gene(s) in SoyNet

Node type = candidate gene
Gene symbol = Glyma.02G157200
Total connection score to query genes(sum of LLS) = 16.93
Locus name = Glyma.02G157200

View this in a new window

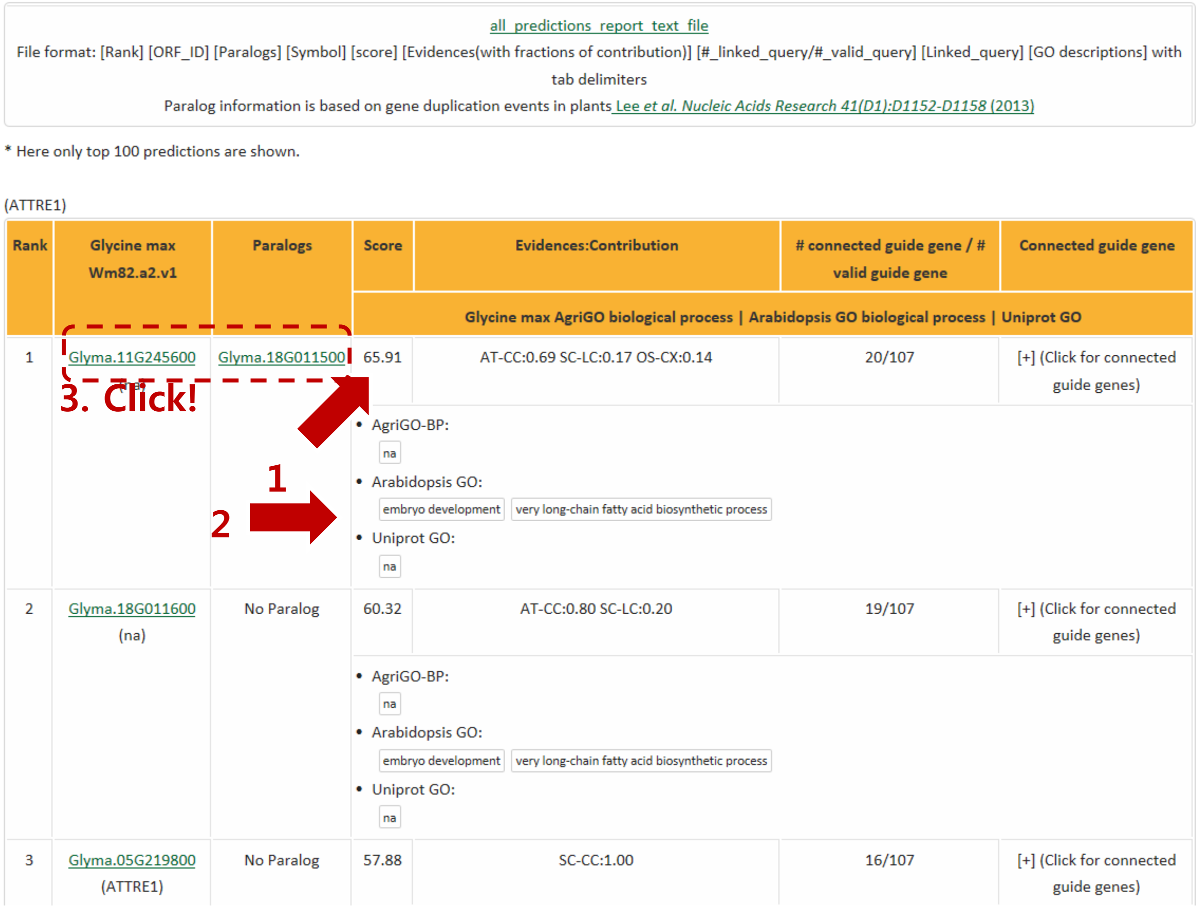
extnw.sif, extnw.GeneID.sif

1. Yellow ovals indicate new candidates of submitted pathway and green ovals indicate guide genes.

2. This plot shows network of new candidates of pathway using Cytoscape Web. Users can click 'Node' or 'Edge' to show additional information of edges or nodes. Additional information of edges of nodes is shown below.

4. Users can view this graph in a new window, or download Cytoscape Network Files to see the network using Cytoscape standard alone program.

F. Result page - A list of new candidates of pathway



all_predictions_report_text_file

File format: [Rank] [ORF_ID] [Paralogs] [Symbol] [score] [Evidences(with fractions of contribution)] [#_linked_query/#_valid_query] [Linked_query] [GO descriptions] with tab delimiters

Paralog information is based on gene duplication events in plants Lee *et al. Nucleic Acids Research 41(D1):D1152-D1158* (2013)

* Here only top 100 predictions are shown.

(ATTRE1)

Rank	Glycine max Wm82.a2.v1	Paralogs	Score	Evidences:Contribution	# connected guide gene / # valid guide gene	Connected guide gene
			Glycine max AgriGO biological process | Arabidopsis GO biological process | Uniprot GO			
1	Glyma.11G245600 (na)	Glyma.18G011500	65.91	AT-CC:0.69 SC-LC:0.17 OS-CX:0.14	20/107	[+] (Click for connected guide genes)
			• AgriGO-BP: na • Arabidopsis GO: embryo development, very long-chain fatty acid biosynthetic process • Uniprot GO: na			
2	Glyma.18G011600 (na)	No Paralog	60.32	AT-CC:0.80 SC-LC:0.20	19/107	[+] (Click for connected guide genes)
			• AgriGO-BP: na • Arabidopsis GO: embryo development, very long-chain fatty acid biosynthetic process • Uniprot GO: na			
3	Glyma.05G219800 (ATTRE1)	No Paralog	57.88	SC-CC:1.00	16/107	[+] (Click for connected guide genes)

This table shows a list of new candidates of input pathway. Those genes are prioritized using network connectivity scores. Therefore higher rank genes could be considered more important for pathways.

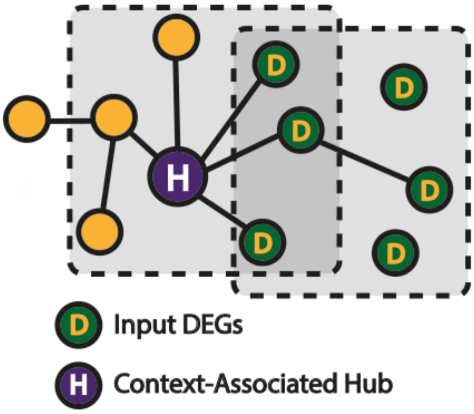
1. Score is calculated by sum of LLS (Log Likelihood Scores) of network links to all other guide genes. Evidence is contribution source of the link.

2. We also present various pathway terms of genes to provide comprehensive analysis of pathways. The annotation of each gene are provided using AgiGO, UniprotGOA and Arabidopsis Gene ontology biological process term.

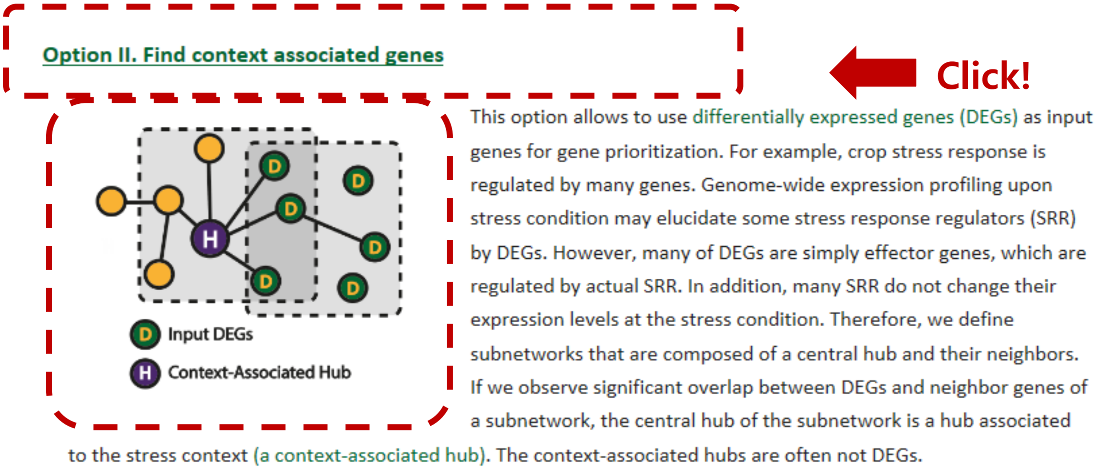
3. Clicking the locus names lead to corresponding annotation databases for detailed descriptions.

Option 2. Find context-associated genes

Overview

	This option allows to use differentially expressed genes (DEGs) as input genes for gene prioritization. For example, crop stress response is regulated by many genes. Genome-wide expression profiling upon stress condition may elucidate some stress response regulators (SRR) by DEGs. However, many of DEGs are simply effector genes, which are regulated by actual SRR. In addition, many SRR do not change their expression levels at the stress condition. Therefore, we define subnetworks that are composed of a central hub and their neighbors. If we observe significant overlap between DEGs and neighbor genes of a subnetwork, the central hub of the subnetwork is a hub associated to the stress context (a context-associated hub). The context-associated hubs are often not DEGs.

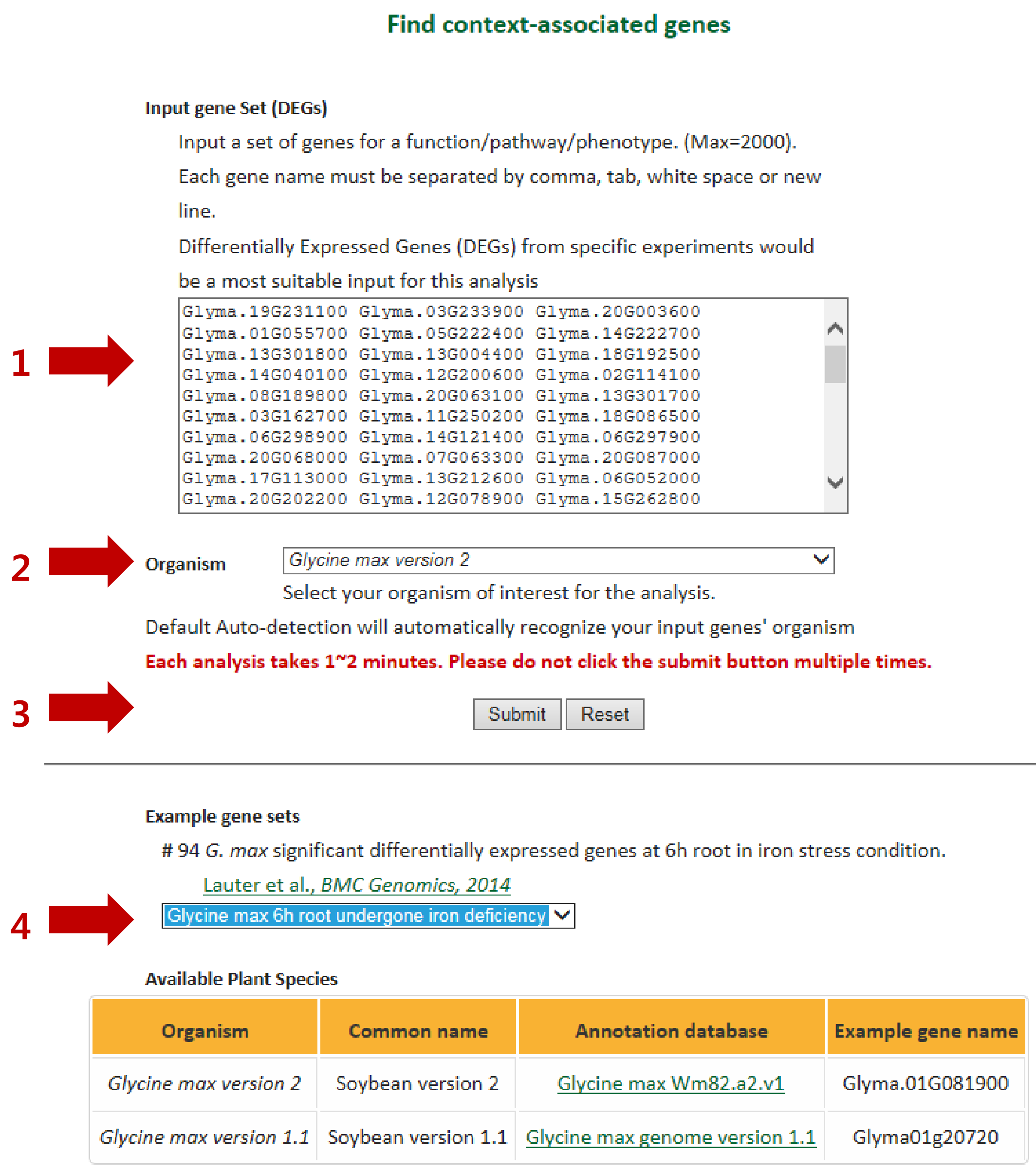
A. Submit your input differentially expressed genes (DEGs).



Option II. Find context associated genes

This option allows to use differentially expressed genes (DEGs) as input genes for gene prioritization. For example, crop stress response is regulated by many genes. Genome-wide expression profiling upon stress condition may elucidate some stress response regulators (SRR) by DEGs. However, many of DEGs are simply effector genes, which are regulated by actual SRR. In addition, many SRR do not change their expression levels at the stress condition. Therefore, we define subnetworks that are composed of a central hub and their neighbors. If we observe significant overlap between DEGs and neighbor genes of a subnetwork, the central hub of the subnetwork is a hub associated to the stress context (a context-associated hub). The context-associated hubs are often not DEGs.

Click “Context-associated genes” on Network-search page.



Find context-associated genes

Input gene Set (DEGs)

Input a set of genes for a function/pathway/phenotype. (Max=2000).

Each gene name must be separated by comma, tab, white space or new line.

Differentially Expressed Genes (DEGs) from specific experiments would be a most suitable input for this analysis

```
Glyma.19G231100 Glyma.03G233900 Glyma.20G003600
Glyma.01G055700 Glyma.05G222400 Glyma.14G222700
Glyma.13G301800 Glyma.13G004400 Glyma.18G192500
Glyma.14G040100 Glyma.12G200600 Glyma.02G114100
Glyma.08G189800 Glyma.20G063100 Glyma.13G301700
Glyma.03G162700 Glyma.11G250200 Glyma.18G086500
Glyma.06G298900 Glyma.14G121400 Glyma.06G297900
Glyma.20G068000 Glyma.07G063300 Glyma.20G087000
Glyma.17G113000 Glyma.13G212600 Glyma.06G052000
Glyma.20G202200 Glyma.12G078900 Glyma.15G262800
```

Organism *Glycine max version 2*

Select your organism of interest for the analysis.

Default Auto-detection will automatically recognize your input genes' organism

Each analysis takes 1~2 minutes. Please do not click the submit button multiple times.

Example gene sets

94 *G. max* significant differentially expressed genes at 6h root in iron stress condition.

Lauter et al., *BMC Genomics, 2014*

Glycine max 6h root undergone iron deficiency

Available Plant Species

Organism	Common name	Annotation database	Example gene name
Glycine max version 2	Soybean version 2	Glycine max Wm82.a2.v1	Glyma.01G081900
Glycine max version 1.1	Soybean version 1.1	Glycine max genome version 1.1	Glyma01g20720

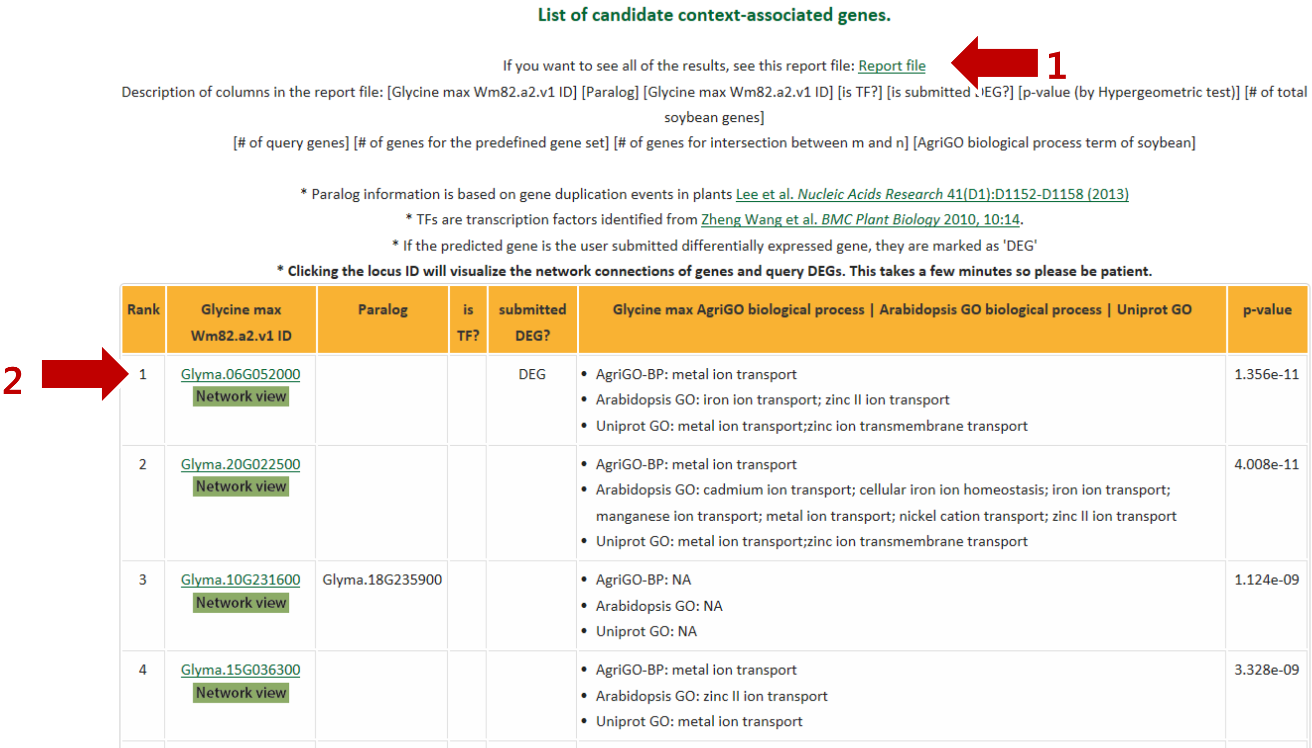
1. Insert your DEGs in the text box. The gene names must be separated by tab, comma, white space or new line. A maximum of 2000 genes are available.

2. Organism selection box. The default setting is auto-detection which option auto-detection will automatically recognize your input gene if the genes are submitted in correct form. Or, User can select between *Glycine max version 2* and *Glycine max version 1.*

3. Click 'submit' button to start the analysis. This may take about 1~2 minutes.

4. You could select our example guide genes for a test run. In this tutorial, *Glycine max* up-regulated genes 8 days after root inoculation of nematode (Race 14, log2 ratio > 2)

B. Result page - Candidate context-associated genes



List of candidate context-associated genes.

If you want to see all of the results, see this report file: Report file

Description of columns in the report file: [Glycine max Wm82.a2.v1 ID] [Paralog] [Glycine max Wm82.a2.v1 ID] [is TF?] [is submitted DEG?] [p-value (by Hypergeometric test)] [# of total soybean genes] [# of query genes] [# of genes for the predefined gene set] [# of genes for intersection between m and n] [AgriGO biological process term of soybean]

* Paralog information is based on gene duplication events in plants Lee et al. *Nucleic Acids Research* 41(D1):D1152-D1158 (2013)

* TFs are transcription factors identified from Zheng Wang et al. *BMC Plant Biology* 2010, 10:14.

* If the predicted gene is the user submitted differentially expressed gene, they are marked as 'DEG'

*** Clicking the locus ID will visualize the network connections of genes and query DEGs. This takes a few minutes so please be patient.**

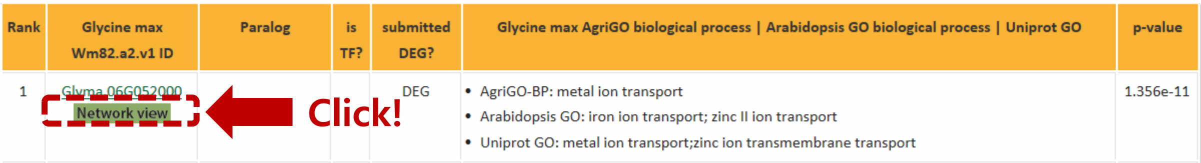
Rank	Glycine max Wm82.a2.v1 ID	Paralog	is TF?	submitted DEG?	Glycine max AgriGO biological process | Arabidopsis GO biological process | Uniprot GO	p-value
1	Glyma.06G052000 Network view			DEG	• AgriGO-BP: metal ion transport • Arabidopsis GO: iron ion transport; zinc II ion transport • Uniprot GO: metal ion transport;zinc ion transmembrane transport	1.356e-11
2	Glyma.20G022500 Network view				• AgriGO-BP: metal ion transport • Arabidopsis GO: cadmium ion transport; cellular iron ion homeostasis; iron ion transport; manganese ion transport; metal ion transport; nickel cation transport; zinc II ion transport • Uniprot GO: metal ion transport;zinc ion transmembrane transport	4.008e-11
3	Glyma.10G231600 Network view	Glyma.18G235900			• AgriGO-BP: NA • Arabidopsis GO: NA • Uniprot GO: NA	1.124e-09
4	Glyma.15G036300 Network view				• AgriGO-BP: metal ion transport • Arabidopsis GO: zinc II ion transport • Uniprot GO: metal ion transport	3.328e-09

SoyNet provides list of candidate context associated genes (hubs) that are ranked by p-value significance. The p-value is obtained by calculating hypergeometric p-value. If the network neighbors of a gene is enriched with DEGs, it is likely to show lower p-value providing evidence that the gene is a good context-associated gene. Maximum 100 predictions are shown on the web

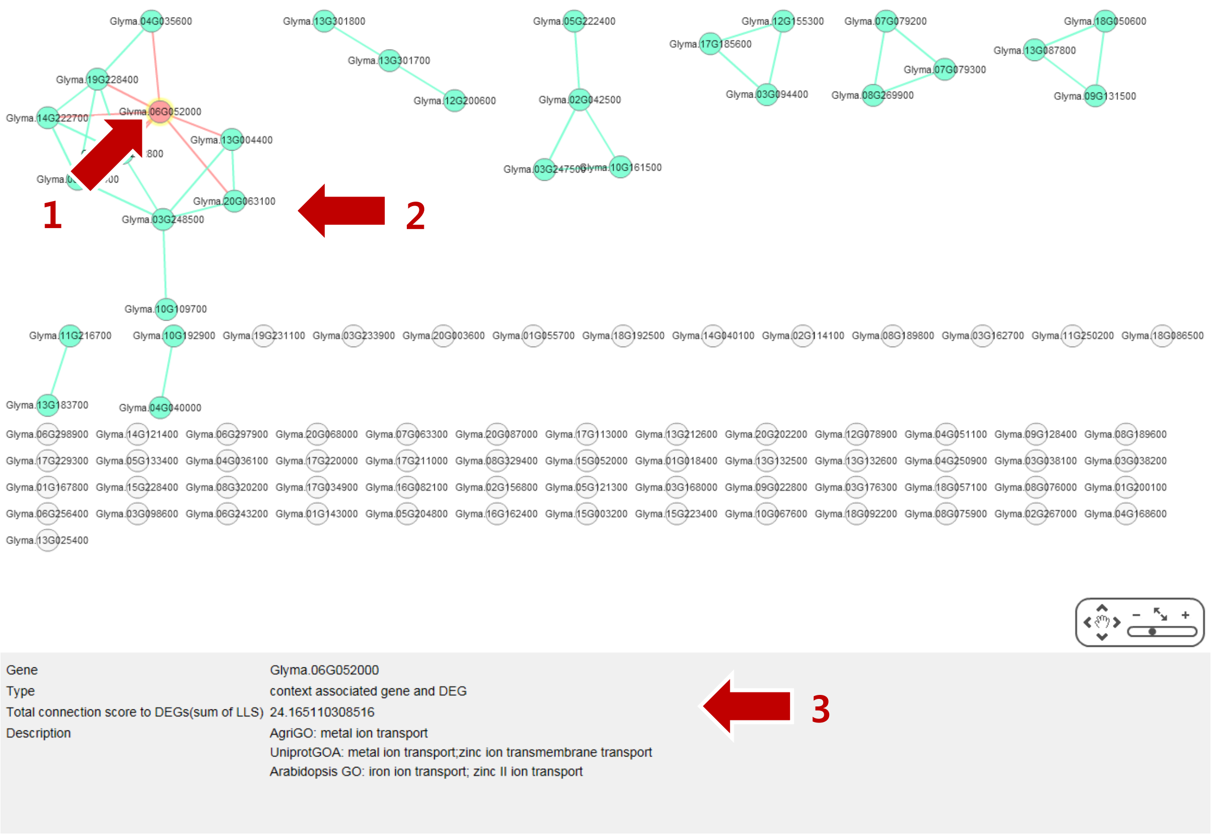
1. You can download the full result list by clicking Report file.

2. Hubs are ranked by their significance of p-values. If the gene is a transcription factor, ‘is TF?’ column is marked as ‘TF’. If the gene is one of the DEGs submitted by the user, ‘submitted DEG?’ is marked as ‘DEG’. P-value of each gene is hypergeometric p-value.

C. Result page - Network visualization of a candidate gene (hub)



Rank	Glycine max Wm82.a2.v1 ID	Paralog	is TF?	submitted DEG?	Glycine max AgriGO biological process | Arabidopsis GO biological process | Uniprot GO	p-value
1	Glyma.06G052000 Network view			DEG	• AgriGO-BP: metal ion transport • Arabidopsis GO: iron ion transport; zinc II ion transport • Uniprot GO: metal ion transport;zinc ion transmembrane transport	1.356e-11



Gene	Glyma.06G052000
Type	context associated gene and DEG
Total connection score to DEGs(sum of LLS)	24.165110308516
Description	AgriGO: metal ion transport UniprotGOA: metal ion transport;zinc ion transmembrane transport Arabidopsis GO: iron ion transport; zinc II ion transport

By clicking the link ‘Network view’, the web tool provides network visualization of the connections between candidate context-associated genes and their DEG partners as below.

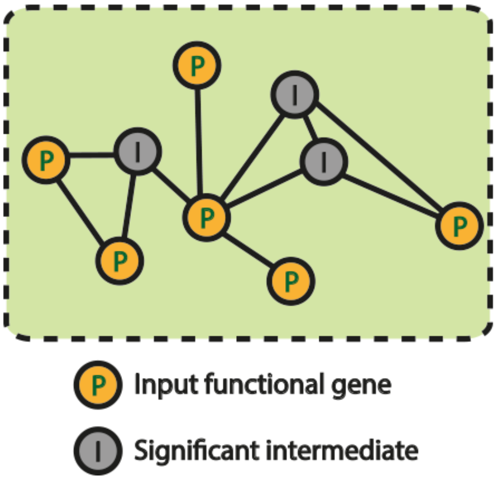
1. Red node: A context-associated gene is denoted in red. The edges to query DEGs are denoted as red lines.

2. Cyan node: Query DEGs are denoted in cyan. Connections of query DEGs are denoted as cyan lines.

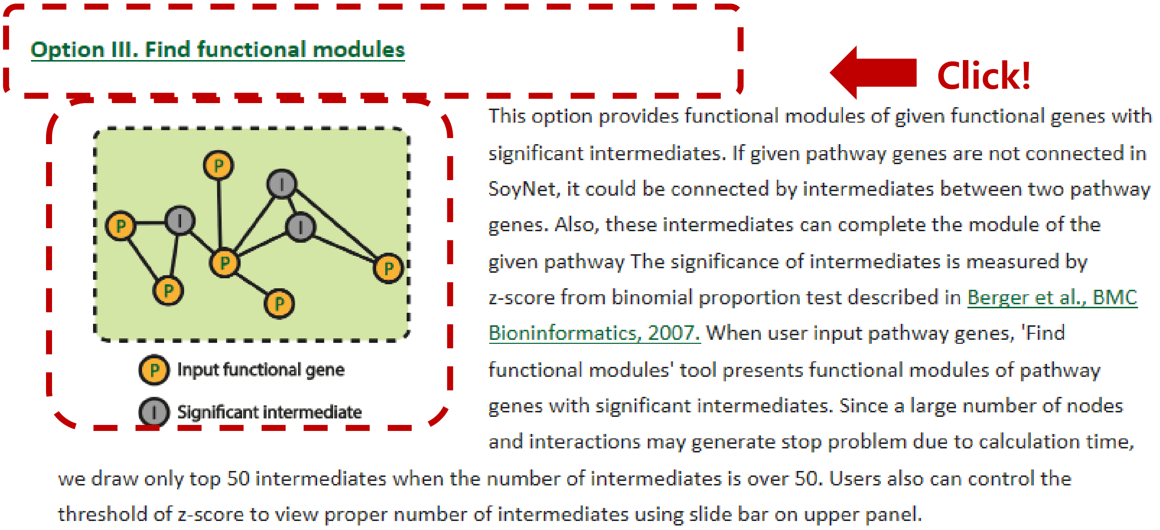
3. You can get the information of the nodes or the links by clicking on them.

Option 3. Find functional modules

Overview

	This option provides functional modules of given functional genes with significant intermediates. If given pathway genes are not connected in SoyNet, it could be connected by intermediates between two pathway genes. Also, these intermediates can complete the module of the given pathway The significance of intermediates is measured by z-score from binomial proportion test described in **Berger et al.** (*BMC Bioinformatics*, 2007). When user input pathway genes, 'Find functional modules' tool presents functional modules of pathway genes with significant intermediates. Since a large number of nodes and interactions may generate stop problem due to calculation time, we draw only top 50 intermediates when the number of intermediates is over 50. Users also can control the threshold of z-score to view proper number of intermediates using slide bar on upper panel.

A. Submit your pathway genes.

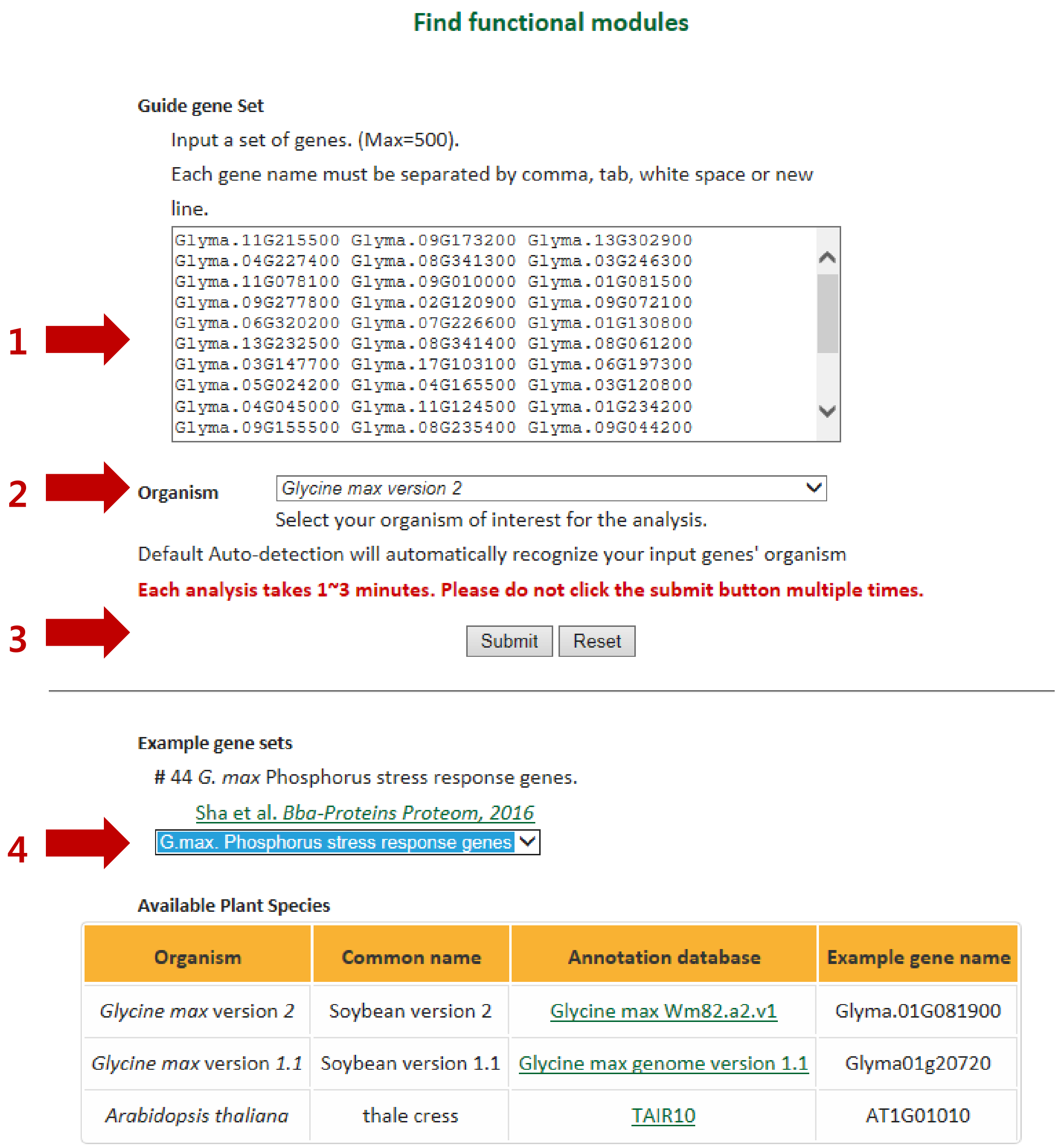


Option III. Find functional modules

Click!

This option provides functional modules of given functional genes with significant intermediates. If given pathway genes are not connected in SoyNet, it could be connected by intermediates between two pathway genes. Also, these intermediates can complete the module of the given pathway The significance of intermediates is measured by z-score from binomial proportion test described in Berger et al., BMC Bioinformatics, 2007. When user input pathway genes, 'Find functional modules' tool presents functional modules of pathway genes with significant intermediates. Since a large number of nodes and interactions may generate stop problem due to calculation time, we draw only top 50 intermediates when the number of intermediates is over 50. Users also can control the threshold of z-score to view proper number of intermediates using slide bar on upper panel.

Click “Find functional modules” on Network-search page.



Find functional modules

Guide gene Set

Input a set of genes. (Max=500).

Each gene name must be separated by comma, tab, white space or new line.

```
Glyma.11G215500 Glyma.09G173200 Glyma.13G302900
Glyma.04G227400 Glyma.08G341300 Glyma.03G246300
Glyma.11G078100 Glyma.09G010000 Glyma.01G081500
Glyma.09G277800 Glyma.02G120900 Glyma.09G072100
Glyma.06G320200 Glyma.07G226600 Glyma.01G130800
Glyma.13G232500 Glyma.08G341400 Glyma.08G061200
Glyma.03G147700 Glyma.17G103100 Glyma.06G197300
Glyma.05G024200 Glyma.04G165500 Glyma.03G120800
Glyma.04G045000 Glyma.11G124500 Glyma.01G234200
Glyma.09G155500 Glyma.08G235400 Glyma.09G044200
```

Organism *Glycine max version 2*

Select your organism of interest for the analysis.

Default Auto-detection will automatically recognize your input genes' organism

Each analysis takes 1~3 minutes. Please do not click the submit button multiple times.

Example gene sets

44 *G. max* Phosphorus stress response genes.

Sha et al. *Bba-Proteins Proteom, 2016*

G.max. Phosphorus stress response genes

Available Plant Species

Organism	Common name	Annotation database	Example gene name
Glycine max version 2	Soybean version 2	Glycine max Wm82.a2.v1	Glyma.01G081900
Glycine max version 1.1	Soybean version 1.1	Glycine max genome version 1.1	Glyma01g20720
Arabidopsis thaliana	thale cress	TAIR10	AT1G01010

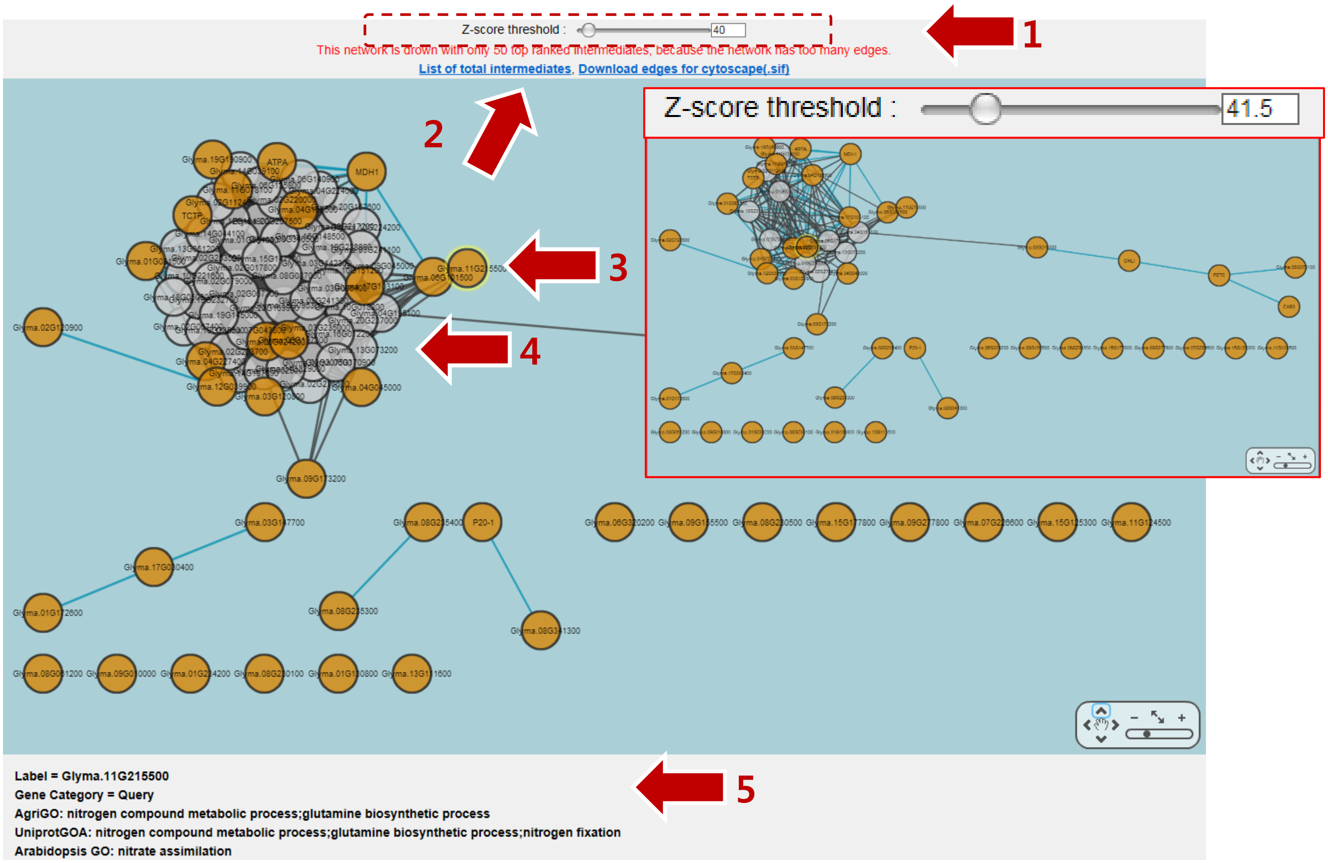
1. Insert your pathway genes in the text box. The gene names must be separated by tab, comma, white space or new line. A maximum of 200 genes are available.

2. Organism selection box. The default setting is auto-detection which option auto-detection will automatically recognize your input gene if the genes are submitted in correct form. Or, User can select between *Glycine max version 2* and *Glycine max version 1.*

3. Click 'submit' button to start the analysis. This may take about 1~2 minutes.

4. You could select our example guide genes for a test run. In this tutorial, *Glycine max* Phosphorus stress response genes are prepared for the test.

B. Result page - Functional modules of input pathway genes



Z-score threshold : 40

This network is drawn with only 50 top ranked intermediates, because the network has too many edges.

List of total intermediates, Download edges for cytoscape(.sif)

Z-score threshold : 41.5

Label = Glyma.11G215500
Gene Category = Query
AgriGO: nitrogen compound metabolic process;glutamine biosynthetic process
UniprotGOA: nitrogen compound metabolic process;glutamine biosynthetic process;nitrogen fixation
Arabidopsis GO: nitrate assimilation

A network figure of functional modules is shown by cytoscape web (Flash player should be need to view the figure).

1. You can control z-score threshold to adjust the number of intermediates using the slide bar or enter z-score in the text box.

2. You can see all intermediates by clicking ‘List of total intermediates’. Also you can download network links as cytoscape sif format.

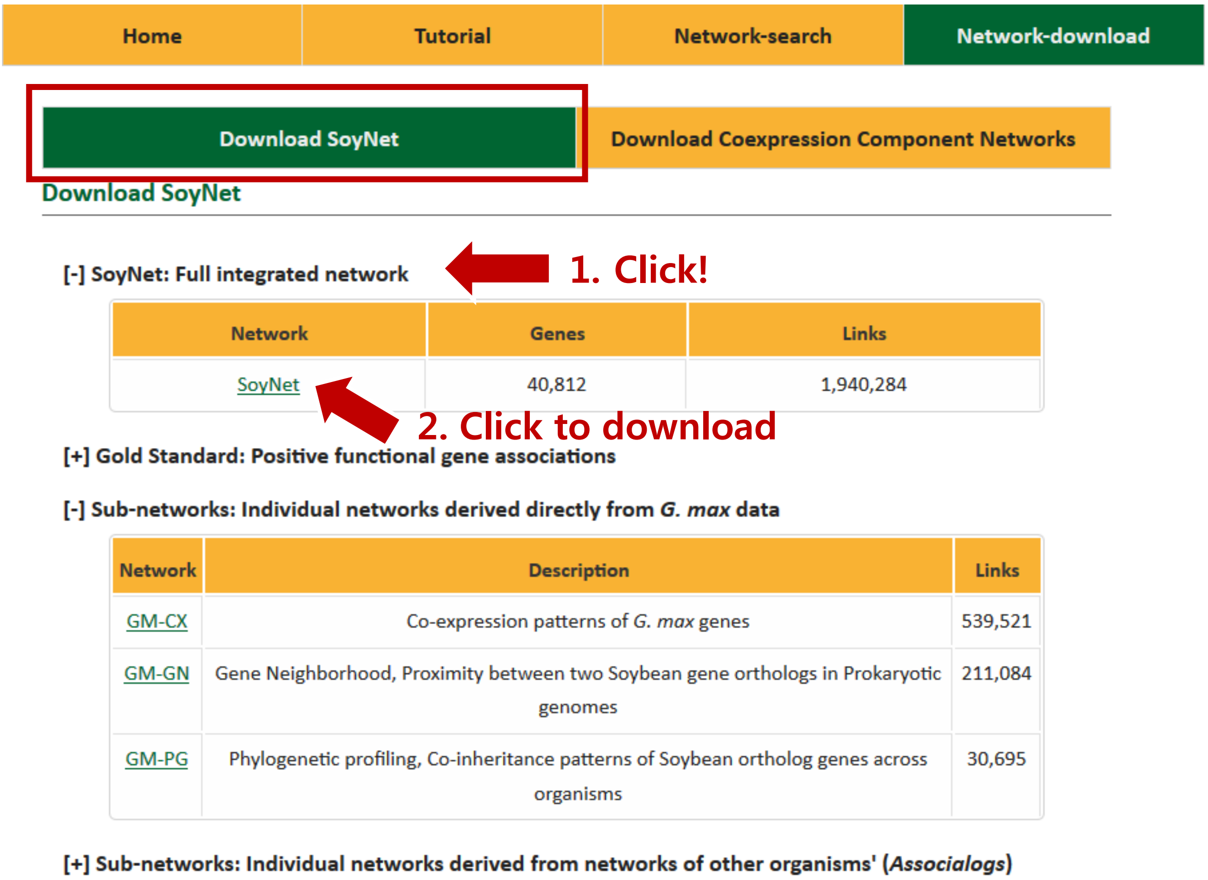
3. Yellow nodes indicate input pathway genes

4. White nodes indicate significant intermediates

5. If you click a node or an edge of interest, detail information will be listed in this panel.

- Network download

Option 1. Download SoyNet



Download SoyNet

[-] SoyNet: Full integrated network

1. Click!

Network	Genes	Links
SoyNet	40,812	1,940,284

2. Click to download

[+] Gold Standard: Positive functional gene associations

[-] Sub-networks: Individual networks derived directly from *G. max* data

Network	Description	Links
GM-CX	Co-expression patterns of *G. max* genes	539,521
GM-GN	Gene Neighborhood, Proximity between two Soybean gene orthologs in Prokaryotic genomes	211,084
GM-PG	Phylogenetic profiling, Co-inheritance patterns of Soybean ortholog genes across organisms	30,695

[+] Sub-networks: Individual networks derived from networks of other organisms' (*Associalogs*)

You can download SoyNet and component networks in the ‘Network-download’ page.

In section ‘Download SoyNet’, SoyNet and its component networks are provided.

1. Click an interest network to show details.

2. Click the network name to download the network.

Descriptions of Downloadable network files

Network file format

Each downloaded network is a plain text file which contains a set of interactions.

Format : Gene A <tab> Gene B <tab> Log-likelihood score (LLS)

Gene A is connected with Gene B in SoyNet.

We trained co-functional gene pairs using gold-standard data with log-likelihood score (LLS) scheme based on Bayesian statistics framework

Log-likelihood score indicates significance of the link. Links with higher score are more significant than lower score.

- SoyNet: Full integrated network

File Name: SoyNet.txt

Description: SoyNet whole network.

- Gold Standard: Positive functional gene associations

File Name: SoyNet_GS.txt

Description: Gold standard pairs used for training SoyNet. The pairs were consturcted using soybean pathway database, KEGG, Mapman, and SoyCyc.

- Sub-networks: Individual networks derived directly from G.max data

File Name: GM-CX.txt

Description: Co-expression patterns of G. max genes

File Name: GM-GN.txt

Description: Gene Neighborhood, Proximity between two Soybean gene orthologs in Prokaryotic genomes

File Name: GM-PG.txt

Description: Phylogenetic profiling, Co-inheritance patterns of Soybean ortholog genes across organisms

- Sub-networks: Individual networks derived from networks of other organisms' (Associalogs)

File Name: AT-CC.txt

Description: PubMed Central full-text Co-citation links of A. thaliana

File Name: AT-CX.txt

Description: Co-expression patterns of A. thaliana genes

File Name: AT-HT.txt

Description: High-throughput PPI data of A. thaliana

File Name: AT-LC.txt

Description: Literature curated PPI of A. thaliana

File Name: CE-CC.txt

Description: Co-citation links in C. elegans network

File Name: CE-CX.txt

Description: Co-expression links in C. elegans network

File Name: DM-CX.txt

Description: Co-expression links in D. melanogaster network

File Name: DM-HT.txt

Description: High-throughput PPI data of D. melanogaster

File Name: DM-LC.txt

Description: Literature curated PPI of D. melanogaster

File Name: DR-CX.txt

Description: Co-expression links in D. rerio network

File Name: HS-HT.txt

Description: High-throughput PPI in H. sapiens network

File Name: HS-LC.txt

Description: Literature curated PPI in H. sapiens network

File Name: SC-CC.txt

Description: Co-citation links in S. cerevisiae network

File Name: SC-CX.txt

Description: Co-expression links in S. cerevisiae network

File Name: SC-GT.txt

Description: Genetic interactions in S. cerevisiae network

File Name: SC-HT.txt

Description: High-throughput PPI in S. cerevisiae network

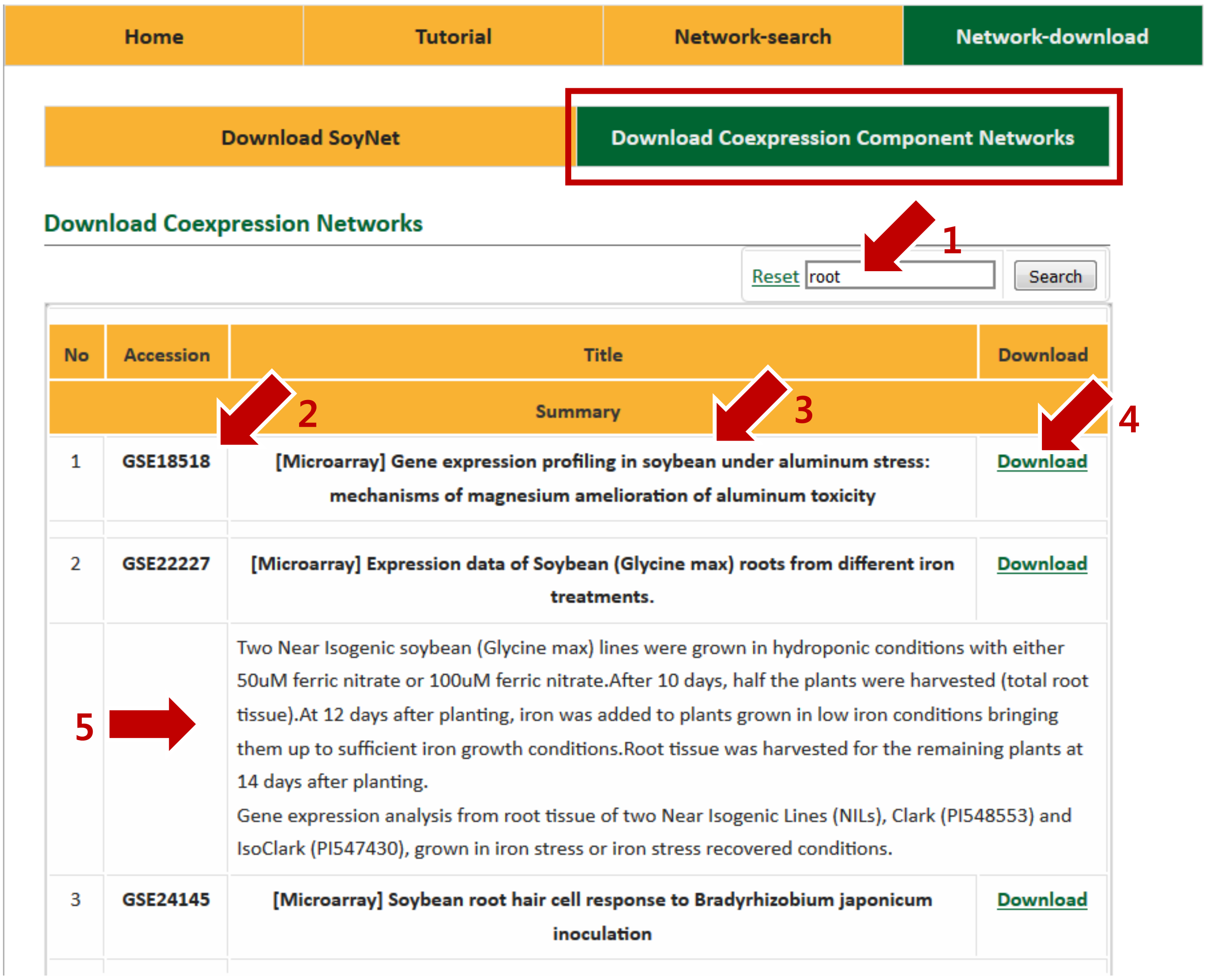
File Name: SC-LC.txt

Description: Literature curated PPI in S. cerevisiae network

File Name: OS-CX.txt

Description: Co-expression links in O. sativa network

Option 2. Download Coexpression Component networks



Download Coexpression Networks

root

No	Accession	Title	Download
		Summary	
1	GSE18518	[Microarray] Gene expression profiling in soybean under aluminum stress: mechanisms of magnesium amelioration of aluminum toxicity	Download
2	GSE22227	[Microarray] Expression data of Soybean (Glycine max) roots from different iron treatments.	Download
		Two Near Isogenic soybean (Glycine max) lines were grown in hydroponic conditions with either 50uM ferric nitrate or 100uM ferric nitrate.After 10 days, half the plants were harvested (total root tissue).At 12 days after planting, iron was added to plants grown in low iron conditions bringing them up to sufficient iron growth conditions.Root tissue was harvested for the remaining plants at 14 days after planting. Gene expression analysis from root tissue of two Near Isogenic Lines (NILs), Clark (PI548553) and IsoClark (PI547430), grown in iron stress or iron stress recovered conditions.	
3	GSE24145	[Microarray] Soybean root hair cell response to Bradyrhizobium japonicum inoculation	Download

In the section of 'Download Coexpression Component Networks', you can download each coexpression network used in SoyNet.

1. You can search network using keywords.

2. Gene Expression Omnibus (GEO) accession of the transcriptome data. http://www.ncbi.nlm.nih.gov/geo/

3. Title of the transcriptome data. Users can click title to show detail description of the study.

4. User can download the network by the link.

5. Detail description of the study

Descriptions of Downloadable network files

Network file format

Each downloaded network is a plain text file which contains a set of interactions.

Format : Gene A <tab> Gene B <tab> Log-likelihood score (LLS)

Gene A is connected with Gene B in SoyNet.

We trained co-functional gene pairs using gold-standard data with log-likelihood score (LLS) scheme based on Bayesian statistics framework

Log-likelihood score indicates significance of the link. Links with higher score are more significant than lower score.

- Component networks for constructing co-expression network

File Name: [Gene Expression Omnibus (GEO) Accession Number].txt

Description: Component co-expression networks