

NetREx

Network-based Rice Expression Analysis

Reference Manual

May 1, 2022

Sanchari Sircar, Mayank Musaddi, Nita Parekh,
Centre for Computational Natural Sciences and Bioinformatics,
International Institute of Information Technology-Hyderabad, India

Contents

1. About App.....	2
2. Feature Overview.....	3
3. Query Option	4
3.1. Query Form	4
3.2. Validation Page.....	5
3.3. Ortholog Network Query	6
4. Network Viewers.....	7
4.1. View.....	8
4.2. Options	10
4.3. Table.....	11
5. Expression Viewer.....	12
6. Browse	13
6.1. Module Wise	13
6.2. Condition Wise	14
6.3. Pathway Wise	14
7. Datasets.....	15

1. About

NetREx is a Network-based Rice Expression Analysis Server hosted as a web application. It allows users to compare and visualize relationships among genes under different abiotic stress conditions/hormone treatments for the *Oryza Sativa L.* cultivar. Information flow between the different “views” in NetREx is represented in **Figure 1** in a user-friendly manner.

One can explore the genes that are differentially co-expressed across four different stress conditions (drought, cold, flood and osmosis) and two hormone treatments (abscisic acid, jasmonic acid) in root and shoot tissues. The correlation between genes is depicted by Pearson correlation coefficient. To reduce network complexity, highest reciprocal rank network, HRR100, is constructed by considering top 100 ranked correlations of a gene as its neighbours.

Tissue-specific weighted gene co-expression network is constructed by considering genes that are differentially expressed for at least two time-points across various stress/treatment conditions. Hierarchical clustering of these signed networks resulted in co-expressed modules that are enriched with up- or down-regulated genes. Thus, GO enrichment of the modules provides information regarding biological processes/pathways that are activated or repressed under the stress conditions. Integrating information like the module membership, fold change across different time-points and associated KEGG pathways augment the network visualization experience making it an integral tool in the analysis of multi-stress tolerance in the rice crop.

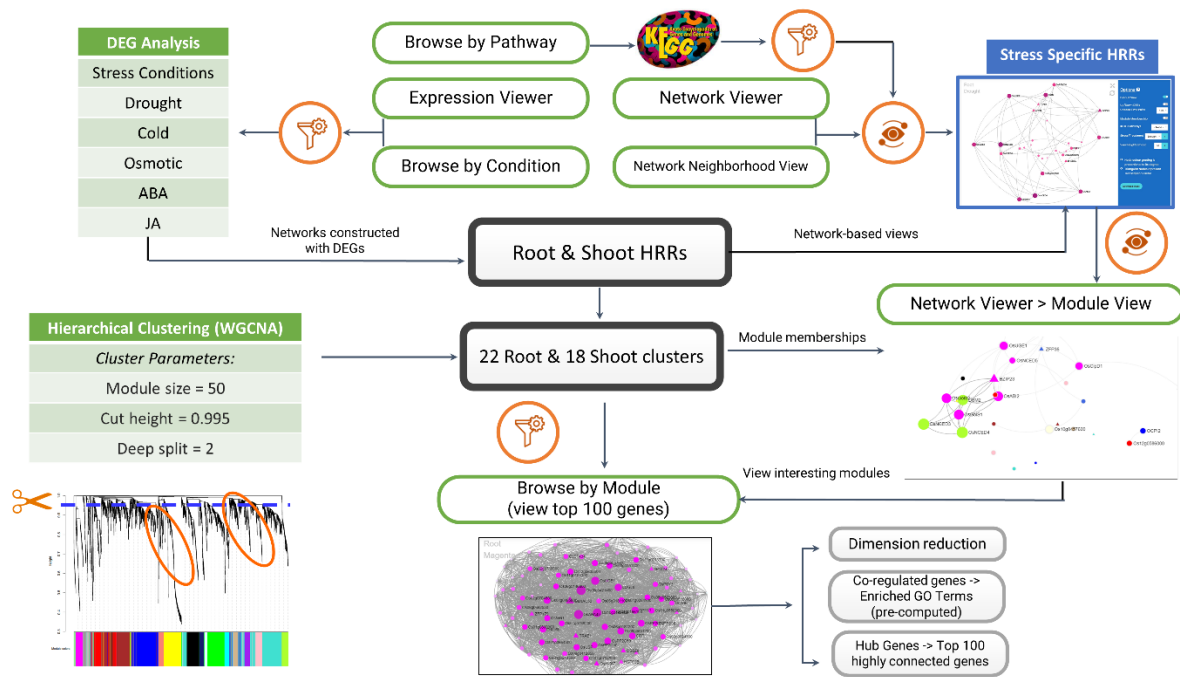


Figure 1: Schematic Representation of NetREx

2. Feature Overview

Four visualization features are provided, viz., Network Viewer, Network Neighborhood Viewer, Expression Viewer and KEGG Pathway Viewer, that are accessible via two options: Query and Browse.

a) Query

Using this option, users can explore interactions between the query genes and their neighbourhood genes, their module membership, associated pathways, and their differential expression across different time-points. The following visualizers provide the above-mentioned features: Network Viewer, Network Neighborhood Viewer, and the Expression Viewer.

b) Browse

This allows the users to explore the complete data in NetREx. It can be accessed from the navigation bar on the Home Page, highlighted in Figure 2. Using the dropdown menu under 'Browse' heading, the user can explore NetREx in three ways: Module-wise, Condition-wise and Pathway-wise. Details of these browse options are discussed in Section 6 below.

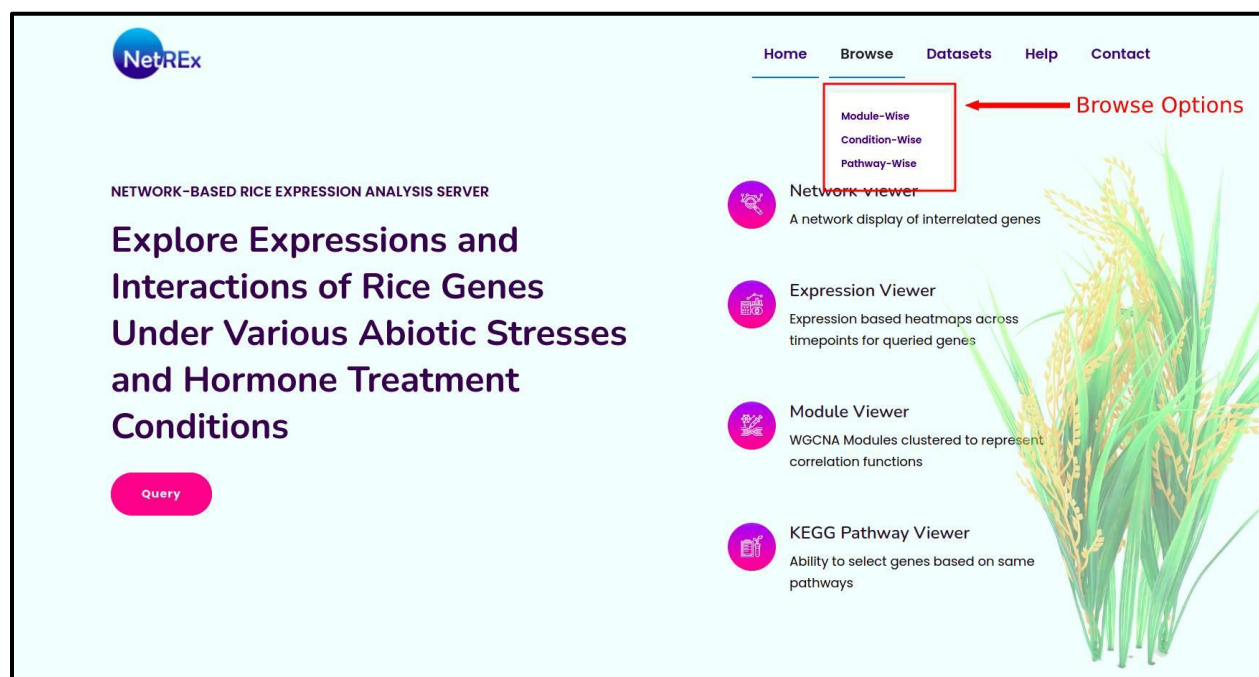


Figure 2: Navigating NetREx using the Browse option.

3. Query Option

3.1. Query Form

To query the network and expression viewers, the user needs to provide following information for query genes using the drop-down menu in **Figure 3**:

- a) **Species** – choose the species query genes belong to
(*Oryza Sativa* / *Arabidopsis Thaliana* / *Triticum aestivum* / *Zea mays* / *Hordeum vulgare* / *Sorghum bicolor*)
- b) **Tissue Type** - choose the tissue in which the user is interested in analyzing the expression of query genes
(*Root* / *Shoot*)
- c) **Condition** – choose the stress or treatment condition in which the user is interested in analyzing the expression of query genes
(Abiotic Stress: *Drought* / *Cold* / *Flood* / *Osmosis*, Phytohormone Treatment: *ABA* / *JA*)
- d) **Gene List** – provide a comma-separated list or a file containing RAPDB Gene IDs of Rice crop or Ensemble Stable Gene IDs of other crops. If the species selected is other than rice, the rice orthologs of the query genes will be fetched from Ensembl Plants Database for querying NetREx.

Alternatively, to explore the features of the application, one can click on the **Load Example** button which will autofill an example set of query genes.

Pressing on the ‘**Send**’ button will lead to the validation/visualizer page.

Pressing on the ‘**Reset**’ button will clear the Gene List entered.

Submit a set of genes to view their network representation at different time-point for different stress/hormone treatments, for root/shoot tissue in rice.

Species : Oryza Sativa

Tissue : Root

Stress / Hormone Treatment : Drought

Enter Comma/Space separated Query Gene List (<300)

Load Example No file chosen

SEND **RESET**

Figure 3: Query Form to submit the genes

3.2. Validation Page

On submitting the query genes, the user is directed to the validation page which displays the Species, Tissue, and Condition selected by the user along with a set of filtered list of genes and invalid genes (Figure 4).

A gene is classified as filtered if it is differentially expressed in at least 2 time-points for the given stress condition. Otherwise, it is termed invalid. For all the analysis in NetREx, only the ‘Filtered’ list is considered for displaying the interactions between genes for the chosen stress/treatment and tissue conditions. One may access the viewers by clicking on the corresponding buttons provided at the top.

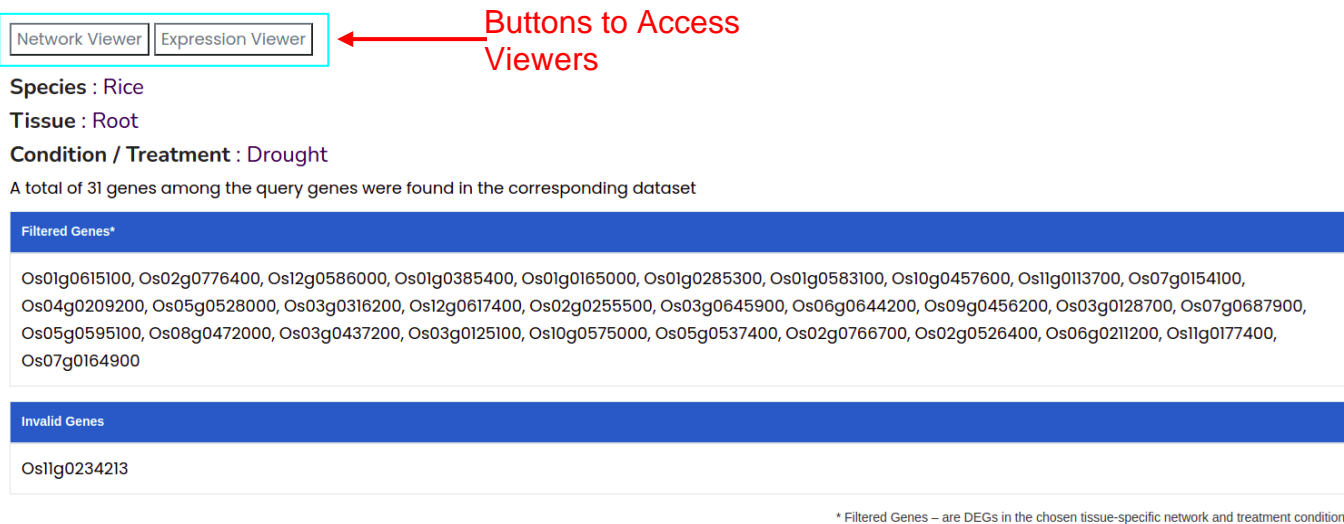


Figure 4: Validation Page

3.3. Ortholog Network Query

To investigate the interaction of genes in other crops under various abiotic stress and phytohormone treatments an option is provided to query the interaction networks of their orthologs in rice. To access this feature, the same query form as specified in section 3.1 is to be used. The appropriate species should be specified from the dropdown menu and a comma separated list of Ensembl Stable Gene IDs (of that species) to be queried should be provided (**Figure 5**). For example, for Arabidopsis: AT2G43490, for Barley: HORVU3Hr1G015620, for Maize: Zm00001d040234, for Sorghum: SORBI_3001G319500, and for Wheat: TraesCS3B02G102400. The genes are mapped to the corresponding rice orthologs based on the Ensembl Plants database. Upon submission, the validation page opens displaying information of the orthologous mapping and the unmapped genes (**Figure 6**).

Submit RAPDB Gene IDs of Rice genes or Ensemble Stable Gene IDs of other crops to be queried and select appropriate tissue and stress/hormone treatments. ?

Species : Arabidopsis Thaliana

Tissue : Root

Stress / Hormone Treatment : Drought

ATIG08880,ATIG10200,ATIG10830,AT2G27700,AT2G40116,AT3G01100,AT3G01420,AT3G02800,AT3G22930,AT3G62270,AT4G19670,AT5G06590,AT5G07920,AT5G13200,AT5G47250,AT5G48160,AT5G57120,AT2G34590,AT4G04850,ATIG19870,AT3G54750,AT5G06590,AT3G03780,AT5G17920,ATIG06590,AT5G45170,AT4G19670,AT4G21231,AT5G12342,ATIG32345

Load Example No file chosen

SEND **RESET**

Figure 5: Example Query for Ortholog Network Query

Network Viewer **Expression Viewer**

Species : Arabidopsis
Tissue : Root
Condition / Treatment : Osmosis

A total of 9 genes among the query genes were found in the corresponding dataset ?

Filtered Genes*

Os02g0526400, Os05g0528000, Os07g0687900, Os10g0575000, Os01g0285300, Os12g0617400, Os03g0437200, Os12g0586000, Os10g0457600

Invalid Genes

AT4G35312, ATIG64445, AT5G07072, Os05g0213500, Os09g0456200, Os09g0456200, Os09g0456200, Os09g0456200, Os07g0608200, Os11g0177400, Os05g0592800, Os05g0592800, Os07g0282500, Os04g0209200, Os02g0530100, Os10g0533500

* Filtered Genes -- are DEGs in the chosen tissue-specific network and treatment condition

Show/Hide Orthologs Mapping ← Button to access ortholog mapping

Valid Mapping **Invalid Mapping** **Unmapped** ← Options to view Valid, Invalid and Unmapped genes

Show entries

Ensembl Stable ID	Rice Ortholog ID	Rice Gene Name
AT1G04710	Os10g0457600	
AT1G57560	Os01g0285300	MYB61, OsMYB55/61
AT4G11230	Os05g0528000	OsRBOH5

Figure 6: Validation Page displaying information of Ortholog Gene Mapping for other crops

4. Network Viewers

The users can access two viewers:

1. **Network Viewer:** It displays the interactions between the filtered set of queried genes, extracted from the HRR network of differentially expressed genes for the selected tissue and stress/treatment condition, shown in **Figure 7**. The nodes correspond to the genes and the edge represents a correlation between the two genes if they have a similar expression profile across time points. It can be accessed by clicking the Network Viewer button on the validation page (**Figure 4**).
2. **Network Neighborhood Viewer:** It displays the connections between the queried genes and their top 50 neighbours (default, max=100), selected based on the top k_{total} values among all the neighbours of the queried genes. It can be accessed from the Network Viewer options panel on the right as shown in **Figure 7**.

The network viewers consist of mainly three panels (**Figure 7**):

1. **View:** Main Panel displaying the network of genes.
2. **Options:** Located to the right of the View panel. It provides button triggers and pull-down menus to access and toggle through various features of the network.
3. **Table:** Accessible through the ‘Show/Hide Table’ button below the View panel. It gives detailed information of all the nodes and edges depicted in the network view.

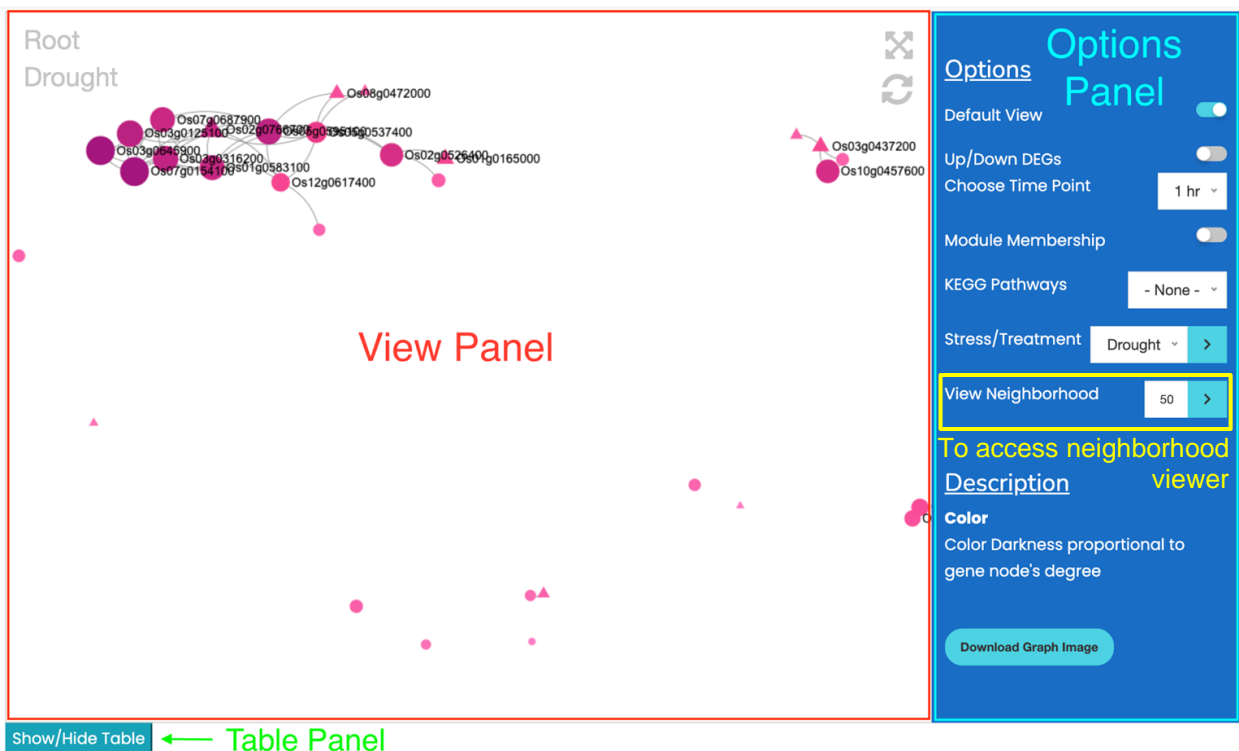


Figure 7: Network Viewer

4.1. View

The network of interacting genes is displayed in the View panel. It allows the user to interact with it, such as zooming, repositioning, and hovering over genes. The nodes in the network correspond to the filtered genes after validation and the interactions between them is extracted from HR100. However, for network neighbourhood viewer, apart from the filtered genes, their top 50 neighbours (max=100) are also shown.

4.1.1. Visuals

Various visual and interaction aspects of these two networks are discussed below.

Layout

The layout of the network is defined by the ForceAtlas2 algorithm, a force-directed algorithm that tends to make connected nodes cluster together and disconnected nodes drift apart.

Size

The size of the gene nodes is proportional to their overall degree in the complete HRR network for the chosen tissue and condition.

Colour

Various options are provided to toggle the colour of gene nodes in the Options panel. By default, the deepness of the colour is proportional to its degree in the complete HRR Network, i.e., higher the degree, darker the colour shade.

In Network Neighborhood Viewer, the query genes are encircled by Lime green colour.

Shape

Genes are represented by filled ‘circles’ while, transcription factors are represented by filled ‘triangles’. The edges are curved shaped for better visibility (**Figure 8**).

Node Label

To avoid overcrowding and improve readability in the network, the labels of all the nodes are not shown all at once. The labels of only those nodes that cross a certain node size threshold across different zoom levels are displayed.

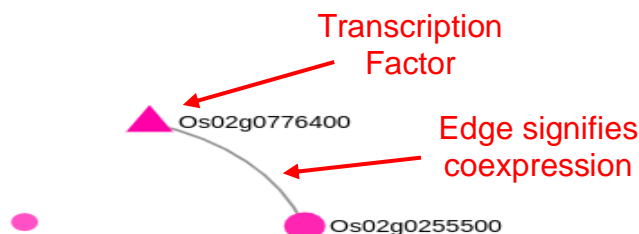


Figure 8: Visualization features of the network

4.1.2. Interaction

The different interactable features are described in this section.

Panning the complete network

Click and drag with the mouse at the stage area of the graph (i.e., the part where there are no nodes and edges present) to pan the network.

Zoom

Double click or scroll to change the zoom level. Zooming also resets the hover highlighted nodes.

Hover

On hovering over a particular node, its label and its first neighbours get highlighted (**Figure 9**).

Move a Node

Clicking over a node and then dragging it moves it relative to the network layout.

Expand ()

For a bigger view, the expand button could be clicked to hide the Options Panel and increase the view size.

Reset ()

If any node is highlighted through hover, or coloured differently using Options Panel, or zoom level has been changed, the reset button can be clicked to go back to the default settings.

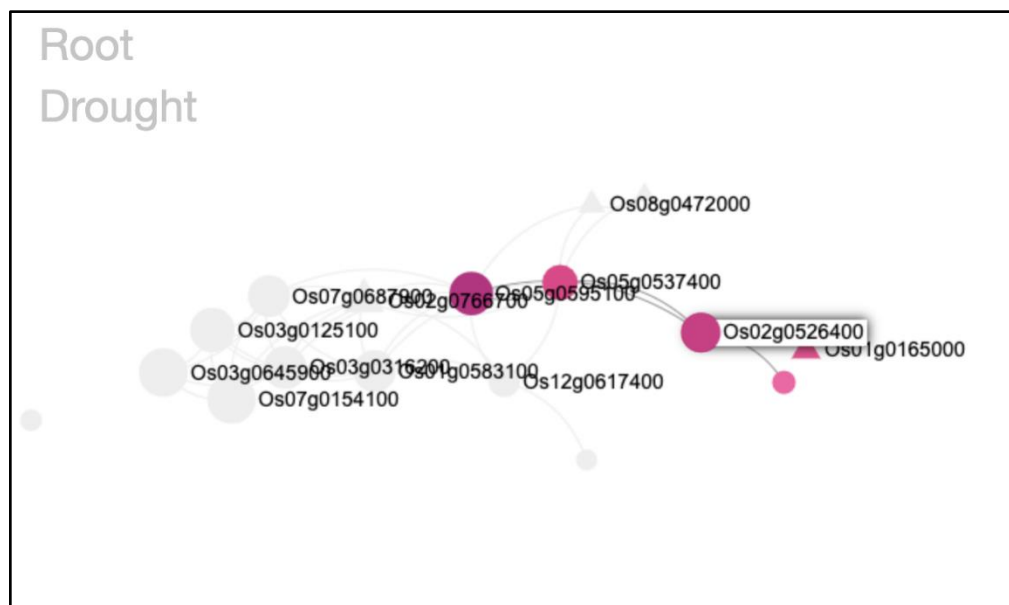


Figure 9: Hovering over a node shown.

4.2. Options

The Options Panel provides a set of features for network visualization such as view genes differentially expressed at different time-points, colouring of nodes, etc. It is placed to the right of the View Panel (**Figure 7**) and its appearance can be toggled by clicking on the expand button in the top right corner of the view panel.

Colouring Switches

The nodes in the graph can be coloured with three different schemes upon toggling the corresponding switches.

- **Default View**

By default, the nodes in the graph are shaded based on their degree in the complete network. A darker shade corresponds to a higher degree and vice versa.

- **Up/Down DEGs**

This feature allows the user to view the genes that are up- and down-regulated at different time-points. From the drop-down menu, the user can select the time-point and switch on the button to enable the colouring. Nodes coloured 'Red' indicates up-regulated genes while those coloured 'Blue' indicates down-regulated genes.

- **Module Membership**

This feature allows the user to view the genes based on their WGCNA module membership. Genes that are co-regulated and belong to the same module are likely to be part of the same biological process. The WGCNA modules are named using colours, which only signifies their membership to a particular module and may be used to associate multiple genes to a common biological process.

KEGG Pathways

Using this feature the user can highlight genes that belong to a specific KEGG pathway. From the drop-down menu the user can select the KEGG pathway. On selecting this option, genes that belong to the selected pathway remain coloured while others are greyed out. This option can be used along with the colouring switches to understand the up/down regulation or module membership in a particular KEGG pathway. However, the colour switches should be toggled before the selection of the pathway.

Stress / Treatment

Using this feature the user can see the network of the queried genes under other stress condition/hormone treatment. From the drop-down menu, the condition can be selected by clicking on the right arrow button. The network for the new stress/treatment condition is loaded in a new tab for comparison.

View Neighborhood

This feature is only provided in the Options Panel of the Network Viewer. In the text box enter the number of neighbours of the queried genes (default = 50, max = 100) the user wishes to see. Clicking on the right arrow button next to the text box loads the neighbourhood viewer (in the same tab).

Download

The user can download the network in the View Panel as a jpg image by clicking the Download Image button.

4.3. Table

The table corresponding to the network could be expanded and collapsed using the “Show/Hide Button” below the View Panel. This table consists of information about the genes in the network view and their interactions, which are listed under the categories, Nodes and Edges.

Edges

Each row corresponds to an edge specifying the source and target gene as well as its PCC and HRR value.

Nodes

The table under the node attribute consists of information about the gene nodes displayed in the network. This information is organized in three categories:

- **Description**
Consists of Gene Name, Transcription Factor (TF) Label, Module Name, General Description, MSU_ID, k_{total} and link to IC4R Expression.
 - On clicking over a module name, the complete network of the corresponding module (Module Viewer) gets loaded in a new tab.
 - On clicking the “Click!” button under IC4R Expression, the IC4R expression profile of that gene gets loaded in a new tab.
- **Function**
It provides functional annotations from GO, Mapman, and KEGG pathway.
- **Fold Change**
Fold change and p-values across different time-points is provided here.

Hover

Upon hovering over a row in the Node attribute table, the corresponding node and its first neighbours get highlighted in the network in the view panel.

One can hover over a particular row in the table to select the corresponding gene. Then move the cursor outside the table panel and scroll up to the view panel to see the selected node and its neighbour.

Search

A search bar is provided on the top right part of the table. A ‘keyword’ search of the entered text is performed across all the categories and only those rows which match the search are displayed.

Download Table

The complete table can be downloaded in different formats (*Copy / Excel / CSV / PDF*) by selecting the corresponding button at the end of the table.

5. Expression Viewer

The Expression Viewer is a heat map-based visualization of the expression profile of queried genes across different time points. It is accessed by clicking on the Expression Viewer button on the Validation Page (**Figure 4**). Two heatmaps are shown, one for fold-change expression values and the other for p -values (**Figure 10**). It is built using ClustergrammerJS plugin and comes with a range of functionalities which are given below.

Sorting

On double-clicking the row or column headers, the heatmap can be sorted based on the expression values of the corresponding row or column.

Zoom

By scrolling up/down over the heatmap, the user can zoom in and out respectively.

Pan

When in a zoomed state, the user can click and drag to pan the heatmap.

Hover

On hovering over a tile in the heatmap, the fold-change value of the respective tile can be seen.

Snapshot ()

To save the image of the heatmap, the user can click on the snapshot button.

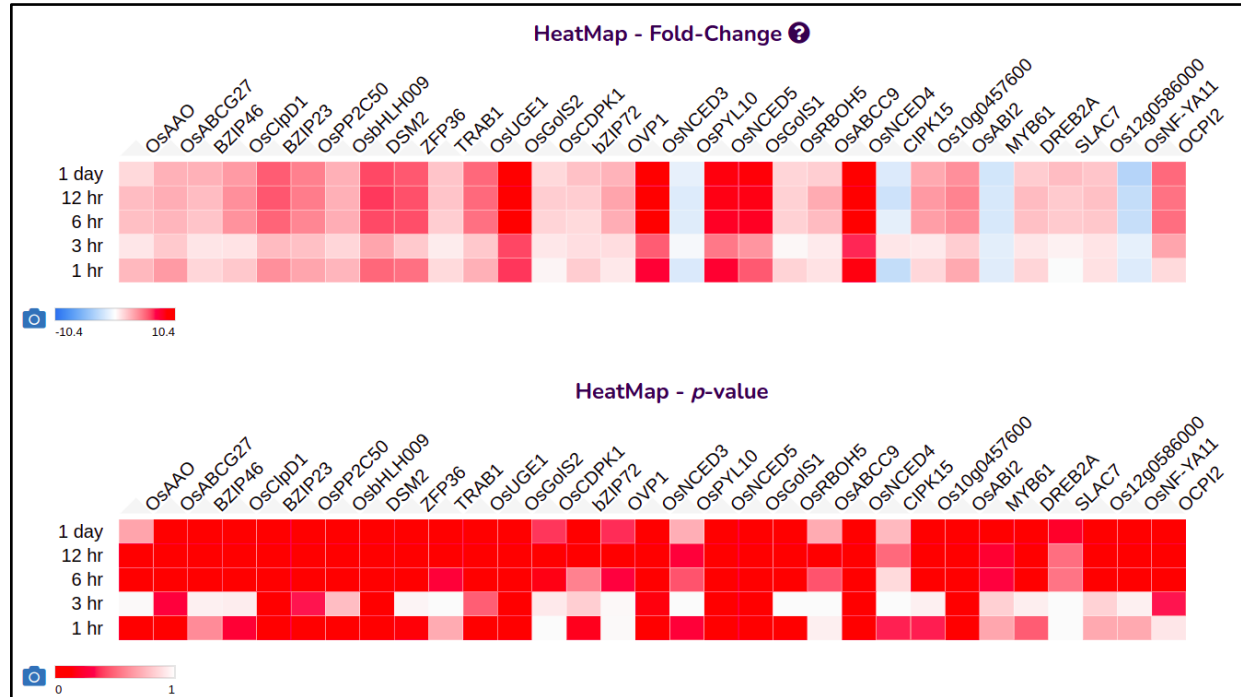


Figure 10: Expression Viewer give the new image with gene names (fixed)

6. Browse

The user can browse the complete data in NetREx using the three different Browse Options which can be accessed from the Navigation bar on the Home page of the application (**Figure 2**).

6.1. Module Wise

In this option, the user can select and view tissue-specific WGCNA modules (top 100 genes) using the Module viewer option. The GO enrichment, node and edge attributes of these genes in a tabular format is also displayed which can be viewed by clicking the respective buttons shown in **Figure 11**. The user needs to select the tissue (*Root / Shoot*) and module colour name as input.

As shown in **Figure 1**, the root and shoot-specific DEGs are clustered (separately) using WGCNA. Hierarchical clustering organizes gene expression profiles into a dendrogram whose branches denote genes with similar expression profiles, a process called “tree cutting”. Here, we set the parameters to adjust the “minimum cluster size” (to get enough genes per cluster to perform GO enrichment), the “branch cut height” (to get optimum number of clusters) and the “deep split” parameter (to fine-tune the sensitivity of cluster detection). The gene clusters or modules are often colour-coded or denoted by numbers. Here, we use ‘colour’ for the purpose of visualization. The co-expressed modules (22 for root and 18 for shoot) are tested for their biological relevance. To that end, the genes of the individual clusters are submitted for over-representation analysis in GO consortium using PANTHER classification system (v14.0) for *Oryza sativa*. Results with Fischer’s Exact Test and Bonferroni correction for multiple testing ($p\text{-value} < 0.05$) are retrieved for each module. These enriched GO terms of the modules can be accessed via table by clicking on “GO Enrichment”. We included only pre-computed GO terms with $FDR < 0.05$. On clicking the Module Viewer button, the network of the top 100 genes of the selected module can be viewed.

Browse Module-wise

Click “Module viewer” to view top 100 highly connected genes based on their within-module connectivity. In this view, the size of the nodes correspond to its degree connectivity while the colour represents the module colour. The table lists the node and edge attributes and function GO enrichment of these genes.

Tissue : Root

Module : Black

[SEND](#)

[Module Viewer](#)

[Nodes](#) [Edges](#) [GO Enrichment](#)

Show 10 entries

[Description](#) [Function](#)

Search:

Label	TF	Description	MSU ID	Ktotal	IC4R Expression
Os01g0121500	-	Conserved hypothetical protein.	-	21.6288395773726	Click!
Os01g01159300	-	Zinc finger, RING/FYVE/PHD-type	-	A7 3377961368288	Click!

Figure 11: Module-wise browse page.

6.2. Condition Wise

In this option, the user can obtain complete details of all the genes differentially expressed in the chosen stress/hormone treatment and tissue in the table format, viz., gene information, gene function and fold-change and p-values at various time-points. The user needs to select for input, the tissue (*Root / Shoot*), and the stress/hormone treatment for the species *Oryza Sativa*. This displays a table containing the complete list of differentially expressed genes for the chosen tissue and stress/hormone treatment. Three tables providing gene information, functional annotation and fold-change values at different time-points is provided.

6.3. Pathway Wise

In this option, the user can select KEGG pathways from the expandable list and visualize the network of differentially expressed genes in the chosen tissue and stress/treatment condition of the selected pathway. The user needs to select for input, the tissue (*Root / Shoot*), and the stress/hormone treatment for the species *Oryza Sativa*. A multi-select tree showing the hierarchically arranged KEGG pathways and the associated genes are displayed. The checkboxes provided next to each pathway, sub pathways and genes may be ticked by the user for selection. Upon selection, the genes from the chosen pathways (max = 100) are displayed in the right-hand side panel (**Figure 12**). On submitting the selected genes, these are sent as query to the network and expression viewers and the user can analyze these genes as discussed above for any set of user-defined genes. A search bar option is also provided for keyword search for KEGG pathways. Thus, this feature allows the user to view genes of the select pathway(s) and their representation in the chosen tissue and stress/treatment condition.

Pathway Browse

Allows the user to analyze the genes of selected pathway and their expression in chosen tissue and stress/treatment condition.

Species :

Oryza Sativa

Tissue :

Root

Stress / Hormone Treatment :

Drought

SEND

The list below displays all the genes, with respect to the selected tissue and stress/hormone treatment, clubbed according to the different KEGG pathways. One may select upto 300 genes and submit them to visualize their relationship from the application's viewers.

Search...

- Global and overview maps
- Carbohydrate metabolism
- Energy metabolism
 - ☒ Oxidative phosphorylation
 - ☒ Os02g0791400
 - ☒ Os05g0438500
 - ☒ Photosynthesis
 - ☒ Carbon fixation in photosynthetic organisms
 - ☒ Nitrogen metabolism
 - Sulfur metabolism
- Lipid metabolism
- Nucleotide metabolism
- Amino acid metabolism
- Metabolism of other amino acids

SUBMIT SELECTED GENES

<input checked="" type="checkbox"/> Os02g0791400	Energy metabolism/Oxidative phosphorylation
<input checked="" type="checkbox"/> Os05g0438500	Energy metabolism/Oxidative phosphorylation
<input checked="" type="checkbox"/> Os01g0773700	Energy metabolism/Photosynthesis
<input checked="" type="checkbox"/> Os03g0784700	Energy metabolism/Photosynthesis
<input checked="" type="checkbox"/> Os05g0443500	Energy metabolism/Photosynthesis
<input checked="" type="checkbox"/> Os07g0141400	Energy metabolism/Photosynthesis
<input checked="" type="checkbox"/> Os09g0481200	Energy metabolism/Photosynthesis
<input checked="" type="checkbox"/> Os01g0188400	Energy metabolism/Carbon fixation in photosynthetic organisms
<input checked="" type="checkbox"/> Os05g0186300	Energy metabolism/Carbon fixation in photosynthetic organisms
<input checked="" type="checkbox"/> Os06g0668200	Energy metabolism/Carbon fixation in photosynthetic organisms
<input checked="" type="checkbox"/> Os01g0639900	Energy metabolism/Nitrogen metabolism
<input checked="" type="checkbox"/> Os03g0223400	Energy metabolism/Nitrogen metabolism
<input checked="" type="checkbox"/> Os03g0712800	Energy metabolism/Nitrogen metabolism

Figure 12: Selecting genes Pathway-wise.

7. Datasets

The complete raw datasets which have been used for constructing NetREx can be downloaded by the user from the Navigation Bar on the Home page in **Figure 2**. It contains data of the Node Lists for Root and Shoot tissues for various stress conditions/hormone treatments along with the fold-change values for different time-points as shown in **Figure 13**.

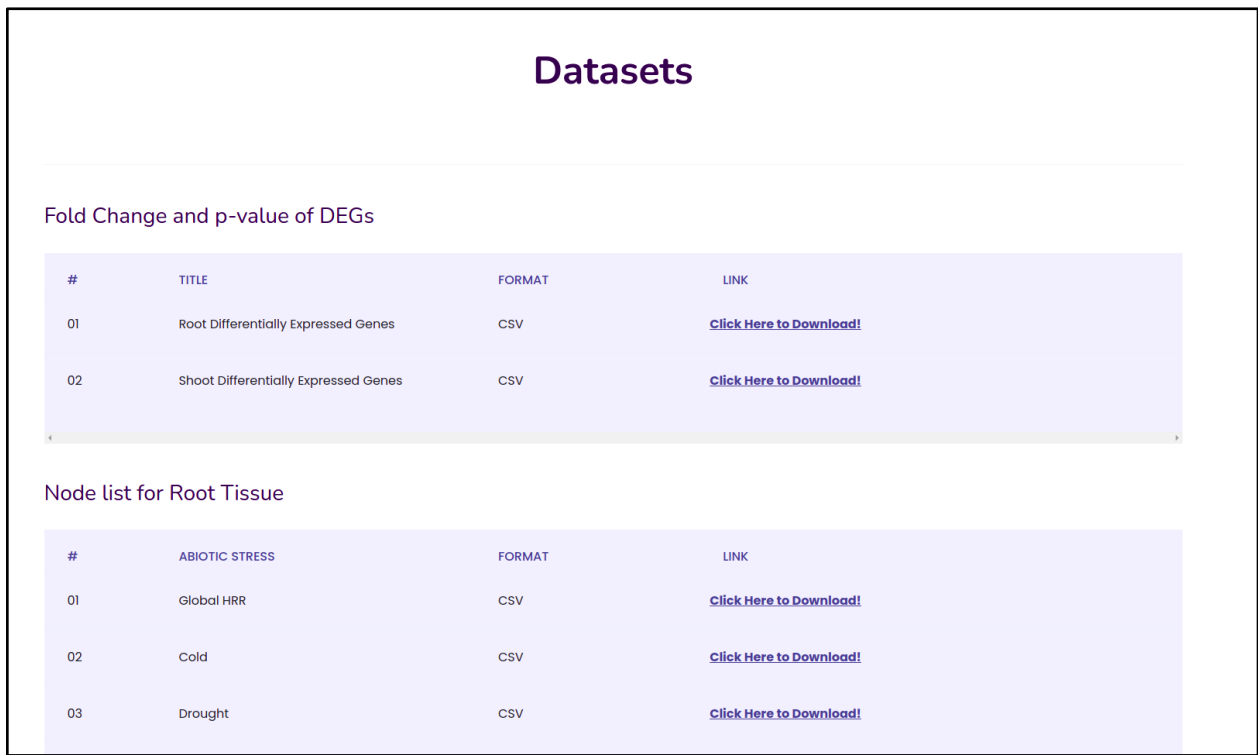


Figure 13: Dataset download Page