

NetREx

Network-based Rice Expression Analysis

Reference Manual

October 1, 2021

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

Contents

About App	2
Home Page of Application	2
Feature Overview	3
Query Option	4
Query Form	4
Validation Page	5
Network Viewers	6
View	7
Options	9
Table	10
Expression Viewer	11
Browse	12
Module Wise	12
Condition Wise	12
Pathway Wise	12
Datasets	13

1. About App

NetREx is a Network-based Rice Expression Analysis Server hosted in the form of a web application that allows users to compare and visualize relationships among different genes under different abiotic stress conditions/ hormone treatments for the *Oryza Sativa* L. cultivar.

One can explore the genes that are differentially co-expressed across four different stress conditions (drought, cold, flood and osmosis) and two hormone treatments (ABA, JA) in different tissues, root and shoot. The correlation between the genes is depicted by the Pearson correlation coefficient and to reduce the complexity of the network further filtering based on top 100 ranked correlations for a gene are represented as its neighbours in HRR100.

Furthermore, from weighted gene co-expression network analysis, different modules have been identified from the network and its GO enrichment has been done enhancing the information of the network. Thus integrating information like the module membership, the fold change across different time points and involvement in KEGG pathways, augments the network visualization experience making it an integral tool in the research of understanding multi stress tolerance in the rice crop.

Link of the application: bioinf.iit.ac.in/netrex/

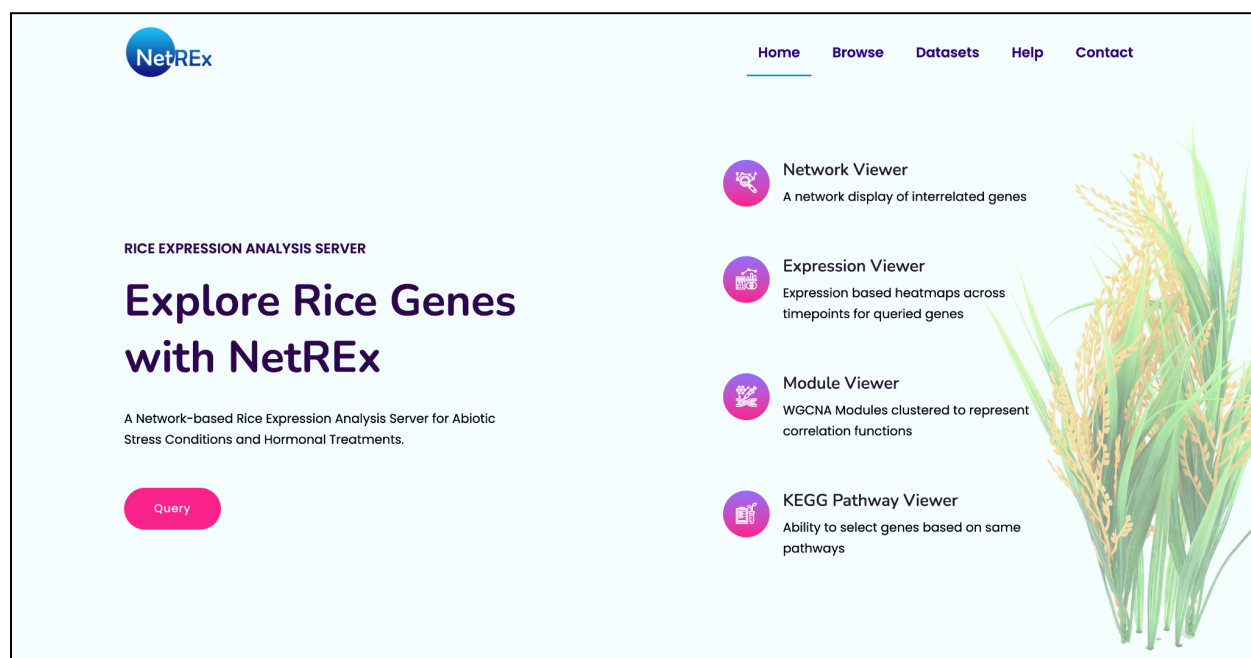


Figure 1: Home Page of Application

2. Feature Overview

All the tools and visualization features are divided into two categories, Query and Browse.

a) Query

This category covers the visualizers that need a user-defined input of genes whose interactions are to be analyzed.

To access this, a query form needs to be filled and submitted on the bioinf.iiit.ac.in/netrex/query.html page, which involves a list of genes. All the visualizers that are displayed pertain to the submitted set of genes. The following visualizers can be accessed using this category:

- i) **Network Viewers:**
 - 1) **Network Viewer**
 - 2) **Network Neighborhood Viewer**
- ii) **Expression Viewer**

b) Browse

This category covers the data browsing and visualization from the complete dataset that is classified based on different metrics.

To access this, one can use the navbar and select the appropriate browse option from the dropdown under the 'Browse' caption. One can also scroll down in the bioinf.iiit.ac.in/netrex/query.html page and click the appropriately labelled buttons leading to the respective browse page. The following data browse option and visualizers are presented under this category:

- i) **WGCNA Module**
 - 1) **Module-wise Data Table**
 - 2) **Module Viewer**
- ii) **Stress Condition / Hormone Treatment**
 - 1) **Condition-wise Data Table**
- iii) **KEGG Pathways**
 - 1) **Condition-wise Gene List clubbed according to different KEGG Pathways**

3. Query Option

3.1. Query Form

To access the network viewers and the expression viewer, this category needs to be selected. A form is presented which requires the user to fill in the following information:

a) Species

(*Oryza Sativa* / *Arabidopsis Thaliana* / *Triticum aestivum* / *Zea mays* / *Hordeum vulgare* / *Sorghum bicolor*)

The plant species corresponding to which the Gene ID list is provided.

b) Tissue Type

(*Root* / *Shoot*)

The tissue type for which the gene differential co-expression is to be observed.

c) Condition

(Abiotic Stress: *Drought* / *Cold* / *Flood* / *Osmosis*, Phytohormone Change: *ABA* / *JA*)

The condition/hormone treatment applied to the rice crop, for which the gene differential co-expression is to be observed

d) Gene List

A comma-separated list or a file containing RAPDB Gene IDs of Rice crop or Ensemble Stable Gene IDs of other crops needs to be provided for which the relationships are to be analyzed corresponding to the tissue and stress condition/hormone treatment.

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

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 Choose file No file chosen

SEND **RESET**

Figure 2: Query Form

3.2. Validation Page

Upon submission of the query, one gets redirected to the validation page which displays the Species, Tissue, and Condition that was specified in the query form along with a filtered list of genes that have been found to be Differentially Expressed corresponding to the query input.

A gene is classified as filtered if it is differentially expressed in at least 2 time-points for a given stress condition. Otherwise, it is termed invalid. All the viewers, display interactions only for the Filtered list of genes. One may access the viewers by clicking on the corresponding buttons provided at the top.

Network Viewers

Expression Viewer

Species : Rice

Tissue : Root

Condition / Treatment : Drought

A total of 31 genes among the query genes were found in the corresponding dataset

Filtered Genes*

Os01g0615100, Os02g0776400, Os02g0586000, Os01g0385400, Os01g0165000, Os01g0285300, Os01g0583100, Os01g0457600, Os01g0113700, Os07g0154100, Os04g0209200, Os05g0528000, Os03g0316200, Os02g0617400, Os02g0255500, Os03g0645900, Os06g0644200, Os09g0456200, Os03g0128700, Os07g0687900, Os05g0595100, Os08g0472000, Os03g0437200, Os03g0125100, Os01g0575000, Os05g0537400, Os02g0766700, Os02g0526400, Os06g0211200, Os01g0177400, Os07g0164900

Invalid Genes

Os04g0332122

* Query Genes present in the list of DEGs for the chosen tissue and condition are referred as "filtered genes". The remaining genes are referred as "invalid genes" for the chosen tissue/condition. Some of the genes may not be DEGs at all time points – so genes which are DEGs in atleast 2 time points for the chosen Tissue/Condition are selected.

Figure 3: Validation Page

4. Network Viewers

Network Viewer displays the differentially co-expressed HRR network of the filtered set of queried genes for the specified tissue and condition. 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.

Under Network Viewers come two networks:

1. **Network Viewer:** Network showing only the connections between the queried genes
2. **Network Neighborhood Viewer:** Network showing the connections between the queried genes as well their top 50 neighbours (default), which can be changed to show up to a maximum of top 100 neighbours of the queried genes

The network viewers consist of mainly three panels:

1. **View:** Panel that displays the network
2. **Options:** Existing on the side of the View panel and contains options to access and toggle different features of the network
3. **Table:** Accessible through the 'Show/Hide Table' button, it displays the information of nodes and edges that constitute the network

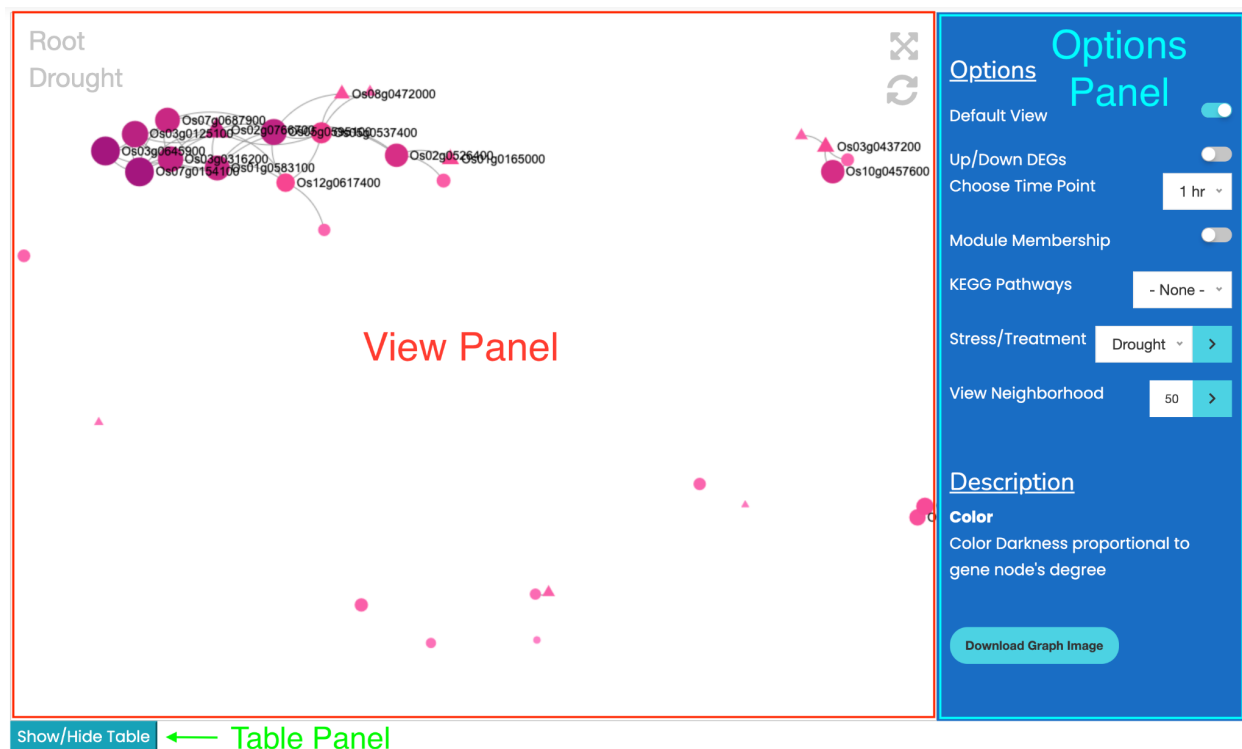


Figure 4: Network Viewer

4.1. View

The View panel constituting the network comprises genes denoted as nodes and differential coexpression relation between two genes denoted by an edge. For simple network viewers, we find that the number of nodes in the network is equal to the number of filtered genes after validation. For network neighbourhood viewer, the number of nodes is equal to the sum of the number of filtered genes and the number of neighbours that had been specified.

4.1.1. Visuals

The different visual aspects related to the network is described in this section.

Layout

The layout of the network is defined by the ForceAtlas2 algorithm, a force-directed algorithm that tends to make related nodes near each other and unrelated nodes drift apart.

Size

The size of the gene nodes are proportional to their overall degree in the complete HRR network for the given tissues and condition

Colour

There are several options to toggle the colours of the gene nodes in the options panel. By default, the deepness of the colour is proportional to its degree in the complete HRR Network.

For Neighborhood Viewer, an outline of Lime green denotes that it is one of the queried gene nodes

Shape

Any general gene has a circular shaped node while a transcription factor is shown as a triangular-shaped node. The edges are curved shaped for better visibility.

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.

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.

Move a Node

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

One may also right-click over a node and then left-click at a place where one wants to move it.

Expand ()

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

Reset ()

If any particular node or pathway is highlighted, the reset button could be clicked to remove the highlights and reset the colouring. Reset also makes the zoom level default where all the nodes are visible.

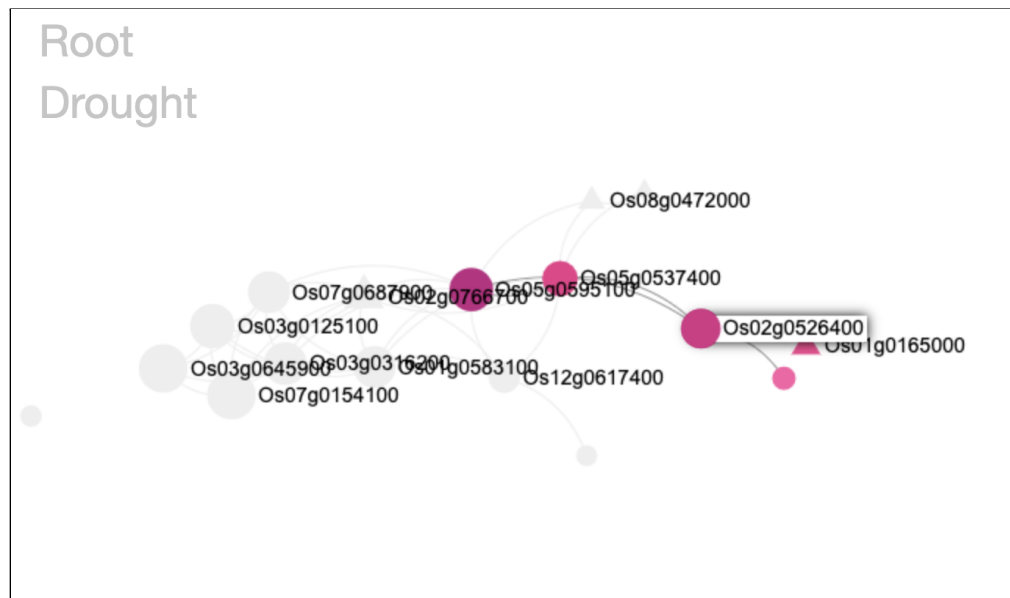


Figure 5: Hovering over a node

4.2. Options

The options panel rests beside the view and its visibility could be toggled by clicking on the expand button in the top right corner of the view panel. This panel hosts a set of options for network interaction and also contains the description of the option based on the options selected.

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 scheme enables to view the up/down-regulated genes at different time points. From the drop-down one could select the time point and switch on the button to enable the colouring. Red indicated up-regulated while blue indicated down-regulated
- **Module Membership**
Switching this will make the nodes coloured on the basis of the WGCNA module to which they belong.

KEGG Pathways

This option enables the user to highlight genes that belong to a specific KEGG pathway. From the drop-down one can select the KEGG pathway and then only those genes which belong to the selected pathway remain coloured while the 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

To see the corresponding viewers for some other stress condition/hormone treatment, one could select that condition from the dropdown and click on the right arrow button next to it, to load the viewers for the same set of genes in a new tab.

View Neighborhood

This option is only provided in the options panel of the network viewer. Here a text box is provided where the user could enter the number of neighbours of the queried genes that he wishes to see. Clicking on the right arrow button next to the text box loads the neighbourhood viewer.

Download

The visualization in the view panel could be downloaded in the form of a jpg image by clicking the Download Image button.

4.3. Table

The table corresponding to the above network could be expanded and collapsed using the “Show/Hide Button” underneath the view panel. This table consists of information about the network classified into two sections, 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 node section consists of gene information displayed in the network. This information has been further divided into three sections for better visibility.

- **Description**

Consists of Transcription Factor (TF) Labelling, Module Name, General Description, MSU_ID, and 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**

Functional annotations like GO, Mapman, and KEGG are listed under this section.

- **Fold Change**

Fold change and p-values across different timepoints are listed under this section

Hover

Upon hovering over a row in the Node Section, 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. They can also now click anywhere in the view panel to move the selected node to that position.

Search

Apart from the above features in the table there also exists a search bar corresponding to both nodes and edges where the entered text is searched for across all the sections.

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 visual of the expression profile of the queried genes across different time points. It is accessed by clicking on the Expression Viewer button on the Validation Page. Two heatmaps are shown, the first corresponding to the fold change expression values, while the second corresponding to the p-values.

Using the plugin ClustergrammerJS for the heatmap, it hosts a range of functionalities.

Sorting

On double-clicking the row or the column headers, the corresponding row or column gets sorted based on the expression values.

Zoom

Scrolling up/down over the heatmap zooms in and out correspondingly

Pan

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

Hover

On hovering over a tile in the heatmap for over a second, the corresponding value of the tile pops up.

Snapshot ()

To save a snapshot of the heatmap, click on this button.



Figure 6: Expression Viewer

6. Browse

The complete dataset could be browsed with the aid of three different browsers that have been provided in the application.

6.1. Module Wise

Accessed from bioinf.iiit.ac.in/netrex/module.html, this browse option takes in the tissue (*Root / Shoot*) and the module colour name input from the user and displays a table containing information about the top 100 genes belonging to the corresponding module. Apart from the nodes and edges, the table also contains GO information about the corresponding module.

Upon clicking on the Module Viewer button, the network of the top 100 genes in the corresponding module is displayed.

6.2. Condition Wise

Accessed from bioinf.iiit.ac.in/netrex/condition.html, this browse option takes in the species (*Oryza Sativa*), the tissue (*Root / Shoot*), and the stress/hormone treatment input from the user and displays a table containing the complete information of the differentially expressed genes based on HRR100 Network constructed for the corresponding tissue and stress/hormone treatment.

Since this data is lengthy, it takes a minute time to load the table after submitting the query. No visualization is provided in this browse view owing to a large amount of data.

6.3. Pathway Wise

Accessed from bioinf.iiit.ac.in/netrex/pathway.html, this browse option takes in the species (*Oryza Sativa*), the tissue (*Root / Shoot*), and the stress/hormone treatment input from the user and displays all the differentially expressed genes for the given query classified based on their participation in a particular KEGG pathway.

One is presented with a multi-select tree showing the hierarchically arranged KEGG pathways and the genes under them. Ticking the checkbox provided left to a gene or a KEGG pathway selects the particular gene. Upon selection of up to a maximum of 100 genes, one can click on the ‘Submit Selected Genes’ option to visualize interaction among them through the use of Network Viewers and Expression Viewer. The multi-select tree under KEGG pathways also has the option for a search bar where the entered text is searched across pathway names and genes.

7. Datasets

The complete datasets using which the graph was constructed could be downloaded under this section bioinf.iiit.ac.in/netrex/dataset.html. It contains data of the Node Lists for the Root and Shoot Tissue for the different stress conditions/hormone treatment. It also has the overall fold change values for the DEGs corresponding to root and shoot tissue.