

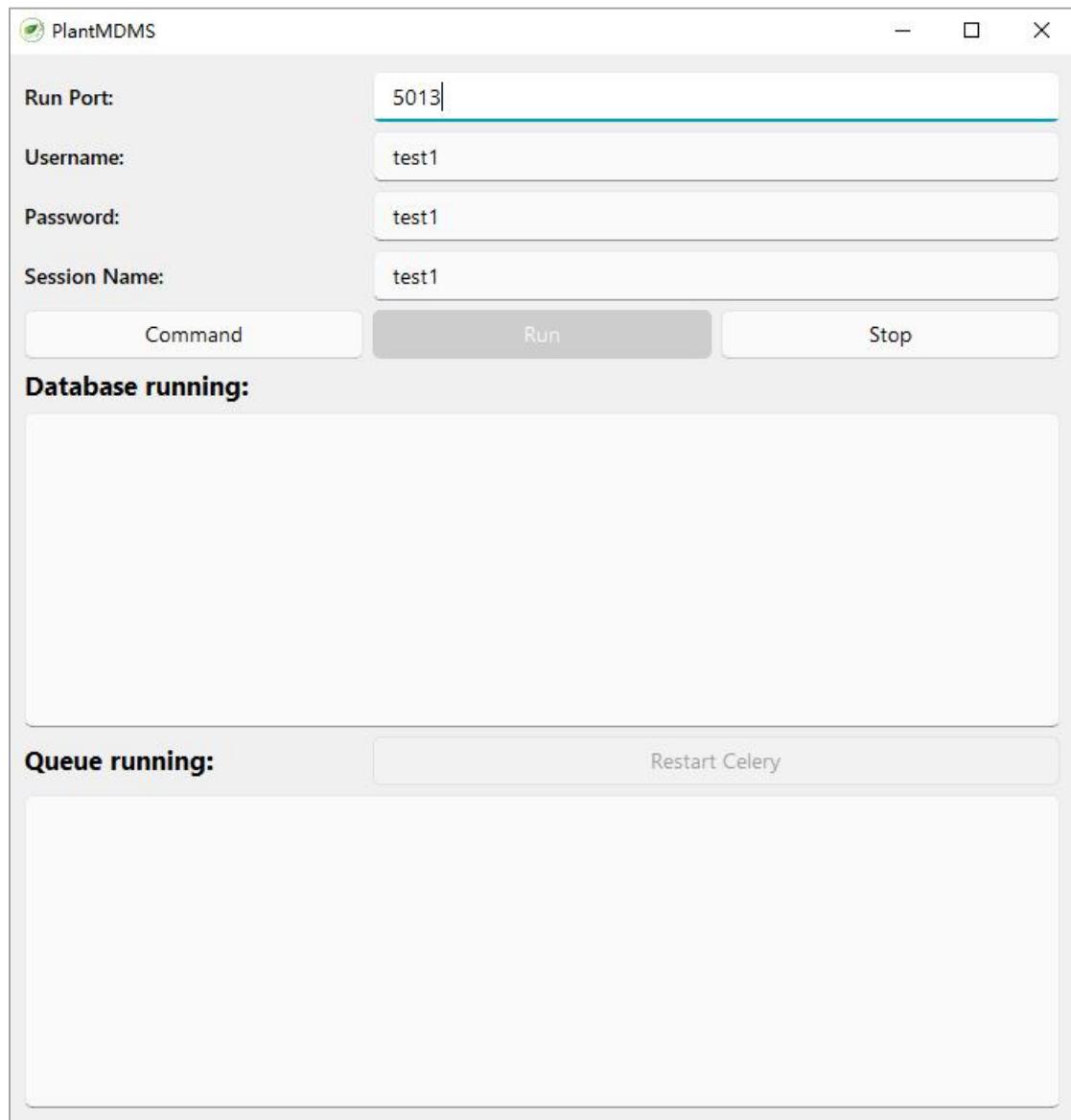
Contents

Quick Start	1
Introduction to Front-End Functions	26
Genome	26
1. Gene Search	26
2. Transcription Factor	27
3. Pathway	29
4. Protein Structure	31
5. Genome Browser	32
Transcriptomics	33
1. Expression Search	33
2. Project information	34
3. DEGs	35
4. Heatmap	36
5. Enrichment:	37
6. Mfuzzy	39
7. PCA	39
8. eFP	42
9. WGCNA	43
10. IGV BAM	49
11. Quality Reports	50
Metabolomics	51
1. Project information	53
2. DEGs	54
3. Mfuzz	55
4. PCA	55

Proteomics	56
1. Protein search.....	56
2. Project information	57
3. DEGs	58
4. Mfuzz	58
5. PCA	58
6. Enrichment	59
Phenotype	59
1. Trait search	59
2. Project information	60
3. Correlation Project.....	61
4. HCA Project.....	62
5. PCA Project.....	63
6. Correlation Traits	64
7. HCA Traits	64
8. PCA Traits	64
Population	64
1. Accession	64
2. Population	65
3. VCF	67
4. GWAS Result.....	68
5. Variation Search	70

Quick Start

Once installed, double-click **PlantMDMS.exe** to launch the initial interface.

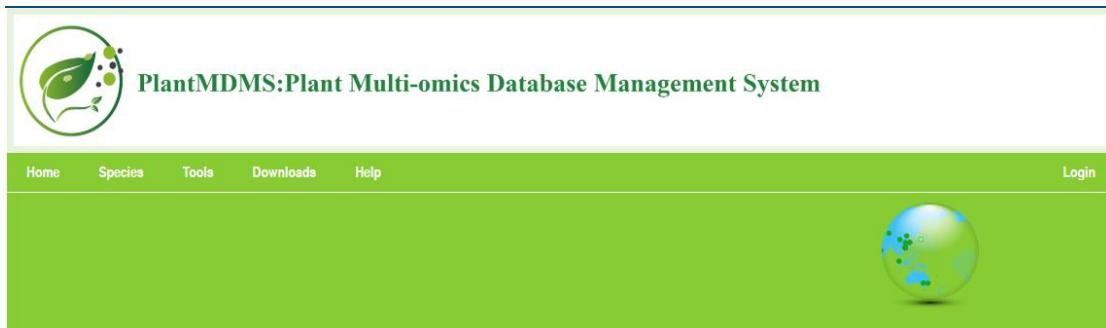


Configure the relevant information, click "**Command**" to activate the "**Run**" button (from gray to white), and then click "**Run**". A database URL will appear at the bottom of the dialog.

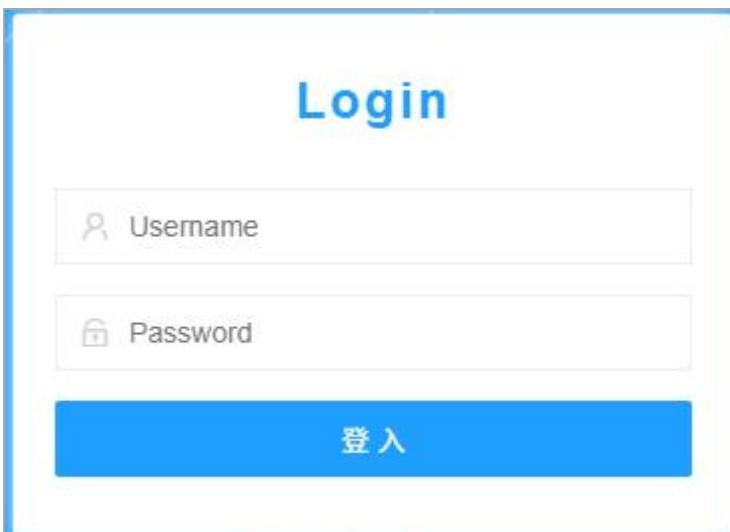
```
Database Run Command:  
D:  
cd D:\PlantMDMS\PlantMDMS  
D:/PlantMDMS/PlantMDMS/omic_flask2/python.exe main.py runserver  
  
Queue Run Command:  
D:  
cd D:\PlantMDMS\PlantMDMS  
D:/PlantMDMS/PlantMDMS/omic_flask2/python.exe -m celery -A app.mycelery.main worker -n Queue_num -E -P eventlet  
-l info --concurrency=2
```

Starting Database
Database running:<http://127.0.0.1:5013>

Copy this URL into your browser and press Enter. The displayed database interface is as follows (this is the front end, responsible for all specific data analysis and visualization).



Click "Login" (Note: Only administrators are authorized to upload data and build the database). Below is the step-by-step process for data upload and database construction.



The default username and password can be found on the initial interface. Next, you'll access the database settings main interface.

The screenshot shows the PlantMDMS interface. On the left is a sidebar with options like 'Genome Upload', 'TF Upload', and 'Protein Feature...'. The main area has tabs for 'Genome', 'Transcriptome', 'Phenotype', 'Variation', 'Metabolome', 'Proteome', and 'Design database'. Below these are sections for 'State' (CPU, GPU, Memory, Disk), 'Statistics' (Genome, Transcriptome, Metabolome, Proteome counts), and 'Version Information' (Name: PlantMDMS-Plant multi omics database management system, Version: v1.0, Features: No threshold / responsive / clean / minimalist, Help: View, Download: GitHub / gitee). A 'System Notice' sidebar on the right lists recent activities.

Click "Design database":

This screenshot shows the same interface as above, but the 'Design database' tab is highlighted with a red box. The rest of the interface remains the same, showing the genome statistics and version information.

Click "Database setting":

This screenshot shows the 'Database Setting' page. It includes sections for 'Set Database' (with fields for port 0-65535, password, start/stop buttons, and delete database), 'Set Queue' (with fields for port, status, start/stop buttons), 'Set WGCNA Queue Num' (with a field for 1 and a set button), 'Set Queue Num' (with a field for 2 and a set button), and 'Run Queue Num' (with a field for 2 and a set button).

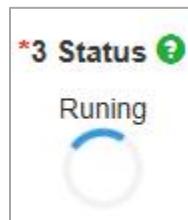
Set the port and password:

Set Database

*1 Set port ?	*2 Set password ?	*3 Status ?	4 Stop ?	5 Delete Database ?
<input type="text" value="3380"/>	<input type="text" value="Password"/>	<input type="button" value="Start"/>	<input type="button" value="Stop"/>	<input type="button" value="Delete Database"/>

For the port number, it can be randomly selected from 0 to 65535. If a port has already been selected by another member in the same group, a corresponding prompt will be generated. Note that the candidate query and analysis are always carried out under this port. Therefore, it is recommended to memorize and save the port number and password at the same time.

Click "Start" to initialize the database creation. Upon successful creation, "Running" will be displayed, as shown below:



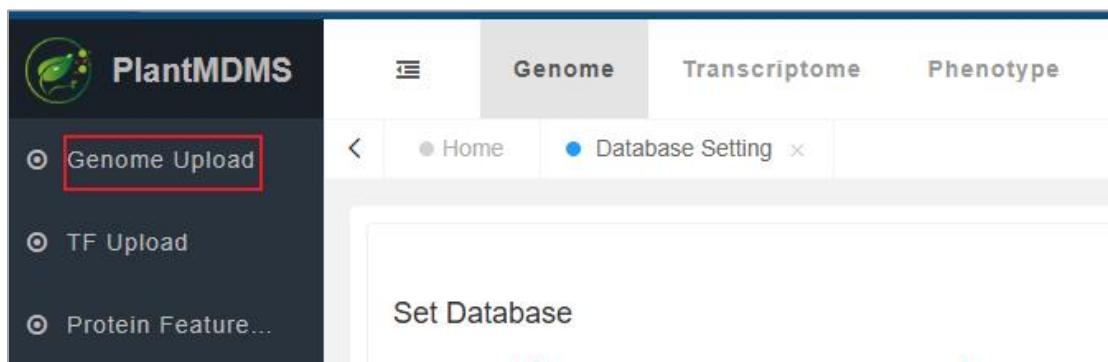
In addition, we have also set up a queue port. This port is mainly designed to prevent program crashes caused by concurrent execution of time-consuming and resource-intensive tasks. The value range of the queue port is the same as that of the database port, but the two are distinct and need to be configured separately.

The screenshot shows a configuration interface with the following sections:

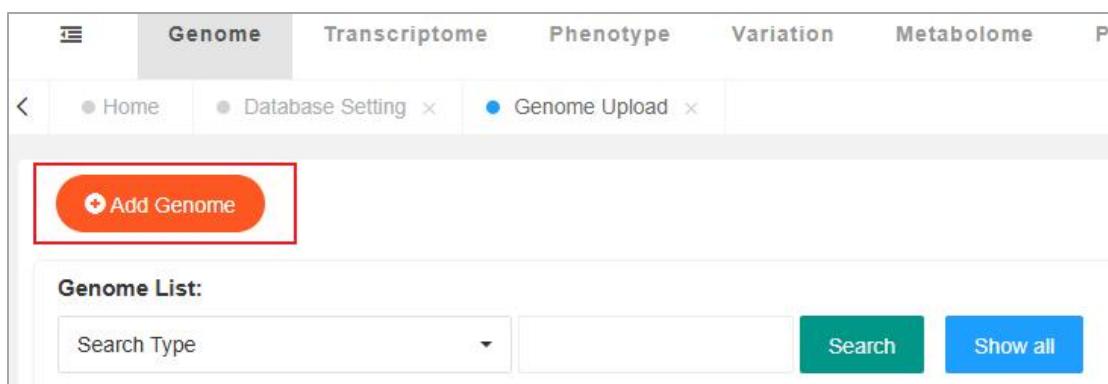
- Set Queue**:
 - *1 Set port**: A text input field labeled "Port".
 - *2 Status**: A blue button labeled "Start".
 - *3 Stop**: An orange button labeled "Stop".
- Set WGCNA Queue Num**:
 - WGCNA Queue Num**: A text input field containing the value "1".
 - Set**: A blue button.
- Set Queue Num**:
 - Run Queue Num**: A text input field containing the value "2".
 - Set**: A blue button.

Subsequently, the corresponding data can be uploaded. Currently, PlantMDMS supports the upload of **Genome**, **Transcriptome**, **Phenotype**, **Variation**, **Metabolome**, and **Proteome** data. Among all types of data, **the Genome data must be uploaded first**. There is no specific requirement for the upload order of the remaining data.

Click the "Genome" tab, then click "**Genome Upload**" on the left side:



Click "Add Genome":



After filling in the corresponding information, click "**Building database**":

Add Genome

*Category: Brassicaceae *Species: Arabidopsis

*Genome: Arabidopsis_t2t N50:

Variety description (not required)

***Genome(Name format:Genome_name_Genome.fasta)**

选择文件: Athaliana_Genome.fasta

***gff(Name format:Variety.gff)**

选择文件: Athaliana.gff

Building database

Once the upload is finished, the interface is shown below:

Genome List:					
Search Type	Search		Show all		
Category	Species	Variety	N50	Description	Delete Genome
Brassicaceae	Arabidopsis	Arabidopsis_t2t		Edit	Delete
< 1 > 到第 1 页 确定 共 1 条 10 条/页					

At this point, the interface will refresh automatically. If it fails to refresh automatically, you can manually refresh the interface. After refreshing, the name of the newly created database will appear in the Gene Annotation section.



Tips: If you do not complete the intended operations on the same day and the PlantMDMS initialization program has been closed (while the webpage remains open), you can restart the initialization program in this case. When it is in the "Running" state, simply refresh the webpage to continue with the unfinished work.

```
Command Running Stop
Database running: http://127.0.0.1:5013
Database Run Command:
D:
cd D:\PlantMDMS\PlantMDMS
D:/PlantMDMS/PlantMDMS/omic_flask2/python.exe main.py runserver
_____
Queue Run Command:
D:
cd D:\PlantMDMS\PlantMDMS
D:/PlantMDMS/PlantMDMS/omic_flask2/python.exe -m celery -A app.mycelery.main worker -n Queue_num -E -P eventlet
-l info --concurrency=2
_____
Starting Database
Database running: http://127.0.0.1:5013
```

Subsequently, select the annotation file and click the "**Upload**" button to upload it:



After the upload is completed, the files will be automatically classified according to different annotation types. Regarding the upload time, based on our test cases, a 50Mb annotation file takes approximately 5 seconds to upload. The upload time for other files can be estimated by the multiple relationship of their sizes. If the time taken significantly exceeds the estimated duration, you may refresh the page to check if the upload has been completed.

Upload Genome Annotation:			
GeneID	GO	KEGG	KEGG_pathway
AT1G01010	--	--	--
AT1G01020	--	K21848 lipid intermediate transporter	--
AT1G01030	GO:0000003 biological_process:obsolete reproduction;GO:0003006 bl...	--	--
AT1G01040	GO:0003674 molecular_function:molecular_function;GO:0003676 mole...	K11592 endonuclease Dicer [EC:3.1.26.-]	ko05206 MicroRNAs in cancer
AT1G01050	GO:0003674 molecular_function:molecular_function;GO:0003824 mole...	K01507:inorganic pyrophosphatase [EC:3.6.1.1]	ko00190 Oxidative phosphorylation
AT1G01060	GO:0000003 biological_process:obsolete reproduction;GO:0001067 m...	K12133 MYB-related transcription factor LHY;K12134 circadian clock associated 1	ko04712 Circadian rhythm - plant
AT1G01070	GO:0000003 biological_process:obsolete reproduction;GO:0003006 bl...	--	--
AT1G01073	--	--	--
AT1G01080	GO:0005575 cellular_component:cellular_component;GO:0005622:cell...	K13126 polyadenylate-binding protein	ko03013 Nucleocytoplasmic transp
AT1G01090	GO:0005575 cellular_component:cellular_component;GO:0005622:cell...	K00161 pyruvate dehydrogenase E1 component subunit alpha [EC:1.2.4.1]	ko00010 Glycolysis / Gluconeogeni

After the upload of genome and annotation data is completed, the next step is to process the transcription factors in accordance with the operation manual:

The screenshot shows the PlantMDMS web application. The main navigation bar includes links for Genome, Transcriptome, Phenotype, Variation, Metabolome, Proteome, and Design database. A sidebar on the left lists options: Genome Upload (radio button selected), TF Upload (radio button selected and highlighted with a red box), and Protein Feature... The main content area is titled 'TF Upload'. It contains several input fields: 'Reference Genome' with a dropdown menu 'Select Variety', 'Family number' with 'Select File' and 'Upload' buttons, 'Search Name' with a search button, 'Family members' with 'Select File' and 'Upload' buttons, 'Search ID' with a search button, 'Family' with a dropdown menu 'Select Family', and 'Family Tree' with 'Select File' and 'Member Annotation: Select File' buttons.

Select the background dataset:

Reference Genome: Arabidopsis_t2t

Upload the transcription factor count file. The format is as follows:

```

Alfin-like 10 Transcription Factors
AP2/ERF-AP2 17 Transcription Factors
AP2/ERF-ERF 148 Transcription Factors
AP2/ERF-RAV 5 Transcription Factors
ARID 13 Transcriptional Regulators
AUX/IAA 40 Transcriptional Regulators
B3 77 Transcription Factors
B3-ARF 37 Transcription Factors
BBR-BPC 17 Transcription Factors
BES1 14 Transcription Factors
bHLH 196 Transcription Factors

```

After selecting the file, click **Upload**:

Family number:	Select File	Upload
Search Name:	Search	
TF_TR_name	TF_TR_number	TF_TR_type
Alfin-like	10	Transcription Factors
AP2/ERF-AP2	17	Transcription Factors
AP2/ERF-ERF	148	Transcription Factors
AP2/ERF-RAV	5	Transcription Factors
ARID	13	Transcriptional Regulators
AUX/IAA	40	Transcriptional Regulators

Select Transcription factor gene family member file, click **Upload**:

Family members: [Select File](#) [Upload](#)

The file format is as follows:

TF_TR_ID	TF_TR_name	TF_TR_type	Sub_Family	Hom_ID	note
AT1G14510.1	Alfin-like	TF	--	--	--
AT2G02470.1	Alfin-like	TF	--	--	--
AT2G02470.2	Alfin-like	TF	--	--	--
AT3G11200.1	Alfin-like	TF	--	--	--
AT3G11200.2	Alfin-like	TF	--	--	--
AT3G42790.1	Alfin-like	TF	--	--	--

After the upload is completed, click on the family name to retrieve the IDs of the members of that family. For example:

TF_TR_ID	TF_TR_name	TF_TR_type	Sub_Family	Hom_ID	note	Operate
AT1G02065.1	SBP	TF	--	--	--	Edit Delete
AT1G02065.2	SBP	TF	--	--	--	Edit Delete
AT1G20980.1	SBP	TF	--	--	--	Edit Delete
AT1G27360.1	SBP	TF	--	--	--	Edit Delete
AT1G27360.2	SBP	TF	--	--	--	Edit Delete

You can click **Edit** to add certain information annotations to family members.

Id:AT1G02065.1 X

Sub_Family:
-+

Hom_ID:
--

note:
--

Edit

Next, you can upload the phylogenetic tree and subfamily information file to view the domains:

Family: BES1 select your interested family

Family Tree: Member Annotation:

Name	Type	Size	Operate
BES1.nwk	Family Tree	0.3kb	<input type="button" value="Delete"/>
BES1_subfamily.csv	Member Annotation	0.5kb	<input type="button" value="Delete"/>

Upload

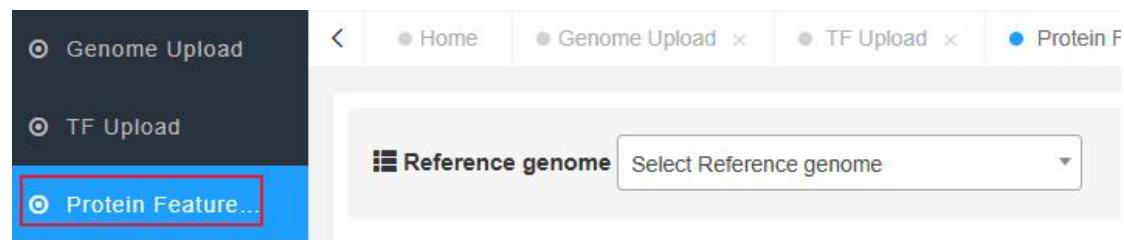
Existing trees:

Family	Operate
BES1	<input type="button" value="Delete"/>

The format of the phylogenetic tree is **nwk**. The format of the subfamily information file (**csv format**) is as follows:

ID	Species	Subfamily
AT1G19350.1	Arabidopsis thaliana	C
AT1G19350.3	Arabidopsis thaliana	C
AT1G19350.4	Arabidopsis thaliana	C
AT1G19350.5	Arabidopsis thaliana	C
AT1G19350.6	Arabidopsis thaliana	C
AT1G75080.1	Arabidopsis thaliana	C
AT1G75080.2	Arabidopsis thaliana	C
AT1G78700.1	Arabidopsis thaliana	B
AT2G45880.1	Arabidopsis thaliana	A
AT3G50750.1	Arabidopsis thaliana	C
AT4G18890.1	Arabidopsis thaliana	B
AT4G36780.1	Arabidopsis thaliana	A
AT5G45300.1	Arabidopsis thaliana	A
AT5G45300.2	Arabidopsis thaliana	A

The Protein Feature Upload module is used to upload protein sequence domains and 3D structures. Click on the module tab:



A screenshot of a web-based application interface. On the left, there is a vertical navigation bar with three items: 'Genome Upload', 'TF Upload', and 'Protein Feature...'. The 'Protein Feature...' item is highlighted with a blue background and a red border. To the right of the navigation bar, the main content area has a header with tabs: 'Home', 'Genome Upload', 'TF Upload', and 'Protein Feature...'. Below the header, there is a section titled 'Reference genome' with a dropdown menu labeled 'Select Reference genome'.

Select the genome:



A screenshot of a dropdown menu for selecting a reference genome. The menu is titled 'Reference genome' and shows a single option 'Arabidopsis_t2t' selected. There is a small downward arrow icon at the end of the dropdown.

Select to upload new data or update data (i.e., replace the original file). The protein structure file is in **csv format**, and its content includes:

Protein_id	Type	Domain	Description	Start	End	Color
AT1G1451	Families	domain1	CPG BINDI	9	249	blue
AT1G1451	domain	domain2	Alfin, N-tei	11	138	blue
AT1G1451	domain	domain3	FYVE/PHD	187	249	red
AT1G1451	Conserved	domain4	CDD - cd1	198	246	red

After that, you can upload the 3D structure file of the protein. Name the protein and click "**Upload**:



A screenshot of a user interface for uploading a protein 3D structure file. It features a text input field labeled 'Upload Protein 3D Structure File:' containing the text 'AT1G14510.1', a blue 'Select File' button, and a green 'Upload' button.

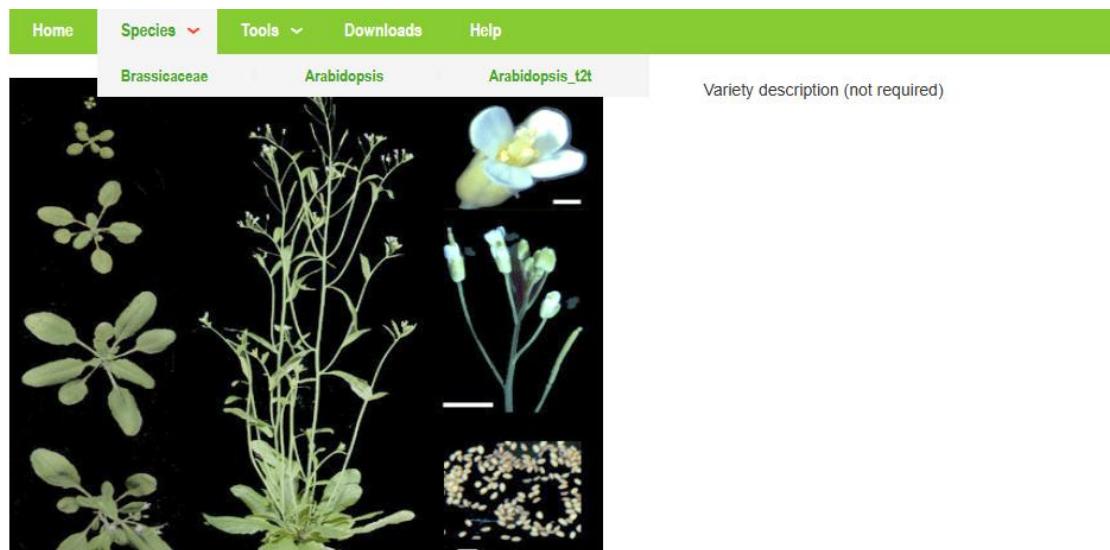
This adds the protein data file just mentioned to the database.

Operations on the front end.

After uploading data of different species and constructing the database, open the front end.



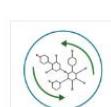
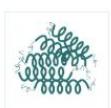
Click "Species" and select the species you wish to analyze.



Welcome to use Arabidopsis_t2t database

The database contains genomics, transcriptomics, population, variant, and metabolomics modules

Select from the following modules based on the content you wish to analyze. For example, if you want to view the protein structure of a specific gene uploaded in the Genome module, select the Genome module:

 <p>Genomics</p> <p>Gene structure, Pathways, Transcription factors, Gene families...</p> <p>Start</p>	 <p>Transcriptomics</p> <p>Sequencing quality, Gene expression patterns under different conditions</p> <p>Start</p>	 <p>Metabolomics</p> <p>Sequencing quality, Gene expression patterns under different conditions</p>
 <p>Proteomics</p> <p>Gene structure, Pathways, Transcription factors, Gene families...</p>	 <p>Phenotype</p> <p>Sequencing quality, Gene expression patterns under different conditions</p>	 <p>Population</p> <p>Sequencing quality, Gene expression patterns under different conditions</p>

Click "**Start**".

After entering the Genome interface:

Gene Search:

ID

Annotation: GO,KEGG,IPR and note

Structure:gene,mRNA,exon,CDS and UTR

Gene list

Gene Sequence:gene,CDS,protein,gene upstream and downstream

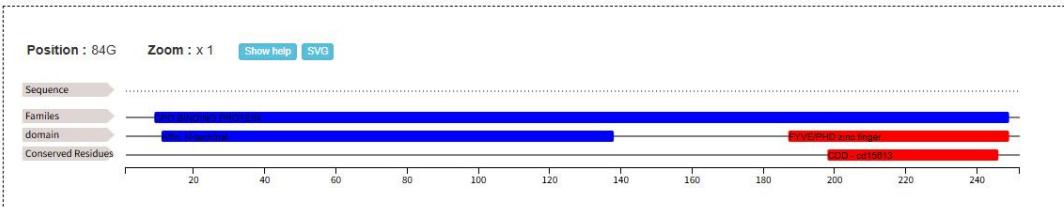
Gene CDS Protein

Upstream the default value is 0 Downstream the default value is 0

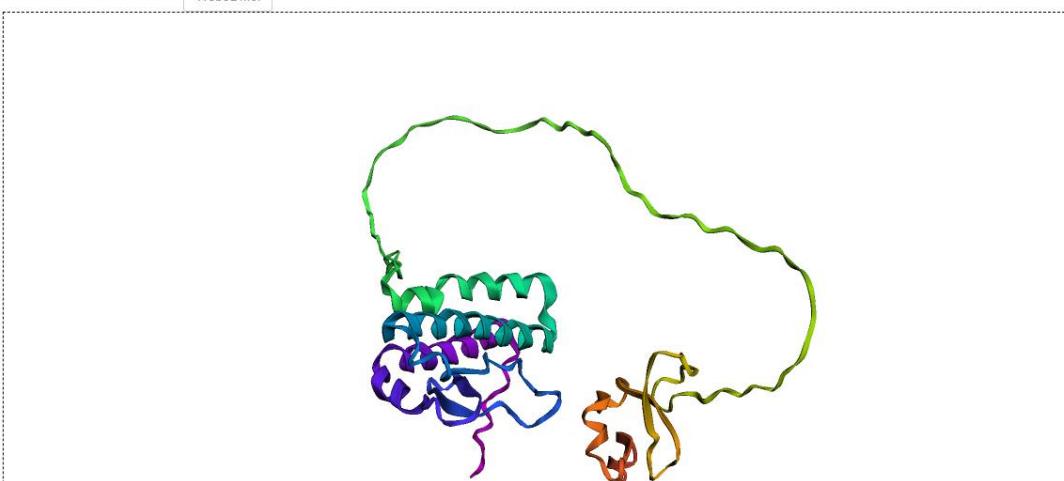
Start

Click "Protein Structure", select the gene of interest from the drop-down list, and click the gene name to display its protein structure.

Protein feature viewer:



Protein 3d viewer:



Next, we proceed with the back-end data upload process. After the genome data, follow the upload of various omics data. First, upload the transcriptome data: switch to the corresponding module and select the reference genome:

Name the data (Naming Rules: Only letters, numbers, and underscores are allowed; Chinese characters and spaces are prohibited). According to the characteristics of the uploaded data, select "FPKM" if the main data type is FPKM, select "count" if it is count, and so on. Then, select and upload the file.

*Project Name	abc	FPKM	Select File	Upload
----------------------	-----	------	-------------	--------

After the upload is completed, the uploaded content will be displayed below.

Upload Sample Information: Select File Upload											
Search: Project Name Filter Show all											
Project_Name	Sample_Group	Sample_Name	Replicate	FPKM	Count	TPM	CPM	Title	Sample_ID	Instrument	Operate
abc		SRR10991444		TRUE							Edit Delete
abc		SRR10991445		TRUE							Edit Delete
abc		SRR10991446		TRUE							Edit Delete
abc		SRR10991447		TRUE							Edit Delete
abc		SRR10991448		TRUE							Edit Delete
abc		SRR10991449		TRUE							Edit Delete

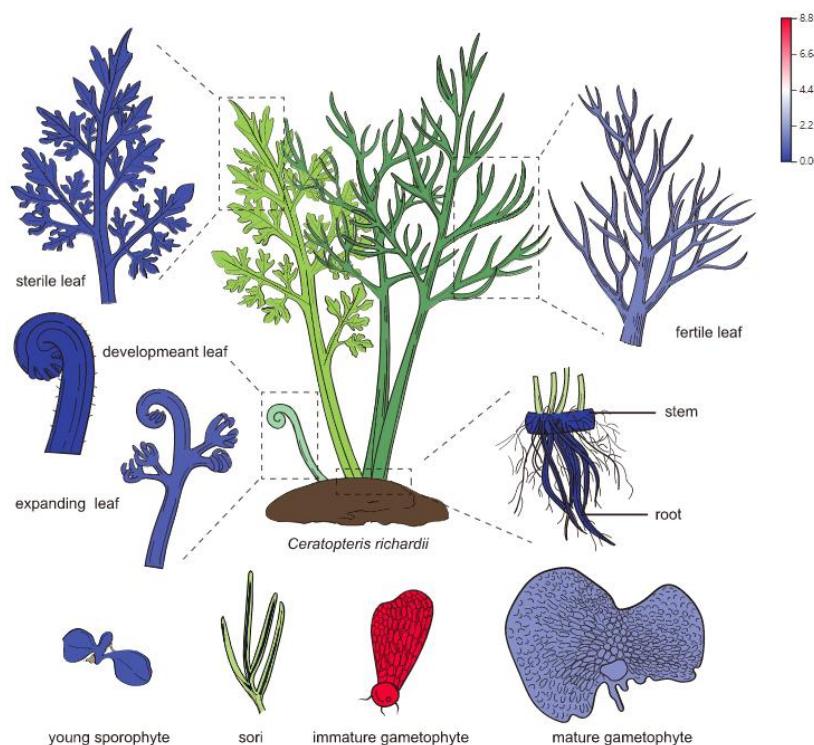
Select "**Select File**" to upload annotation information related to the transcriptome, including and limited to: Project_Name, Sample_Group, Sample_Name, Replicate, Title, Sample_ID, Instrument, RNA_Seq_Number, Tissue, Sample_Information, Germplasm, Germplasm_Characteristics, Geographic_Location, Data_Sources, Note.

Note that users may construct the database with data from different species or tissues. Therefore, it is necessary to strictly define the data composition during upload to prevent data confusion. Here, the format of the uploaded file must be CSV, and the headers must be the above keywords—their order cannot be changed and none can be missing. For items with no corresponding values, leave them blank, but the headers must not be omitted. Meanwhile, Project_Name must be consistent with the "Project Name" on this page.

After successful upload, the corresponding data in the file will be automatically populated into the database as shown in the figure below.

Project_Name	Sample_Group	Sample_Name	Replicate	FPKM	Count	TPM	CPM	Title	Sample_ID	Instrument	Operate
abc	Mito_col	SRR10991444	biologica...	TRUE				Mito_col_1			Edit Delete
abc	Mito_col	SRR10991445	biologica...	TRUE				Mito_col_2			Edit Delete
abc	Mito_col	SRR10991446	biologica...	TRUE				Mito_col_3			Edit Delete
abc	Mito_otp90	SRR10991447	biologica...	TRUE				Mito_otp90_1			Edit Delete
abc	Mito_otp90	SRR10991448	biologica...	TRUE				Mito_otp90_2			Edit Delete

eFP is mainly used for uploading cartoons related to tissues and organs. Examples of eFP are as follows:



In the back-end of this section, the main task is to establish the corresponding relationship between colors and samples, such as:

 #000000	SRR10991444	▼
 #9bc82e	SRR10991445	▼
 #040000	SRR10991446	▼
 #58b760	SRR10991448	▼
 #40b8a9	SRR10991445	▼
 #d2c12e	SRR10991447	▼

After establishing the corresponding relationship, click "**Setting finished**". Then set the color gradient for measuring expression abundance.

Expression Legend:

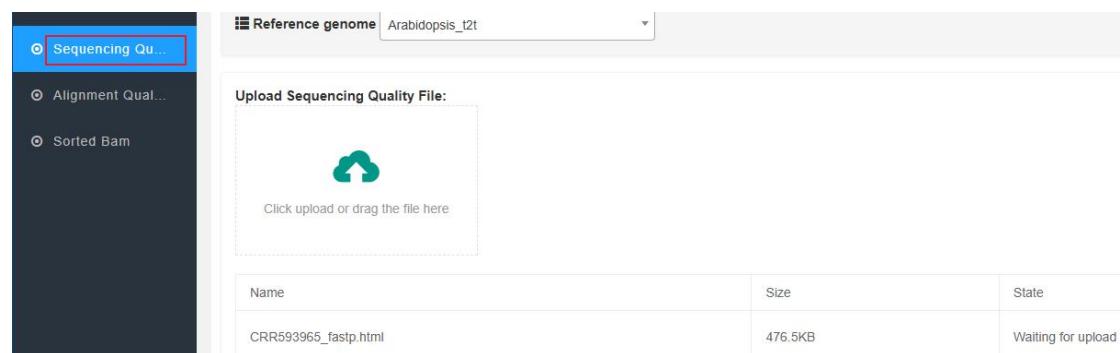
Low:  Middle:  High: 

Use the following settings to specify the size and position of the legend:

Expression Legend Show:

Position: Size: Display style: Horizontal Vertical

After completing the settings, click "**Legend_show**" to display the legend on the figure for flexible adjustment. Once the adjustment is finished, click "**Update SVG**" to save the legend.



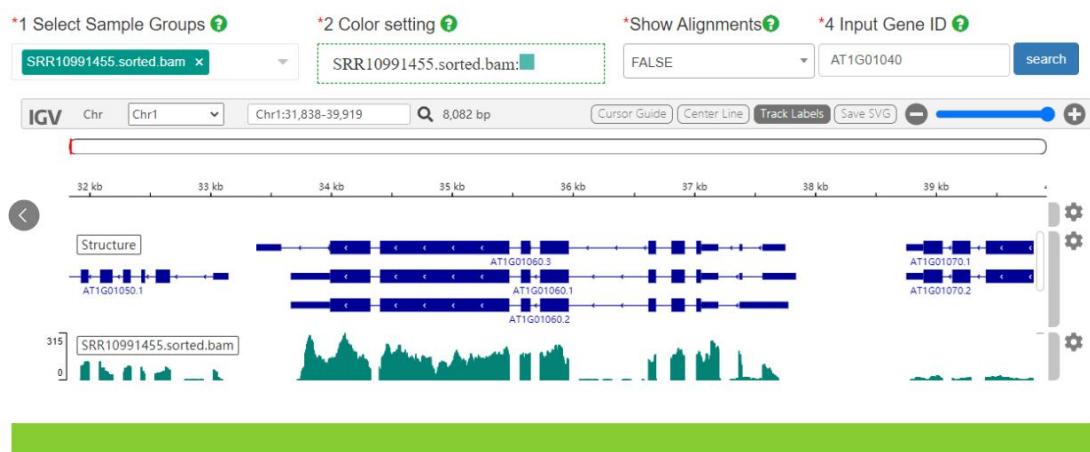
Name	Size	State
CRR593965_fastp.html	476.5KB	Waiting for upload

There are strict requirements for the file extension, which must be _fastp.html (refer to the sample file for the specific format); in addition, you can upload the alignment quality file via the Alignment Quality module—this file must be generated by the HiSAT2 alignment program, with the basic process being aligning the transcriptome sequencing data to the reference genome, which will produce a SAM file and a corresponding alignment quality report simultaneously—and then name this report with the suffix _hisat2_summary.txt before uploading it to this module.

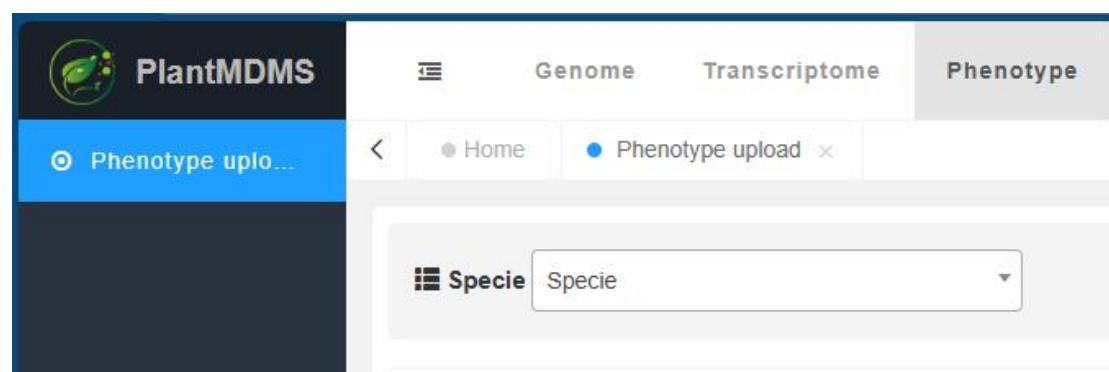
Next, upload the alignment file itself via Sorted Bam: the result file needs to be sorted, and in addition, a .bai file must be generated using tabix software; after naming these two files with the suffixes .sorted.bam and .sorted.bam.bai respectively, upload them to the database separately.

The analysis graphs shown below can be generated on the front end using the uploaded data.

Gene Search:



Regarding the upload of phenotypic data:



After selecting the species, name the project (following the same naming rules as above); once the naming is completed, you can upload the phenotypic data. It should be emphasized here that the phenotypic data must meet the following specific format requirements:

Trait_replace_number	Trait_name	Unit	Tissue	Description	13_24	13_26	13_49	13_60	13_64	14_2	14_5	14_6	14_10
frut_weight_1	frut_weight	mm	frut		73.66	78.01	63	72.91	80.12	73.08	70.55	82.05	65.15
frut_weight_2	frut_weight	mm	frut		78.26	85.81	68.53	85.67	82.07	81.19	73.84	88.96	65.13
frut_weight_3	frut_weight	mm	frut		74.84	78.68	60.66	78.52	81.95	71.06	80.67	82.64	64.2
frut_weight_4	frut_weight	mm	frut		71.01	80.99	57.49	68.23	79.2	66.45	71.27	76.7	64.95
frut_weight_5	frut_weight	mm	frut		67.18	86.05	59.36	73.65	78.24	82.39	77.17	71.74	65.49
frut_weight_6	frut_weight	mm	frut		61.41	85.16	49.49	75.05	83.13	83.48	76.26	63.6	66.22
frut_weight_7	frut_weight	mm	frut		65.37	84.31	61.67	73.83	83.44	80.09	77.31	86.67	64.56
frut_weight_8	frut_weight	mm	frut		75.57	85.23	59.86	80.82	82.17	78.17	76.18	75.34	67.47

To standardize data and prevent runtime errors, the headers of the first five columns in this file must not be modified in any way, and no data is allowed to be missing—the headers of the first five columns are respectively: Trait_replace_number, Trait_name, Unit, Tissue, Description. Following these columns are the IDs of different individuals (similarly, Chinese characters and spaces are prohibited, while letters, numbers, and underscores are allowed), and different values are used under each individual to represent quantitative traits.

It should be noted here that in addition to quantitative traits, there is also a qualitative trait; when representing a qualitative trait, "**quality**" must be indicated in the Unit column. For example:

water_content_1	water_content	mm	frut	
water_content_2	water_content	mm	frut	
water_content_3	water_content	mm	frut	
petal_type	petal_type	quality	flower	1:a;2:b;3:c;4:d

In different samples, a fixed value is used to represent a single trait—for example, for traits of "blooming red flowers" and "blooming yellow flowers", set 1 to represent red flowers and 2 to represent yellow flowers. There are no specific regulations on the exact values chosen, but the same value must be consistently used for the same trait.

Next, you need to upload the sample information file, in which:

Project_Name	Sample_Group	Sample_Name	Cultivar	Geographic_Location
Project_1		1 13_24	name	83.438354,43.534702
Project_1		1 13_26	name	83.438354,43.534703
Project_1		1 13_49	name	83.438354,43.534704
Project_1		2 13_60	name	83.438354,43.534705
Project_1		2 13_64	name	83.438354,43.534706
Project_1		2 14_2	name	83.438354,43.534707

Project_Name must be consistent with the name set in this module; Sample_Group represents grouping information, where samples in the same group share the same number; Sample_Name follows the same naming rules as above; Cultivar refers to the specific variety name; Geographic_location indicates the longitude and latitude of sample collection (longitude first, then latitude), which must be entered in English with a comma separating the two values.

From the above process, it is believed that everyone has gradually understood the basic logic of database upload. For the following modules (including Resequence, Metabolome, Proteome, and Design database), we will focus on introducing the specific format requirements for uploaded data; as for the upload operation rules, they are the same as those mentioned earlier, and you can follow the previous steps.

Resequence-Accessions

Resequencing Sample Information Table: The format is as follows:

Sample_Name	Sample_Group	Germplasm	Geographic_Location	Note
88 a		Germplasm2	83.438354,43.534702	
108 a		Germplasm2	83.438354,43.534702	
139 a		Germplasm2	83.438354,43.534702	
159 a		Germplasm2	83.438354,43.534702	
265 a		Germplasm2	83.438354,43.534702	
350 a		Germplasm2	83.438354,43.534702	
351 a		Germplasm2	83.438354,43.534702	
403 a		Germplasm2	83.438354,43.534702	
410 a		Germplasm1	83.487713,43.52555	

The headers include: Sample_Name, Sample_Group, Germplasm, Geographic_Location, and Note.

They respectively represent the sample name, the group to which the sample belongs, the type of germplasm resources, longitude and latitude, and precautions.

Resequence-Population-PCA

Sample	PCA1	PCA2	PCA3
88	0.140068	-0.0838189	-0.0838189
108	0.140619	-0.0852322	-0.0852322
139	0.139991	-0.0838636	-0.0838636
159	0.139906	-0.0844351	-0.0844351
265	0.140726	-0.0859495	-0.0859495
350	0.14084	-0.0859659	-0.0859659
351	0.140838	-0.0852597	-0.0852597
403	0.140826	-0.0853783	-0.0853783
410	0.140879	-0.0852168	-0.0852168

The sample names in PCA must correspond to those in the Resequencing Sample Information Table and must not exceed the number of samples in the table—specifically, the samples in PCA should be a subset of the samples listed in the Resequencing Sample Information Table.

Resequence-Population-Structure

Similarly, for the structural data here:

Sample	a	b	c	d	e
88	0.00001	0.000017	0.999949	0.000014	0.00001
108	0.00001	0.000017	0.999953	0.00001	0.00001
139	0.000017	0.00001	0.999945	0.000018	0.00001
159	0.00001	0.000016	0.99995	0.000014	0.00001
265	0.00001	0.000017	0.999953	0.00001	0.00001
350	0.000011	0.000017	0.999952	0.00001	0.00001
351	0.000014	0.000017	0.999949	0.00001	0.00001
403	0.00001	0.000017	0.999945	0.000018	0.00001
410	0.00001	0.000017	0.999953	0.00001	0.00001
424	0.000013	0.000017	0.99995	0.00001	0.00001
428	0.00001	0.000017	0.999953	0.00001	0.00001

The sample names must correspond to those in Resequence-Accessions, and the column names must correspond to Sample_Group (note: corrected the possible typo "Sample_Group" to the standard "Sample_Group" as per previous context).

Resequence-Population-Structure Tree

A phylogenetic tree can be constructed based on population SNPs, and the tree must be in nwk format.

Resequence-Variation-vcf

This module requires the VCF file generated during the resequencing process; it is essential to ensure that the sequence IDs in the VCF file (mostly chromosome sequence IDs) are consistent with those in the genome and GFF files uploaded to this system. If there is an inconsistency, you can use the following command to correct the IDs: **bctools.exe annotate --rename-chrs chr_name_change.txt [VCF file] -Oz -o [output result]**. The chr_name_change file contains two columns separated by a tab: the first column is the ID from the VCF file, and the second column is the corresponding chromosome ID of the genome.

Resequence-Variation-tbi

The tbi file can be generated using the tabix program with the command line: **tabix -p vcf input.vcf.gz**.

Resequence-GWAS_Result

After obtaining the original GWAS file, you need to extract information from it and organize it into the format required by this program—the format includes five columns with fixed headers: CHR, BP, Ref, Var, and p_value. An example is as follows:

CHR	BP	Ref	Var	p_value
Chr1	4654	G	A	7.109694e-01
Chr1	29345	A	G	2.807430e-01
Chr1	57383	A	T	4.068318e-01
Chr1	57686	C	A	4.623254e-01
Chr1	59016	G	C	4.732450e-01

Resequence-FST_Result

The basic format of this file is as follows:

CHROM	POS	WEIR_AND_COCKERHAM_FST
Chr1	2338	0.121264
Chr1	2449	0.569683
Chr1	2529	0.184219
Chr1	2551	0.0330317
Chr1	2563	0.191599
Chr1	2565	0.191599

Among them, the third column can be obtained by calculating the FST selection signal using the vcftools program.

Metabolome-Metabolome upload-metabolome data

This section is mainly used for uploading metabolomic data. The first six columns are respectively Project_Index, Compounds, Class I, Class II, Formula, and Molecular weight (Da). Similarly, the column names and their expressions must not be modified to avoid program recognition issues. The example format of these six columns is as follows:

Project_Index	Compounds	Class I	Class II	Formula	Molecular weight (Da)
Cmcp004764	16R-Sitsirkine	Alkaloids	Plumerane	C21H26N2O3	354.1943
Cmhp003670	Secologanin	Terpenoids	Monoterpenoids	C17H24O10	388.1369
Cmsp002787	3,4'-Dihydroxy-3',5' Others		Ketone compounds	C11H14O5	226.0841
Cmvp005062	Salireposide	Phenolic acids	Phenolic acids	C20H22O9	406.1264
Cmvp005063	Grandidentatin	Phenolic acids	Phenolic acids	C21H28O9	424.1733
HJAP005	Laricitrin-3-O-gluc	Flavonoids	Flavonols	C22H22O13	494.106

In addition to the six columns mentioned above, the subsequent columns contain the specific metabolic values of each sample obtained from the metabolome. For example:

CK_3d_T_1	CK_3d_T_2	CK_3d_T_3	CK_3d_S_1	CK_3d_S_2
24516.252	27589.788	74280.739	24535.665	44438.724
409102.902	327840.4	172440.2983	92416.89735	108292.932
128417.38	66392.627	93696.66	146649.6554	206911.488
38510.5	58815.4344	42918.096	89405.96933	78788.484
154769.17	189930.7575	151128.792	171967.84	195538.4587
51346.412	61865.894	93805.8048	55832.736	46785.904
56785.18	146407.515	73362.336	175031.4195	165716.024

Metabolome-Metabolome upload-sample information

This file contains the following columns: Project_Name, Sample_Group, Sample_Name, Replicate, Sample_Information, Germplasm,

Germplasm_Characteristics, Instrument, Title, Tissue, Geographic_Location, Data_Sources, and Note. As mentioned earlier, the column names must not be modified, and Project_Name must be consistent with the Project_Name set in this module.

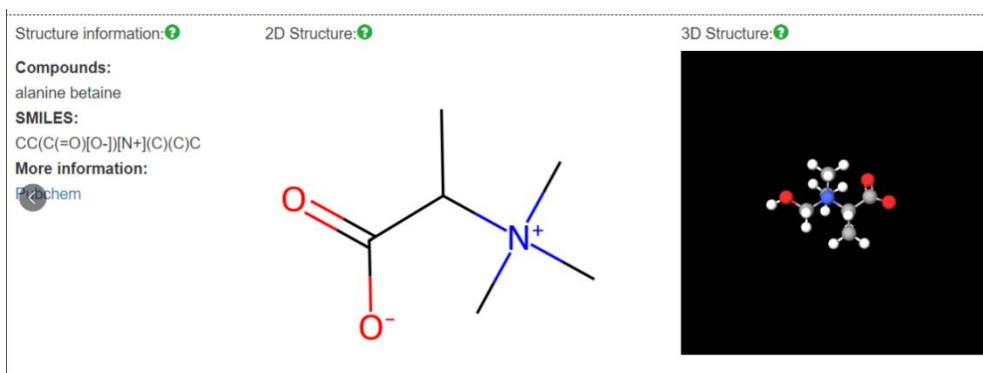
The screenshot shows a user interface for managing metabolome data. At the top, there is a dropdown menu labeled "Specie" with "Specie" selected. Below it, a red-highlighted field is labeled "*Project Name". Further down, there are two sections for file uploads: "Upload metabolome data:" with "Select File" and "Upload" buttons, and "Upload sample information:" with "Select File" and "Upload" buttons. At the bottom, there is a search bar labeled "Search: Project Name" with "Filter" and "Show all" buttons.

Metabolome-Compounds information

The uploaded file includes three columns in total: the first column is the component; the second column is the website hyperlink (official website: <https://pubchem.ncbi.nlm.nih.gov>); you need to query the corresponding chemical formula of the component through this homepage and fill it in as the third column. An example is as follows:

Compounds		
16k-Sitirkine	https://pubchem.ncbi.nlm.nih.gov/compound/3050539	COC(=O)[C@H](CO)[C@H](C[C@H]2C=C(CCN2C)C@H)C4=CC=CC=C4N3
Secologanin	https://pubchem.ncbi.nlm.nih.gov/compound/161276	COC(=O)C1=C(O)C2=C1[C@H](C[C@H]1CC=O)C=O[C@H]2C[C@H]1C[C@H](C[C@H](O)C)O)O
3,4'-Dihydroxy-3',5'-diethoxypropiophenone	https://pubchem.ncbi.nlm.nih.gov/compound/54353927	CCOC1=CC=C(C=C1O)C2=C(O)C3=C(O)C4=C(O)C5=C(O)C6=C(O)C7=C(O)C8=C(O)C9=C(O)C10=C(O)C11=C(O)C12=C(O)C13=C(O)C14=C(O)C15=C(O)C16=C(O)C17=C(O)C18=C(O)C19=C(O)C20=C(O)C21=C(O)C22=C(O)C23=C(O)C24=C(O)C25=C(O)C26=C(O)C27=C(O)C28=C(O)C29=C(O)C30=C(O)C31=C(O)C32=C(O)C33=C(O)C34=C(O)C35=C(O)C36=C(O)C37=C(O)C38=C(O)C39=C(O)C40=C(O)C41=C(O)C42=C(O)C43=C(O)C44=C(O)C45=C(O)C46=C(O)C47=C(O)C48=C(O)C49=C(O)C50=C(O)C51=C(O)C52=C(O)C53=C(O)C54=C(O)C55=C(O)C56=C(O)C57=C(O)C58=C(O)C59=C(O)C60=C(O)C61=C(O)C62=C(O)C63=C(O)C64=C(O)C65=C(O)C66=C(O)C67=C(O)C68=C(O)C69=C(O)C70=C(O)C71=C(O)C72=C(O)C73=C(O)C74=C(O)C75=C(O)C76=C(O)C77=C(O)C78=C(O)C79=C(O)C80=C(O)C81=C(O)C82=C(O)C83=C(O)C84=C(O)C85=C(O)C86=C(O)C87=C(O)C88=C(O)C89=C(O)C90=C(O)C91=C(O)C92=C(O)C93=C(O)C94=C(O)C95=C(O)C96=C(O)C97=C(O)C98=C(O)C99=C(O)C100=C(O)C101=C(O)C102=C(O)C103=C(O)C104=C(O)C105=C(O)C106=C(O)C107=C(O)C108=C(O)C109=C(O)C110=C(O)C111=C(O)C112=C(O)C113=C(O)C114=C(O)C115=C(O)C116=C(O)C117=C(O)C118=C(O)C119=C(O)C120=C(O)C121=C(O)C122=C(O)C123=C(O)C124=C(O)C125=C(O)C126=C(O)C127=C(O)C128=C(O)C129=C(O)C130=C(O)C131=C(O)C132=C(O)C133=C(O)C134=C(O)C135=C(O)C136=C(O)C137=C(O)C138=C(O)C139=C(O)C140=C(O)C141=C(O)C142=C(O)C143=C(O)C144=C(O)C145=C(O)C146=C(O)C147=C(O)C148=C(O)C149=C(O)C150=C(O)C151=C(O)C152=C(O)C153=C(O)C154=C(O)C155=C(O)C156=C(O)C157=C(O)C158=C(O)C159=C(O)C160=C(O)C161=C(O)C162=C(O)C163=C(O)C164=C(O)C165=C(O)C166=C(O)C167=C(O)C168=C(O)C169=C(O)C170=C(O)C171=C(O)C172=C(O)C173=C(O)C174=C(O)C175=C(O)C176=C(O)C177=C(O)C178=C(O)C179=C(O)C180=C(O)C181=C(O)C182=C(O)C183=C(O)C184=C(O)C185=C(O)C186=C(O)C187=C(O)C188=C(O)C189=C(O)C190=C(O)C191=C(O)C192=C(O)C193=C(O)C194=C(O)C195=C(O)C196=C(O)C197=C(O)C198=C(O)C199=C(O)C200=C(O)C201=C(O)C202=C(O)C203=C(O)C204=C(O)C205=C(O)C206=C(O)C207=C(O)C208=C(O)C209=C(O)C210=C(O)C211=C(O)C212=C(O)C213=C(O)C214=C(O)C215=C(O)C216=C(O)C217=C(O)C218=C(O)C219=C(O)C220=C(O)C221=C(O)C222=C(O)C223=C(O)C224=C(O)C225=C(O)C226=C(O)C227=C(O)C228=C(O)C229=C(O)C230=C(O)C231=C(O)C232=C(O)C233=C(O)C234=C(O)C235=C(O)C236=C(O)C237=C(O)C238=C(O)C239=C(O)C240=C(O)C241=C(O)C242=C(O)C243=C(O)C244=C(O)C245=C(O)C246=C(O)C247=C(O)C248=C(O)C249=C(O)C250=C(O)C251=C(O)C252=C(O)C253=C(O)C254=C(O)C255=C(O)C256=C(O)C257=C(O)C258=C(O)C259=C(O)C260=C(O)C261=C(O)C262=C(O)C263=C(O)C264=C(O)C265=C(O)C266=C(O)C267=C(O)C268=C(O)C269=C(O)C270=C(O)C271=C(O)C272=C(O)C273=C(O)C274=C(O)C275=C(O)C276=C(O)C277=C(O)C278=C(O)C279=C(O)C280=C(O)C281=C(O)C282=C(O)C283=C(O)C284=C(O)C285=C(O)C286=C(O)C287=C(O)C288=C(O)C289=C(O)C290=C(O)C291=C(O)C292=C(O)C293=C(O)C294=C(O)C295=C(O)C296=C(O)C297=C(O)C298=C(O)C299=C(O)C300=C(O)C301=C(O)C302=C(O)C303=C(O)C304=C(O)C305=C(O)C306=C(O)C307=C(O)C308=C(O)C309=C(O)C310=C(O)C311=C(O)C312=C(O)C313=C(O)C314=C(O)C315=C(O)C316=C(O)C317=C(O)C318=C(O)C319=C(O)C320=C(O)C321=C(O)C322=C(O)C323=C(O)C324=C(O)C325=C(O)C326=C(O)C327=C(O)C328=C(O)C329=C(O)C330=C(O)C331=C(O)C332=C(O)C333=C(O)C334=C(O)C335=C(O)C336=C(O)C337=C(O)C338=C(O)C339=C(O)C340=C(O)C341=C(O)C342=C(O)C343=C(O)C344=C(O)C345=C(O)C346=C(O)C347=C(O)C348=C(O)C349=C(O)C350=C(O)C351=C(O)C352=C(O)C353=C(O)C354=C(O)C355=C(O)C356=C(O)C357=C(O)C358=C(O)C359=C(O)C360=C(O)C361=C(O)C362=C(O)C363=C(O)C364=C(O)C365=C(O)C366=C(O)C367=C(O)C368=C(O)C369=C(O)C370=C(O)C371=C(O)C372=C(O)C373=C(O)C374=C(O)C375=C(O)C376=C(O)C377=C(O)C378=C(O)C379=C(O)C380=C(O)C381=C(O)C382=C(O)C383=C(O)C384=C(O)C385=C(O)C386=C(O)C387=C(O)C388=C(O)C389=C(O)C390=C(O)C391=C(O)C392=C(O)C393=C(O)C394=C(O)C395=C(O)C396=C(O)C397=C(O)C398=C(O)C399=C(O)C400=C(O)C401=C(O)C402=C(O)C403=C(O)C404=C(O)C405=C(O)C406=C(O)C407=C(O)C408=C(O)C409=C(O)C410=C(O)C411=C(O)C412=C(O)C413=C(O)C414=C(O)C415=C(O)C416=C(O)C417=C(O)C418=C(O)C419=C(O)C420=C(O)C421=C(O)C422=C(O)C423=C(O)C424=C(O)C425=C(O)C426=C(O)C427=C(O)C428=C(O)C429=C(O)C430=C(O)C431=C(O)C432=C(O)C433=C(O)C434=C(O)C435=C(O)C436=C(O)C437=C(O)C438=C(O)C439=C(O)C440=C(O)C441=C(O)C442=C(O)C443=C(O)C444=C(O)C445=C(O)C446=C(O)C447=C(O)C448=C(O)C449=C(O)C450=C(O)C451=C(O)C452=C(O)C453=C(O)C454=C(O)C455=C(O)C456=C(O)C457=C(O)C458=C(O)C459=C(O)C460=C(O)C461=C(O)C462=C(O)C463=C(O)C464=C(O)C465=C(O)C466=C(O)C467=C(O)C468=C(O)C469=C(O)C470=C(O)C471=C(O)C472=C(O)C473=C(O)C474=C(O)C475=C(O)C476=C(O)C477=C(O)C478=C(O)C479=C(O)C480=C(O)C481=C(O)C482=C(O)C483=C(O)C484=C(O)C485=C(O)C486=C(O)C487=C(O)C488=C(O)C489=C(O)C490=C(O)C491=C(O)C492=C(O)C493=C(O)C494=C(O)C495=C(O)C496=C(O)C497=C(O)C498=C(O)C499=C(O)C500=C(O)C501=C(O)C502=C(O)C503=C(O)C504=C(O)C505=C(O)C506=C(O)C507=C(O)C508=C(O)C509=C(O)C510=C(O)C511=C(O)C512=C(O)C513=C(O)C514=C(O)C515=C(O)C516=C(O)C517=C(O)C518=C(O)C519=C(O)C520=C(O)C521=C(O)C522=C(O)C523=C(O)C524=C(O)C525=C(O)C526=C(O)C527=C(O)C528=C(O)C529=C(O)C530=C(O)C531=C(O)C532=C(O)C533=C(O)C534=C(O)C535=C(O)C536=C(O)C537=C(O)C538=C(O)C539=C(O)C540=C(O)C541=C(O)C542=C(O)C543=C(O)C544=C(O)C545=C(O)C546=C(O)C547=C(O)C548=C(O)C549=C(O)C550=C(O)C551=C(O)C552=C(O)C553=C(O)C554=C(O)C555=C(O)C556=C(O)C557=C(O)C558=C(O)C559=C(O)C560=C(O)C561=C(O)C562=C(O)C563=C(O)C564=C(O)C565=C(O)C566=C(O)C567=C(O)C568=C(O)C569=C(O)C570=C(O)C571=C(O)C572=C(O)C573=C(O)C574=C(O)C575=C(O)C576=C(O)C577=C(O)C578=C(O)C579=C(O)C580=C(O)C581=C(O)C582=C(O)C583=C(O)C584=C(O)C585=C(O)C586=C(O)C587=C(O)C588=C(O)C589=C(O)C590=C(O)C591=C(O)C592=C(O)C593=C(O)C594=C(O)C595=C(O)C596=C(O)C597=C(O)C598=C(O)C599=C(O)C600=C(O)C601=C(O)C602=C(O)C603=C(O)C604=C(O)C605=C(O)C606=C(O)C607=C(O)C608=C(O)C609=C(O)C610=C(O)C611=C(O)C612=C(O)C613=C(O)C614=C(O)C615=C(O)C616=C(O)C617=C(O)C618=C(O)C619=C(O)C620=C(O)C621=C(O)C622=C(O)C623=C(O)C624=C(O)C625=C(O)C626=C(O)C627=C(O)C628=C(O)C629=C(O)C630=C(O)C631=C(O)C632=C(O)C633=C(O)C634=C(O)C635=C(O)C636=C(O)C637=C(O)C638=C(O)C639=C(O)C640=C(O)C641=C(O)C642=C(O)C643=C(O)C644=C(O)C645=C(O)C646=C(O)C647=C(O)C648=C(O)C649=C(O)C650=C(O)C651=C(O)C652=C(O)C653=C(O)C654=C(O)C655=C(O)C656=C(O)C657=C(O)C658=C(O)C659=C(O)C660=C(O)C661=C(O)C662=C(O)C663=C(O)C664=C(O)C665=C(O)C666=C(O)C667=C(O)C668=C(O)C669=C(O)C670=C(O)C671=C(O)C672=C(O)C673=C(O)C674=C(O)C675=C(O)C676=C(O)C677=C(O)C678=C(O)C679=C(O)C680=C(O)C681=C(O)C682=C(O)C683=C(O)C684=C(O)C685=C(O)C686=C(O)C687=C(O)C688=C(O)C689=C(O)C690=C(O)C691=C(O)C692=C(O)C693=C(O)C694=C(O)C695=C(O)C696=C(O)C697=C(O)C698=C(O)C699=C(O)C700=C(O)C701=C(O)C702=C(O)C703=C(O)C704=C(O)C705=C(O)C706=C(O)C707=C(O)C708=C(O)C709=C(O)C710=C(O)C711=C(O)C712=C(O)C713=C(O)C714=C(O)C715=C(O)C716=C(O)C717=C(O)C718=C(O)C719=C(O)C720=C(O)C721=C(O)C722=C(O)C723=C(O)C724=C(O)C725=C(O)C726=C(O)C727=C(O)C728=C(O)C729=C(O)C730=C(O)C731=C(O)C732=C(O)C733=C(O)C734=C(O)C735=C(O)C736=C(O)C737=C(O)C738=C(O)C739=C(O)C740=C(O)C741=C(O)C742=C(O)C743=C(O)C744=C(O)C745=C(O)C746=C(O)C747=C(O)C748=C(O)C749=C(O)C750=C(O)C751=C(O)C752=C(O)C753=C(O)C754=C(O)C755=C(O)C756=C(O)C757=C(O)C758=C(O)C759=C(O)C760=C(O)C761=C(O)C762=C(O)C763=C(O)C764=C(O)C765=C(O)C766=C(O)C767=C(O)C768=C(O)C769=C(O)C770=C(O)C771=C(O)C772=C(O)C773=C(O)C774=C(O)C775=C(O)C776=C(O)C777=C(O)C778=C(O)C779=C(O)C780=C(O)C781=C(O)C782=C(O)C783=C(O)C784=C(O)C785=C(O)C786=C(O)C787=C(O)C788=C(O)C789=C(O)C790=C(O)C791=C(O)C792=C(O)C793=C(O)C794=C(O)C795=C(O)C796=C(O)C797=C(O)C798=C(O)C799=C(O)C800=C(O)C801=C(O)C802=C(O)C803=C(O)C804=C(O)C805=C(O)C806=C(O)C807=C(O)C808=C(O)C809=C(O)C810=C(O)C811=C(O)C812=C(O)C813=C(O)C814=C(O)C815=C(O)C816=C(O)C817=C(O)C818=C(O)C819=C(O)C820=C(O)C821=C(O)C822=C(O)C823=C(O)C824=C(O)C825=C(O)C826=C(O)C827=C(O)C828=C(O)C829=C(O)C830=C(O)C831=C(O)C832=C(O)C833=C(O)C834=C(O)C835=C(O)C836=C(O)C837=C(O)C838=C(O)C839=C(O)C840=C(O)C841=C(O)C842=C(O)C843=C(O)C844=C(O)C845=C(O)C846=C(O)C847=C(O)C848=C(O)C849=C(O)C850=C(O)C851=C(O)C852=C(O)C853=C(O)C854=C(O)C855=C(O)C856=C(O)C857=C(O)C858=C(O)C859=C(O)C860=C(O)C861=C(O)C862=C(O)C863=C(O)C864=C(O)C865=C(O)C866=C(O)C867=C(O)C868=C(O)C869=C(O)C870=C(O)C871=C(O)C872=C(O)C873=C(O)C874=C(O)C875=C(O)C876=C(O)C877=C(O)C878=C(O)C879=C(O)C880=C(O)C881=C(O)C882=C(O)C883=C(O)C884=C(O)C885=C(O)C886=C(O)C887=C(O)C888=C(O)C889=C(O)C890=C(O)C891=C(O)C892=C(O)C893=C(O)C894=C(O)C895=C(O)C896=C(O)C897=C(O)C898=C(O)C899=C(O)C900=C(O)C901=C(O)C902=C(O)C903=C(O)C904=C(O)C905=C(O)C906=C(O)C907=C(O)C908=C(O)C909=C(O)C910=C(O)C911=C(O)C912=C(O)C913=C(O)C914=C(O)C915=C(O)C916=C(O)C917=C(O)C918=C(O)C919=C(O)C920=C(O)C921=C(O)C922=C(O)C923=C(O)C924=C(O)C925=C(O)C926=C(O)C927=C(O)C928=C(O)C929=C(O)C930=C(O)C931=C(O)C932=C(O)C933=C(O)C934=C(O)C935=C(O)C936=C(O)C937=C(O)C938=C(O)C939=C(O)C940=C(O)C941=C(O)C942=C(O)C943=C(O)C944=C(O)C945=C(O)C946=C(O)C947=C(O)C948=C(O)C949=C(O)C950=C(O)C951=C(O)C952=C(O)C953=C(O)C954=C(O)C955=C(O)C956=C(O)C957=C(O)C958=C(O)C959=C(O)C960=C(O)C961=C(O)C962=C(O)C963=C(O)C964=C(O)C965=C(O)C966=C(O)C967=C(O)C968=C(O)C969=C(O)C970=C(O)C971=C(O)C972=C(O)C973=C(O)C974=C(O)C975=C(O)C976=C(O)C977=C(O)C978=C(O)C979=C(O)C980=C(O)C981=C(O)C982=C(O)C983=C(O)C984=C(O)C985=C(O)C986=C(O)C987=C(O)C988=C(O)C989=C(O)C990=C(O)C991=C(O)C992=C(O)C993=C(O)C994=C(O)C995=C(O)C996=C(O)C997=C(O)C998=C(O)C999=C(O)C999=C(O)

The visualization result of this function on the front end is as follows:



Proteome-Proteome upload-Proteome Data

The first eight columns of this file are fixed as: Accession, Gene, GO, KEGG_pathways, KEGG_orthologs, IPR, Description, and note. All this information can be extracted from the proteome file and organized into the following format:

Accession	Gene	GO	KEGG_pathways	KEGG_orthologs	IPR	Description	note
A0A5N5EU86	D8674_036	GO:000004	ko05322:Systemic I	K11090:lupus La protein	IPR01155: NA 12-oxophytodienoate reductase 11 OS=Pyrus uss		
A0A5N5EU88	D8674_036	GO:000004	ko05322:Systemic I	K11090:lupus La protein	--	Uncharacterized protein OS=Pyrus ussuriensis x Py	
A0A5N5EU94	D8674_036	GO:000032	ko05231:Choline m	K08202:MFS transporter, CIPRO29063: S-ε-Methyltransferase OS=Pyrus ussuriensis x Pyrus cor	CIPRO29063: S-ε-Methyltransferase OS=Pyrus ussuriensis x Pyrus cor		
A0A5N5EUAT	D8674_012	GO:000000	ko05231:Choline m	K15377:solute carrier famil	IPR036901: Asp-aspartate carbamoyltransferase OS=Pyrus ussurier		
A0A5N5EUC7	D8674_036	--	--	ko05231:Choline m	K08202:MFS transporter, CIPRO09056: Cyt Cytochrome c OS=Pyrus ussuriensis x Pyrus comm		
A0A5N5EUE4	D8674_036	GO:000032	ko05231:Choline m	K08200:MFS transporter, CIPRO36628: Clp Clp protease-related protein OS=Pyrus ussuriensis	CIPRO36628: Clp Clp protease-related protein OS=Pyrus ussuriensis		
A0A5N5EU6G	D8674_036	GO:000032	ko05231:Choline m	K08200:MFS transporter, CIPRO28226: Prc Uncharacterized protein OS=Pyrus ussuriensis x Py	CIPRO28226: Prc Uncharacterized protein OS=Pyrus ussuriensis x Py		
A0A5N5EUH6	D8674_036	GO:000032	ko05231:Choline m	K08200:MFS transporter, CIPRO13819: Lip Lipoygenase OS=Pyrus ussuriensis x Pyrus comm	CIPRO13819: Lip Lipoygenase OS=Pyrus ussuriensis x Pyrus comm		

The other columns are the corresponding quantitative values of the samples. For example:

L_1	L_2	L_3	LD_1	LD_2	LD_3
1.078733	1.062195	0.946504	0.736678	0.689896	1.283411
2.05725	2.072127	2.036453	1.949878	1.907849	2.000581
1.002007	1.015497	0.999638	1.106532	1.043764	1.021021
4.609982	4.461109	4.389813	4.274224	4.262235	4.204973
0.909062	1.098809	0.942233	0.934172	0.814113	1.175757
5.81071	5.438964	5.037357	5.711435	5.220466	5.994661
0.115391	0	0.073444	0.154881	0.109644	0

Proteome-Proteome upload-Sample Information

The project information file is similar to that of the metabolome. The column headers and their content are as follows: Project_Name, Sample_Group, Sample_Name, Replicate, Sample_Information, Germplasm, Germplasm_Characteristics, Instrument, Title, Tissue, Geographic_Location, Data_Sources, and Note. An example is as follows:

Project_Name	Sample_Group	Sample_Name	Replicate	Sample_Information	Germplasm	Germplasm_Characteristics	Instrument	Title	Tissue	Geographic_Location	Data_Sources	Note
Project_1	CK_3d_T	CK_3d_T_1							Leaf			
Project_1	CK_3d_T	CK_3d_T_2							Leaf			
Project_1	CK_3d_T	CK_3d_T_3							Leaf			
Project_1	CK_3d_S	CK_3d_S_1							Leaf			
Project_1	CK_3d_S	CK_3d_S_2							Leaf			
Project_1	CK_3d_S	CK_3d_S_3							Leaf			

Proteome-Proteome upload- proteome sequence

The sequence uploaded here corresponds to the proteome (in Fasta format). Note that the IDs must be consistent.

Introduction to Front-End Functions

Genome

1. Gene Search

On the gene search page, users can enter a gene list (①) to query gene annotation information, gene structure information, and sequence information. Users can view the GO, KEGG, and IPR annotation information of genes, click the annotation ID (②) to check the detailed description of the annotation, use IGV for interactive viewing of gene structure information (③), quickly obtain the gene sequence, CDS sequence, and protein sequence of the target gene (④), and independently design and acquire the upstream and downstream sequence information of the gene (⑤).

Gene Search:

Psin01G000010
Psin01G000020
Psin01G000030
Psin01G000040
Psin01G000050

Demo search

Annotation: GO,KEGG,IPR and note

GeneID	GO	KEGG
Psin01G000010	--	--
Psin01G000020	--	--
Psin01G000030	GO 0000003 biological_process obsolete reproduction GO 0000226 biological_process microtubule cytoskeleton organization GO 000220 cellular_component nuclear chromosome GO 0000276 biological_process mitotic cell cycle GO 0000280 biological_process nuclear division GO 0000281 biological_process mitotic cytokinesis GO 0000724 biological_process double-strand break repair via homologous recombination GO 0000723 biological_process telomere maintenance GO 0000725 biological_process recombinational repair	K08775
Psin01G000040	--	--
Psin01G000050	--	K15255

Showing 1 to 5 of 5 entries Previous 1 Next

Structure:gene,mRNA,exon,CDS and UTR

Psin01G000030

IGV Chr GWHTVS00 GWHTVS00000001:146,345 Q 6,611 bp Cursor Guide Center Line Track Labels Save SVG +

147 kb 148 kb 149 kb 150 kb 151 kb 152 kb 1!

Structure

Gene Sequence:gene,CDS,protein,gene upstream and downstream

Gene CDS Protein

Upstream the default value is 0 Downstream the default value is 0 Start

Psin01G000010

```

ATGCCCTCCACTACTGTGAAGAAAGTTGGTGTGCAATAAAGTAGATTGGTGTGGAAACCCGTTACGTGGGGATAAAAGCCTAAAGAGATTATTCCGAAATTTCG-G-  

TGCGTAGTTGGQAGATTGTGGATGCTACAGGTTTCTGGGTTATTTGGCTTTATGQATAAGACTAAAGTTAAAGTGAAATTGGTGTGGCTTTGTAAGCAAGGTTACTGT  

TCAAATTTTATGATGTTAACGTTAATGGTATTTCCTACTTATTGTCAGCCTTGATATAAAGGAGAAATGCCTTATGGGAATGTTCTATAATTAAGCTAGATGTGATCT  

TCTCTGGCTCTCTGGTGAAGATTAAAGTTGGTGTGGCTGTGATGAAAGAAGGGGGCTCTAGAGAGGAGTAGTTGGTGTGTTAAAGATTTGGTCTATGGCTAATTTTATG  

GTGGATTGGGATTTGGTGTCTTCACTTGGATGGTTAAGAGCTTGAAGGCTTGTAAAGGCTTGTAAAGGCTTGTAAAGGCTTGTAAAGGCTTGTAAAGGCTTGTAAAGGCTTGT  

TTCTTTAA
  
```

Psin01G000020

```

ATGACATCAAAGAAAGGATCAACGCTGTCCTGCAATCAAAGCAAAGACAAGAACATCATCGCCCAAGTGTGATATTGGGCCACATAACCCGAGTAAGGCAAGAGCTT  

TTGGCCGGCTCCGCCCTGTATTAACTGCAAGGAAAGACAAGACCTAACGGCAGCTTGGCAGGAGCTGTGACCTTGGCTGGCTAGGGCATCAAAGGGGAAGGCTGA  

GGAAATACTGAGCTTACTCTCCGAGCGCTGATTCAAGCGAGCTGACCATGCAACCATGCACTGAGTAACCTTCAATCGAGGAGCAGCTAGCTCATGATGATGAA  

GCAATGCGAAAGGCTCACACGGACTGAAAGGAGAAAGGAGGAGCTGATGGCTGACGGCTGAGCTCATCACGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT  

CTAAAGAAGCCGAAGAACCTTAAACCAAATGGCTACGAAACCTGATGGCTGAGCTTGGCTCTCTTCACATTTGGAAAGAAGGTTAACACCGCTCAATG  

ACAAGAACCTGAAAGAACCTGAGCTGAGCTGAGCTTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT  

AGCAAGGTTGAGAAAGGCTAGTGTCTCACACATCGAGCTGAGCTTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT  

GCTCCAGAAATTGAGCTTAACTGCAAGGGTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
  
```

← 3 ← 4 ← 5

2. Transcription Factor

The Transcription Factor page displays the number of all transcription factor family members in the species for users. Users can obtain detailed information about individual transcription factor family members on this page (①); clicking the transcription factor family name (②) will redirect to the detailed information page of the gene family members, including member information, phylogenetic tree, and chromosome localization. Users can view the member's ID, subfamily (③), and homologous ID (④) in the member information table, check and locate the family (⑤) and member classification information (⑥) in the phylogenetic tree, and quickly

locate the family (⑧) and member position information (⑦) in the chromosome localization. It should be noted that Blast, Muscle, and FastTree tools are embedded in the system; if the user does not upload the phylogenetic tree and subfamily information in the back end, PlantMDMS will automatically complete the identification of subfamily members and the construction of the phylogenetic tree using *Arabidopsis thaliana* as a reference.

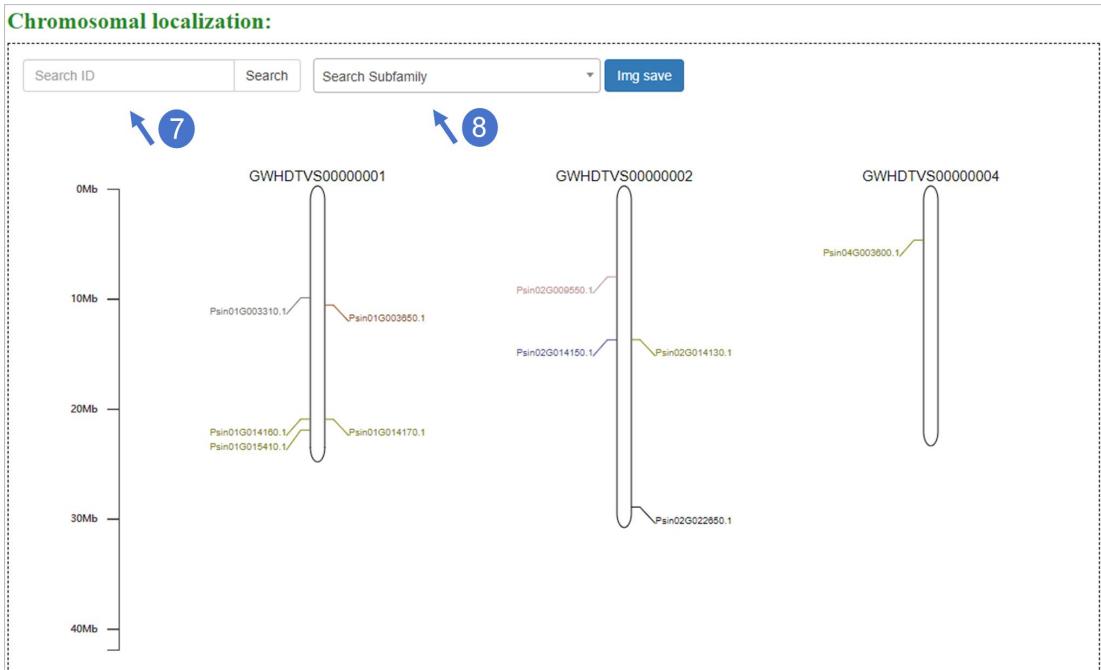
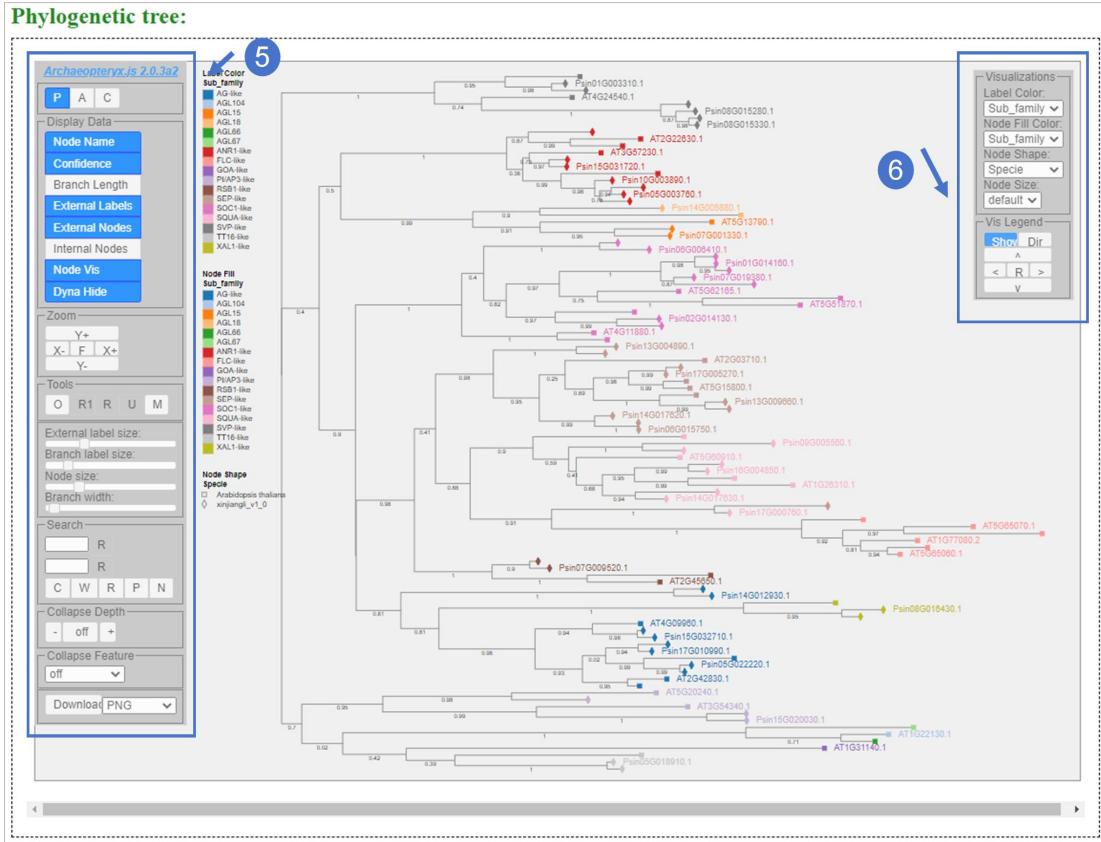
The screenshot shows a search results page for the query "Psin01G014160.1". At the top, there is a search bar with the value "Psin01G014160.1" and a "Search" button. Below the search bar, the ID "Psin01G014160.1", family "MADS-MIKC", subfamily "SOC1-like", and best homolog "AT5G62165.1" are displayed. A blue arrow labeled 1 points from the text "the user does not upload the phylogenetic tree and subfamily information in the back end, PlantMDMS will automatically complete the identification of subfamily members and the construction of the phylogenetic tree using *Arabidopsis thaliana* as a reference." to the subfamily information. Below this, a table titled "Transcription Factors" lists various families and their counts. A blue arrow labeled 2 points from the text "the user does not upload the phylogenetic tree and subfamily information in the back end, PlantMDMS will automatically complete the identification of subfamily members and the construction of the phylogenetic tree using *Arabidopsis thaliana* as a reference." to the HRT(2) entry in the table. Another blue arrow labeled 3 points from the text "the user does not upload the phylogenetic tree and subfamily information in the back end, PlantMDMS will automatically complete the identification of subfamily members and the construction of the phylogenetic tree using *Arabidopsis thaliana* as a reference." to the MADS-MIKC(72) entry in the table. A fourth blue arrow labeled 4 points from the text "the user does not upload the phylogenetic tree and subfamily information in the back end, PlantMDMS will automatically complete the identification of subfamily members and the construction of the phylogenetic tree using *Arabidopsis thaliana* as a reference." to the MADS-M-type(74) entry in the table.

Alfin-like(10)	AP2/ERF-AP2(17)	AP2/ERF-ERF(148)	AP2/ERF-RAV(5)	B3(77)
B3-ARF(37)	BBR-BPC(17)	BES1(14)	bHLH(196)	BSD(2)
bZIP(120)	C2C2-CO-like(21)	C2C2-Dof(47)	C2C2-GATA(41)	C2C2-LSD(12)
C2C2-YABBY(8)	C2H2(124)	C3H(76)	CAMTA(9)	CPP(9)
CSD(4)	DBB(7)	DBP(7)	DDT(6)	E2F-DP(16)
EIL(6)	FAR1(26)	GARP-ARR-B(15)	GARP-G2-like(61)	GeBP(21)
GRAS(37)	GRF(9)	HB-BELL(23)	HB-HD-ZIP(52)	HB-KNOX(10)
HB-other(12)	HB-PHD(3)	HB-WOX(18)	HRT(2)	HSF(25)
LFY(1)	LIM(11)	LOB(50)	MADS-MIKC(72)	MADS-M-type(74)
MYB(165)	MYB-related(99)	NAC(137)	NF-X1(2)	NF-YA(21)
NF-YB(27)	NF-YC(21)	NOZZLE(1)	OPF(12)	PLATZ(15)
RWP-RK(17)	S1Fa-like(4)	SAP(1)	SBP(30)	SRS(15)
STAT(4)	TCP(33)	Tify(27)	Trihelix(31)	TUB(19)
ULT(3)	VOZ(3)	Whirly(4)	WRKY(91)	zf-HD(18)

The screenshot shows a table titled "Members information:" with columns: ID, Sub_Family, Hom_ID, and Note. The table contains 59 entries. A blue arrow labeled 3 points from the text "the user does not upload the phylogenetic tree and subfamily information in the back end, PlantMDMS will automatically complete the identification of subfamily members and the construction of the phylogenetic tree using *Arabidopsis thaliana* as a reference." to the SOC1-like entry in the table. A blue arrow labeled 4 points from the text "the user does not upload the phylogenetic tree and subfamily information in the back end, PlantMDMS will automatically complete the identification of subfamily members and the construction of the phylogenetic tree using *Arabidopsis thaliana* as a reference." to the AGL15 entry in the table.

ID	Sub_Family	Hom_ID	Note
Psin01G003310.1	SVP-like	AT2G22540.1	--
Psin01G003650.1	ANR1-like	AT4G37940.1	--
Psin01G014160.1	SOC1-like	AT5G62165.1	--
Psin01G014170.1	SOC1-like	AT5G62165.1	--
Psin01G015410.1	SOC1-like	AT5G62165.1	--
Psin02G009550.1	PI/AP3-like	AT3G54340.1	--
Psin02G014130.1	SOC1-like	AT2G45660.1	--
Psin02G014150.1	RSB1-like	AT2G45650.1	--
Psin02G022650.1	AGL15	AT5G13790.1	--
Psin04G003600.1	SOC1-like	AT4G22950.1	--

Showing 1 to 10 of 59 entries



3. Pathway

The Pathway section includes plant-related metabolic pathways; users can click the ID to access the corresponding metabolic pathway page (①). On the metabolic pathway page, users can obtain all gene IDs annotated to this pathway in the table

(②), and when hovering the mouse over the green areas in the metabolic pathway, the corresponding gene list will be interactively displayed (③). Users can copy these genes for batch search on other functional pages.

Pathway Annotation

1. Metabolism

1.1 Carbohydrate metabolism

00010 Glycolysis / Gluconeogenesis
 00020 Citrate cycle (TCA cycle)
 00030 Pentose phosphate pathway
 00040 Pentose and glucuronate interconversions
 00051 Fructose and mannose metabolism
 00052 Galactose metabolism
 00053 Ascorbate and aldarate metabolism
 00500 Starch and sucrose metabolism
 00520 Amino sugar and nucleotide sugar metabolism
 00620 Pyruvate metabolism
 00630 Glyoxylate and dicarboxylate metabolism
 00640 Propanoate metabolism
 00650 Butanoate metabolism
 00660 C5-Branched dibasic acid metabolism
 00562 Inositol phosphate metabolism

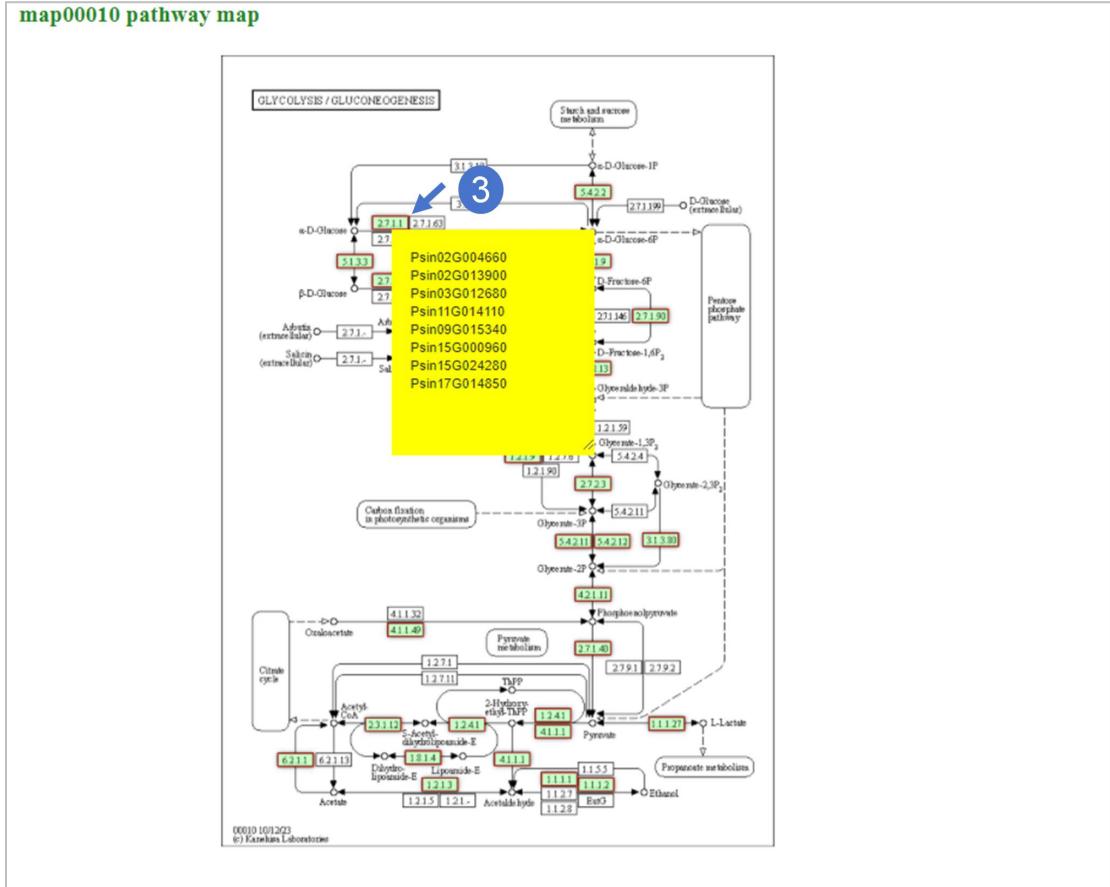
map00010 information table

Gene ID	KEGG_orthologs	KEGG_pathways
Psin01G002250	K01834:2,3-bisphosphoglycerate-dependent phosphoglycerate mutase [EC:5.4.2.11]	ko00010:Glycolysis / Gluconeogenesis ko00260:Glycine, serine and threonine metabolism ko00020:Citrate cycle (TCA cycle)
Psin01G003190	K00850:6-phosphofructokinase 1 [EC:2.7.1.11]	ko00010:Glycolysis / Gluconeogenesis ko00030:Pentose phosphate pathway ko00051:Fructose and mannose metabolism
Psin01G003880	K01610:phosphoenolpyruvate carboxykinase (ATP) [EC:4.1.1.49]	ko00010:Glycolysis / Gluconeogenesis ko00020:Citrate cycle (TCA cycle) ko00030:Pentose phosphate pathway
Psin01G006600	K01792:glucose-6-phosphate 1-epimerase [EC:5.1.3.15]	ko00010:Glycolysis / Gluconeogenesis ko01100:Metabolic pathways ko00020:Citrate cycle (TCA cycle)
Psin01G007700	K00161:pyruvate dehydrogenase E1 component subunit alpha [EC:1.2.4.1]	ko00010:Glycolysis / Gluconeogenesis ko00020:Citrate cycle (TCA cycle) ko00030:Pentose phosphate pathway
Psin01G008390	K00627:pyruvate dehydrogenase E2 component (dihydrolipoylysine-residue acetyltransferase) [EC:2.3.1.12]	ko01120:Microbial metabolism in diverse environments ko01130:-- ko00020:Citrate cycle (TCA cycle)
Psin01G010750	K01803:triosephosphate isomerase (TIM) [EC:5.3.1.1]	ko00010:Glycolysis / Gluconeogenesis ko00051:Fructose and mannose metabolism ko00020:Citrate cycle (TCA cycle)
Psin01G012510	K03103:multiple inositol-polyphosphate phosphatase / 2,3-bisphosphoglycerate 3-phosphatase [EC:3.1.3.62 3.1.3.80]	ko00010:Glycolysis / Gluconeogenesis ko00562:Inositol phosphate metabolism ko00020:Citrate cycle (TCA cycle)
Psin01G013860	K00850:6-phosphofructokinase 1 [EC:2.7.1.11]	ko00010:Glycolysis / Gluconeogenesis ko00030:Pentose phosphate pathway ko00051:Fructose and mannose metabolism
Psin01G014220	K01835:phosphoglucomutase [EC:5.4.2.2]	ko00010:Glycolysis / Gluconeogenesis ko00030:Pentose phosphate pathway ko00051:Fructose and mannose metabolism

Showing 1 to 10 of 328 entries

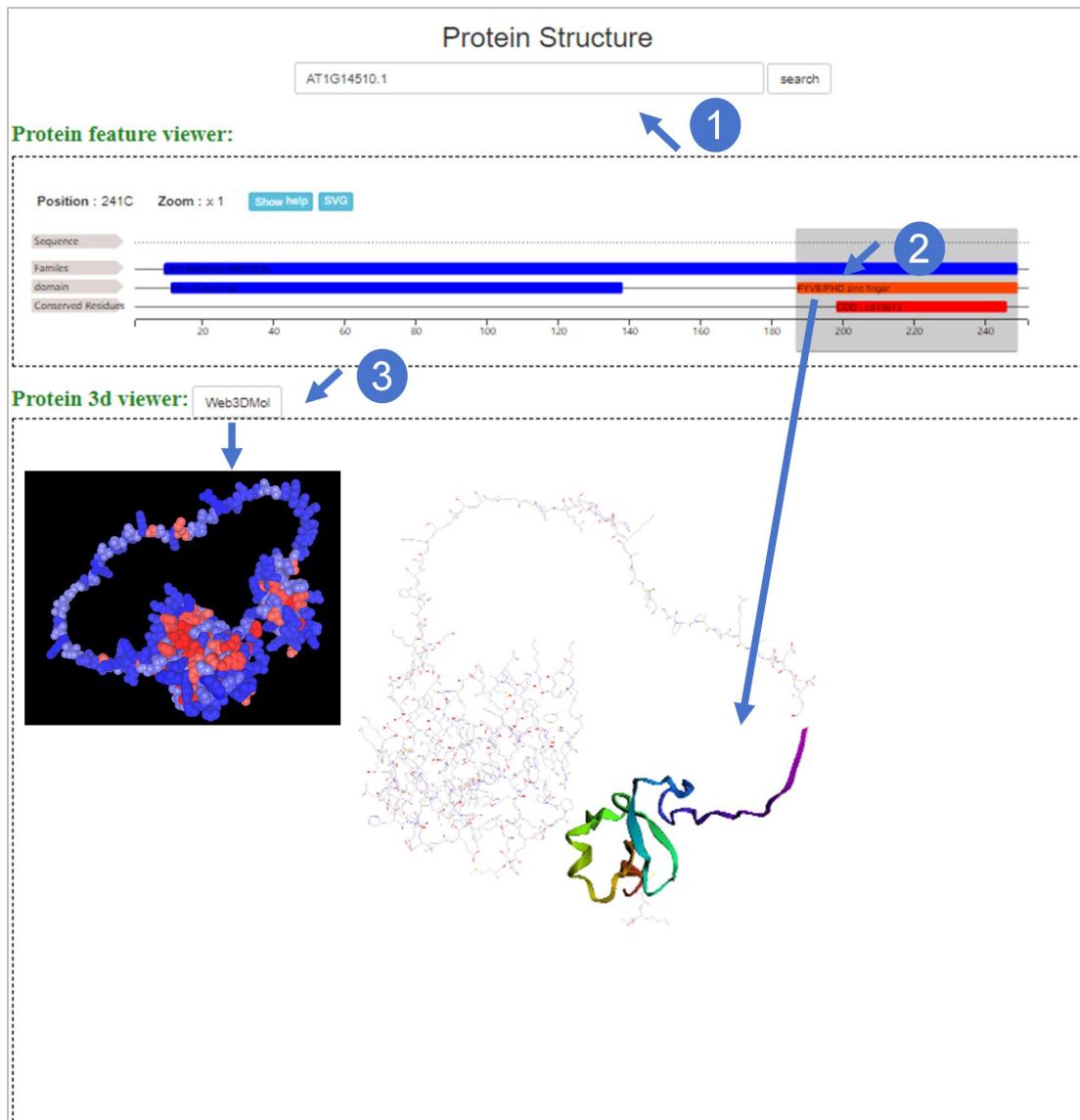
Previous 1 2 3 4 5 ... 33 Next

map00010 pathway map



4. Protein Structure

The Protein Structure page is used to search for and display information about the functional domains and 3D structures of proteins. Users can enter the protein ID for retrieval (①) and interactively view the corresponding 3D structure by clicking on the functional domain (②). In addition, users can click the Web3DMol button to view and edit the 3D structure of the protein (③).



5. Genome Browser

The Genome Browser page is used to view gene structure information; users can use the search tool to retrieve the transcript ID corresponding to a gene ID and the gene corresponding to a transcript.

Members information:

IGV Chr2 Chr2:16,211,050-16,249,522 38 kb Cursor Guide Center Line Track Labels Save SVG - +

16,220 kb 16,230 kb 16,240 kb 16

Structure AT2G38780.1 AT2G38800.1 AT2G38820.2 AT2G38830.1 AT2G38840.1 AT2G38860.2 AT2 AT2G38860.1 AT2G38860.3

AT2G38790.1 AT2G38810.1 AT2G38820.1 AT2G38810.3

AT2G38860.3

type	mRNA
start	16233315
end	16235234
ID	AT2G38860.3
Name	AT2G38860.3
pacid	19638158
longest	0
Parent	AT2G38860

type	CDS
start	162334471
end	16234941
ID	AT2G38860.3.CDS.2
Parent	AT2G38860.3
pacid	19638158

Gene information:

Select Type:	Gene ID	Enter Gene ID:	Input ID	SEARCH
Gene ID	Gene Position	Transcript ID	Chr	Direction
AT1G01010	3631-5899	AT1G01010.1	Chr1	+
AT1G01020	5928-8737	AT1G01020.1	Chr1	-
AT1G01020	5928-8737	AT1G01020.2	Chr1	-
AT1G01030	11649-13714	AT1G01030.1	Chr1	-
AT1G01040	23146-31227	AT1G01040.2	Chr1	+
AT1G01040	23146-31227	AT1G01040.1	Chr1	+
AT1G01050	31170-33153	AT1G01050.1	Chr1	-
AT1G01060	33379-37871	AT1G01060.1	Chr1	-
AT1G01060	33379-37871	AT1G01060.2	Chr1	-
AT1G01060	33379-37871	AT1G01060.3	Chr1	-

Showing 0 to 0 of 0 entries (filtered from 35,386 total entries)

First Previous **1** 2 3 4 5 ... Next Last

Transcriptomics

1. Expression Search

Expression Search is used to display the expression level information and sample information of transcriptome data. Users can enter a gene list (①) and select samples (②) to obtain three parts of content: a heatmap, an expression level data table, and sample information. Users can copy the expression level and sample information data to a local device for subsequent research.

Gene Search:

*1 Select Sample

SRR10498806
SRR10498808
SRR10498807

SRR10498805
SRR10498809

*2 Select Sample

Psin01G000010
Psin01G000020
Psin01G000030
Psin01G000040
Psin01G000050

1 2

Heatmap:

Expression Table:

Gene_ID	SRR10498805	SRR10498806	SRR10498807	SRR10498808	SRR10498809
Psin01G000010	0	0	0	0	0
Psin01G000020	0.0838726952811361	0.0349514400866962	0	0.0176324603049792	0.0177810291475402
Psin01G000030	0.111343327395927	0.184049387862974	0.260957616825668	0.273088429984392	0.296043645817258
Psin01G000040	0	0	0	0	0
Psin01G000050	0	0	0	0	0
Psin01G000060	0.0235594270072486	0	0	0	0
Psin01G000070	5.93199456555869	6.02242556254692	6.28999617227695	7.34949803567566	6.47285935440854
Psin01G000080	2.83710218324614	3.88435391118521	3.22690208026286	2.4831655405395	3.17467475210661
Psin01G000090	0	0	0	0.103047238049476	0
Psin01G000100	0.285188579946717	0.251420705271644	0.288330388609352	0.321757787822219	0.347981115060522

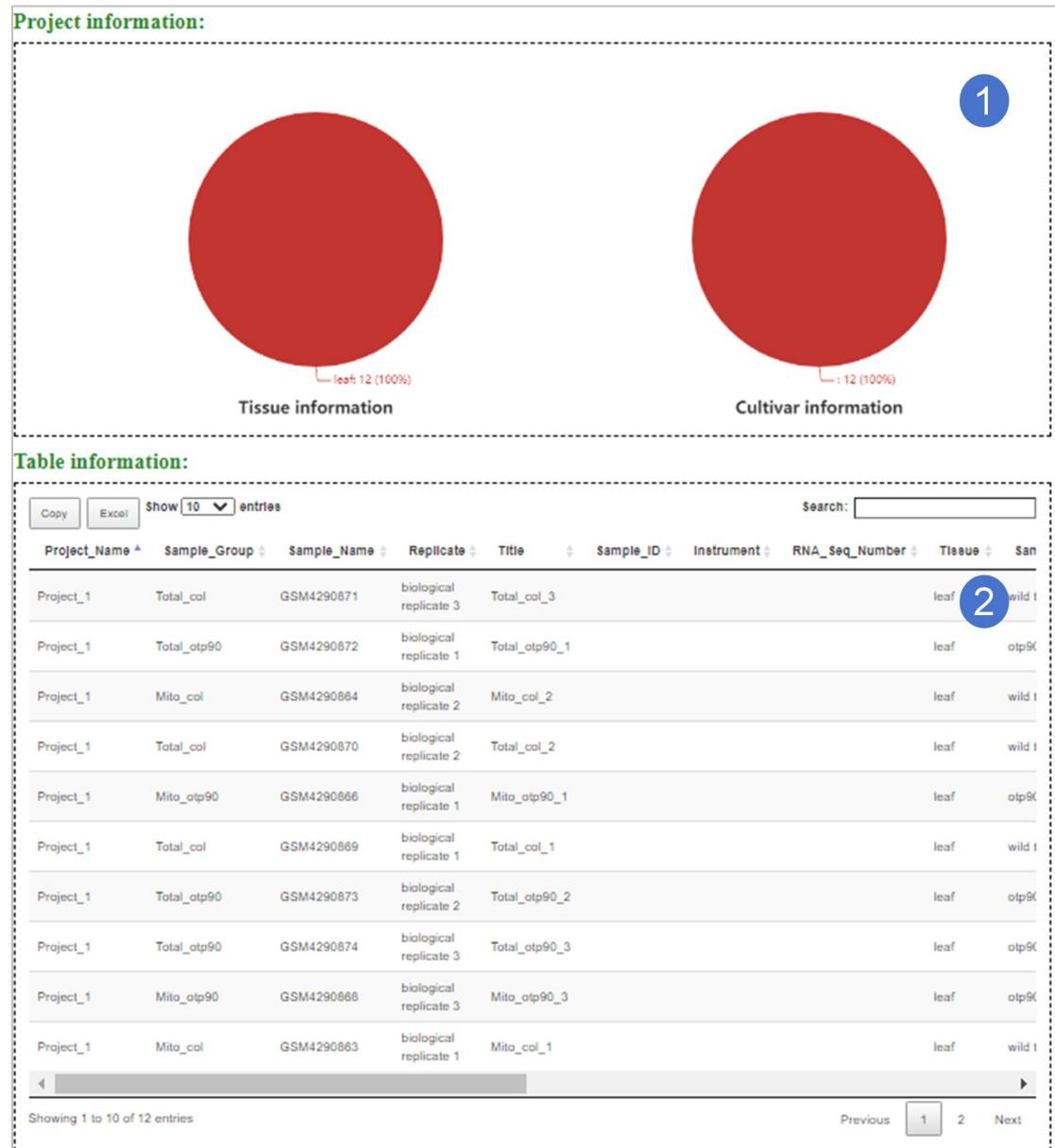
Sample Information Table:

Project_Name	Sample_Group	Sample Name	Replicate	Tissue	Sample Information	Cultivar	Cultivar Information
Project_1	1	SRR10498808					
Project_1	1	SRR10498806					
Project_1	1	SRR10498807					
Project_1	2	SRR10498805					
Project_1	2	SRR10498809					

2. Project information

Project Information is used to display detailed information about projects and samples.

The pie chart presents statistical information of different tissues and germplasm resources across all related projects (①), and users can also query and download sample information (②).



3. DEGs

In DEGs analysis, users can select sample groups for analysis and click the "All

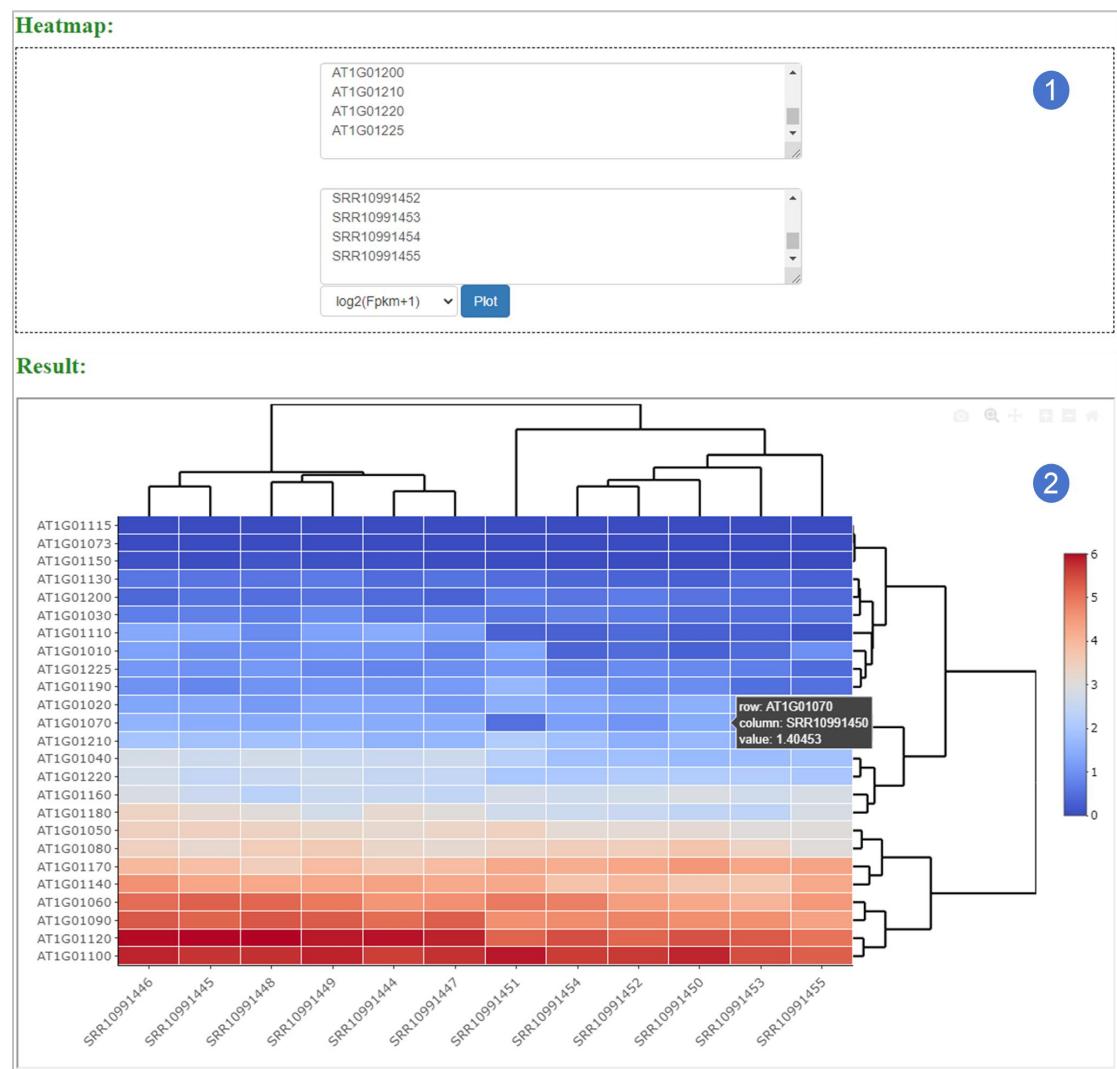
"Sample Details" button to view the group information of samples (①). The results are divided into two parts: a volcano plot (②) and a result information table (③). Clicking on the points in the volcano plot allows for interactive display of the annotation information of the corresponding gene, and users can also download the result files of the differential analysis.



4. Heatmap

In the Heatmap function, users can enter a gene list and a sample list, then select the

data type to generate a clustered heatmap. The heatmap is plotted using the heatmaply package and supports interactive zoom-in and zoom-out operations.



5. Enrichment:

Enrichment provides GO and KEGG enrichment analysis functions. Users need to enter a gene list (①), select the analysis type (②), and set analysis parameters to perform the enrichment analysis (③). The analysis results are presented to users in the form of a bubble chart and a table: the bubble chart displays the top 20% of enriched results, and users can obtain all enrichment result information in the table.

Enrichment:

KEGG GO

All

1 2 3

pvalueCutoff

qvalueCutoff

pAdjustMethod

Bubble Chart:

Result Top20:

Pathway	Gene number	Rich Ratio
negative regulation of endopeptidase activity	3	~0.002
generation of precursor metabolites and energy	3	~0.003
response to copper ion	3	~0.004
photosynthesis	5	~0.006
cellular response to toxic substance	5	~0.007
response to toxic substance	5	~0.008
cellular detoxification	5	~0.009
cellular oxidant detoxification	5	~0.010
photosynthetic membrane	7	~0.011
thylakoid membrane	7	~0.012
detoxification	7	~0.013
plastid thylakoid membrane	9	~0.014
chloroplast thylakoid membrane	9	~0.015
obsolete thylakoid part	11	~0.016
plastid thylakoid	11	~0.017
chloroplast thylakoid	13	~0.018

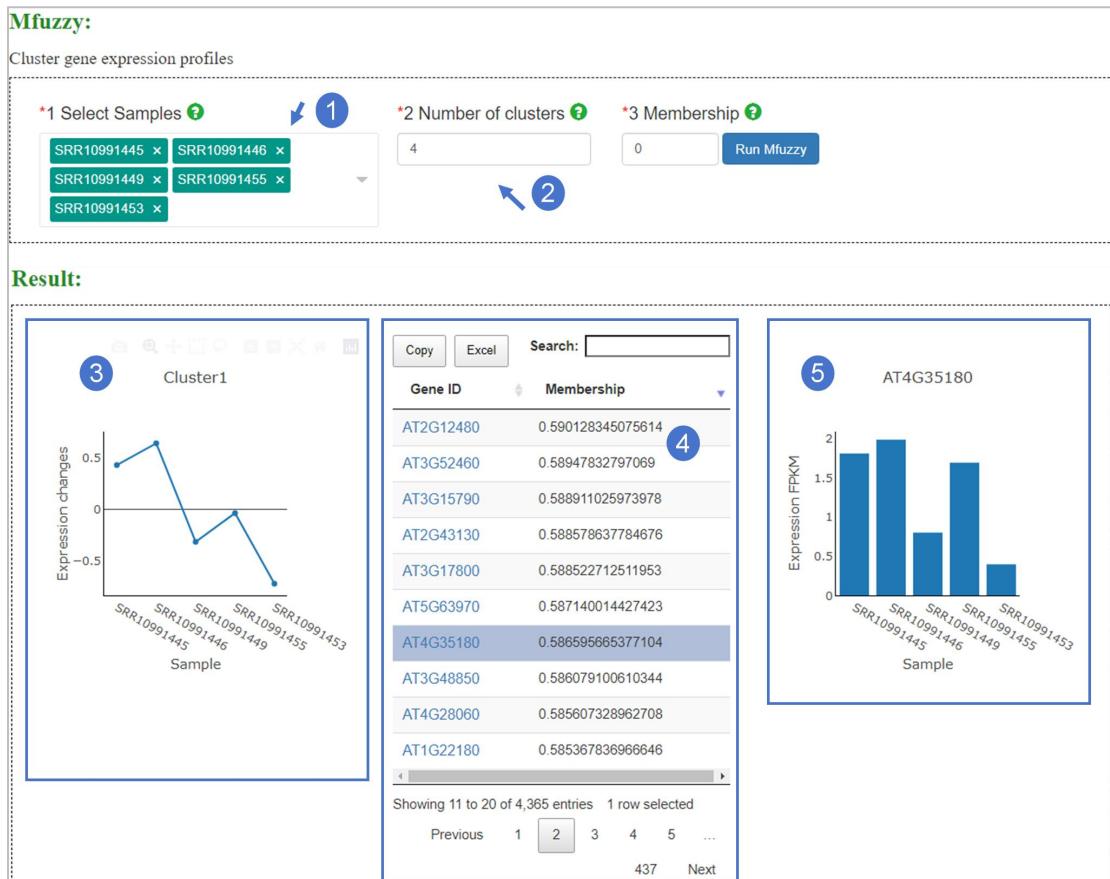
Result Data:

Show 10 entries									Search: <input type="text"/>
ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue	Count	geneID	
GO:0006091	generation of precursor metabolites and energy	7/85	278/15415	0.000856057131710696	0.0421750813556136	0.0396489618897586	7	ATCG00420/AT	
GO:0009534	chloroplast thylakoid	13/85	413/15415	3.89690421962832e-07	0.000160653496458023	0.000151030991613554	13	AT1G52230/AT	
GO:0009535	chloroplast thylakoid membrane	11/85	328/15415	1.79523417587417e-06	0.000282006095258079	0.0002651150528742	11	AT1G52230/AT	
GO:0009636	response to toxic substance	10/85	379/15415	4.34867877221643e-05	0.0026853896619698	0.00252454551228555	10	AT1G08830/AT	
GO:0010951	negative regulation of endopeptidase activity	3/85	35/15415	0.000932311589889904	0.0430611415560399	0.0404819506136406	3	AT5G02380/AT	
GO:0015979	photosynthesis	6/85	157/15415	0.000228592165093302	0.0129945853849192	0.0122162614543789	6	AT1G14150/AT	
GO:0031976	plastid thylakoid	13/85	417/15415	4.34786187978412e-07	0.000160653496458023	0.000151030991613554	13	AT1G52230/AT	
GO:0034357	photosynthetic membrane	11/85	350/15415	3.3606062479779e-06	0.000310436002156959	0.000291842121534923	11	AT1G52230/AT	
GO:0042651	thylakoid membrane	11/85	349/15415	3.2695061187095e-06	0.000310436002156959	0.000291842121534923	11	AT1G52230/AT	
GO:0044436	obsolete thylakoid part	12/85	374/15415	9.32025904423201e-07	0.000229589047789582	0.000215837577866426	12	AT1G52230/AT	

Showing 1 to 10 of 16 entries Previous 1 2 Next

6. Mfuzzy

This page analyzes the temporal trend of transcriptional changes using the Mfuzzy package. Users first select samples (①), then set the number of clusters to perform the analysis (②). The results for each cluster include three parts: the cluster trend plot (③), the list of genes included in the cluster (④), and the expression trend plot of individual genes (⑤).



7. PCA

The PCA page provides principal component analysis functionality. Users can perform principal component analysis by selecting different groups (①), setting group colors (②), choosing the expression level data type (③), and configuring normalization (④). The results include three parts: the principal component 2D plot (⑤), the principal component 3D plot (⑥), and the principal component result table (⑦).

PCA:

*1 Select Sample Groups  *2 Color setting  *3 select type  *scale  *Run 

Mito_col  Mito_otp90  Total_otp90 
Total_col 

Mito_col:  FPKM  TRUE  Start

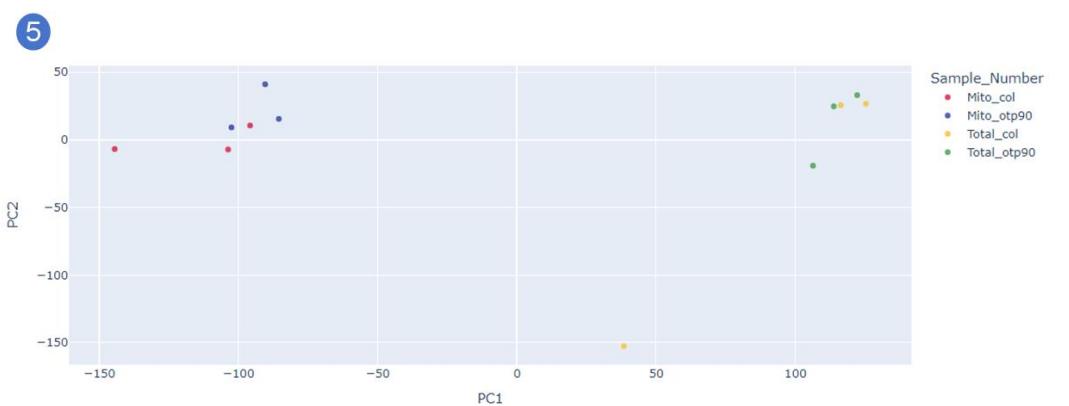
Mito_otp90:   3  4

Total_otp90:   2

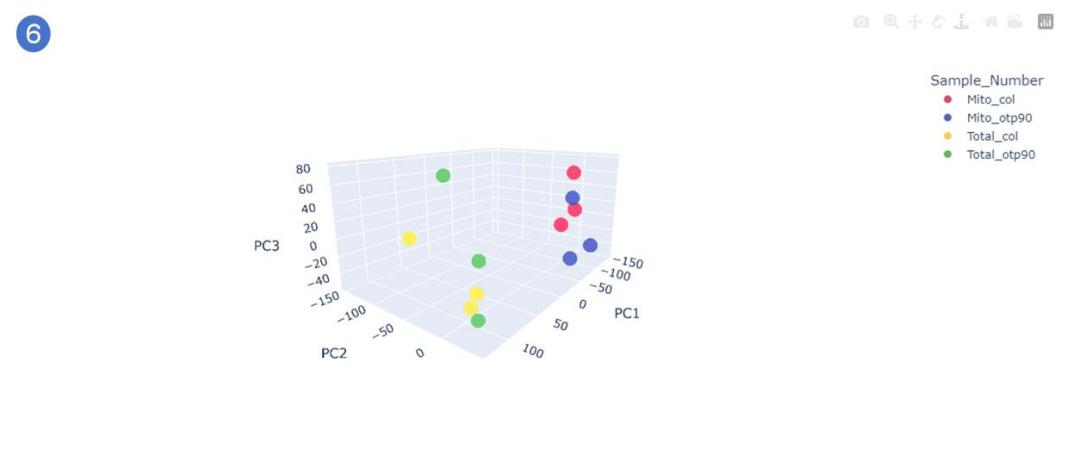
Total_col:   3

1

PCA 2D:



PCA 3D:



PCA Table:

7

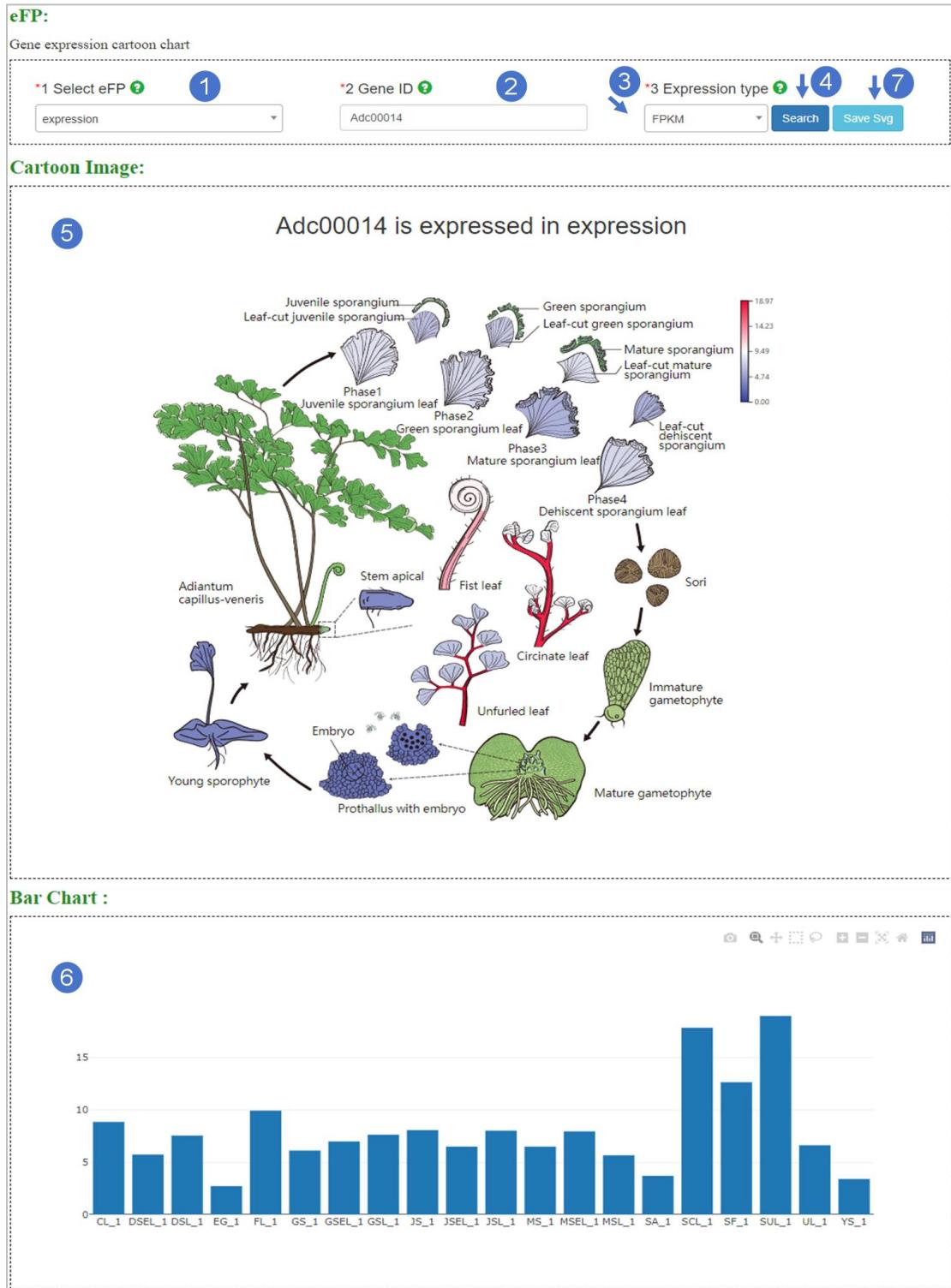
Sample_Name	Sample_Number	PC1	PC2	PC3	PC4	PC5	PC6
SRR10991444	Mito_col	-95.8125315260707	10.5211961905155	62.9213964389421	6.4855499036352	-21.3954509119268	31.297714590
SRR10991445	Mito_col	-103.70805944273	-7.20170040995781	-10.5011387576005	-19.7102031439596	55.4720311544545	50.934460598
SRR10991446	Mito_col	-144.453814916878	-6.84281631030036	3.09718731289097	-88.8512043066645	-22.5767218578983	-26.27119826
SRR10991447	Mito_otp90	-85.5005250274691	15.3970898767369	32.6467483755336	61.1185419566078	-31.4881689293614	28.342989790
SRR10991448	Mito_otp90	-102.505277726766	9.12524350611831	-55.2806363140622	27.4420382694051	57.0365465379655	-9.892268402
SRR10991449	Mito_otp90	-90.4014062113119	41.0103328089421	-26.1766746784755	31.6118665029094	-32.6483753657433	-66.89293451
SRR10991450	Total_col	116.301278554851	25.4755884015269	-15.4425007631826	-37.0944610274947	-29.0655786986606	25.249879888
SRR10991451	Total_col	38.4167431902495	-152.525737820698	-22.1831926483947	16.4287561985731	-24.207172545334	-1.773659423
SRR10991452	Total_col	125.326698172461	26.5894031984932	-24.2961814950824	0.451538415130594	-19.7637588409316	13.216881598
SRR10991453	Total_otp90	122.207060673325	32.9678036248174	17.7944032809331	1.50219089535226	2.89457134071981	0.0773659978

Showing 1 to 10 of 12 entries

Previous 1 2 Next

8. eFP

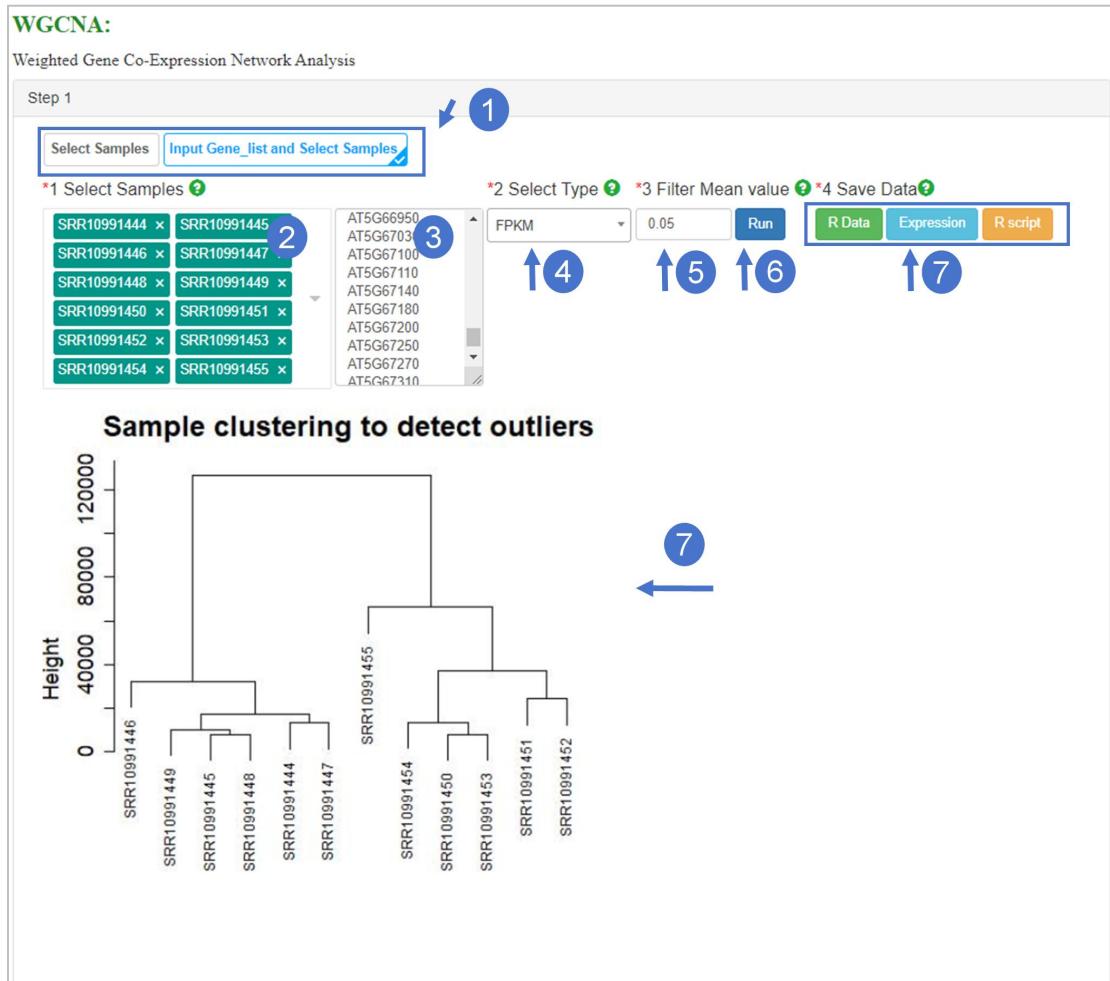
Electronic Fluorescence Pictograph (eFP) provides an interactive tool. By selecting a project (①), entering a gene ID (②), choosing the gene expression type (③), and clicking the Search button (④), users will see the specific expression of genes in various tissues within the project. The results include two parts: the eFP (⑤) and the bar chart of gene expression (⑥). Users can directly download the eFP image to their local device for further research (⑦).



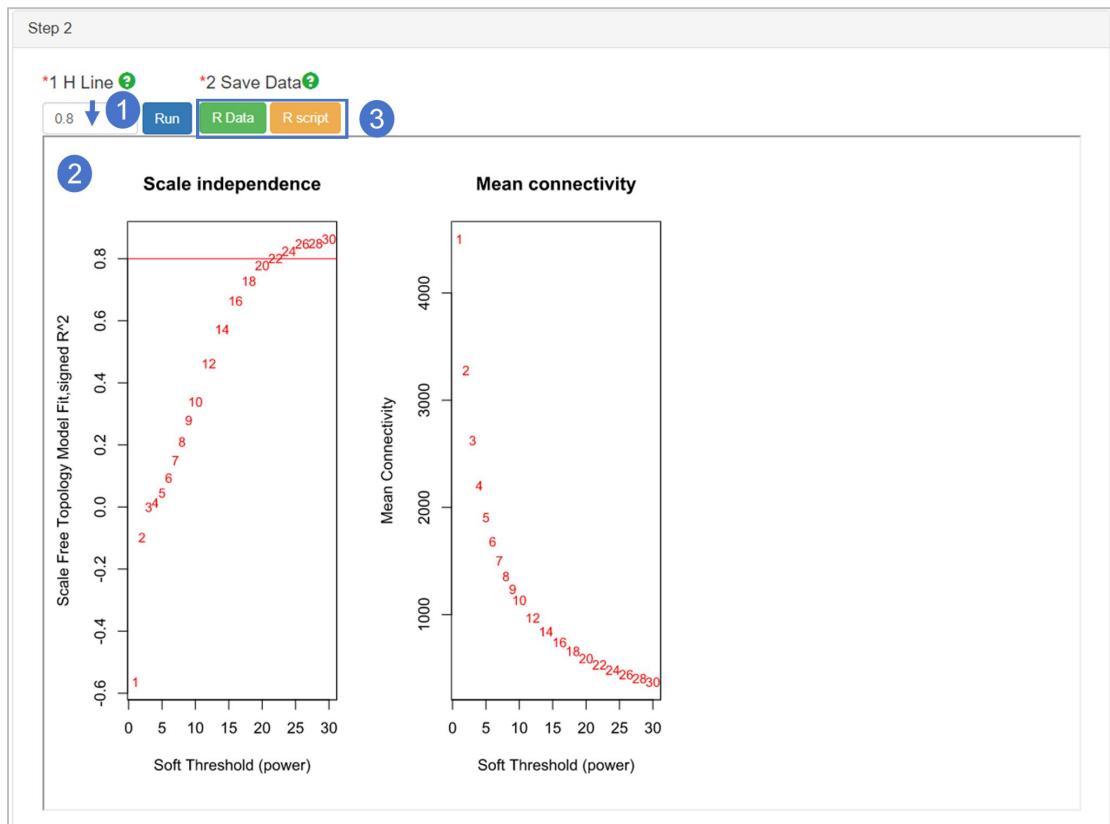
9. WGCNA

WGCNA consists of 6 steps. Step 1 is used for gene filtering and sample clustering, including two types: all genes and gene list (①). Users select the samples to be analyzed (②), enter a gene list (optional for all-gene analysis) (③), choose the expression level type (④), set the gene filtering parameters (⑤), and click Run to

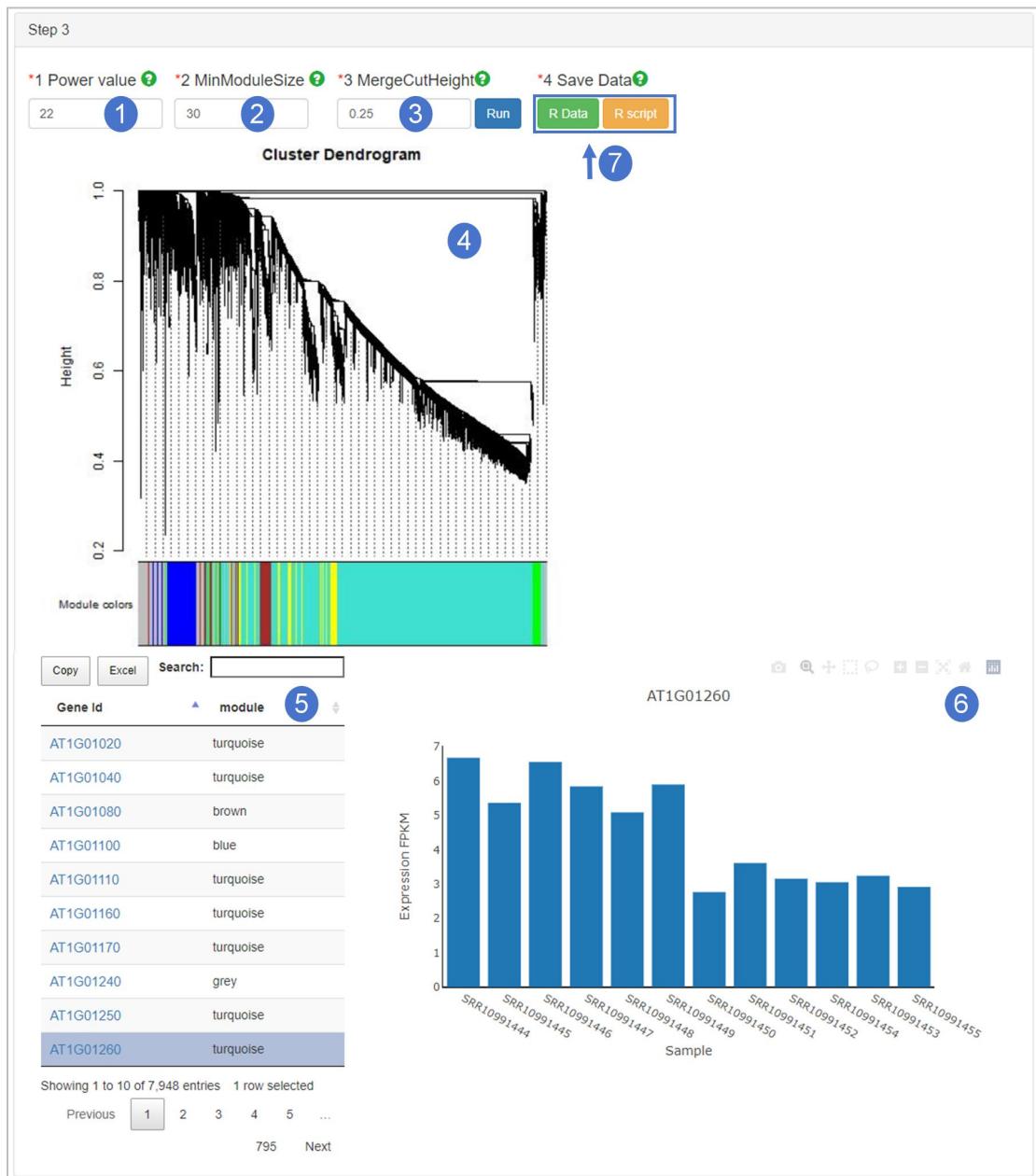
perform calculations (⑥). The results are presented to users in graphical form (⑦). In Step 1, users can download the gene expression level data, datasets generated by R language operation, and R scripts to their local device for local computation (⑧).



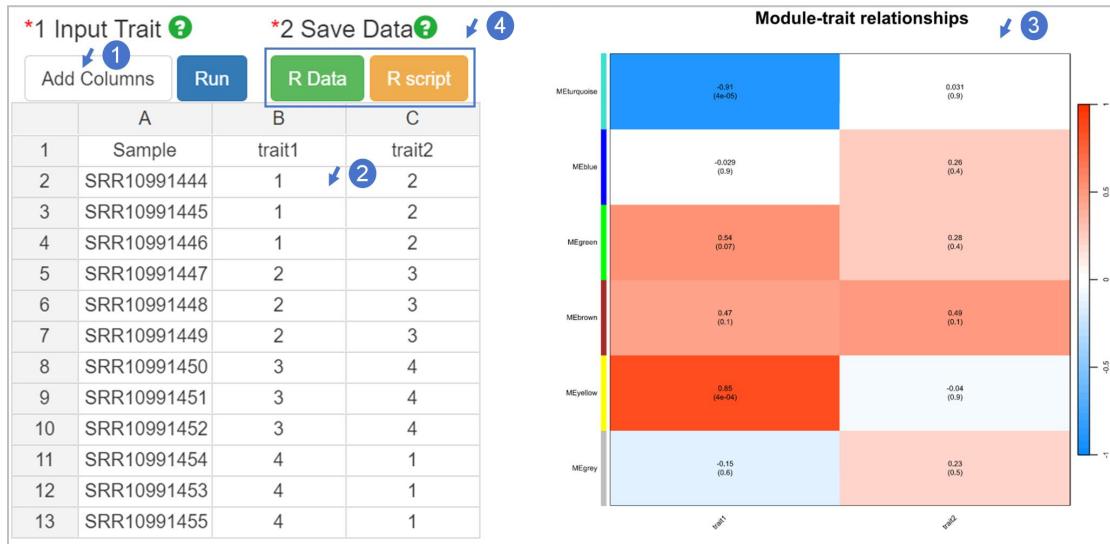
Step 2 is used to calculate the power value. Users need to set the dashed Y-axis value (①) and determine the power value by observing the intersection point with the curve (②). In Step 2, users can download the datasets generated by R language operation and R scripts to their local device for local computation (③).



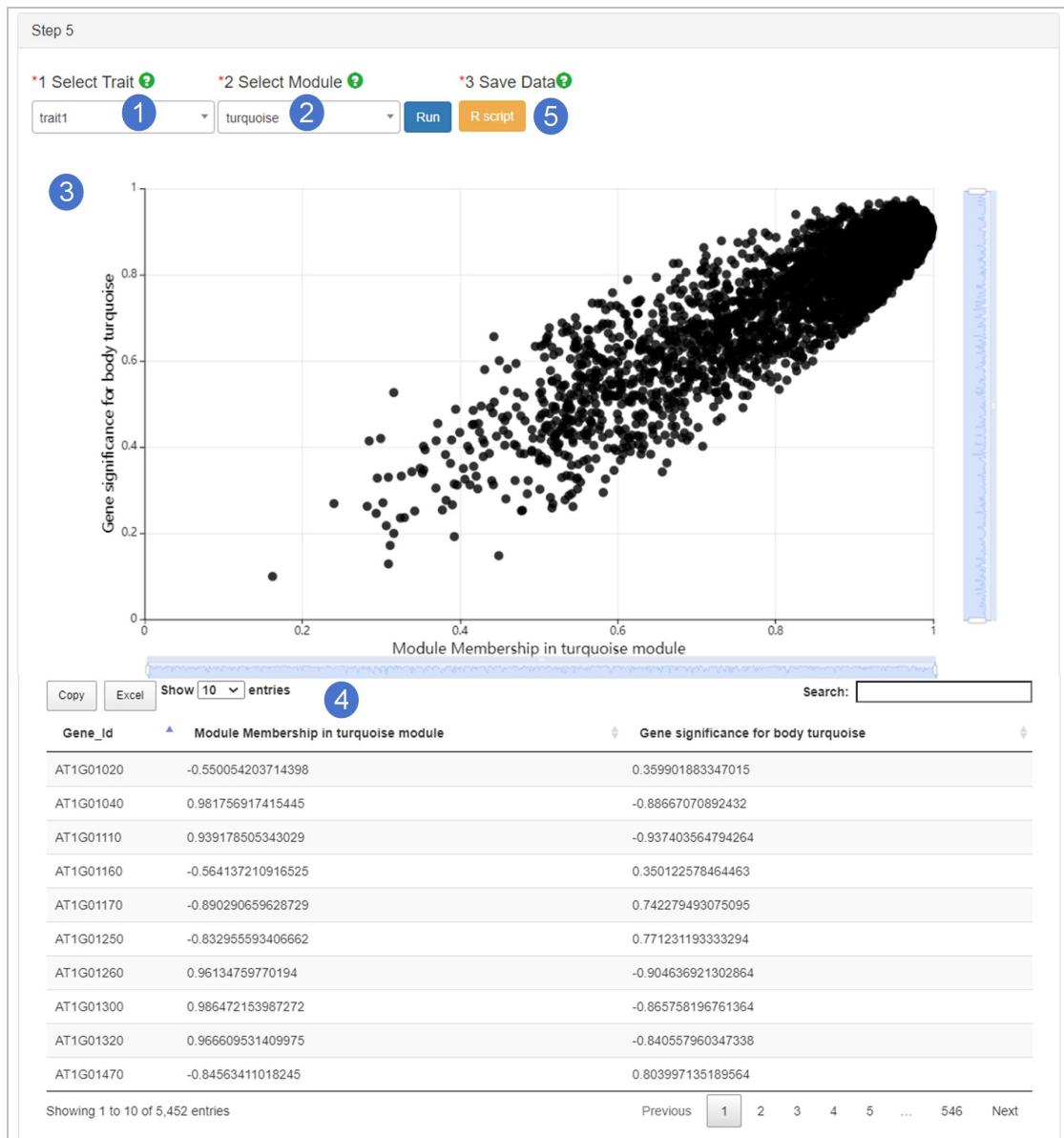
Step 3 requires users to input the power value determined in Step 2 (①), set the minimum number of module genes (②) and the threshold for module merging (③), then click Run to generate modules. The results include a module clustering plot (④) and a gene information table (⑤). Users can click on gene names to interactively view their expression trend plots (⑥). In Step 3, users can download the datasets generated by R language operation and R scripts to their local device for local computation (⑦).



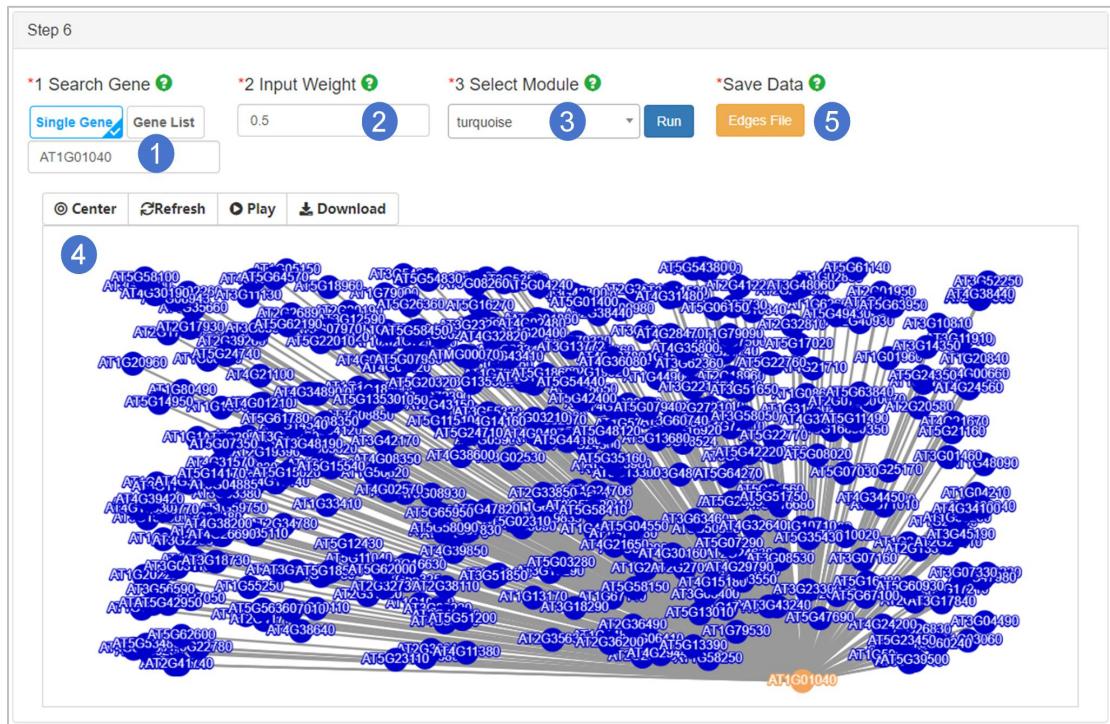
Step 4 performs correlation analysis between the module ME values generated in Step 3 and phenotypes. Users need to click the "Add Columns" button (①), add phenotypic data (②), and run the analysis. The results are presented to users in graphical form (③). In Step 4, users can download the datasets generated by R language operation and R scripts to their local device for local computation (④).



Step 5 is used to view the correlation between genes and key modules (Module Membership, MM) as well as the correlation between genes and traits (Gene Significance, GS). After users select traits (①) and modules (②), they can click Run to perform the analysis. The results are presented to users in the form of graphs (③) and tables (④). In Step 5, users can use R scripts combined with the datasets generated in Step 4 for local computation (⑤).



Step 6 is used to query genes correlated with the target gene in the module. Gene search includes two types: single-gene search and correlation calculation among multiple genes (①). Users enter the genes to be searched (①), set the threshold (②), select the module (③) to obtain a correlation network graph (④), and can click the "Edges Files" button to download relevant result files (⑤).



10. IGV BAM

The IGV BAM page allows viewing of transcriptome sequencing BAM files. Users can select BAM files from multiple samples (①), set the color for each sample's BAM file (②), choose the display type (③), and enter the target gene (④) to view the transcriptome sequencing results (⑤).



11. Quality Reports

On the Quality Reports page, users can select samples (①) to view the quality control data of their transcriptome sequencing and the alignment data with the reference genome (②). The results are presented to users in both web page and text formats (③).

Quality Reports:

① *1 Select Sample [?](#) SRR10498808

② *2 Select type [?](#) Fastp report

③ *3 Run [?](#) Run Search

fastp report ②

Summary

General

fastp version:	0.28.2 (https://github.com/OpenGene/fastp)
sequencing:	paired end (150 cycles + 150 cycles)
mean length before filtering:	148bp, 148bp
mean length after filtering:	148bp, 148bp
duplication rate:	27.986127%
Insert size peak:	240

Before filtering

total reads:	61.304082 M
total bases:	9.096059 G
Q20 bases:	8.868206 G (97.494607%)
Q30 bases:	8.496746 G (93.410879%)
GC content:	47.806937%

After filtering

total reads:	60.452410 M
total bases:	8.955220 G
Q20 bases:	8.777371 G (98.014011%)
Q30 bases:	8.423362 G (94.060911%)
GC content:	47.765446%

Filtering result

reads passed filters:	60.452410 M (98.610742%)
reads corrected:	2.488567 M (4.059382%)
bases corrected:	4.863156 M (0.050166%)
reads with low quality:	757.148000 K (1.235069%)
reads with too many N:	1.694000 K (0.002763%)
reads too short:	92.830000 K (0.151425%)

Adapters

Adapter or bad ligation of read1
The input has little adapter percentage (~0.189627%), probably it's trimmed before.

Sequence	Occurrences
A	78266
AG	79136
AA	78726

Metabolomics

Metabolite Search is used to view metabolite-related information in samples, including metabolite expression levels, categories, 2D structures, and 3D structures. Users select the project name (①), confirm the sample name (②), enter a list of metabolite names (③), and click the "Run Search" button to view the results. The results include two parts: metabolite content and compound information. Metabolite content is presented to users in the form of a heatmap (④) and a table (⑤), and users can directly download the table data. Compound information is displayed to users in a table, including metabolite categories, structural formulas, and molecular formulas (⑨); when users click on a compound name, the source of the compound information, SMILES formula (⑥), 2D structure diagram (⑦), and 3D structure diagram (⑧) are interactively shown.

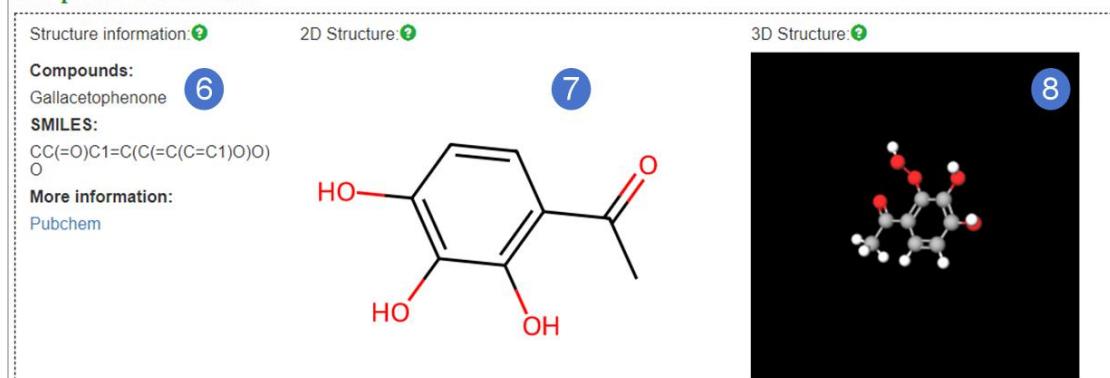
Metabolite search:

*1 Select Project ?	*3 Input Metabolites ?
Project_1 1	3 Run Search
*2 Select Sample ?	
300_3d_S_3 x 300_3d_T_2 x 300_3d_T_1 x 300_3d_T_3 x 300_5d_S_2 x 300_5d_S_3 x 300_5d_T_2 x 300_5d_T_1 x 300_7d_S_1 x 300_7d_S_2 x	2

Metabolite content:



Compounds information:



9

Project_Index	Compounds	Class I	Class II	Formula	Molecular weight (Da)
Hahp000801	alanine betaine	Alkaloids	Alkaloids	C5H11NO2	117.079
Hmbp001276	Gallacetophenone	Others	Ketone compounds	C8H8O4	168.0423
Hmcp002539	Isorhamnetin-3-O-rhamnoside	Flavonoids	Flavonols	C22H22O11	462.1162
HX1367	coumarin	Lignans and Coumarins	Coumarins	C9H6O2	146.0368

Showing 1 to 4 of 4 entries 1 row selected Previous 1 Next

1. Project information

Project Information consists of three parts: the project information table (①), the compound information table (②), and the compound search function (③). Users can view detailed information for each sample in the project in the project information table, including 13 data points such as sample name, sample group, sample details, and sequencing platform. In the compound information table, users can check the SMILES formula and corresponding links of each metabolite. In addition, users can search for the project(s) where a specific compound is located using the compound search function.

Project information:

Project_Name	Sample_Group	Sample_Name	Sample_Information	Tissue	Replicate	Instrument	Germplasm	Germplasm_Character
Project		300_3d_S_3						
Project		300_3d_S_2						
Project		300_3d_S_1						
Project		300_3d_T_3						
Project		300_3d_T_2						
Project		300_3d_T_1						
Project		CK_3d_S_3						
Project		CK_3d_S_2						
Project		CK_3d_S_1						
Project		CK_3d_T_3						

Showing 1 to 10 of 12 entries Previous [1] [2] Next

Compounds information:

Compounds	SMILES	Link	Note
coumarin	C1=CC=C2C(=C1)C=CC(=O)O2	https://pubchem.ncbi.nlm.nih.gov/compound/323	
Isorhamnetin-3-O-rhamnoside	C[C@H]1[C@H]([C@H]([C@H]([C@H](O1)OC2=C(OC3=CC(=CC(=C3C2=O)O)O)C4=CC(=C(C=C4)O)OC)O)O	https://pubchem.ncbi.nlm.nih.gov/compound/23634491	
Gallacetophenone	CC(=O)C1=C(C(=C=C1)O)O	https://pubchem.ncbi.nlm.nih.gov/compound/10706	
alanine betaine	CC(C(=O)[O-])[N+](C)C	https://pubchem.ncbi.nlm.nih.gov/compound/11355498	
3,4'-Dihydroxy-3',5'-dimethoxypropiphenone	COC1=CC(=CC(=C1O)OC)C(=O)CCO	https://pubchem.ncbi.nlm.nih.gov/compound/54353627	

Showing 1 to 5 of 5 entries 1 row selected Previous [1] Next

Search compounds in project:

Compounds	Project_name
Gallacetophenone	Project
Isorhamnetin-3-O-rhamnoside	Project

Run Search Search: _____

Showing 1 to 2 of 2 entries Previous [1] Next

2. DEGs

DEGs is used for differential analysis between different groups in metabolomics. After users select the project (①), groups, and samples (②) to perform the analysis, the background will automatically round the metabolite expression levels and use the DESeq2 package for differential analysis. The results are presented to users in two forms: a volcano plot (④) and a table (⑤). Users can click on the points in the volcano plot to interactively view the difference value and type information of the corresponding metabolite, and download the differential analysis results from the table to their local device.

Metabolite DEGs:

*1 Select Project ①
Project

*2 Select sample group ②
300_3d_S
300_3d_S_3 300_3d_S_2 300_3d_S_1

*3 Select sample group ③
CK_3d_S
CK_3d_S_3 CK_3d_S_2 CK_3d_S_1

*4 Run ④
Run DEG

Volcano plot:

300_3d_S VS CK_3d_S Result:

Log10(pvalue)

Log2(FoldChange)

Compound: Sucrose-6-phosphate

log2FoldChange: 4.04

-log10(pvalue): 143.01

Class I: Others

Class II: Saccharides

Table Result:

Compounds	baseMean	log2FoldChange	IfcSE	stat	pvalue	padj
3'-Glucosyl-6,7-dihydroxy-N-methylbenzyltetrahydroisoquinoline	128800.974223093	-0.699317106619782	0.165145390457811	-4.2345541990676	2.2900549700771e-05	0.00050736018264
3,4,5-Tricaffeoylquinic acid	321898.298202549	2.07954140818561	0.215465071745131	9.651408422454	4.84851800080124e-22	3.81335940763016
3,4-Dimethoxycinnamic acid	84116.1450692194	-1.49074257933775	0.24589718631903	-6.06246294092868	1.34052702328355e-09	5.02059287529769
3-[4-(Beta-D-Glucopyranosyloxy)Phenyl]Propionic Acid	446808.582754942	0.909562121602942	0.277441254250642	3.27839536358655	0.00104399061187669	0.01440523888142
3-Carbamyl-1-methylpyridinium;(1-Methylnicotinamide)	3356679.05179081	0.697540548028578	0.0924593564260244	7.54429378476878	4.54744970361286e-14	2.4665994426838e
3-Oxopomolic acid	34938.7838192387	-0.820300866875456	0.245929506638236	-3.33551218838544	0.00085142421403708	0.01216255828347
3a-Dihydrocadambine	23296.0445546381	-5.20178723838547	0.447387789875921	-11.6270210231445	3.00398796681082e-31	3.9377275598278e
3 β ,6 β ,19 α ,24-Tetrahydroxyurs-12-en-28-oic acid	21834.415442502	2.57025164031568	0.159847130798717	16.0794355674247	3.55604065390648e-58	6.2151688317721e
4-Hydroxy-L-glutamic acid	29863.5677330432	0.746063128235055	0.205813445639612	3.62494843772958	0.00028901911011361	0.00483645808732
4-Hydroxyacetophenone	202104.175388743	-0.789954040805983	0.172921688015342	-4.56827625194069	4.91751634828101e-06	0.00012086333149

3. Mfuzz

For metabolome Mfuzz analysis, users need to first select a project name; the remaining steps are the same as those for transcriptome Mfuzz analysis.

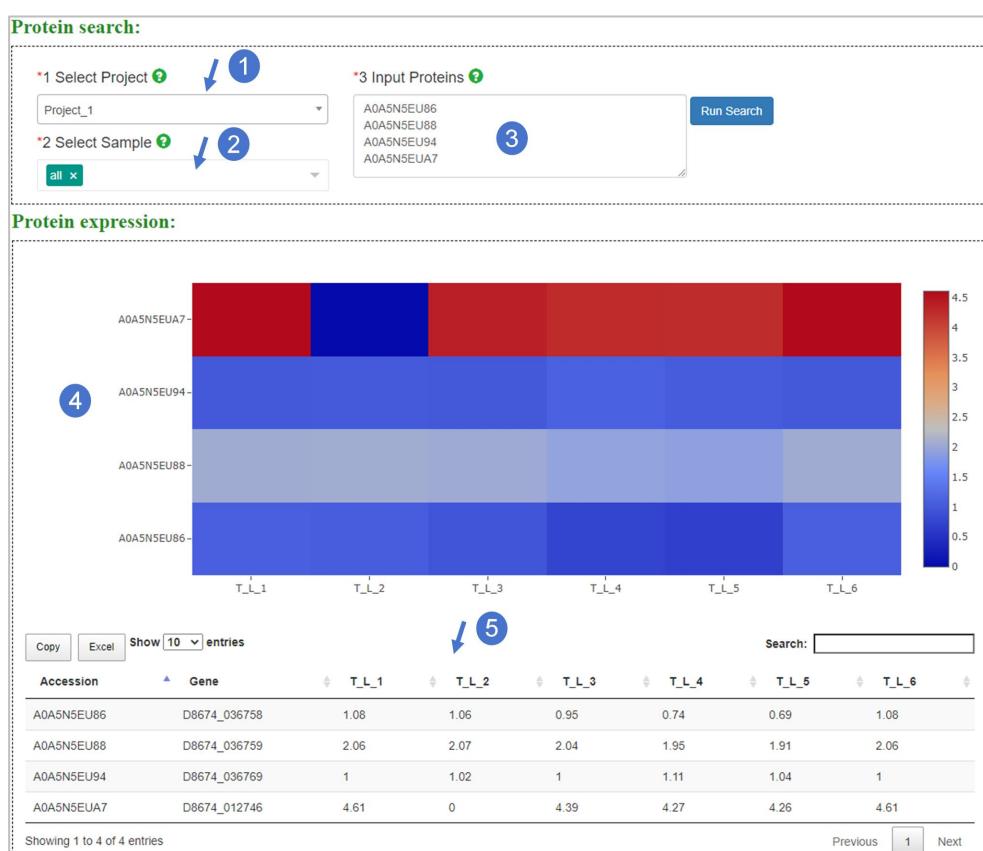
4. PCA

For metabolome PCA analysis, users need to first select a project name; the remaining steps are the same as those for transcriptome Mfuzz analysis.

Proteomics

1. Protein search

Protein Search is used to retrieve protein-related information, including protein expression levels, annotation information, and protein sequences. Users select the project name (①), sample number (②), enter the protein ID (③), and click the "Run Search" button to view the results. The results consist of three parts: protein content (with a table), protein annotation information, and protein sequences. Users can download gene expression levels and annotation information from the table to their local device. Protein content is presented to users in the form of a heatmap (④) and a table (⑤), and users can directly download the table data. Protein annotation information is displayed to users in a table (⑥), including GO, KEGG, IPR, and Description. In addition, users can directly copy the protein sequence to their local device (⑦).



Protein Annotation: 6

Copy
Excel
Show 10 entries
Search:

Accession	Gene	GO	KEGG_orthologs	KEGG_pathways	IPR	Description	note
A0A5N5EU86	D8674_036758	GO:0010181 FMN binding; GO:0016491 oxidoreductase activity	K05894	ko00592: alpha-Linolenic acid metabolism; ko01100: Metabolic pathways; ko01110: Biosynthesis of secondary metabolites	IPR001155: NADH:flavin oxidoreductase/NADH oxidase, N-terminal; IPR013785: Aldolase-type TIM barrel	12-oxophytodienoate reductase 11 OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_036758 PE=3 SV=1	--
A0A5N5EU87	D8674_012746	GO:0004070 aspartate carbamoyltransferase activity; GO:0016597 amino acid binding	K00609	ko00240: Pyrimidine metabolism; ko00250: Alanine, aspartate and glutamate metabolism; ko01100: Metabolic pathways; ko01240: Biosynthesis of cofactors	IPR036901: Aspartate/ornithine carbamoyltransferase superfamily; IPR006130: Aspartate/ornithine carbamoyltransferase; IPR002082: Aspartate carbamoyltransferase; IPR006132: Aspartate/ornithine carbamoyltransferase, carbamoyl-P binding; IPR006131: Aspartate/ornithine carbamoyltransferase, Asp/Orn-binding domain	aspartate carbamoyltransferase OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_012746 PE=3 SV=1	--
A0A5N5EU94	D8674_036769	GO:0008168 methyltransferase activity	K18162	--	IPR029063: S-adenosyl-L-methionine-dependent methyltransferase superfamily; IPR013216: Methyltransferase type 11	Methyltransferase OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_036769 PE=4 SV=1	--
A0A5N5EU88	D8674_036759	--	--	--	--	Uncharacterized protein OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_036759 PE=4 SV=1	--

Showing 1 to 4 of 4 entries

Previous 1 Next

Protein Sequence: 7

```
>A0A5N5EU86
MATTQTCVGGPTSPLLAPYKLGNFNLSHRVVLAFLTRQRSYGNVPQPHAVLYSQRTSNGGLVSEATGVSDTAQGYPDTPIWTKTEQVEAWKPIVDAVAKGGIFCQIWHGRVSN
TGSPQNGQAPICTDKSLTSETPNSGVYAIIFSPPRLRTDEVPIQVNDFLRAARNAMEAGFDGVEIHGAHGLVDQFLKDHNDRDQYGGSVENRCRFALEIVEANEIGADKVGIRL
SPFADFMDSGDSNPKEGLCMANSLNKYIGLYCHMVPERITIAAEKRDYSHSLVPMRKAFTGNTIAAGGYKEDGNNALAEGRADIAYGRIFLANDPLKRFLETVPLNKVSQVGLIGPDL
PFTYNYGVWEDEFEGVLYLDSTVESIVEASNGISHVARGHNIVVSCRYKGITLATVASAAAASGKLLOYKVGGPKVAVQTAYGIEVEDW/SWIPVGGSPLNTEOKNLAFAAACMVHPATAIVG
GVSHANQIDRLARTVCCDLHILSAHSLTCKEYTRKKGNALFGLALIQQDTEGIMTFCTFFCLSTWMWQGFLGSTLSSAVLIGFALSRSENGLTAYNRSMI
>A0A5N5EU88
MMGRRVGSAYVVLYLVCLFLVSNLGDYVSDSESGEYTSIAGDPMGRNPNVRALEAWNFCNEVGSEAPGMGSPLRADCADIHCPTVTTEPMDRISLDVSKCVCVLRHVNESDNILR
AGDMFPLRDFKSYTDPDLVATEKELYACLCLEVHNSSDPW/GFVMVMLKNGNFDKNTLCPENGKHLISKVTDNRPFCFEGGCMNQPLVYHNSRSLVSDGYQKLSVTGFFYGT/DLDAD
LTKIGIRNSYFSVSWQKNLSTSWSWIFTSRQLTSSKYPWLMLYLRADATKGFNGGYHYDGRGVMRKLPVSPHFKVRLLTDVKRGGGPNQSFYMLDIGSCW/KNNNGKPCDGNVLTDVTRYSE
EMINPETTSWCRPDNLVSCPYPHTVNGDITYNETSRFPYSAHYCAPGNAEYLEKPYDICDPYSPNQAEQELVQILPHPEWAVHGPGKQGNGWVGDPRTWELDVGAASSRLFYQ
DPGTPKAERVWSSINVGETIYVSNSGETAEWTVSDFDVLV/PEGVEGGYSSY
>A0A5N5EU94
MRGLYRRCSLWRTGRAPHSLLFSNACLTDTNNIDGVQGSRIKVFRHRFRKQLDRAAWLMRSKDSFVDAVAENILLDRLEDCKKTFPTALCLGGQLEAVRRLLHGRGAIKEKLMMSASY
DMVKLKRINAERELQTNETISFMVGEELFPLVKESVSDLVLSCGLHW/TNDLPGMAMQCRLLALKPDGLFLAAILGGETLKELRIACTVAQMEREGRISPRVSPLAQVRDAGNLTRAGFTL
PGVGDVDFEVRYPSALDLIELHRLAMGETNALLQRNTILKRDATALAAYDSMFATEDGTIPATFOVIVYMTGW/RDHPSQRRAKRGSATVSFQDIQKQFGSET
>A0A5N5EU87
MAASSSLFSSSSFOGVMSFQSTKLCASSSNFVYKQPMHLKRSSQNGNTVSCRALEVDKSSSSYVGKEFOVTDVIEAQQFDRDLGAIIFDVAREMEKIEKKSPGSQILKGYLMLTFYEP
STRTRLSFESAMRRLGGEVLTTENAREFSSAAKGETLEDTIRTEVGYSDIVVMRHFESEAARKAAATAGIPVIAGDGPQHQHPTQALLDVYTIERIEGKLDGKVALVGDLANGRTVSLAYL
LARYNDVKVYFVSPDVVKMNDIKDYLTSRNQIWEESADLMEVASKCDVYYQTRIQRERGERIDLYEEARGKYIVDSRVLEVQMKAHVMMHPLRDLDEITVVDGDTRAAYFRQAQKGLYI
RMALLKLLLGVW
```

2. Project information

Project Information consists of two parts: the project information table (①) and the protein annotation information table (②). Users can view detailed information for each sample in the project in the project information table, including 13 data points such as sample name, sample group, sample details, and sequencing platform. Clicking on the project name allows interactive viewing of the corresponding protein

annotation information.

Project information:

Project_Name	Sample_Group	Sample_Name	Sample_Information	Tissue	Replicate	Instrument	Germplasm	Germplasm Characteristics	G
Project_1		T_L_4							L
Project_1		T_L_6							
Project_1		T_L_5							
Project_1		T_L_3							
Project_1		T_L_2							
Project_1		T_L_1							

Showing 1 to 6 of 6 entries Previous **1** Next

Protein Annotation:

Accession	Gene	GO	KEGG_pathways	KEGG_orthologs	IPR	Description	not
A0A5N5EU88	D8674_036758	GO:0010181 FMN binding;GO:0016491 oxidoreductase activity	ko00592: alpha-Linolenic acid metabolism;ko01100: Metabolic pathways;ko01110: Biosynthesis of secondary metabolites	K05894	IPR001155: NADH:flavin oxidoreductase/NADH oxidase, N-terminal;IPR013785: Aldolase-type TIM barrel	12-oxophytodienoate reductase 11 OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_036758 PE=3 SV=1	--
A0A5N5EU88	D8674_036759	--	--	--	--	Uncharacterized protein OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_036759 PE=4 SV=1	--
A0A5N5EU94	D8674_036769	GO:0008168 methyltransferase activity	--	K18162	IPR029063: S-adenosyl-L-methionine-dependent methyltransferase superfamily;IPR013216: Methyltransferase type 11	Methyltransferase OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_036769 PE=4 SV=1	--
A0A5N5EUA7	D8674_012746	GO:0004070 aspartate carbamoyltransferase activity;GO:0016597 amino acid binding	ko00240: Pyrimidine metabolism;ko00250: Alanine, aspartate and glutamate metabolism;ko01100: Metabolic pathways;ko01240: Biosynthesis of cofactors	K00609	IPR036901: Aspartate/ornithine carbamoyltransferase superfamily;IPR006130: Aspartate/ornithine carbamoyltransferase;IPR002082: Aspartate carbamoyltransferase;IPR006132: Aspartate/ornithine carbamoyltransferase, carbamoyl-P binding;IPR006131: Aspartate/ornithine carbamoyltransferase, Asp/Orn-binding domain	aspartate carbamoyltransferase OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_012746 PE=3 SV=1	--
A0A5N5EUC7	D8674_036798	GO:0009055 electron transfer activity;GO:0020037 heme binding;GO:0046872 metal ion binding	ko00190: Oxidative phosphorylation;ko01100: Metabolic pathways	K08738	IPR009056: Cytochrome c-like domain;IPR036909: Cytochrome c-like domain superfamily;IPR002327: Cytochrome c, class IA/ IB	Cytochrome c OS=Pyrus ussuriensis x Pyrus communis OX=2448454 GN=D8674_036798 PE=3 SV=1	--

Showing 0 to 0 of 0 entries (filtered from 10,989 total entries) First Previous **1** 2 3 4 5 ... Next Last

3. DEGs

For DEGs analysis, refer to Section 4.3

4. Mfuzz

For Mfuzz analysis, refer to Section 4.4

5. PCA

For PCA analysis, refer to Section 5.5

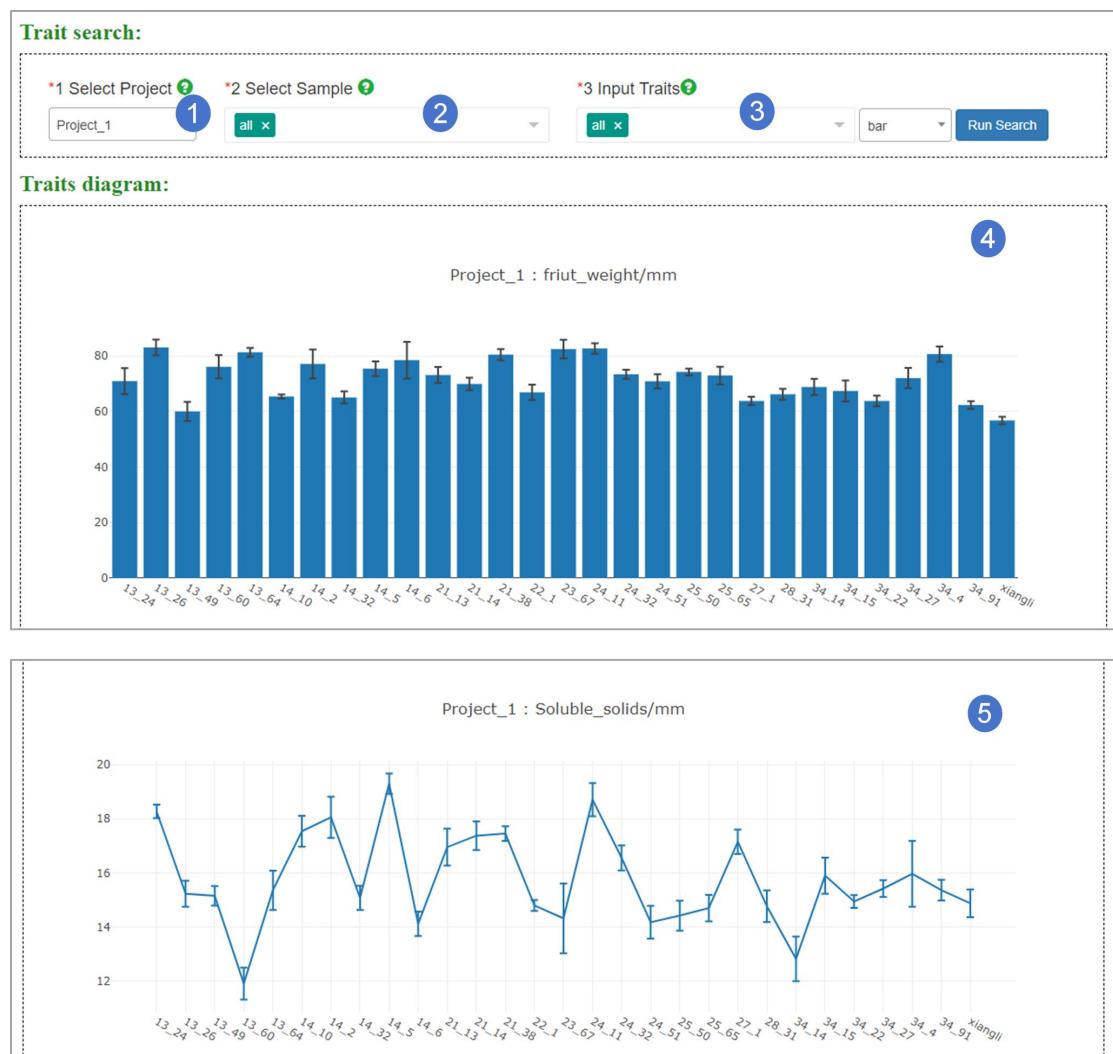
6. Enrichment

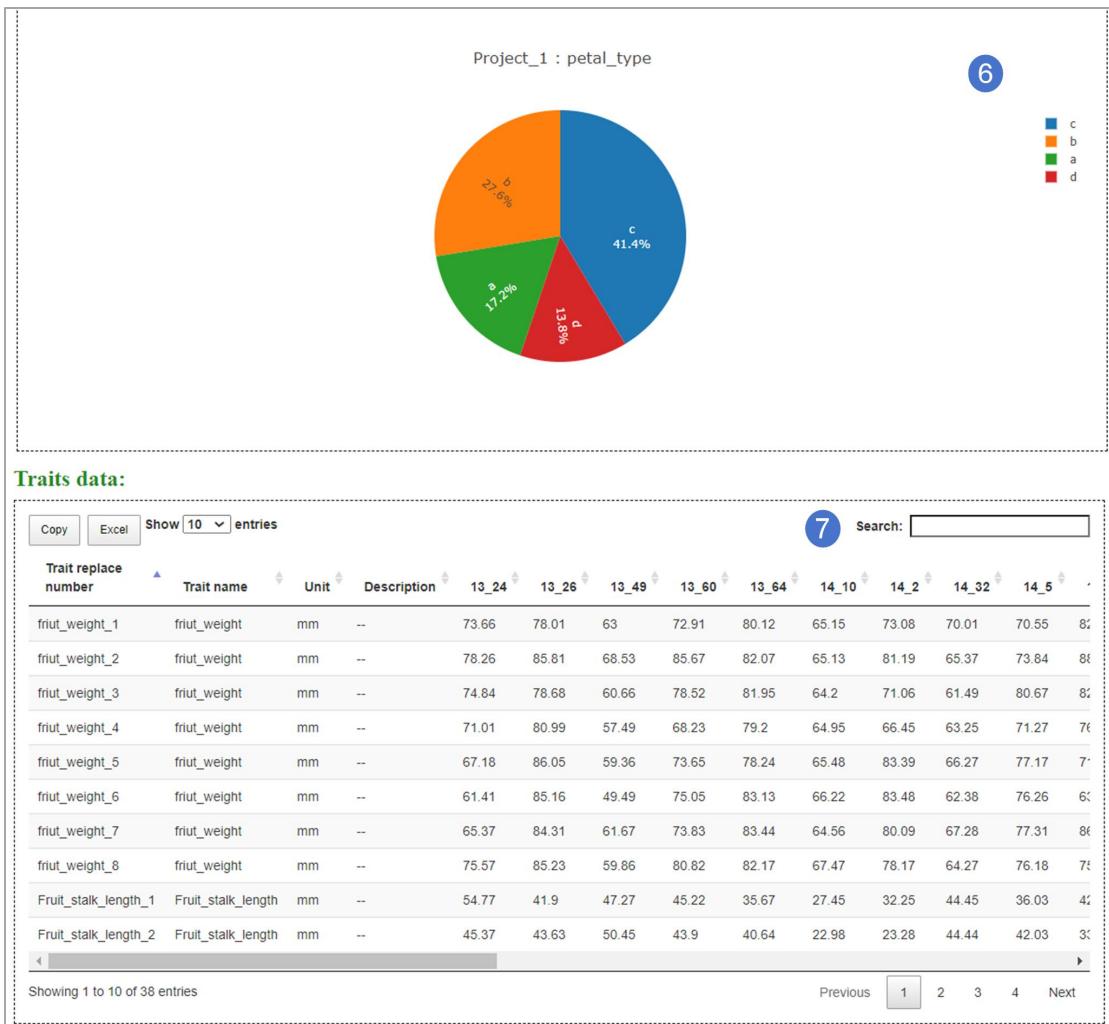
Protein enrichment analysis is similar to that in transcriptome (Section 3.3), except that users need to select the project name in advance.

Phenotype

1. Trait search

Trait Search is used to retrieve and view the phenotypic data of a project. Users select the project name (①), sample number (②), and phenotype (③), then click the "Run Search" button to view the results. The results include two parts: graphs and a table. In the graph results, quantitative traits are displayed in the form of bar charts (④) and line charts (⑤), while qualitative traits are displayed as pie charts (⑥). Users can download the phenotypic data from the table (⑦).





2. Project information

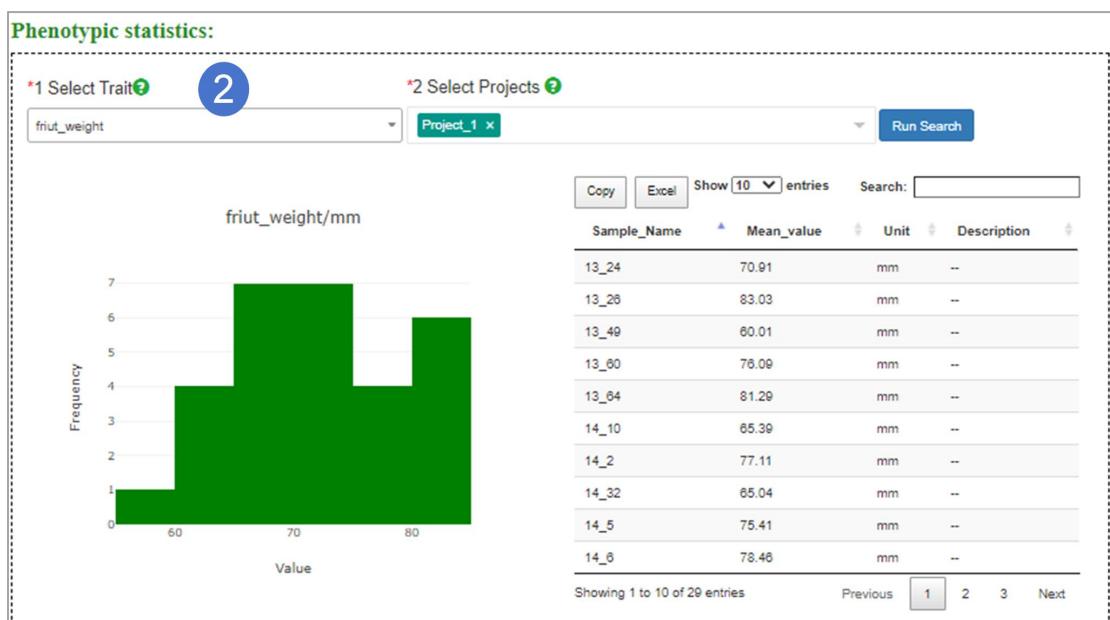
Project Information consists of two parts: the project information table (①) and phenotypic statistics (②). Users can view detailed information for each sample in the project through the project information table, including sample name, sample group, germplasm resources, sample details, and location information. Additionally, users can check the frequency distribution charts of quantitative traits they are interested in across single or multiple projects, as well as the pie charts of qualitative traits across single or multiple projects.

Project information:

Project_Name	Sample_Group	Sample_Name	Cultivar	Tissue	Geographic_Location	note
Project_1		xiangli		fruit;flower		
Project_1		34_14		fruit;flower		
Project_1		27_1		fruit;flower		
Project_1		23_87		fruit;flower		
Project_1		21_38		fruit;flower		
Project_1		21_14		fruit;flower		
Project_1		14_32		fruit;flower		
Project_1		14_10		fruit;flower		
Project_1		14_5		fruit;flower		
Project_1		34_91		fruit;flower		

Showing 1 to 10 of 20 entries

Previous 1 2 3 Next



3. Correlation Project

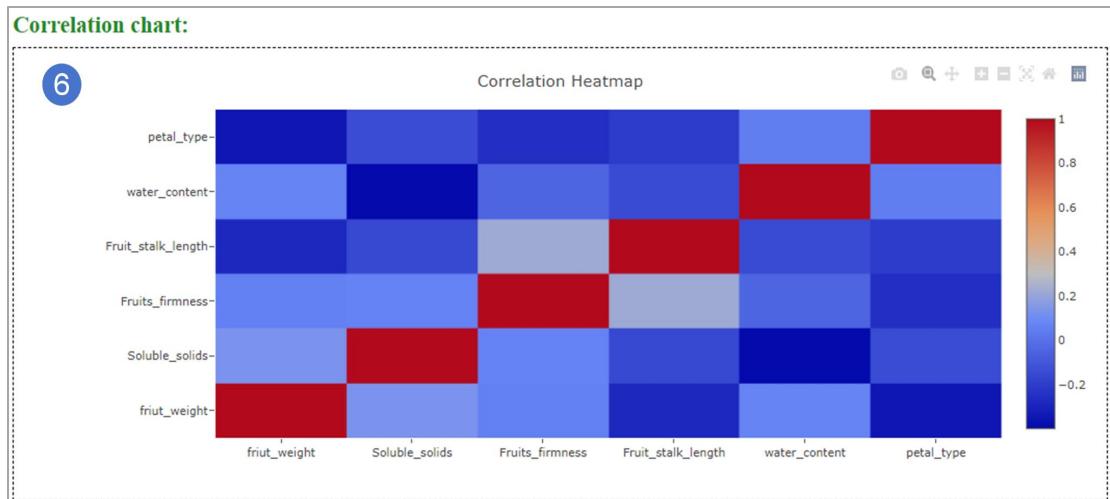
Correlation Project is used for correlation analysis between phenotypic data within a single project. Users select the project name (①), samples (②), and traits (③), set standardization (④) and the correlation calculation model (⑤), then click Run to obtain the results. The results include three parts: a correlation heatmap (⑥), the mean values of sample phenotypic data (⑦), and a correlation coefficient matrix (⑧).

Users can download the relevant data from the table.

Correlation:

*1 Select Project 1 *2 Select Sample 2 *3 Input Traits 3 *4 Z-Score 4 *5 Select type 5

Project_1 1 all x 2 all x 3 4 NO S pearson 5 Run



Traits mean data:

7

Trait	Unit	13_24	13_26	13_49	13_60	13_64	14_10	14_2	14_32	14_5	14_6	21_13	21_14	21_38
fruit_weight	mm	70.91	83.03	60.01	76.08	81.29	65.40	77.11	65.04	75.41	78.46	73.11	69.88	80.46
Fruit_stalk_length	mm	51.55	42.97	48.21	42.43	35.25	28.52	27.68	40.28	37.85	37.78	45.82	47.30	39.43
Fruits_firmness	mm	6.46	5.73	3.80	3.79	3.33	3.27	3.71	4.37	3.75	4.51	2.79	3.48	4.23
petal_type	quality	1.00	1.00	1.00	1.00	1.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
Soluble_solids	mm	18.28	15.23	15.16	11.91	15.36	17.54	18.06	15.08	19.30	14.12	16.96	17.38	17.46
water_content	mm	78.45	84.14	82.66	84.98	83.95	81.47	82.57	84.98	79.56	83.06	80.60	83.09	83.34

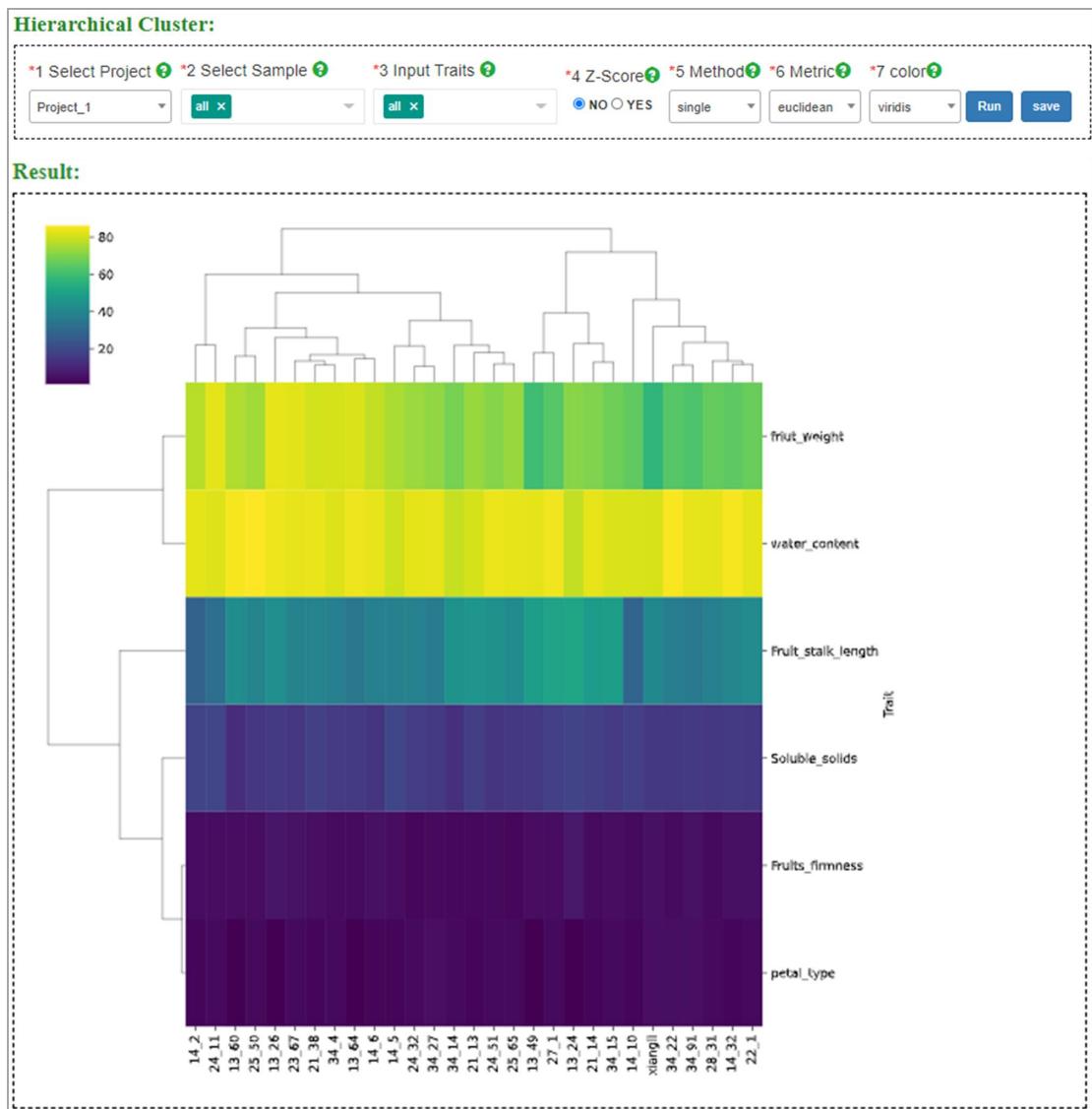
Correlation matrix:

8

Trait	fruit_weight	Soluble_solids	Fruits_firmness	Fruit_stalk_length	water_content	petal_type
fruit_weight	1.00	0.13	0.06	-0.28	0.07	-0.35
Fruit_stalk_length	-0.28	-0.16	0.22	1.00	-0.15	-0.20
Fruits_firmness	0.06	0.07	1.00	0.22	-0.05	-0.25
petal_type	-0.35	-0.14	-0.25	-0.20	0.05	1.00
Soluble_solids	0.13	1.00	0.07	-0.16	-0.40	-0.14
water_content	0.07	-0.40	-0.05	-0.15	1.00	0.05

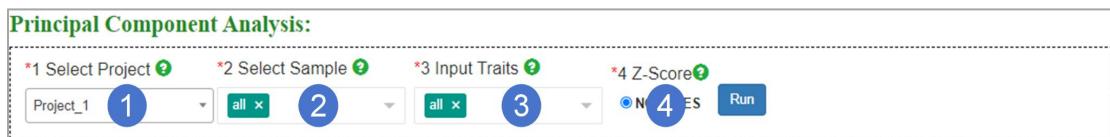
4. HCA Project

HCA Project is used for cluster analysis of phenotypic data within a single project. Users select the project name (①), samples (②), and traits (③), set standardization (④), the model (⑤), the Metric parameter (⑥), and colors (⑦), then click Run to obtain the phenotypic cluster diagram as the result.



5. PCA Project

PCA Project is used for principal component analysis (PCA) of phenotypic data within a single project. Users select the project name (①), samples (②), and traits (③), set standardization (④), then click Run to obtain the results. The results include three parts: the contribution rate and eigenvalues of principal components (⑤), and heatmaps and table data showing the contribution of different traits to principal components (⑥). Users can download the relevant data from the table.

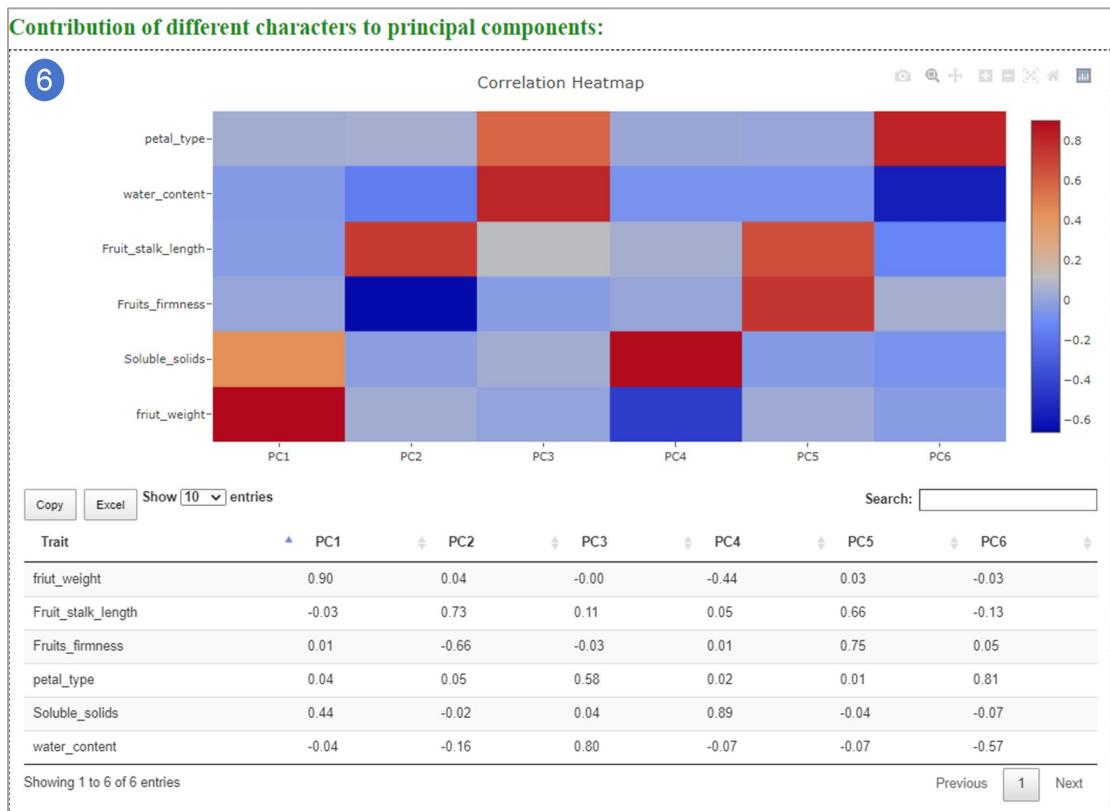


Eigenvalues, contribution rate: 5

Item	PC1	PC2	PC3	PC4	PC5	PC6
contribution rate	0.62	0.30	0.05	0.02	0.01	0.01
eigenvalues	58.78	28.74	4.45	1.75	0.82	0.59

Showing 1 to 2 of 2 entries 1 row selected

Previous 1 Next



6. Correlation Traits

For Correlation Traits analysis, refer to Section 43

7. HCA Traits

For HCA Traits analysis, refer to Section 44

8. PCA Traits

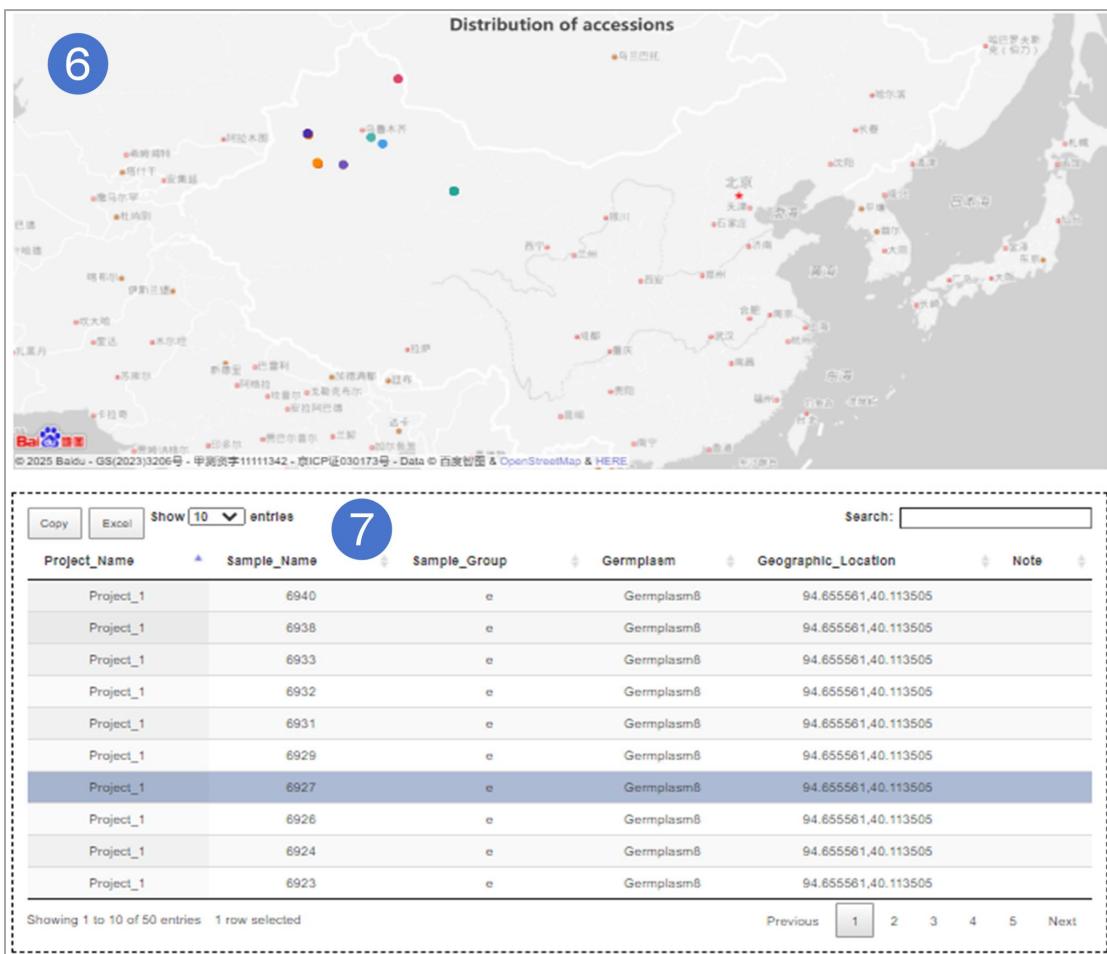
For PCA Traits analysis, refer to Section 45

Population

1. Accession

Accession is used to search and display the geographical distribution map of different resources in the project. Users select the type (①), project (②), and data (③), set the color (④) and graph size (⑤), then click Run to obtain the results. The results include graphic information (⑥) and table information (⑦) of the geographical distribution of

resources.



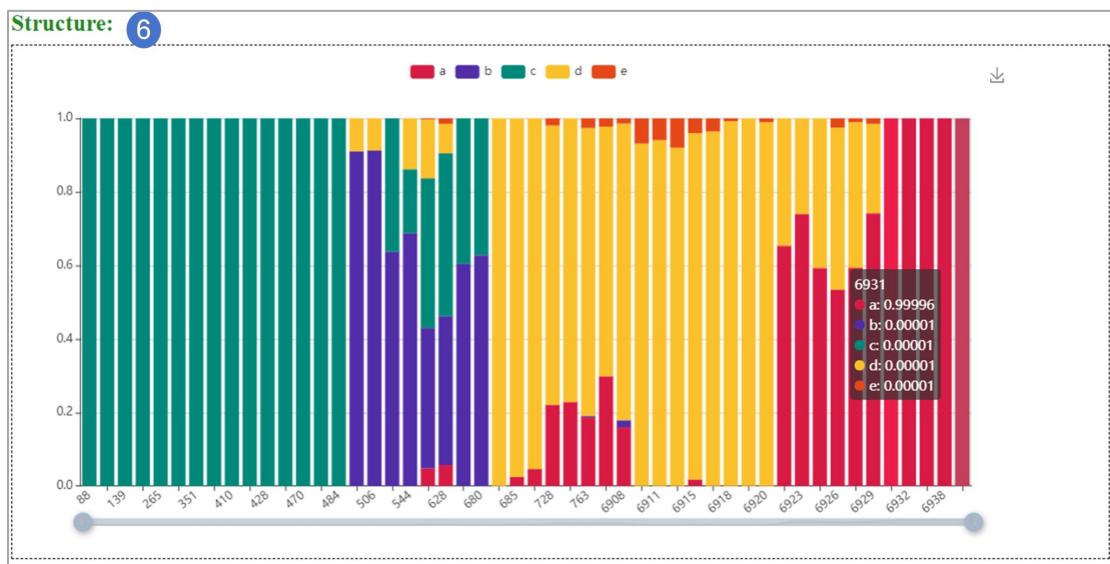
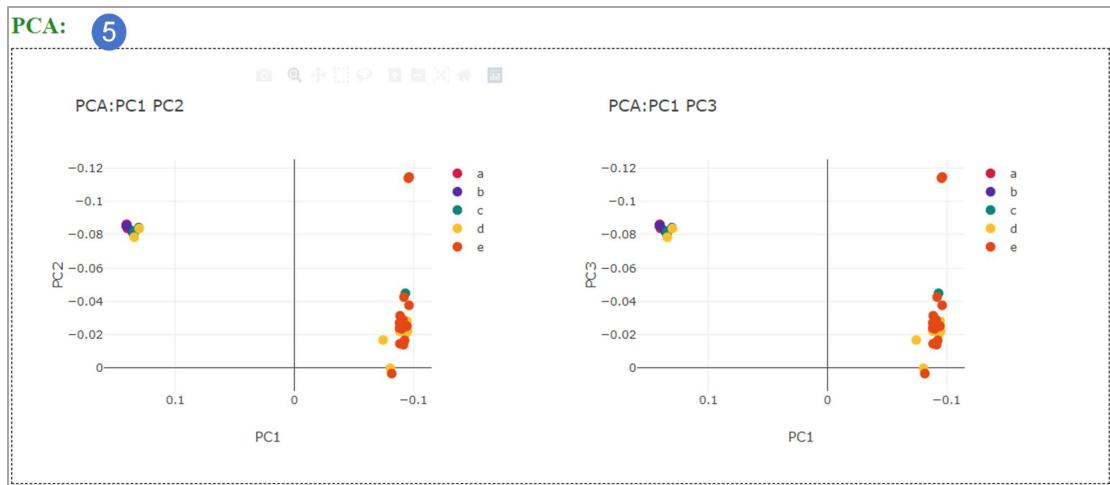
2. Population

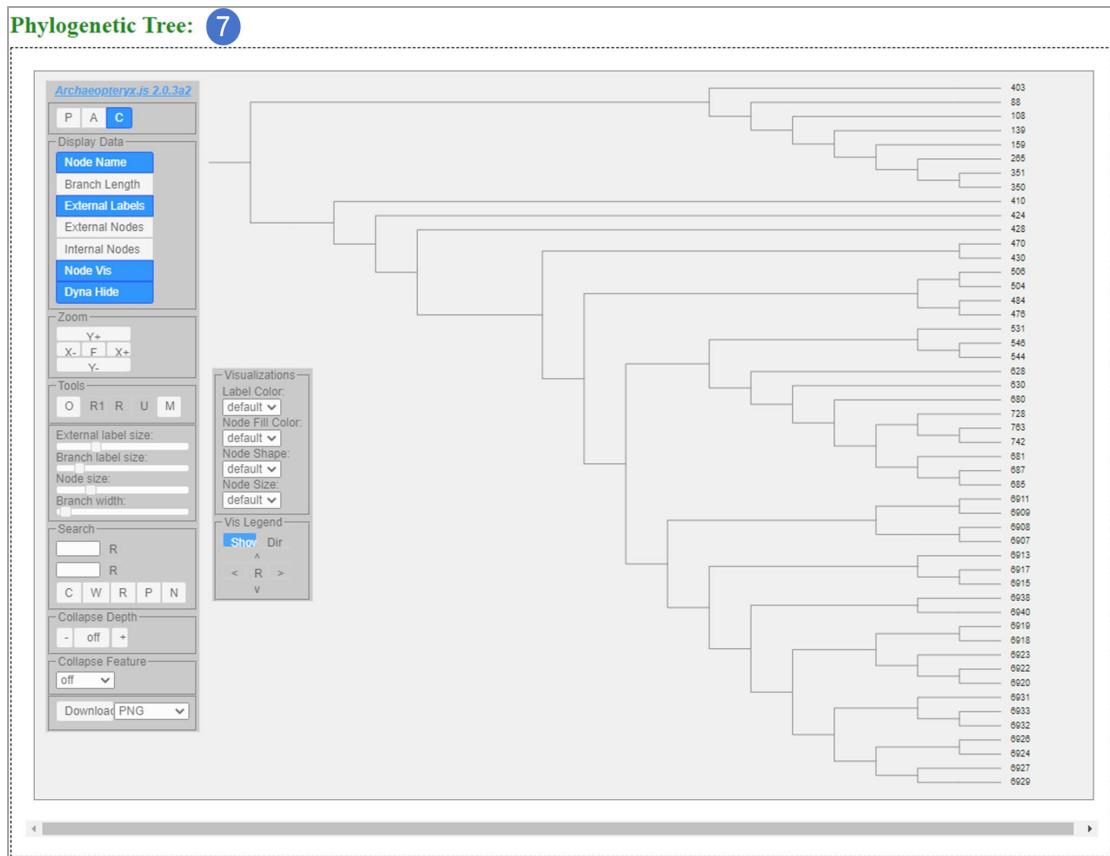
Population is used to search and display the population structure diagram in the project. Users select the project name (①), set the color (②), graph size (③), and visualization range (④), then click Run to perform the analysis and obtain the results. The results include three parts: the population cluster diagram (⑤), the population structure diagram (⑥), and the phylogenetic tree (⑦).

Population:

*1 Select Project  *2 Color setting  *3 Set Size  *4 Scope Expansion 

Project_1	1	Run
2	10	0.05
e: 	3	4
d: 		
c: 		
b: 		
a: 		





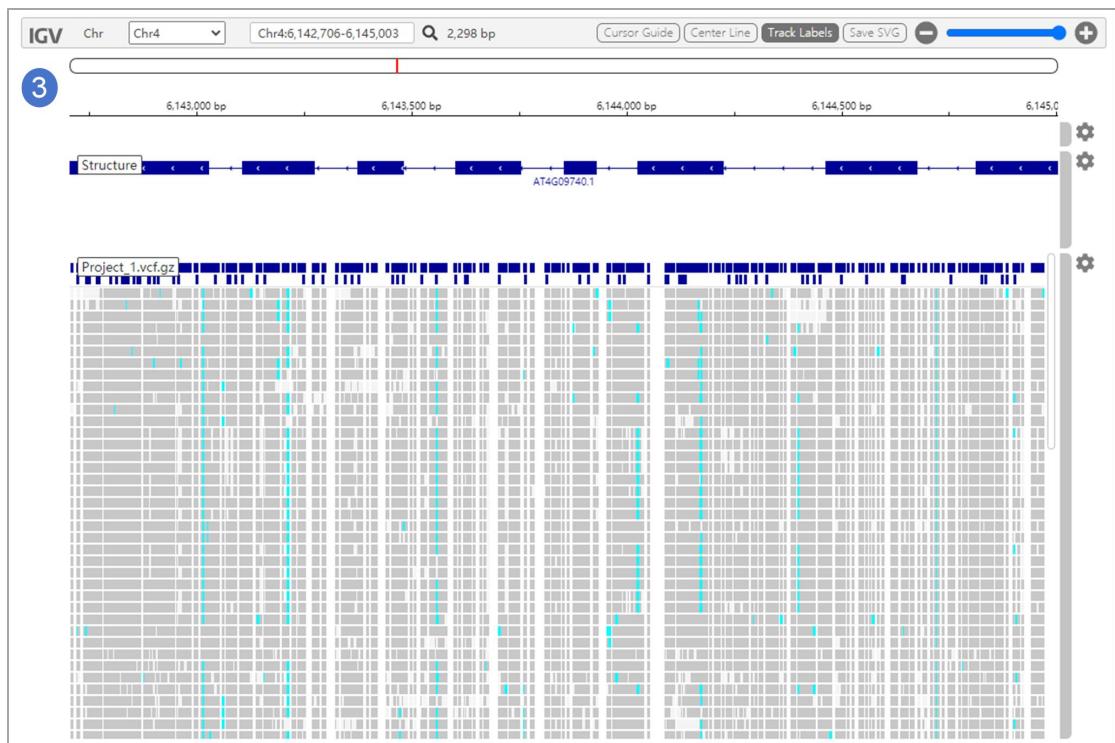
3. VCF

VCF is used to search for variation information of target genes. Users select the project name (①), enter the gene ID (②), and click Search to obtain the results. The results are interactively displayed to users in graphic form (③).

Reference Genome: Athaliana

*1 Select Project ① *2 Input Gene ID ②

Project_1.vcf.gz AT4G09740 search



4. GWAS Result

GWAS Result is used to interactively view GWAS analysis result data. Users select the project name (①), the GWAS analysis result file (②), and the chromosome number (③), set the color (④), threshold (⑤), and the upstream and downstream regions of the target locus (⑥, ⑦) (for calculating linkage disequilibrium), then click Run to obtain the results. The results include three parts: the Manhattan plot of GWAS analysis results (⑧), the linkage disequilibrium plot (⑨), and the GWAS result table (⑩). When users click on a point in the Manhattan plot, the corresponding linkage disequilibrium plot for that locus will be interactively generated and displayed.

GWAS View:

*1 Select Project ? Project_1 1

*2 GWAS Result Name ? Project_1_GWAS_Result 2

*3 Select Chr ? Chr1 ✕ Chr2 ✕ Chr3 ✕ 3

Chr1 ✕ Chr2 ✕ Chr3 ✕
Chr4 ✕ Chr5 ✕ ChrM ✕
ChrC ✕

*4 Color setting ? Chr1:
Chr2:
Chr3:
Chr4:
Chr5:
ChrM:
ChrC: 4

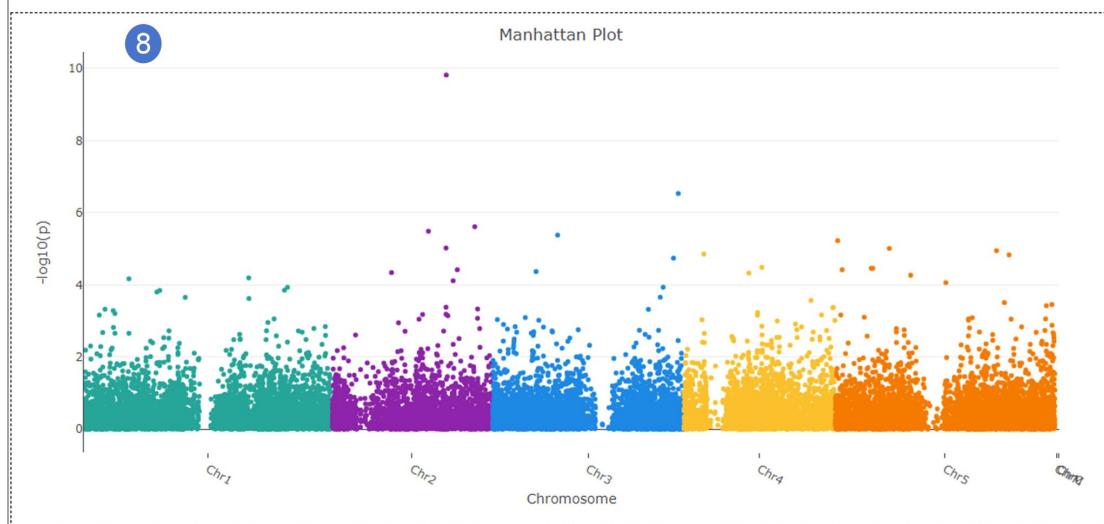
*5 -log₁₀(P) ? 0 5

*6 Upstream? 2000 6

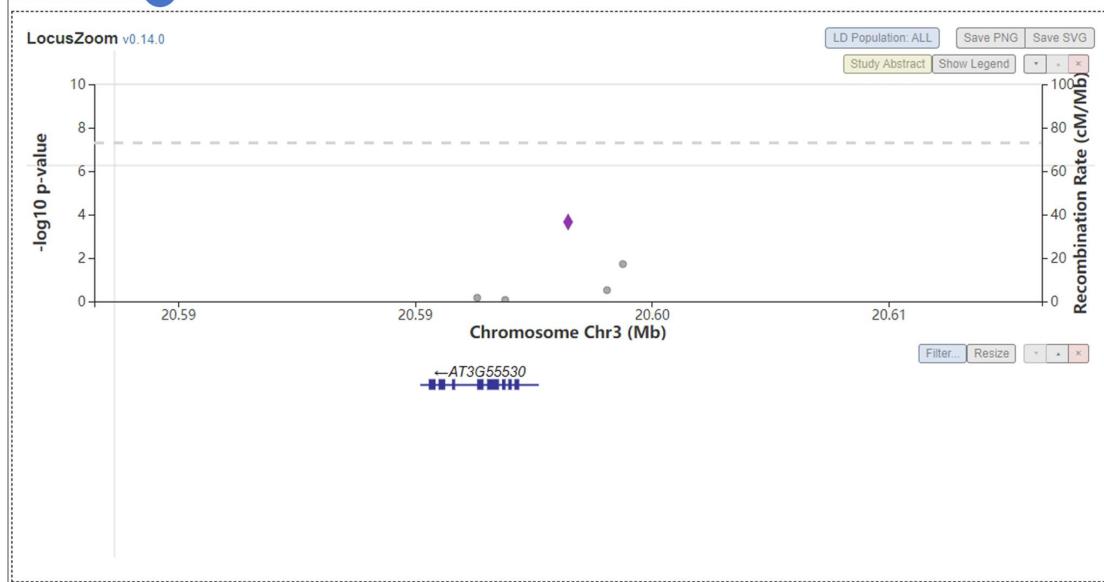
*7 Downstream? 2000 7

Run

Manhattan Plot:



LocusZoom: 9



Search GWAS Result:

10

CHR	BP	Ref	Var	-log10_p_value
Chr1	4654	G	A	0.14814909082737734
Chr1	29345	A	G	0.5516910635590182
Chr1	57383	A	T	0.3905851078024906
Chr1	57686	C	A	0.33505224591555677
Chr1	59016	G	C	0.3249139658139927
Chr1	59876	G	A	0.49946796511369745
Chr1	64304	A	C	1.1362456774584186
Chr1	65247	G	T	0.17577598886367957
Chr1	67282	G	C	0.5291295827983218
Chr1	71348	T	C	0.09026846995683122

Showing 1 to 10 of 21,467 entries

Previous 1 2 3 4 5 ... 2147 Next

5. Variation Search

Variation Search is used to retrieve variation information. Users select the project name (①), chromosome number (②), set the start position (③) and end position (④), then click Run to view the results. The results include two parts: a variation information table (⑤) and statistics of single variation loci (⑥). When users click on the position information of a single variation locus, the statistical results will be interactively displayed in the form of pie charts and tables.

Variation Search:

*1 Select Project ① *2 Select Chr ② *3 Start ③ *4 End ④ Run

Project_1 Chr2 1000 20000

Variation Table:

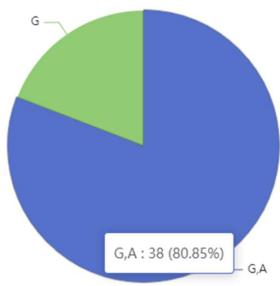
5

CHR	BP	Ref	Var
Chr2	10550	G	A
Chr2	10551	T	C
Chr2	10679	T	C
Chr2	10797	A	C
Chr2	10816	G	A
Chr2	10828	A	G
Chr2	10839	G	C
Chr2	10841	C	T
Chr2	10891	C	T
Chr2	10892	T	C

Showing 41 to 50 of 158 entries 1 row selected

Previous 1 ... 4 5 6 ... 16 Next

Single variant site statistics: 6



pos:Chr2:10550

G,A	108,139,159,265,350,351,410,430,470,476,484,504,506,531,628,630,685,687,728,742,763,6907,6908,6911,6913,6915,6917,6918,6920,6922,6923,6924,6926,6927,6929,6931,6933,6938
G	424,428,544,546,680,681,6909,6919,6932