Computational and Structural Biotechnology Journal 20 (2022) 3234–3247

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

REGENOMICS: A web-based application for plant REGENeration-associated transcriptOMICS analyses

Soon Hyung Bae [a], Yoo-Sun Noh [b,c], Pil Joon Seo [a,d,e,*]

[a] Department of Chemistry, Seoul National University, Seoul 08826, South Korea
[b] School of Biological Sciences, Seoul National University, Seoul 08826, South Korea
[c] Research Center for Plant Plasticity, Seoul National University, Seoul 08826, South Korea
[d] Plant Genomics and Breeding Institute, Seoul National University, Seoul 08826, South Korea
[e] Research Institute of Basic Sciences, Seoul National University, Seoul 08826, South Korea

ARTICLE INFO

Article history:
Received 31 March 2022
Received in revised form 13 June 2022
Accepted 13 June 2022
Available online 16 June 2022

Keywords:
Plant regeneration
Co-expression network
Gene regulatory network
Trajectory inference
Gene expression browser
Single-cell analysis

ABSTRACT

In plants, differentiated somatic cells exhibit an exceptional ability to regenerate new tissues, organs, or whole plants. Recent studies have unveiled core genetic components and pathways underlying cellular reprogramming and *de novo* tissue regeneration in plants. Although high-throughput analyses have led to key discoveries in plant regeneration, a comprehensive organization of large-scale data is needed to further enhance our understanding of plant regeneration. Here, we collected all currently available transcriptome datasets related to wounding responses, callus formation, *de novo* organogenesis, somatic embryogenesis, and protoplast regeneration to construct REGENOMICS, a web-based application for plant REGENeration-associated transcriptOMICS analyses. REGENOMICS supports single- and multi-query analyses of plant regeneration-related gene-expression dynamics, co-expression networks, gene-regulatory networks, and single-cell expression profiles. Furthermore, it enables user-friendly transcriptome-level analysis of REGENOMICS-deposited and user-submitted RNA-seq datasets. Overall, we demonstrate that REGENOMICS can serve as a key hub of plant regeneration transcriptome analysis and greatly enhance our understanding on gene-expression networks, new molecular interactions, and the crosstalk between genetic pathways underlying each mode of plant regeneration. The REGENOMICS web-based application is available at http://plantregeneration.snu.ac.kr.

1. Introduction

Being sessile organisms, plants are vulnerable to a plethora of environmental stimuli, some of which can cause wounds or the complete loss of organs. To recover from environmental damage, plant cells acquire pluripotency and give rise to new undifferentiated stem cells [49]. The stem cell niche exhibits continuous cell division and subsequent cell differentiation, leading to the formation of various cell types and replenishing the damaged tissues [41,69]. In addition, owing to the remarkable tissue regeneration potential of plants, several *in vitro* tissue-culture techniques have been developed [60]. Exogenous hormone treatment with a certain auxin to cytokinin ratio allows pluripotency acquisition and tissue regeneration [60].

Four main modes of plant regeneration have been investigated to date: wound-induced tissue repair, hormone-induced plant regeneration, somatic embryogenesis, and protoplast regeneration [1,9,51,57]. Under natural conditions, wounding causes rapid cell proliferation at injury site to replenish damaged cells and tissues [38]. Upon the excision of *Arabidopsis* root tips as a model of wound-induced tissue repair, jasmonic acid (JA) and auxin accumulation leads to the activation of *CYCLIN D6* (*CYCD6;1*) and *ETHYLENE RESPONSE FACTOR 115* (*ERF115*), which in turn activates the RETINOBLASTOMA-RELATED (RBR)-SCARECROW (SCR)-SHORTROOT (SHR) module, regenerating the stem cell niche and subsequently the root tip [57,85]. In some cases, wounding leads to the formation of a dedifferentiated cell mass, called callus [24]. Callus cells are pluripotent and therefore are capable of regenerating tissues, organs, or even plantlets [1]. Wound-responsive WOUND INDUCED DEDIFFERENTIATION (WIND) transcription factors are key players in wound-induced tissue repair and callus development, especially at the injury site [24]. WIND-

* Corresponding author at: Department of Chemistry, Seoul National University, Seoul 08826, South Korea.
E-mail address: pjseo1@snu.ac.kr (P.J. Seo).

https://doi.org/10.1016/j.csbj.2022.06.033

induced callus formation is dependent on ENHANCER OF SHOOT REGENERATION 1 (ESR1)/DORNRÖSCHEN (DRN) and type-B ARABIDOPSIS RESPONSE REGULATORs (ARRs), which activate cytokinin signaling-associated cell proliferation and tissue regeneration [22].

Plant regeneration can be facilitated through a two-step *in vitro* tissue-culture method [26,83,34]. Pluripotent callus can be obtained from differentiated plant tissues by culturing explants on auxin-rich callus-inducing medium (CIM) [1,63]. CIM-induced callus formation resembles the lateral root initiation process [1,63]. Consistently, *ABERRANT LATERAL ROOT FORMATION 4* (*ALF4*), *WUSCHEL-RELATED HOMEOBOX 11* (*WOX11*), and *LATERAL ORGAN BOUNDARIES DOMAIN 16/17/18/29* (*LBD16/17/18/29*) genes, which play crucial roles in lateral root initiation, are also essential for callus formation [63,35]. In the middle of the callus formation process, pluripotency acquisition is enabled with the activation of root meristem regulators such as *PLETHORA 1* (*PLT1*), *PLT2*, *SCR*, and *WOX5* [1,63,27]. Following pluripotency acquisition, calli are usually transferred to cytokinin-rich shoot-inducing medium (SIM), which promotes *de novo* shoot organogenesis [26,34]. The cytokinin-inducible type-B ARR transcription factors directly bind to the *WUSCHEL* (*WUS*) promoter, specify the shoot stem cell niche in callus, and promote *de novo* shoot regeneration [83].

Somatic embryogenesis can be induced by the application of high concentrations of auxin to somatic cells or somatic cell-derived callus [75,9]. Somatic embryogenesis is stochastically initiated during *in vitro* tissue culture, and an emerged somatic embryo gives rise to a new plant [9]. Because somatic embryogenesis is similar to normal embryogenesis, both processes share common genetic components [14], including *BABY BOOM* (*BBM*), *LEAFY COTYLEDON 1* (*LEC1*), *LEC2*, *ABSCISIC ACID INSENSITIVE 3* (*ABI3*), *FUSCA 3* (*FUS3*), and *AGAMOUS-LIKE 15* (*AGL15*) [14,71].

Notably, a single protoplast, which is a cell dissociated from its cell wall, has the ability to proliferate to form microcallus and subsequently regenerate into a whole plant [64]. This mode of plant regeneration involves dedifferentiation from a fully differentiated somatic cell, which is essential for callus formation and tissue regeneration. Protoplast regeneration shares some features of other types of plant regeneration, as crucial molecular factors, such as *WINDs*, *WUS* and *ESR1/DRN* [5,76], are pervasively involved, although much is unknown about the regulatory mechanisms underlying protoplast regeneration.

Recent studies on plant regeneration have elucidated key genetic players and signaling pathways underlying pluripotency acquisition and plant regeneration, and have generated enormous amounts of high-throughput RNA-seq data relevant to plant regeneration processes [6,13,15,17,19,20,27,29,30,33,37,40,47,52,59,54,71,73,76,78,81]. However, despite the numerous RNA-seq datasets available publicly, extracting meaningful information from these data remains challenging. A few web applications, such as ePlant [70] and ATTED [44], allow users to easily explore gene-expression patterns and co-expression networks. However, none of these web applications are solely dedicated to plant regeneration studies or allow customized transcriptome-level analyses. Furthermore, because distinct genetic pathways underlie different modes of plant regeneration, it is important to analyze the gene expression patterns and co-expression networks in the context of each plant regeneration mode. In addition, although studies on plant regeneration have been mainly conducted in *Arabidopsis thaliana* to date due to its advantages as a model species and a high capacity of cellular reprogramming [18,39], plant regeneration is an important issue in other plant species [10,21,43,19]. Therefore, efforts to integrate transcriptome data generated from various plant species are required to investigate the universal or specific regulators of plant regeneration across the plant kingdom, highlighting the necessity of a transcriptome analytic hub for regeneration studies on various plant species.

A large volume of single-cell RNA-seq data has also been generated recently from a diverse range of plant species and tissues [36,45,58,72,76,80,82]. Accordingly, tools for single-cell data analysis are being newly developed and improved at an astonishing speed, allowing us to interpret important biological processes at a single-cell level [53,79]. One of the major developments is trajectory inference, which uses single-cell datasets to computationally predict developmental trajectories within the dataset and order cells according to a pseudotime, an abstract measure of biological progress [66]. High-throughput single-cell datasets also allow the construction of high-confidence gene regulatory networks (GRNs). As single-cell datasets provide massive amounts of transcriptome data with little technical bias, sophisticated gene regulations can be predicted with higher reliability. However, GRNs inferred with plant single-cell RNA-seq data are currently unavailable [67]. Another limitation is that many recently developed tools for single-cell data analysis are not yet readily accessible in the form of web applications.

Here, we developed a web-based database and analytic platform, REGENOMICS, which integrates all currently available transcriptome datasets related to the four main modes of plant regeneration, and enables integrative analysis to gain new insights into the crucial connections between various genetic components and signaling pathways underlying the plant regeneration process. REGENOMICS provides a user-friendly interface and is freely available at http://plantregeneration.snu.ac.kr.

2. Results and discussion

2.1. Overview of REGENOMICS

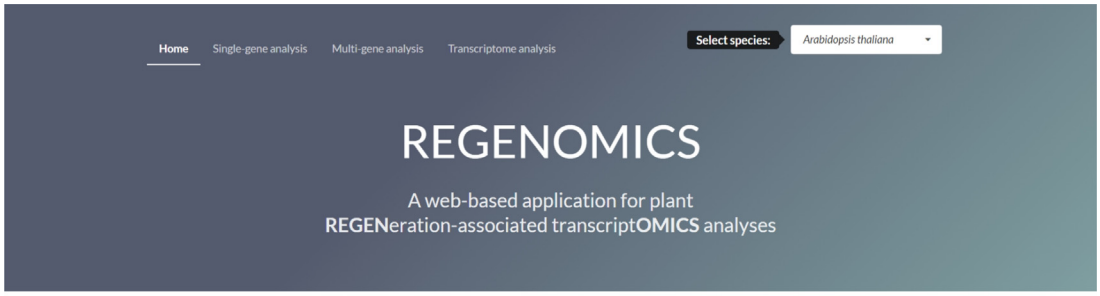
REGENOMICS is a plant regeneration-dedicated database that contains all currently available transcriptomic data related to regeneration, specifically wound-induced tissue repair, hormone-induced plant regeneration, somatic embryogenesis, and protoplast regeneration (Table 1), along with an analytical pipeline that processes user-submitted queries. The REGENOMICS analytical pipeline is divided into three analysis modes (Single-gene analysis, Multi-gene analysis, and Transcriptome analysis) depending on the type of query (a single gene of interest [GOI], a list of GOIs, or a whole transcriptome) (Fig. 1). In the Single-gene analysis mode, a user can view the expression pattern, co-expression network constructed using bulk RNA-seq, and single-cell RNA-seq-derived GRNs of any GOI involved in plant regeneration. The Multi-gene analysis mode supports the comparison of expression patterns, construction of co-expression networks and GRNs, and differentially expressed gene (DEG) enrichment analysis for a list of GOIs. The Transcriptome analysis mode enables users to select REGENOMICS-deposited and user-submitted transcriptomes for correlation analysis, dimensionality reduction analysis, and differential expression analysis.

The REGENOMICS application has five main features. First, gene expression profiles of any GOI(s) involved in any mode of plant regeneration can be easily accessed by searching for either the gene name or the gene locus code. The mode of plant regeneration can be selected from the dropdown menu. Second, REGENOMICS contains five different co-expression networks: one built with all REGENOMICS-deposited bulk RNA-seq datasets, and four constructed using transcriptome datasets corresponding to wound-induced tissue repair, hormone-induced plant regeneration, somatic embryogenesis, or protoplast regeneration. A sub-network can be drawn with the submitted query gene(s) to visualize its (their) connections with closely co-expressed genes in the context of each mode of plant regeneration. Gene ontology (GO) and DEG enrichment analyses are performed using genes compris-

Table 1
Datasets Collected and Processed for the REGENOMICS Database.

Category	Tissue (origin)	Plant Materials	Accession Number	Reference
Wound-induced tissue repair	Hypocotyl	*Arabidopsis thaliana* Col-0	GSE101422	[17]
Wound-induced tissue repair	Leaf	*Arabidopsis thaliana* Col-0, *coi1-2, sdg8-2*	GSE120418	[81]
Wound-induced tissue repair	Leaf	*Arabidopsis thaliana* Col-0, *pSPL10:rSPL10-GR*	PRJCA001184	[78]
Wound-induced tissue repair	Leaf	*Arabidopsis thaliana* Col-0	GSE108253	[47]
Wound-induced tissue repair	Root	*Arabidopsis thaliana* Col-0	GSE168385	[13]
Wound-induced tissue repair	Root	*Arabidopsis thaliana* Col-0	GSE145488	[40]
Wound-induced tissue repair	Root	*Arabidopsis thaliana* Col-0	E-MTAB-7609	[52]
Hormone-induced plant regeneration	Hypocotyl	*Arabidopsis thaliana* Col-0, *siz1-2*	GSE141188	[6]
Hormone-induced plant regeneration	Hypocotyl	*Arabidopsis thaliana* Col-0, *plt1plt2*	GSE178354	[80]
Hormone-induced plant regeneration	Hypocotyl	*Arabidopsis thaliana* Col-0, *wox5-1wox7-1*	GSE156990	[80]
Hormone-induced plant regeneration	Hypocotyl	*Arabidopsis thaliana* Col-0	GSE176161	[29]
Hormone-induced plant regeneration	Root	*Arabidopsis thaliana* Col-0, *ldl3-1*	DRP004816	[20]
Hormone-induced plant regeneration	Root	*Arabidopsis thaliana* Col-0, *hag1-6*	GSE100966	[27]
Hormone-induced plant regeneration	Root	*Arabidopsis thaliana* Col-0, *atprmt5*	E-MTAB-5044	[33]
Hormone-induced plant regeneration	Root	*Arabidopsis thaliana* Col-0, *arr1-4*, $35S_{pro}$:ARR1	GSE146690	[37]
Protoplast regeneration	Mesophyll cells	*Arabidopsis thaliana* Col-0	PRJNA739879	[54]
Protoplast regeneration	Mesophyll cells	*Arabidopsis thaliana* Col-0	PRJNA648028	[76]
Somatic embryogenesis	Immature embryo and seedling	*Arabidopsis thaliana* Col-0, *p35S::LEC2-GR*	PRJCA002620	[71]
Somatic embryogenesis	Immature embryo	*Arabidopsis thaliana* Col-0	E-MTAB-2403	[73]
Root tip cell types and developmental zones	Root tip	*Arabidopsis thaliana* Col-0	PRJNA323955	[32]
Wound-induced tissue repair	Hypocotyl	*Solanum lycopersicum* Micro-Tom	PRJNA731333	[30]
Somatic embryogenesis	Seed	*Oryza sativa* TNG67, IR64	GSE82138	[15]
Somatic embryogenesis	Seed	*Oryza sativa* ssp. *japonica*	PRJNA804674	[59]
Wound-induced tissue repair	Thallus	*Marchantia polymorpha* Tak-1	PRJDB12610	[19]
Wound-induced tissue repair	Thallus	*Marchantia polymorpha* Tak-1, *Mperf15ko*	GSE196912	Unpublished
Single-cell RNA-sequencing	Hypocotyl explant	*Arabidopsis thaliana* Col-0	GSE156991	[80]
Single-cell RNA-sequencing	Leaf explant	*Arabidopsis thaliana* Col-0	GSE147289	[36]
Single-cell RNA-sequencing	Shoot-borne roots	*Solanum lycopersicum cv M82*	GSE159055	[45]
Single-cell RNA-sequencing	Root tip	*Arabidopsis thaliana* Col-0	GSE141730	[72]
Single-cell RNA-sequencing	Shoot apex and leaf	*Arabidopsis thaliana* Col-0	PRJCA003094	[82]
Single-cell RNA-sequencing	Protoplasts	*Arabidopsis thaliana* Col-0	PRJNA648028	[76]
Single-cell RNA-sequencing	Lateral root primordium	*Arabidopsis thaliana* Col-0	GSE161970	[58]

ing the resulting output sub-network to estimate biological features of the sub-network. Third, DEG enrichment analysis allows the estimation of the biological and functional aspects of a group of genes. This analysis can be performed with genes comprising the co-expression network as well as a user-submitted list of GOIs. Fourth, the REGENOMICS database contains single-cell transcriptome data of tissues undergoing regeneration as well as root and shoot apical meristems. Users can explore cell type-specific expression patterns of GOI(s) and also perform trajectory inference analysis for each dataset. Furthermore, genes that regulate and are regulated by the GOI(s) can be viewed from the sub-networks drawn from the GRNs inferred from each single-cell RNA-seq dataset. Lastly, users can input any *Arabidopsis* transcriptome profile (RNA-seq data) in comma-separated values (CSV) format to analyze its transcriptome-level correlation using the selected REGENOMICS-deposited transcriptome data and user-deposited



Single-gene analysis

Find a gene by its name or gene locus code for single-gene analysis.

Go

Multi-gene analysis

Enter a list of genes (in gene locus code) to analyze multiple genes together.

Go

Transcriptome analysis

Select among deposited transcriptomes to analyze them with user-submitted transcriptomes.

Go

	Bulk RNA-seq transcriptomes	Single-cell RNA-seq datasets	Single-gene analysis	Multi-gene analysis	Transcriptome analysis
Arabidopsis thaliana	302	6	✓	✓	✓
Solanum lycopersicum	20	1	✓	✓	
Oryza sativa	18	0	✓	✓	
Marchantia polymorpha	51	0	✓	✓	

Click here to download a table of datasets used in REGENOMICS.

Click here to download a glossary of terms used in REGENOMICS.

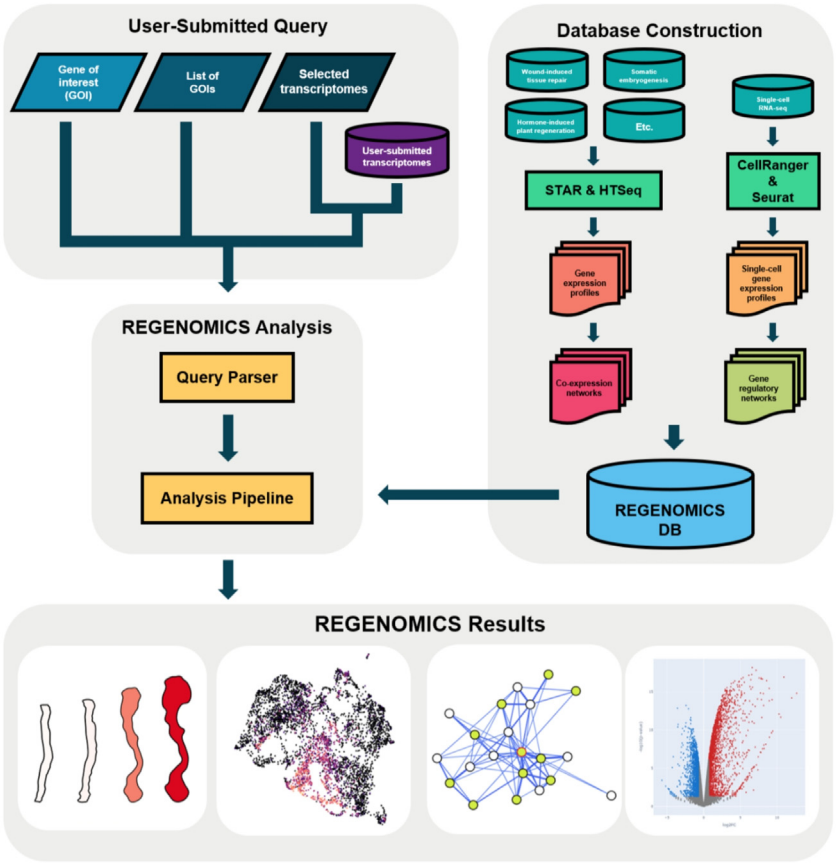


Fig. 1. Workflow for REGENOMICS as shown in the REGENOMICS web interface. Transcriptomic datasets related to plant regeneration are collected and processed according to the REGENOMICS analytical pipeline. The pipeline is divided into three sections, Single-gene analysis, Multi-gene analysis, and Transcriptome analysis, each of which accepts a gene of interest (GOI), a list of GOIs, and transcriptomes as the input query, respectively. Then, REGENOMICS outputs the gene expression profile(s), co-expression network and gene-regulatory network for input gene(s). Transcriptome analysis supports correlation, dimensionality reduction, and differential expression analyses for user-defined transcriptome datasets.

RNA-seq data. In addition, users can select two transcriptome datasets and conduct differential expression analysis to obtain a table of DEGs as well as the corresponding GO terms.

It should be noted that REGENOMICS is not limited to *Arabidopsis* data, but includes data from other species. Users can select the plant species in the dropdown menu located at the top of each page, and perform the same analysis for each species. However, while all analyses shown here are available for use with *Arabidopsis* studies, some of the analyses provided in REGENOMICS are not yet ready for use for other species, due to the limited availability of RNA-seq data. With future researches, we plan to support all features of REGENOMICS for all plant species to promote regeneration studies covering a wide spectrum of plant species. Overall, REGENOMICS is designed to enable easy access to plant regeneration-related high-throughput transcriptome data to gain a comprehensive insight into the molecular connections between the genetic components and signaling pathways underlying each mode of plant regeneration in various species (Fig. 1).

2.2. Single-gene analysis

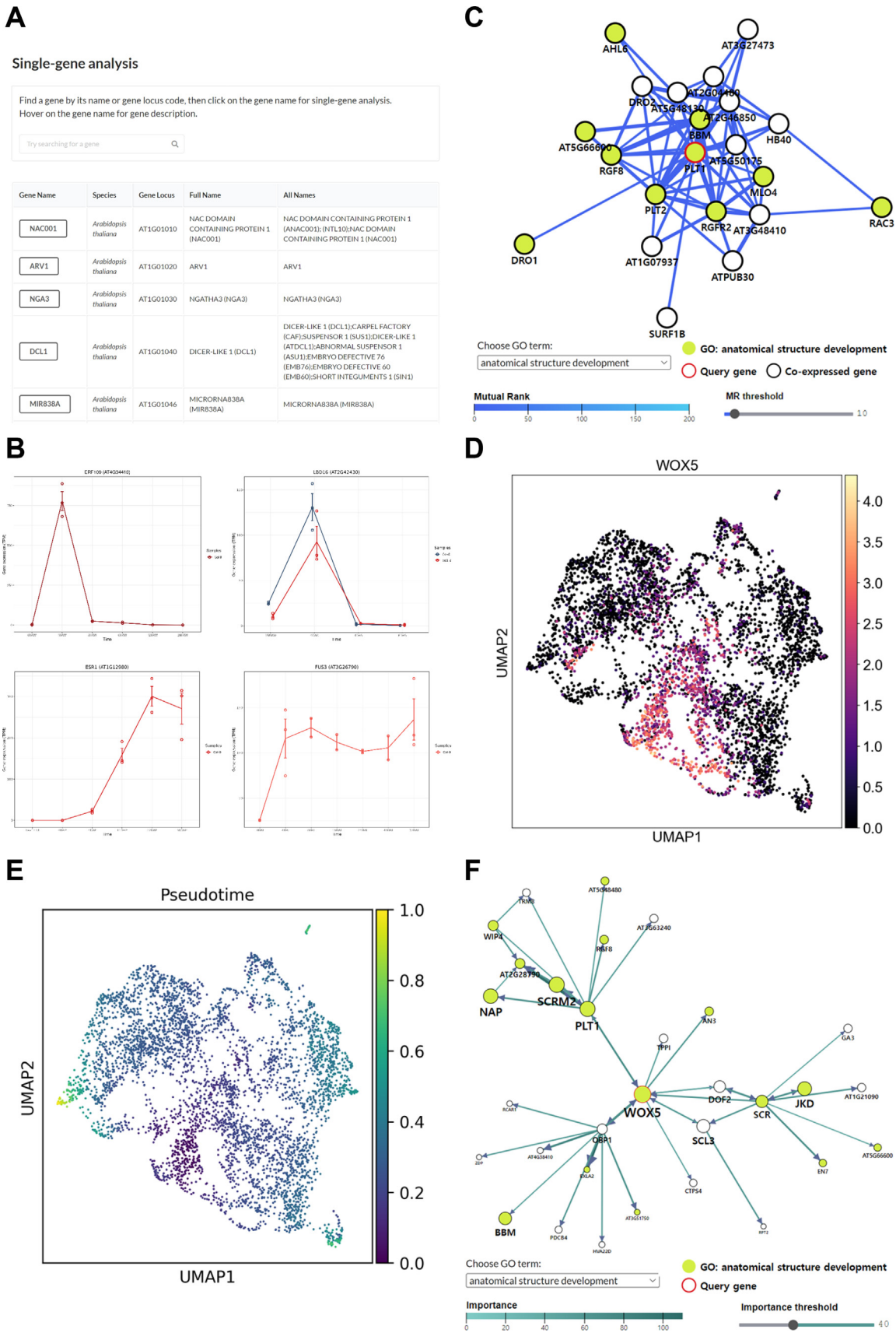
To demonstrate each analysis supported in REGENOMICS, we used *Arabidopsis* studies as representative examples. While some features are currently available only for *Arabidopsis* due to a lack of published data for other species (Fig. 1), all analyses in REGENOMICS will become available for every species in the same format as *Arabidopsis* through future updates. In the Single-gene analysis mode, the REGENOMICS web application supports three different analyses (Gene expression browser, Co-expression network, and Single-cell analysis) using all currently available transcriptome data related to wound-induced tissue repair, hormone-induced plant regeneration, somatic embryogenesis, and protoplast regeneration. Users can retrieve information for any gene by searching for a gene name (e.g., *WOX5*) or the gene locus code (e.g., AT3G11260) (Fig. 2A). In the Gene expression browser, the GOI's expression values normalized in transcripts per million (TPM) are displayed as line plots or bar plots for each study under the selected mode of plant regeneration (Fig. 2B). To validate the gene expression browser of REGENOMICS, we queried several genes involved in wound-induced tissue regeneration, hormone-induced plant regeneration, somatic embryogenesis, or protoplast regeneration. The results obtained using REGENOMICS showed that the wound-responsive *ERF109* was transiently up-regulated shortly after wounding (Fig. 2B) as previously reported [81]. Additionally, the expression patterns of *LBD16*, *ESR1/DRN*, and *FUS3* in REGENOMICS, which were reported to be essential in hormone-induced regeneration, protoplast regeneration, and somatic embryogenesis, respectively, matched the patterns reported previously (Fig. 2B) [14,34,71,76].

REGENOMICS generates a co-expression network that connects functionally-related genes based on correlated expression patterns. Co-expression networks were built by collecting the bulk RNA-seq data of each mode of plant regeneration and then analyzing gene-to-gene expression correlation via the Mutual Rank (MR) [44]. Users can access a sub-network constructed with the user-submitted GOI and its close neighbor genes that have significantly correlated expression levels (Fig. 2C). By selecting a mode of plant regeneration, users can draw a sub-network built with datasets specific for that mode of plant regeneration. Users can interact with the output network by dragging the nodes, which represent the query gene and its co-expressed genes. Edges between nodes are colored, according to the MR gradient, and can be filtered out by changing the MR threshold with the slider (Fig. 2C). A downloadable table including co-expressed genes and their MR values is provided below the sub-network (Table S1). Additionally, the biological processes overrepresented in the output sub-network can be viewed by selecting a GO term, which colors the gene nodes corresponding to the GO term (Fig. 2C). These sub-network genes are also used for DEG enrichment analysis. We generated sets of up- or down-regulated genes (DEGs) from pair-wise comparisons of transcriptomes for each study deposited in the REGENOMICS database. Enrichment analysis allows the identification of DEG sets enriched in the output sub-network (Fig. S1), thus providing an insight into the potential biological impact of the sub-network on plant regeneration processes.

We conducted a case study using *PLT1*, a gene known to be related to the pluripotency acquisition of callus cells [26], as query. Co-expression network analysis using hormone-induced plant regeneration datasets (Fig. 2C) identified *PLT2*, another known regulator in establishing pluripotency during callus formation [26]. Similarly, root-meristem regulator genes, such as *WOX5* and *RGF8* [16,27,46], were also shown to be closely co-expressed with *PLT1* (Table S1). Consistently, genes with a GO term, 'anatomical structure development', were predominantly found in the output sub-network (yellow-green color in Fig. 2C). Furthermore, DEG enrichment analysis showed that the sub-network was enriched with genes up-regulated in the hypocotyl explant-derived calli incubated on CIM (DACs in Fig. S1), relative to samples from wounding or *de novo* shoot regeneration [6], consistent with the role of *PLT1* in callus formation [26].

Moreover, a query gene can be analyzed by the Single-cell analysis, which contains single-cell RNA-seq datasets for different modes of plant regeneration as well as meristem (and meristem-like) tissues (Table 1). Users can view the gene expression dynamics of the query gene occurring during the course of plant regeneration as well as the expression patterns in various cell types found in pluripotent cells and stem-cell derived differentiated cells. We analyzed single-cell RNA-seq datasets according to our unified analysis pipeline and annotated cell types using cell type-specific marker genes. Transcript accumulation of the GOI in each cell type was visualized as a violin plot (Fig. S2) and Uniform Manifold Approximation and Projection (UMAP) plots for each cell (Fig. 2D). To validate the data processing and cell-type annotation, we used *PLT1* as query in the single-cell RNA-seq dataset from hypocotyl-derived callus, and confirmed that the expression of each gene was enriched in QC-like-cells (Figs. S2 and 2D) as reported in the original study [80]. Additionally, trajectory inference analysis allows the understanding of the cell identity transition dynamics in terms of pseudotime (Fig. 2E). We were able to confirm that QC-like cells differentiate into different cell types

Fig. 2. Interface and features of Single-gene analysis. (A) Interface and search page for Single-gene analysis. (B) Gene expression browser. Gene expression browser returns gene expression patterns in line plots for a single query gene. Expression patterns for *ERF109*, *LBD16*, *ESR1/DRN*, and *FUS3* are shown in different modes of plant regeneration. (C) Co-expression network analysis. The *PLT1* gene (node circled in red) was used as an input query gene. Co-expressed genes include key regulators of the root meristem such as *PLT2*, *BBM*, *WOX5*, *RGF8*, and *RGFR2*. Genes with the GO term 'anatomical structure development' are colored yellow-green. The network was built with hormone-induced plant regeneration bulk RNA-seq datasets. (D) Uniform Manifold Approximation and Projection (UMAP) for Single-cell analysis. UMAP plot colored according to the expression of *WOX5* in single-cell transcriptome of hypocotyl callus is shown. Each dot represents a single cell (see also Fig. S3). (E) Trajectory inference analysis for Single-cell analysis. UMAP plot shows the trajectory inference analysis results performed with the single-cell transcriptome of hypocotyl callus. Each cell is colored according to the calculated pseudotime. (F) Gene regulatory network for Single-cell analysis. The sub-network was drawn using *WOX5* gene (node circled in red) as an input query gene. Key regulators of the pluripotency acquisition, such as *PLT1* and *SCR*, were included in the sub-network. Genes with the GO term 'anatomical structure development' are colored yellow-green. The network was inferred from single-cell transcriptome of hypocotyl callus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)


A
Single-gene analysis
Find a gene by its name or gene locus code, then click on the gene name for single-gene analysis.
Hover on the gene name for gene description.
Try searching for a gene
Gene Name
Species
Gene Locus
Full Name
All Names
NAC001
Arabidopsis thaliana
AT1G01010
NAC DOMAIN CONTAINING PROTEIN 1 (NAC001)
NAC DOMAIN CONTAINING PROTEIN 1 (ANAC001); (NTL10);NAC DOMAIN CONTAINING PROTEIN 1 (NAC001)
ARV1
Arabidopsis thaliana
AT1G01020
ARV1
ARV1
NGA3
Arabidopsis thaliana
AT1G01030
NGATHA3 (NGA3)
NGATHA3 (NGA3)
DCL1
Arabidopsis thaliana
AT1G01040
DICER-LIKE 1 (DCL1)
DICER-LIKE 1 (DCL1);CARPEL FACTORY (CAF);SUSPENSOR 1 (SUS1);DICER-LIKE 1 (ATDCL1);ABNORMAL SUSPENSOR 1 (ASU1);EMBRYO DEFECTIVE 76 (EMB76);EMBRYO DEFECTIVE 60 (EMB60);SHORT INTEGUMENTS 1 (SIN1)
MIR838A
Arabidopsis thaliana
AT1G01046
MICRORNA838A (MIR838A)
MICRORNA838A (MIR838A)
C
AHL6
AT3G27473
DRD2
AT2G04450
AT5G48130
AT2G48850
HB40
AT5G66600
BBM
PLT1
AT5G50175
RGF8
MLO4
PLT2
RGFR2
AT3G48410
RAC3
DRO1
AT1G07937
ATPUB30
SURF1B
Choose GO term:
anatomical structure development
GO: anatomical structure development
Query gene
Co-expressed gene
Mutual Rank
0
50
100
150
200
MR threshold
1.0
B
ERF109 (AT4G34410)
LBD16 (AT2G42430)
ESR1 (AT1G12980)
FUS3 (AT3G26790)
D
WOX5
UMAP2
UMAP1
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
E
Pseudotime
UMAP2
UMAP1
1.0
0.8
0.6
0.4
0.2
0.0
F
AT5G48480
TMO6
AT3G61240
WIP4
RGF8
AT2G28790
NAP
SCRM2
PLT1
AN3
TPPI
GA3
DOF2
JKD
AT1G21090
WOX5
SCR
RCAR1
OBP1
SCL3
AT5G66600
AT4G38410
EN7
CTPS4
BBM
PDC84
AT3G51750
Choose GO term:
anatomical structure development
GO: anatomical structure development
Query gene
Importance
0
20
40
60
80
100
Importance threshold
4.0

and diverge into epidermis-like cells and vascular-initial-like cells [80]. Furthermore, using single-cell RNA-seq data that provides numerous unbiased samples of transcriptomes at single-cell resolution, we constructed unique GRNs from each single-cell RNA-seq dataset to predict high-confidence gene regulations between transcription factors and their target genes. Upon query input, sub-networks are built with nodes for the query gene and genes having direct regulatory connections with the query (Fig. 2F).

3. Multi-gene analysis

Multi-gene analysis allows the simultaneous analysis of multiple queries. The Multi-gene analysis mode enables four different analyses (Gene expression heatmap, Co-expression network, DEG enrichment, and Single-cell analysis) for each mode of plant regeneration simply by selecting the menu tab and then submitting a list of genes by their gene locus codes (Fig. 3A). In the Gene expression heatmap analysis, users should submit multi-query genes and select a transcriptome dataset (mode of plant regeneration) for heatmap visualization (Fig. 3B). Rows are ordered by hierarchical clustering based on gene expression patterns, and subsets of the dendrogram can be magnified for a closer look. In this study, as an example, we queried *CUC1*, *CUC2*, *ESR1/DRN*, *ESR2/DRNL*, *PLTs*, *STM*, *WOX5*, and *WUS*, key regulators of hormone-induced plant regeneration, and showed that their expression patterns can be simultaneously compared during the process of hormone-induced plant regeneration (Fig. 3B). The results were consistent with previous reports [1,14,26,27,41,63,71].

Users can also submit multiple query genes to draw sub-networks that show correlations in gene expression between input query genes and other co-expressed genes with the selection of plant regeneration mode (Fig. 3C). Compared with Single-gene analysis, the construction of sub-networks containing multi-query genes employs a higher threshold of Mutual Rank to obtain a reasonable number of close neighbor genes for each query gene, generating a concise output sub-network. Genes consisting this co-expression network are used as input for GO enrichment analysis, and the results of analysis are generated as a table (Fig. 3D) (see below for validation of co-expression networks). Similar to Single-gene analysis, the genes in the network can also be used to perform DEG enrichment analysis.

In addition, users can investigate their GOIs by immediately running the DEG enrichment analysis in Multi-gene analysis. Over-enrichment of DEGs reveals the biological processes overrepresented in the input genes. To validate this approach, we queried top 200 PLT2-activated genes (Table S2) obtained from the PLT2-YFP chromatin immunoprecipitation sequencing (ChIP-seq) analysis [55]. The results showed that genes up-regulated in callus samples relative to their expression in wounding or shoot-regenerated samples (Fig. 3E) [6,20] were enriched with the query genes, consistent with the pivotal role of PLT2 in callus formation and pluripotency acquisition. By contrast, upon querying 96 genes targeted and repressed by PLT2 (Table S2), the enriched sets of DEGs included genes down-regulated in callus samples compared to other samples (Fig. 3F).

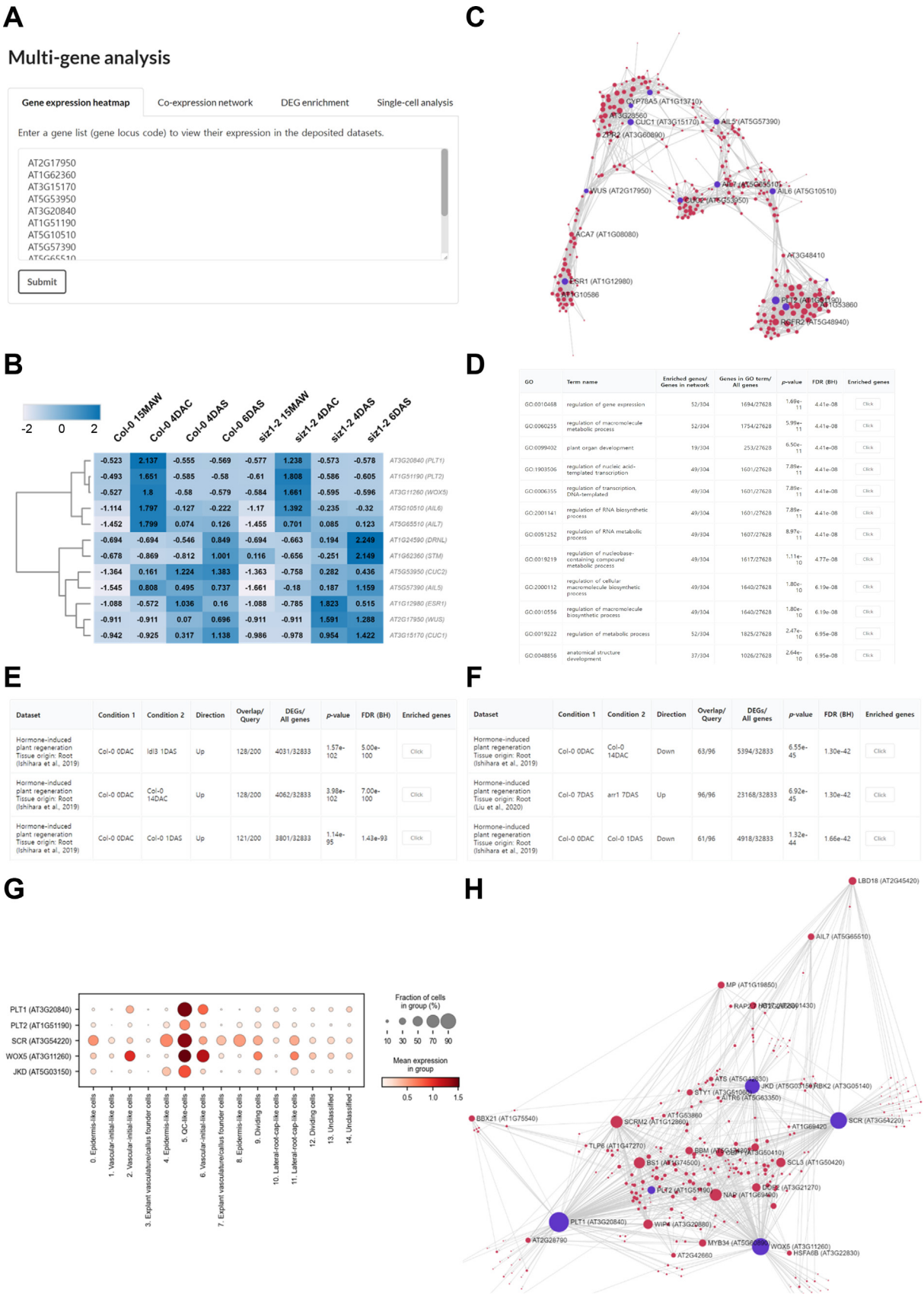
Another feature of Multi-gene analysis is Single-cell analysis. Upon submission of multi-query genes, a dot plot displays the gene expression patterns of query genes in each cell type, the percentage of cells expressing the GOI, and the mean expression values of the GOIs. For instance, when we input several root stem-cell regulators, including *PLT1*, *PLT2*, *SCR*, *WOX5*, and *JKD*, the results showed that these genes are primarily expressed in QC-like cells (Fig. 3G), consistent with previous reports [26,27,63,80]. In addition to browsing the expression patterns of query genes in single-cell RNA-seq datasets, users can draw GRNs with these genes (Fig. 3H). GO and DEG enrichment analyses can also be performed with the genes in this output GRN.

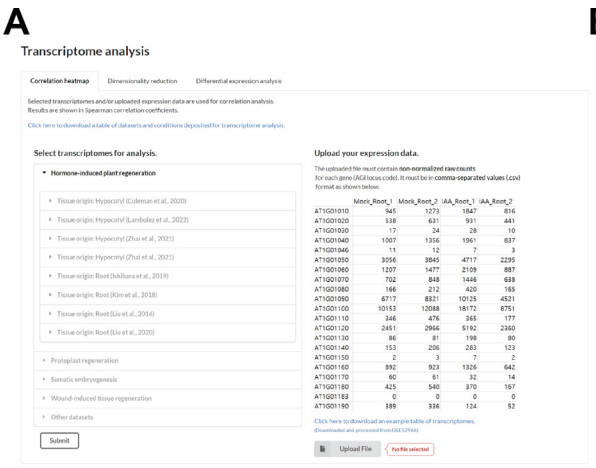
3.1. Transcriptome-level analysis

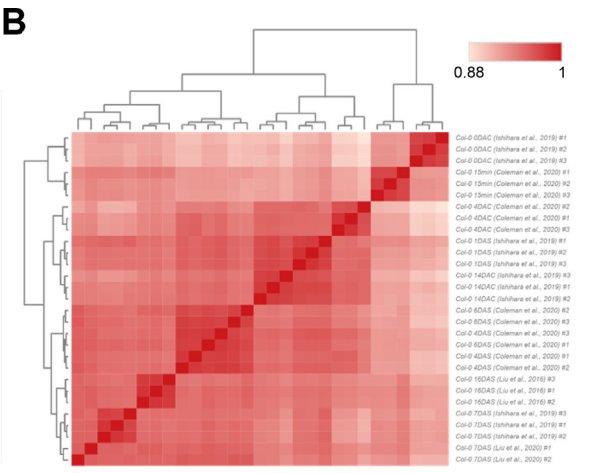
REGENOMICS provides transcriptome-level analyses as well. Users can select some or all of the *Arabidopsis* bulk RNA-seq datasets deposited in the REGENOMICS database (Table 1) to measure and visualize correlations between transcriptomes (Fig. 4A and B). Transcriptome-level correlation analysis is shown as an interactive heatmap generated with Spearman's rank correlation and hierarchical clustering (Fig. 4B). In addition, different dimensionality reduction methods, such as principal component analysis (PCA; Fig. 4C) or multidimensional scaling (MDS; Fig. 4D), can be performed to allow the visualization of transcriptomes on two-dimensional (2D) or 3D scatter plots.

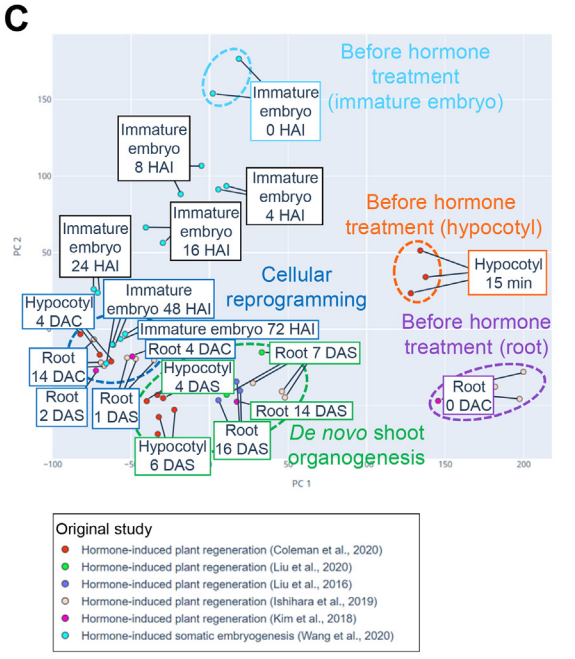
Users can run these analyses with their own transcriptome data (in CSV format) and compare with transcriptomes from the REGENOMICS transcriptome database (Fig. 4A). Although the analysis takes any transcriptome data in the gene-sample matrix format, it is suggested that users follow the analysis pipeline used to process bulk RNA-seq data for REGENOMICS. We conducted correlation analysis using REGENOMICS-deposited bulk RNA-seq datasets of hormone-induced plant regeneration (Fig. 4B). Interestingly, transcriptome datasets were clustered into the stages of two-step plant regeneration, even though the datasets come from different studies (Fig. 4B), indicating that batch effects were resolved for transcriptome-level analysis. We also compared bulk RNA-seq datasets of hormone-induced plant regeneration with somatic embryogenesis transcriptomes using dimensionality reduction methods, PCA (Fig. 4C), and MDS (Fig. 4D). Notably, initial states of samples for hormone-induced plant regeneration and somatic embryogenesis were distant due to the different types of explants (root and hypocotyl explants for hormone-induced plant regeneration versus immature embryo for somatic embryogenesis). However, transcriptome data from later stages of callus formation (4 $\sim$ 14 DAC) and somatic embryogenesis (48 $\sim$ 72 HAI) showed high degree of correlation. This result was consistent with the fact that both processes involve high-level auxin-dependent cellular reprogramming. In addition, PCA analysis was performed for the comparison of hormone-induced plant regeneration datasets together with bulk RNA-seq datasets of different cell types and developmental zones of the root tip (Fig. S4). Notably, the transcriptomes from the root meristem (meristematic zone and quiescent center) were closely related to those of hormone-induced hypocotyl- and root-derived calli on CIM (Fig. S4), consistent with

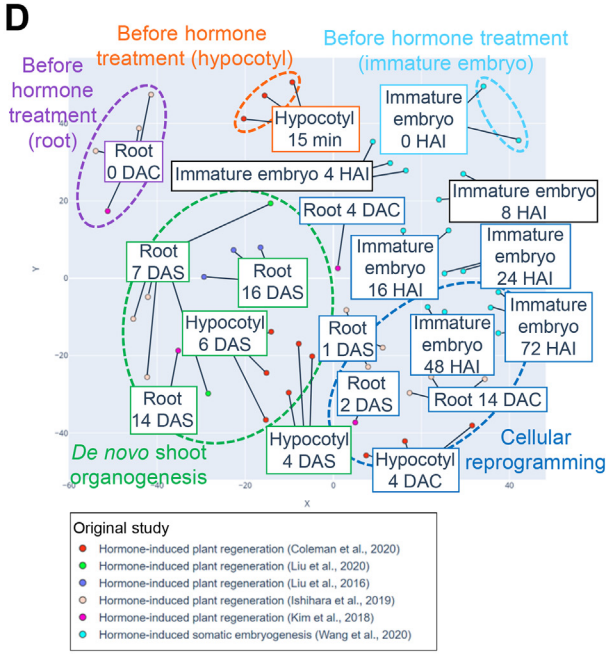
Fig. 3. Interface and features of Multi-gene analysis. (A) Interface and input page for Multi-gene analysis. (B) Gene expression heatmap. Gene expression heatmap visualizes the expression patterns of the submitted genes in each dataset. Gene expression patterns of *CUC1*, *CUC2*, *ESR1/DRN*, *ESR2/DRNL*, *PLTs*, *STM*, *WOX5*, and *WUS* during hormone-induced plant regeneration are shown. (C) Sub-network constructed using key regulators of hormone-induced plant regeneration. The *CUC1, CUC2, ESR1/DRN, ESR2/DRNL, PLTs, STM, WOX5,* and *WUS* genes were submitted as query. Query genes and co-expressed genes are each colored violet and magenta, respectively. All plant regeneration-related bulk RNA-seq datasets were used to generate the sub-network. (D) GO enrichment analysis. GO analysis was conducted with all genes shown in (C). (E and F) DEG enrichment analysis. The DEG enrichment analyses were performed with top 200 genes targeted and activated by PLT2 (E; see also Table S2), and top 96 genes targeted and repressed by PLT2 (F; see also Table S2). (G and H) Dot plot and gene regulatory network (GRN) drawn with root stem cell regulators. The *PLT1*, *PLT2*, *SCR*, *WOX5*, and *JKD* genes were submitted as input queries to the Single-cell analysis in the Multi-gene analysis. Query genes in the GRN are colored violet, while other genes are colored magenta (H). The single-cell RNA-seq dataset was derived from hypocotyl callus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

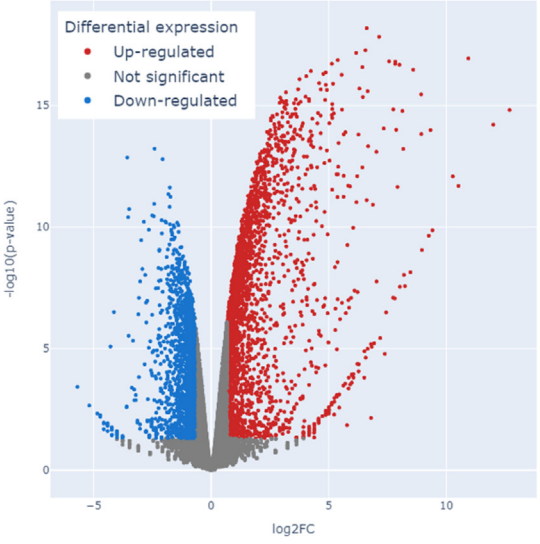

A
Multi-gene analysis
Gene expression heatmap
Co-expression network
DEG enrichment
Single-cell analysis
Enter a gene list (gene locus code) to view their expression in the deposited datasets.
AT2G17950
AT1G62360
AT3G15170
AT5G53950
AT3G20840
AT1G51190
AT5G10510
AT5G57390
Submit
B
-2 0 2
Col-0 15MAW
Col-0 4DAC
Col-0 4DAS
Col-0 6DAS
siz1-2 15MAW
siz1-2 4DAC
siz1-2 4DAS
siz1-2 6DAS
-0.523 2.137 -0.555 -0.569 -0.577 1.238 -0.573 -0.578 AT3G20840 (PLT1)
-0.493 1.651 -0.585 -0.58 -0.61 1.808 -0.586 -0.605 AT1G51190 (PLT2)
-0.527 1.8 -0.58 -0.579 -0.584 1.661 -0.595 -0.596 AT3G11260 (WOX5)
-1.114 1.797 -0.127 -0.222 -1.17 1.392 -0.235 -0.32 AT5G10510 (AIL6)
-1.452 1.799 0.074 0.126 -1.455 0.701 0.085 0.123 AT5G65510 (AIL7)
-0.694 -0.694 -0.546 0.849 -0.694 -0.663 0.194 2.249 AT1G24590 (DRNL)
-0.678 -0.869 -0.812 1.001 0.116 -0.656 -0.251 2.149 AT1G62360 (STM)
-1.364 0.161 1.224 1.383 -1.363 -0.758 0.282 0.436 AT5G53950 (CUC2)
-1.545 0.808 0.495 0.737 -1.661 -0.18 0.187 1.159 AT5G57390 (AIL5)
-1.088 -0.572 1.036 0.16 -1.088 -0.785 1.823 0.515 AT1G12980 (ESR1)
-0.911 -0.911 0.07 0.696 -0.911 -0.911 1.591 1.288 AT2G17950 (WUS)
-0.942 -0.925 0.317 1.138 -0.986 -0.978 0.954 1.422 AT3G15170 (CUC1)
C
CYP78A5 (AT1G13710)
AT3G28560
CUC1 (AT3G15170)
AIL5 (AT5G57390)
ZHD2 (AT3G60890)
WUS (AT2G17950)
AIL7 (AT5G65510)
AIL6 (AT5G10510)
CUC2 (AT5G53950)
ACA7 (AT1G08080)
AT3G48410
ESR1 (AT1G12980)
AT1G10586
PLT2 (AT1G51190)
AT1G53860
RGFR2 (AT5G48940)
D
GO | Term name | Enriched genes/ Genes in network | Genes in GO term/ All genes | p-value | FDR (BH) | Enriched genes
GO:0010468 | regulation of gene expression | 52/304 | 1694/27628 | 1.69e-11 | 4.41e-08 | Click
GO:0060255 | regulation of macromolecule metabolic process | 52/304 | 1754/27628 | 5.99e-11 | 4.41e-08 | Click
GO:0099402 | plant organ development | 19/304 | 253/27628 | 6.50e-11 | 4.41e-08 | Click
GO:1903506 | regulation of nucleic acid-templated transcription | 49/304 | 1601/27628 | 7.89e-11 | 4.41e-08 | Click
GO:0006355 | regulation of transcription, DNA-templated | 49/304 | 1601/27628 | 7.89e-11 | 4.41e-08 | Click
GO:2001141 | regulation of RNA biosynthetic process | 49/304 | 1601/27628 | 7.89e-11 | 4.41e-08 | Click
GO:0051252 | regulation of RNA metabolic process | 49/304 | 1607/27628 | 8.97e-11 | 4.41e-08 | Click
GO:0019219 | regulation of nucleobase-containing compound metabolic process | 49/304 | 1617/27628 | 1.11e-10 | 4.77e-08 | Click
GO:2000112 | regulation of cellular macromolecule biosynthetic process | 49/304 | 1640/27628 | 1.80e-10 | 6.19e-08 | Click
GO:0010556 | regulation of macromolecule biosynthetic process | 49/304 | 1640/27628 | 1.80e-10 | 6.19e-08 | Click
GO:0019222 | regulation of metabolic process | 52/304 | 1825/27628 | 2.47e-10 | 6.95e-08 | Click
GO:0048856 | anatomical structure development | 37/304 | 1026/27628 | 2.64e-10 | 6.95e-08 | Click
E
Dataset | Condition 1 | Condition 2 | Direction | Overlap/ Query | DEGs/ All genes | p-value | FDR (BH) | Enriched genes
Hormone-induced plant regeneration Tissue origin: Root (Ishihara et al., 2019) | Col-0 0DAC | Idl3 1DAS | Up | 128/200 | 4031/32833 | 1.57e-102 | 5.00e-100 | Click
Hormone-induced plant regeneration Tissue origin: Root (Ishihara et al., 2019) | Col-0 0DAC | Col-0 14DAC | Up | 128/200 | 4062/32833 | 3.98e-102 | 7.00e-100 | Click
Hormone-induced plant regeneration Tissue origin: Root (Ishihara et al., 2019) | Col-0 0DAC | Col-0 1DAS | Up | 121/200 | 3801/32833 | 1.14e-95 | 1.43e-93 | Click
F
Dataset | Condition 1 | Condition 2 | Direction | Overlap/ Query | DEGs/ All genes | p-value | FDR (BH) | Enriched genes
Hormone-induced plant regeneration Tissue origin: Root (Ishihara et al., 2019) | Col-0 0DAC | Col-0 14DAC | Down | 63/96 | 5394/32833 | 6.55e-45 | 1.30e-42 | Click
Hormone-induced plant regeneration Tissue origin: Root (Liu et al., 2020) | Col-0 7DAS | arr1 7DAS | Up | 96/96 | 23168/32833 | 6.92e-45 | 1.30e-42 | Click
Hormone-induced plant regeneration Tissue origin: Root (Ishihara et al., 2019) | Col-0 0DAC | Col-0 1DAS | Down | 61/96 | 4918/32833 | 1.32e-44 | 1.66e-42 | Click
G
PLT1 (AT3G20840)
PLT2 (AT1G51190)
SCR (AT3G54220)
WOX5 (AT3G11260)
JKD (AT5G03150)
Fraction of cells in group (%)
10 30 50 70 90
Mean expression in group
0.5 1.0 1.5
0. Epidermis-like cells
1. Vascular-initial-like cells
2. Vascular-initial-like cells
3. Explant vasculature/callus founder cells
4. Epidermis-like cells
5. QC-like cells
6. Vascular-initial-like cells
7. Explant vasculature/callus founder cells
8. Epidermis-like cells
9. Dividing cells
10. Lateral-root-cap-like cells
11. Lateral-root-cap-like cells
12. Dividing cells
13. Undassified
14. Undassified
H
LBD18 (AT2G45420)
AIL7 (AT5G65510)
MP (AT1G19850)
ATS (AT5G42630)
JKD (AT5G03150)
RBK2 (AT3G05140)
AITR6 (AT5G63350)
BBX21 (AT1G75540)
SCRM2 (AT1G12890)
SCR (AT3G54220)
AT1G69420
TLP6 (AT1G47270)
BS1 (AT1G74500)
SCL3 (AT1G50420)
PLT2 (AT1G51190)
NAP (AT1G69490)
PLT1 (AT3G20840)
WIP4 (AT3G20880)
AT2G28790
MYB34 (AT5G60890)
WOX5 (AT3G11260)
HSFA6B (AT3G22830)
AT2G42660











F

GO	Term name	Differential expression	Enriched genes/ DEGs	Genes in GO term/ All genes	p-value	FDR (BH)	Enriched genes
GO:0050896	response to stimulus	Upregulated	452/1766	3159/27628	3.99e-66	1.37e-62	Click
GO:1901700	response to oxygen-containing compound	Upregulated	182/1766	725/27628	3.67e-60	6.30e-57	Click
GO:0006950	response to stress	Upregulated	297/1766	1918/27628	1.34e-48	1.54e-45	Click
GO:0042221	response to chemical	Upregulated	228/1766	1358/27628	9.54e-43	8.20e-40	Click
GO:0010033	response to organic substance	Upregulated	171/1766	874/27628	1.18e-40	8.11e-38	Click
GO:1901698	response to nitrogen compound	Upregulated	63/1766	150/27628	1.16e-35	6.63e-33	Click
GO:0010200	response to chitin	Upregulated	58/1766	127/27628	2.15e-35	1.06e-32	Click
GO:0010243	response to organonitrogen compound	Upregulated	58/1766	135/27628	1.37e-33	5.88e-31	Click
GO:0009611	response to wounding	Upregulated	55/1766	132/27628	4.40e-31	1.68e-28	Click
GO:0006952	defense response	Upregulated	123/1766	701/27628	6.93e-25	2.38e-22	Click

the fact that pluripotent calli predominantly express root stem-cell regulators [26,27]. This analysis showed that transcriptomes of *WOX5*-expressing QC cells [32] are close to those of hormone-induced hypocotyl- and root-derived calli as well as hormone-induced immature embryos (Fig. S4). In contrast, transcriptomes of root maturation zones and differentiated cell types, such as those expressing *S18* (mature xylem pole), *COR* (mature cortex), and *COBL9* (root hair) [32], were the most distantly positioned to the transcriptomes of regenerating tissues (Fig. S4), which is consistent with the pluripotent nature of callus cells undergoing cellular reprogramming.

In addition, REGENOMICS supports the differential expression (DE) analysis using the deposited bulk RNA-seq data. Users can select two transcriptomes from a dataset deposited in REGENOMICS for the comparison. REGENOMICS identified DEGs between two samples, and the results were visualized as a downloadable volcano plot (Fig. 4E). GO analysis was also performed using each up-regulated and down-regulated genes, and the enriched GO terms could be displayed as a table (Fig. 4F).

4. Validation of co-expression networks in each mode of plant regeneration

One of the main features of REGENOMICS is to present co-expression networks, each of which is constructed based on the bulk RNA-seq transcriptome datasets related to a specific mode or to all modes of plant regeneration (wound-induced tissue repair, hormone-induced plant regeneration, somatic embryogenesis, and/or protoplast regeneration). To confirm whether the co-expression networks produced from different regeneration datasets are valid, we generated sub-networks using query genes that have been previously established as key regulators of each mode of plant regeneration.

ALF4, ERF115, WIND1, ASA1, YUC4, WOX11, ARF7, ARF19, and *LBD16*, that are known as key regulators of cellular reprogramming and callus formation at wound sites [17,22,23,24,35,81], were selected as query genes to build a sub-network using wound-induced tissue repair dataset. Interestingly, the sub-network constructed using only wound-induced tissue repair transcriptome dataset displayed close connections among these query genes along with additional key regulators of wound responses (Fig. 5A). This co-expression sub-network was enriched with GO terms related to wounding responses and *de novo* root-regeneration processes, such as 'auxin biosynthetic process', and 'lateral root formation' (Fig. 5B). However, when datasets of other plant regeneration modes were used, the sub-networks were fragmented without higher-order assembly and showed no enrichment of wound-related GO terms (Fig. S5A and S5B).

Similar results were obtained by submitting key regulators of hormone-induced *de novo* shoot regeneration, including *WOX5*, *WUS*, *STM*, *ESR1*/*DORNRÖSCHEN* (*DRN*), and *ESR2*/*DRN-LIKE* (*DRNL*) (Fig. 5C, S5C and S5D), as input query genes. The sub-network, which was constructed only with the hormone-induced plant regeneration datasets (Fig. 5C), showed the enrichment of GO terms such as 'anatomical structure development', 'stem cell population maintenance', and 'shoot system development' (Fig. 5D). Additionally, submission of key genes involved in somatic embryogenesis, including *ABI3*, *BBM*, *LEC1*, *LEC2*, *FUS3*, and *AGL15*, resulted in a sub-network enriched with genes associated with GO terms such as 'developmental process', 'regeneration', and 'somatic embryogenesis' (Fig. 5E and F). Sub-networks constructed with the same query genes using different datasets displayed different sub-networks with missing links between the key players (Fig. S5). Taken together, co-expression network analyses indicated that co-expressed genes depend heavily on the samples used as inputs and reflect the connections within a gene regulatory pathway specific to a given mode of plant regeneration.

Currently, several publicly accessible web applications containing extensive amounts of transcriptomes of diverse plant species are available for the examination of co-expression patterns of plant genes. In this regard, REGENOMICS enables plant regeneration-specific co-expression analyses and provides users with functional insights into the plant regeneration process. Notably, the co-expressions derived from the wound-induced tissue repair dataset were distinct from those derived from the co-expression networks publicly available at ATTED (https://atted.jp/) (Fig. S6A). Similarly, the sub-networks generated with hormone-induced plant regeneration and somatic embryogenesis datasets included co-expressions that could not be derived from co-expression network built with other transcriptome data (Fig. S6B and S6C).

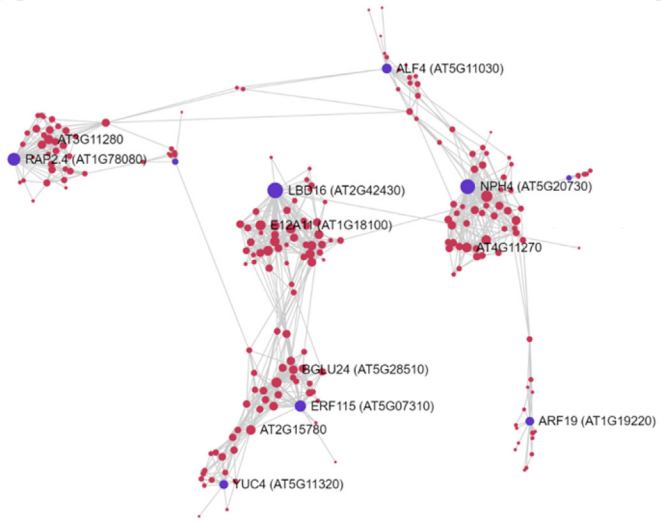
As most of the molecular genetic studies conducted on plant regeneration to date have been performed using the model plant *Arabidopsis*, currently REGENOMICS database is mainly composed of *Arabidopsis* regeneration-specific RNA-seq datasets. However, we aim to annually expand and update the REGENOMICS web application by including transcriptome datasets obtained from other plant species. Thus, multi-species analyses might also be available in a near future to help finding key genetic factors conserved in plant regeneration. Further, we will attempt to perform multi-omics analyses by including genomics and epigenomics data (e.g. dynamics on DNA methylation and histone modifications during plant regeneration) to provide a comprehensive insight into the plant regeneration process. Through these updates, REGENOMICS will develop as a key resource in exploring the genetic components and pathways underlying plant regeneration.

Fig. 4. Interface and features of Transcriptome analysis. (A) Interface of Transcriptome analysis. Users can select transcriptomes deposited in the REGENOMICS database (Table 1) and/or submit other custom transcriptome data in CSV format. (B) Heatmap of sample-to-sample correlation with hierarchical clustering. Correlation between samples are represented with colors according to the color scale. (C and D) Dimensionality reduction analysis. Principal component analysis (C) and multidimensional scaling analysis (D) were performed with transcriptomes datasets for hormone-induced plant regeneration transcriptomes and somatic embryogenesis transcriptomes. Each replicate is expressed as a dot on the scatter plot and colored according to its dataset. DAC, days after incubation on CIM; DAS, days after incubation on SIM; min, minutes after excision from seedling; HAI, hours after induction on E5 media. (E) Volcano plot of differentially expressed genes (DEGs). RNA-seq data of root samples harvested 0 h (0 h) and 1 h after wounding were used for DE analysis. Each gene is represented as a dot. Red and blue dots represent genes up-regulated and down-regulated, respectively, in 1 h after wounding compared with 0 h after wounding. (F) GO enrichment analysis performed with DEGs. The DEGs were obtained from (E). The analysis was performed separately for up-regulated genes. Top 10 GO terms were shown in the table. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Conclusions

Determining gene expression profiles and molecular networks provides important insights into gene functions in a given biological context. REGENOMICS is a collection of all available plant regeneration-related transcriptome datasets and implements easily accessible tools for the exploration of novel genetic components and connections underlying plant regeneration process. In addition, this web-based application permits user-friendly transcriptome-level analysis by using the databases constructed within REGENOMICS and customized user-submitted datasets to obtain systematic views on plant regeneration.

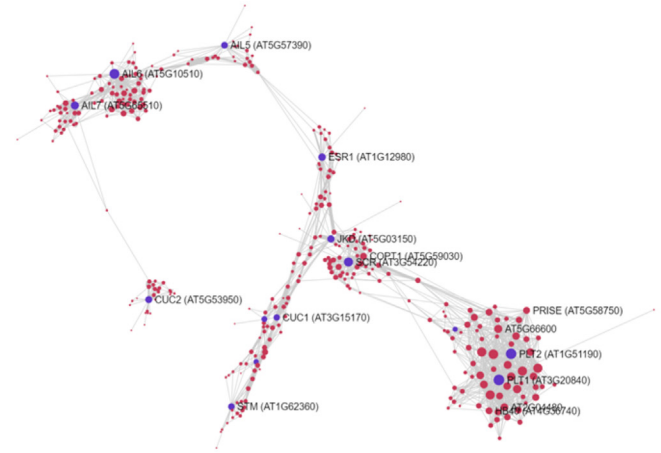
A



B

GO	Term name	Enriched genes/ Genes in network	Genes in GO term/ All genes	*p*-value	FDR (BH)	Enriched genes
GO:0000162	tryptophan biosynthetic process	5/215	27/27628	1.91e-06	2.51e-03	Click
GO:0046219	indolalkylamine biosynthetic process	5/215	27/27628	1.91e-06	2.51e-03	Click
GO:0009073	aromatic amino acid family biosynthetic process	6/215	49/27628	2.19e-06	2.51e-03	Click
GO:0006586	indolalkylamine metabolic process	5/215	31/27628	3.93e-06	2.70e-03	Click
GO:0006568	tryptophan metabolic process	5/215	31/27628	3.93e-06	2.70e-03	Click
GO:0009850	auxin metabolic process	5/215	33/27628	5.41e-06	2.84e-03	Click
GO:0019748	secondary metabolic process	10/215	208/27628	5.81e-06	2.84e-03	Click
GO:0009072	aromatic amino acid family metabolic process	6/215	59/27628	6.62e-06	2.84e-03	Click
GO:0010311	lateral root formation	4/215	17/27628	7.84e-06	2.99e-03	Click
GO:1901605	alpha-amino acid metabolic process	9/215	179/27628	1.20e-05	4.12e-03	Click
GO:0009851	auxin biosynthetic process	4/215	21/27628	1.92e-05	5.45e-03	Click

C



D

GO	Term name	Enriched genes/ Genes in network	Genes in GO term/ All genes	*p*-value	FDR (BH)	Enriched genes
GO:0032502	developmental process	65/463	1570/27628	1.73e-11	3.54e-08	Click
GO:0007389	pattern specification process	16/463	104/27628	2.06e-11	3.54e-08	Click
GO:0050789	regulation of biological process	101/463	3170/27628	1.43e-10	1.64e-07	Click
GO:0099402	plant organ development	22/463	253/27628	3.81e-10	3.28e-07	Click
GO:0048856	anatomical structure development	46/463	1026/27628	1.68e-09	9.47e-07	Click
GO:0050794	regulation of cellular process	93/463	2948/27628	1.84e-09	9.47e-07	Click
GO:0065007	biological regulation	105/463	3492/27628	1.93e-09	9.47e-07	Click
GO:0032501	multicellular organismal process	43/463	1022/27628	5.93e-08	2.55e-05	Click
GO:0019827	stem cell population maintenance	7/463	24/27628	9.60e-08	3.30e-05	Click
GO:0098727	maintenance of cell number	7/463	24/27628	9.60e-08	3.30e-05	Click

E

F

GO	Term name	Enriched genes/ Genes in network	Genes in GO term/ All genes	*p*-value	FDR (BH)	Enriched genes
GO:0019915	lipid storage	6/421	18/27628	1.92e-07	3.67e-05	Click
GO:0051235	maintenance of location	6/421	31/27628	6.44e-06	1.17e-03	Click
GO:0010262	somatic embryogenesis	3/421	4/27628	1.39e-05	2.17e-03	Click
GO:0010344	seed oilbody biogenesis	3/421	4/27628	1.39e-05	2.17e-03	Click
GO:0031099	regeneration	3/421	4/27628	1.39e-05	2.17e-03	Click
GO:0022414	reproductive process	32/421	927/27628	2.53e-05	3.76e-03	Click
GO:0032502	developmental process	46/421	1570/27628	2.62e-05	3.76e-03	Click
GO:0048609	multicellular organismal reproductive process	6/421	40/27628	2.99e-05	4.11e-03	Click
GO:0050794	regulation of cellular process	72/421	2948/27628	6.12e-05	8.00e-03	Click
GO:0048519	negative regulation of biological process	14/421	265/27628	6.28e-05	8.00e-03	Click
GO:0003006	developmental process involved in reproduction	26/421	753/27628	1.14e-04	1.40e-02	Click

6. Methods

6.1. Data collection and processing

To conduct a comprehensive analysis of various plant regeneration processes, all currently available bulk RNA-seq datasets, generated from a total of 391 wound-induced tissue regeneration, hormone-induced plant regeneration, somatic embryogenesis, and protoplast regeneration experiments (Table 1), were downloaded from the NCBI Sequence Read Archive (SRA), DDBJ, EMBL-EBI, and NGDC databases. The raw RNA-seq data were processed using a customized pipeline (Fig. 1). The TAIR10 reference genome sequence and Araport11 annotations (https://www.arabidopsis.org/) for *Arabidopsis*, the ITAG4.0 genome and annotation (https://solgenomics.net/) for tomato, the MSU7.0 (http://rice.uga.edu/) genome and annotation for rice, and MpTak1_v5.1r2 (https://marchantia.info/) genome and annotation for *Marchantia* were used as reference for mapping reads with STAR aligner [7] and quantified and normalized with the transcripts per million (TPM) method with RSEM [31].

In addition, single-cell transcriptomes were collected and processed. Standardized workflows, including the 10X Genomics CellRanger pipeline, the Seurat package in R [56], and Scanpy in Python [74], were used for raw data processing and subsequent single-cell analysis. For single-cell transcriptomes generated with the Smart-Seq2 protocol, STAR aligner was used. UMI-tools was used to analyze single-cell transcriptomes generated by the BD Rhapsody system [62]. The TAIR10 and ITAG4.0 reference genomes were used for *Arabidopsis* and tomato datasets, respectively. Cell type-specific marker gene lists were prepared, based on previous studies [3,4,8,32,65,72,77,82,84], and were used with the cluster-specific marker genes given in the original study for each dataset. These marker genes were then used as input to score and annotate each cluster for each cell type using scanpy.tl.score_genes provided in Scanpy.

6.2. Gene expression browser and heatmap

Gene counts normalized with the TPM normalization method were displayed on the Gene expression browser in Single-gene analysis. Plotting was done with the ggplot2 package in R. For the Gene expression heatmap in Multi-gene analysis, TPM values were scaled with StandardScaler and MinMaxScaler from the scikit-learn Python module [48], and displayed as Z-scores.

6.3. Gene co-expression network analysis and Sub-network generation

Gene counts normalized with the TPM normalization method were log-transformed as $\log_2(x + 1)$ to create gene expression matrices for each type of plant regeneration. Next, the resulting matrix was used to calculate Pearson correlation coefficients between each pair of gene. The ranks of correlation coefficients were then measured and used to obtain the geometric mean, the Mutual Rank (MR) [44]. Co-expression networks were first generated with all transcriptomes in the REGENOMICS database as input for detecting gene correlations common to a variety of regeneration methods. Then, datasets corresponding to the four plant regeneration modes were used separately to create a co-expression network specific to each mode of regeneration. Sub-networks were generated based on the query gene(s), connecting the top 20 closest neighbor genes (based on the MR values) with significant co-expression values. When constructing sub-networks with multiple query genes, only close neighbor genes with strong co-expression (MR $\leq$ 30) were used to build the networks in order to avoid creating an overly dense co-expression network. Sub-networks were generated using the Pandas, NetworkX [11], and ForceAtlas2 Python packages, and visualized with d3.js (https://d3js.org/), ObservableHQ (https://observablehq.com), and Sigma.js (https://www.sigmajs.org/).

Fig. 5. Validation of co-expression networks in each mode of plant regeneration. (A, C and E) Sub-networks generated with key regulators of each mode of plant regeneration. The sub-networks were drawn using bulk RNA-seq datasets related to wound-induced tissue repair (A), hormone-induced plant regeneration (C), and somatic embryogenesis (E). Input query genes (A, *ALF4, ERF115, WIND1, ASA1, YUC4, WOX11, ARF7, ARF19,* and *LBD16*; C, *WUS, STM, CUC1, CUC2, PLT1, PLT2, PLT3, PLT5, PLT7, WOX5, ESR1/DRN, ESR2/DRNL, SCR,* and *JKD*; E, *ABI3, BBM, LEC1, LEC2, FUS3, AGL15, SERK1, GA2OX6, GA3OX1, IAA30, MYB118, WUS, YUC2, YUC4* and *YUC10*) are colored violet and co-expressed genes are colored magenta. (B, D and F) GO enrichment analysis. GO terms enriched in each output sub-network (A, C and E) are shown. A part of the significant GO term list is shown in B, D and F. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6.4. Gene regulatory network inference and sub-network generation

Gene regulatory network inference for each scRNA-seq dataset was performed using the SCENIC workflow [68]. Pre-processing was performed in Scanpy, and the network was inferred using GRNBoost2 [42]. Transcription factor gene list was obtained from PlantTFDB 4.0 [25]. The directed edges between nodes were ranked by importance, and the top 10 edges connected with the query gene were first taken to build the sub-network in Single-gene analysis. Directly linked neighbor genes were further taken to build an extensive gene regulatory network. For Multi-gene analysis, only gene connections with an importance score of 10 or greater were used to build the sub-networks. Visualization of the sub-networks were done using the same tools and packages as co-expression networks.

6.5. Single-cell RNA-seq visualization and trajectory inference analysis

Expression of query gene(s) in Single cell analysis was visualized with Uniform Manifold Approximation and Projection (UMAP) plots, violin plots, and dot plots, which were drawn using the Scanpy toolkit. Trajectory inference was also performed in Scanpy, using scanpy.tl.dpt to compute pseudotime for each gene. The selection of root cell type was done as shown in each original study, and the visualization of the pseudotime was give as a UMAP plot.

6.6. GO and DEG enrichment analyses

GO enrichment analysis was conducted using the GOATOOLS Python library [28]. Only protein-coding genes were used as input for analysis. The table of enriched GO terms was filtered, showing terms with a p-value <0.05 and those categorized as biological processes. To perform DEG enrichment analysis, genes with fold-change >2 and p-value <0.05 were identified as DEGs with edgeR [50]. The significance of enrichment was tested with Fisher's exact test, and p-values were corrected using the Benjamini and Hochberg method [2].

6.7. Correlation heatmap and dimensionality reduction

Spearman's rank correlation coefficients were calculated with the Pandas Python package, and visualized with InCHlib [61] and ObservableHQ. Dimensionality reduction was conducted as PCA and MDS with the Pandas, NumPy [12], and scikit-learn Python packages. Results were visualized as 2D and 3D scatter plots with the Plotly Python package.

6.8. Differential expression analysis

Pre-processed RNA-seq data were used for differential expression (DE) analysis. Using edgeR, the reads from the selected RNA-seq data were normalized with CPM, and DEGs were identified if the fold-change >2 and p-value <0.05. DEGs were visualized as a volcano plot with Plotly.

6.9. Application architecture

REGENOMICS was built based on the Django web framework (https://www.djangoproject.com) on a Linux platform. Python and R scripts were implemented for the analysis pipeline. STAR, RSEM, CellRanger, UMI-tools, Seurat, and Scanpy were used for pre-processing datasets. The user-interface for REGENOMICS was built with HTML, JavaScript, and Semantic UI (https://semantic-ui.com/). The results were visualized mainly with d3.js, sigma.js, and plotly.js (https://plotly.com/). The co-expression network (https://observablehq.com/@tnsgud996/co-expression-network) and gene regulatory network (https://observablehq.com/@tnsgud996/gene-regulatory-network) for Single-gene analysis, and colored heatmap (https://observablehq.com/@tnsgud996/correlation-heatmap) for Gene expression heatmap in Multi-gene analysis were built with and are available on ObservableHQ.

Funding

This work was supported by the Samsung Science and Technology Foundation under Project Number SSTF-BA2001-10. S.H.B. is grateful for financial support from Hyundai Motor Chung Mong-Koo Foundation.

Author contributions

P.J.S. and Y.S.N. conceived the project; S.H.B. constructed the database and designed the web application; P.J.S. revised the web application; S.H.B. wrote the first draft of the manuscript; P.J.S. and Y.S.N. revised the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.06.033.

References

[1] Atta R, Laurens L, Boucheron-Dubuisson E, Guivarc'h A, Carnero E, Giraudat-Pautot V, et al. Pluripotency of Arabidopsis xylem pericycle underlies shoot regeneration from root and hypocotyl explants grown in vitro. Plant J 2009;57:626–44.
[2] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodological) 1995;57:289–300.
[3] Brady SM, Orlando DA, Lee JY, Wang JY, Koch J, Dinneny JR, et al. A high-resolution root spatiotemporal map reveals dominant expression patterns. Science 2007;318:801–6.
[4] Bruex A, Kainkaryam RM, Wieckowski Y, Kang YH, Bernhardt C, Xia Y, et al. A gene regulatory network for root epidermis cell differentiation in arabidopsis. PLoS Genet 2012;8:e1002446.
[5] Chupeau M-C, Granier F, Pichon O, Renou J-P, Gaudin V, Chupeau Y. Characterization of the early events leading to totipotency in an Arabidopsis protoplast liquid culture by temporal transcript profiling. Plant Cell 2013;25:2444–63.
[6] Coleman D, Kawamura A, Ikeuchi M, Favero DS, Lambolez A, Rymen B, et al. The SUMO E3 ligase SIZ1 negatively regulates shoot regeneration. Plant Physiol 2020;184:330–44.
[7] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
[8] Efroni I, Mello A, Nawy T, Ip P-L, Rahni R, DelRose N, et al. Root regeneration triggers an embryo-like sequence guided by hormonal interactions. Cell 2016;165:1721–33.
[9] Gaj MD. Direct somatic embryogenesis as a rapid and efficient system for in vitro regeneration of Arabidopsis thaliana. Plant Cell, Tissue Organ Cult 2001;64:39–46.
[10] Green C, Phillips R. Plant regeneration from tissue cultures of maize 1. Crop Sci 1975;15:417–21.
[11] Hagberg A, Swart PS, Chult D. Exploring network structure, dynamics, and function using networkx. *In*, United States; 2008.
[12] Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature 2020;585:357–62.
[13] Hernández-Coronado M, Araujo PCD, Ip P-L, Nunes CO, Rahni R, Wudick MM, et al. Plant glutamate receptors mediate a bet-hedging strategy between regeneration and defense. Dev Cell 2022;57(451–465):e456.
[14] Horstman A, Li M, Heidmann I, Weemen M, Chen B, Muino JM, et al. The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. Plant Physiol 2017;175:848–57.
[15] Hsu F-M, Gohain M, Allishe A, Huang Y-J, Liao J-L, Kuang L-Y, et al. Dynamics of the methylome and transcriptome during the regeneration of rice. Epigenomes 2018;2:14.
[16] Hu X, Xu L. Transcription factors WOX11/12 directly activate WOX5/7 to promote root primordia initiation and organogenesis. Plant Physiol 2016;172:2363–73.
[17] Ikeuchi M, Iwase A, Rymen B, Lambolez A, Kojima M, Takebayashi Y, et al. Wounding triggers callus formation via dynamic hormonal and transcriptional changes. Plant Physiol 2017;175:1158–74.
[18] Ikeuchi M, Ogawa Y, Iwase A, Sugimoto K. Plant regeneration: cellular origins and molecular mechanisms. Development 2016;143:1442–51.
[19] Ishida S, Suzuki H, Iwaki A, Kawamura S, Yamaoka S, Kojima M, et al. Diminished auxin signaling triggers cellular reprogramming by inducing a regeneration factor in the Liverwort Marchantia polymorpha. Plant Cell Physiol 2022;63:384–400.
[20] Ishihara H, Sugimoto K, Tarr PT, Temman H, Kadokura S, Inui Y, et al. Primed histone demethylation regulates shoot regenerative competency. Nat Commun 2019;10:1786.
[21] Ishikawa M, Murata T, Sato Y, Nishiyama T, Hiwatashi Y, Imai A, et al. Physcomitrella cyclin-dependent kinase A links cell cycle reactivation to other cellular changes during reprogramming of leaf cells. Plant Cell 2011;23:2924–38.
[22] Iwase A, Harashima H, Ikeuchi M, Rymen B, Ohnuma M, Komaki S, et al. WIND1 promotes shoot regeneration through transcriptional activation of ENHANCER OF SHOOT REGENERATION1 in arabidopsis. Plant Cell 2017;29:54–69.
[23] Iwase A, Kondo Y, Laohavisit A, Takebayashi A, Ikeuchi M, Matsuoka K, et al. WIND transcription factors orchestrate wound-induced callus formation, vascular reconnection and defense response in Arabidopsis. New Phytol 2021;232:734–52.
[24] Iwase A, Mitsuda N, Koyama T, Hiratsu K, Kojima M, Arai T, et al. The AP2/ERF transcription factor WIND1 controls cell dedifferentiation in arabidopsis. curr biol 2011;21:508–14.
[25] Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res 2016;45:D1040–5.
[26] Kareem A, Durgaprasad K, Sugimoto K, Du Y, Pulianmackal AJ, Trivedi ZB, et al. PLETHORA genes control regeneration by a two-step mechanism. Curr Biol 2015;25:1017–30.
[27] Kim JY, Yang W, Forner J, Lohmann JU, Noh B, Noh YS. Epigenetic reprogramming by histone acetyltransferase HAG1/AtGCN5 is required for pluripotency acquisition in Arabidopsis. EMBO J 2018;37.
[28] Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, Naldi A, et al. GOATOOLS: A Python library for gene ontology analyses. Sci Rep 2018;8:10872.
[29] Lambolez A, Kawamura A, Takahashi T, Rymen B, Iwase A, Favero DS, et al. Warm temperature promotes shoot regeneration in Arabidopsis thaliana. Plant Cell Physiol 2022;63:618–34.
[30] Larriba E, Sánchez-García AB, Martínez-Andújar C, Albacete A, Pérez-Pérez JM. Tissue-specific metabolic reprogramming during wound-induced organ formation in tomato hypocotyl explants. Int J Mol Sci 2021;22:10112.
[31] Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinf 2011;12:1–16.
[32] Li S, Yamada M, Han X, Ohler U, Benfey PN. High-resolution expression map of the Arabidopsis root reveals alternative splicing and lincRNA regulation. Dev Cell 2016;39:508–22.
[33] Liu H, Ma X, Han Hua N, Hao YJ, Zhang Xian S. AtPRMT5 regulates shoot regeneration through mediating histone H4R3 Dimethylation on KRPs and Pre-mRNA Splicing of RKP in Arabidopsis. Mol Plant 2016;9:1634–46.
[34] Liu J, Hu X, Qin P, Prasad K, Hu Y, Xu L. The WOX11-LBD16 pathway promotes pluripotency acquisition in callus cells during de novo shoot regeneration in tissue culture. Plant Cell Physiol 2018;59:734–43.

[35] Liu J, Sheng L, Xu Y, Li J, Yang Z, Huang H, et al. WOX11 and 12 are involved in the first-step cell fate transition during de novo root organogenesis in Arabidopsis. Plant Cell 2014;26:1081–93.
[36] Liu W, Zhang Y, Fang X, Tran S, Zhai N, Yang Z, et al. Transcriptional landscapes of de novo root regeneration from detached Arabidopsis leaves revealed by time-lapse and single-cell RNA sequencing analyses.. Plant Commun 2022;100306.
[37] Liu Z, Dai X, Li J, Liu N, Liu X, Li S, et al. The Type-B cytokinin response regulator ARR1 inhibits shoot regeneration in an ARR12-dependent manner in Arabidopsis. Plant Cell 2020;32:2271–91.
[38] Marhava P, Hoermayer L, Yoshida S, Marhavý P, Benková E, Friml J. Re-activation of stem cell pathways for pattern restoration in plant wound healing. Cell 2019;177:957–969.e913.
[39] Mathew MM, Prasad K. Model systems for regeneration: Arabidopsis. Development 2021;148:dev195347.
[40] Matosevich R, Cohen I, Gil-Yarom N, Modrego A, Friedlander-Shani L, Verna C, et al. Local auxin biosynthesis is required for root regeneration after wounding. Nat Plants 2020;6:1020–30.
[41] Mayer KF, Schoof H, Haecker A, Lenhard M, Jurgens G, Laux T. Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. Cell 1998;95:805–15.
[42] Moerman T, Aibar Santos S, Bravo González-Blas C, Simm J, Moreau Y, Aerts J, et al. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. Bioinformatics 2018;35:2159–61.
[43] Nishihama R, Ishizaki K, Hosaka M, Matsuda Y, Kubota A, Kohchi T. Phytochrome-mediated regulation of cell division and growth during regeneration and sporeling development in the liverwort Marchantia polymorpha. J Plant Res 2015;128:407–21.
[44] Obayashi T, Aoki Y, Tadaka S, Kagaya Y, Kinoshita K. ATTED-II in 2018: A plant coexpression database based on investigation of the statistical property of the mutual rank index. Plant Cell Physiol 2018;59:440.
[45] Omary M, Gil-Yarom N, Yahav C, Steiner E, Hendelman A, Efroni I. A conserved superlocus regulates above- and belowground root initiation. Science 2022;375:eabf4368.
[46] Ou Y, Tao B, Wu Y, Cai Z, Li H, Li M, et al. Essential roles of SERKs in the ROOT MERISTEM GROWTH FACTOR-mediated signaling pathway. Plant Physiol 2022.
[47] Pan J, Zhao F, Zhang G, Pan Y, Sun L, Bao N, et al. Control of de novo root regeneration efficiency by developmental status of Arabidopsis leaf explants. J Genet Genomics 2019;46:133–40.
[48] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn. Machine learning in Python 2011;12:2825–30.
[49] Reinhardt D, Frenz M, Mandel T, Kuhlemeier C. Microsurgical and laser ablation analysis of interactions between the zones and layers of the tomato shoot apical meristem. Development 2003;130:4073–83.
[50] Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–40.
[51] Roest S, Gilissen L. Plant regeneration from protoplasts: a literature review. Acta Botan Neerl 1989;38:1–23.
[52] Rymen B, Kawamura A, Lambolez A, Inagaki S, Takebayashi A, Iwase A, et al. Histone acetylation orchestrates wound-induced transcriptional activation and cellular reprogramming in Arabidopsis. Commun Biol 2019;2:404.
[53] Saelens W, Cannoodt R, Todorov H, Saeys Y. A comparison of single-cell trajectory inference methods. Nat Biotechnol 2019;37:547–54.
[54] Sakamoto Y, Kawamura A, Suzuki T, Segami S, Maeshima M, Polyn S, De Veylder L, Sugimoto K (2021) Transcriptional activation of auxin biosynthesis drives developmental reprogramming of differentiated cells. bioRxiv.
[55] Santuari L, Sanchez-Perez GF, Luijten M, Rutjens B, Terpstra I, Berke L, et al. The PLETHORA gene regulatory network guides growth and cell differentiation in Arabidopsis roots. Plant Cell 2016;28:2937–51.
[56] Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 2015;33:495–502.
[57] Sena G, Wang X, Liu H-Y, Hofhuis H, Birnbaum KD. Organ regeneration does not require a functional stem cell niche in plants. Nature 2009;457:1150–3.
[58] Serrano-Ron L, Perez-Garcia P, Sanchez-Corrionero A, Gude I, Cabrera J, Ip P-L, et al. Reconstruction of lateral root formation through single-cell RNA sequencing reveals order of tissue initiation. Mol Plant 2021;14:1362–78.
[59] Shim S, Kim HK, Bae SH, Lee H, Lee HJ, Jung YJ, et al. Transcriptome comparison between pluripotent and non-pluripotent calli derived from mature rice seeds. Sci Rep 2020;10:1–12.
[60] Skoog F, Miller CO. Chemical regulation of growth and organ formation in plant tissues cultured in vitro. Symp Soc Exp Biol 1957;11:118–30.
[61] Škuta C, Bartůněk P, Svozil D. InCHlib – interactive cluster heatmap for web applications. J Cheminf 2014;6:44.
[62] Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res 2017;27:491–9.
[63] Sugimoto K, Jiao Y, Meyerowitz EM. Arabidopsis regeneration from multiple tissues occurs via a root development pathway. Dev Cell 2010;18:463–71.
[64] Takebe I, Labib G, Melchers G. Regeneration of whole plants from isolated mesophyll protoplasts of tobacco. Naturwissenschaften 1971;58:318–20.
[65] Tian C, Wang Y, Yu H, He J, Wang J, Shi B, et al. A gene expression map of shoot domains reveals regulatory mechanisms. Nat Commun 2019;10:141.
[66] Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 2014;32:381–6.
[67] Tripathi RK, Wilkins O. Single cell gene regulatory networks in plants: Opportunities for enhancing climate change stress resilience. Plant, Cell Environ 2021;44:2006–17.
[68] Van de Sande B, Flerin C, Davie K, De Waegeneer M, Hulselmans G, Aibar S, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. Nat Protoc 2020;15:2247–76.
[69] van den Berg C, Willemsen V, Hendriks G, Weisbeek P, Scheres B. Short-range control of cell differentiation in the Arabidopsis root meristem. Nature 1997;390:287–9.
[70] Waese J, Fan J, Pasha A, Yu H, Fucile G, Shi R, et al. ePlant: visualizing and exploring multiple levels of data for hypothesis generation in plant biology. Plant Cell 2017;29:1806–21.
[71] Wang FX, Shang GD, Wu LY, Xu ZG, Zhao XY, Wang JW. Chromatin accessibility dynamics and a hierarchical transcriptional regulatory network structure for plant somatic embryogenesis. Dev Cell 2020;54(742–757):e748.
[72] Wendrich JR, Yang B, Vandamme N, Verstaen K, Smet W, Van de Velde C, et al. Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions. Science 2020;370:eaay4970.
[73] Wickramasuriya AM, Dunwell JM. Global scale transcriptome analysis of Arabidopsis embryogenesis in vitro. BMC Genomics 2015;16:301.
[74] Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 2018;19:15.
[75] Wu Y, Haberland G, Zhou C, Koop HU. Somatic embryogenesis, formation of morphogenetic callus and normal development in zygotic embryos ofArabidopsis thaliana in vitro. Protoplasma 1992;169:89–96.
[76] Xu M, Du Q, Tian C, Wang Y, Jiao Y. Stochastic gene expression drives mesophyll protoplast regeneration. Sci Adv 2021;7:eabg8466.
[77] Yadav RK, Tavakkoli M, Xie M, Girke T, Reddy GV. A high-resolution gene expression map of the Arabidopsis shoot meristem stem cell niche. Development 2014;141:2735–44.
[78] Ye BB, Shang GD, Pan Y, Xu ZG, Zhou CM, Mao YB, et al. AP2/ERF transcription factors integrate age and wound signals for root regeneration. Plant Cell 2020;32:226–41.
[79] Zappia L, Theis FJ. Over 1000 tools reveal trends in the single-cell RNA-seq analysis landscape. Genome Biol 2021;22:301.
[80] Zhai N, Xu L. Pluripotency acquisition in the middle cell layer of callus is required for organ regeneration. Nat Plants 2021;7:1453–60.
[81] Zhang G, Zhao F, Chen L, Pan Y, Sun L, Bao N, et al. Jasmonate-mediated wound signalling promotes plant regeneration. Nat Plants 2019;5:491–7.
[82] Zhang T-Q, Chen Y, Wang J-W. A single-cell analysis of the Arabidopsis vegetative shoot apex. Dev Cell 2021.
[83] Zhang TQ, Lian H, Zhou CM, Xu L, Jiao Y, Wang JW. A two-step model for de novo activation of WUSCHEL during plant shoot regeneration. Plant Cell 2017;29:1073–87.
[84] Zhang TQ, Xu ZG, Shang GD, Wang JW. A single-cell RNA sequencing profiles the developmental landscape of Arabidopsis root. Mol Plant 2019;12:648–60.
[85] Zhou W, Lozano-Torres JL, Blilou I, Zhang X, Zhai Q, Smant G, et al. A jasmonate signaling network activates root stem cells and promotes regeneration. Cell 2019;177(942–956):e914.