

EMBL-EBI workshop: Exploring plant gene expression in Expression Atlas

Expression Atlas: integrating and displaying plant gene expression

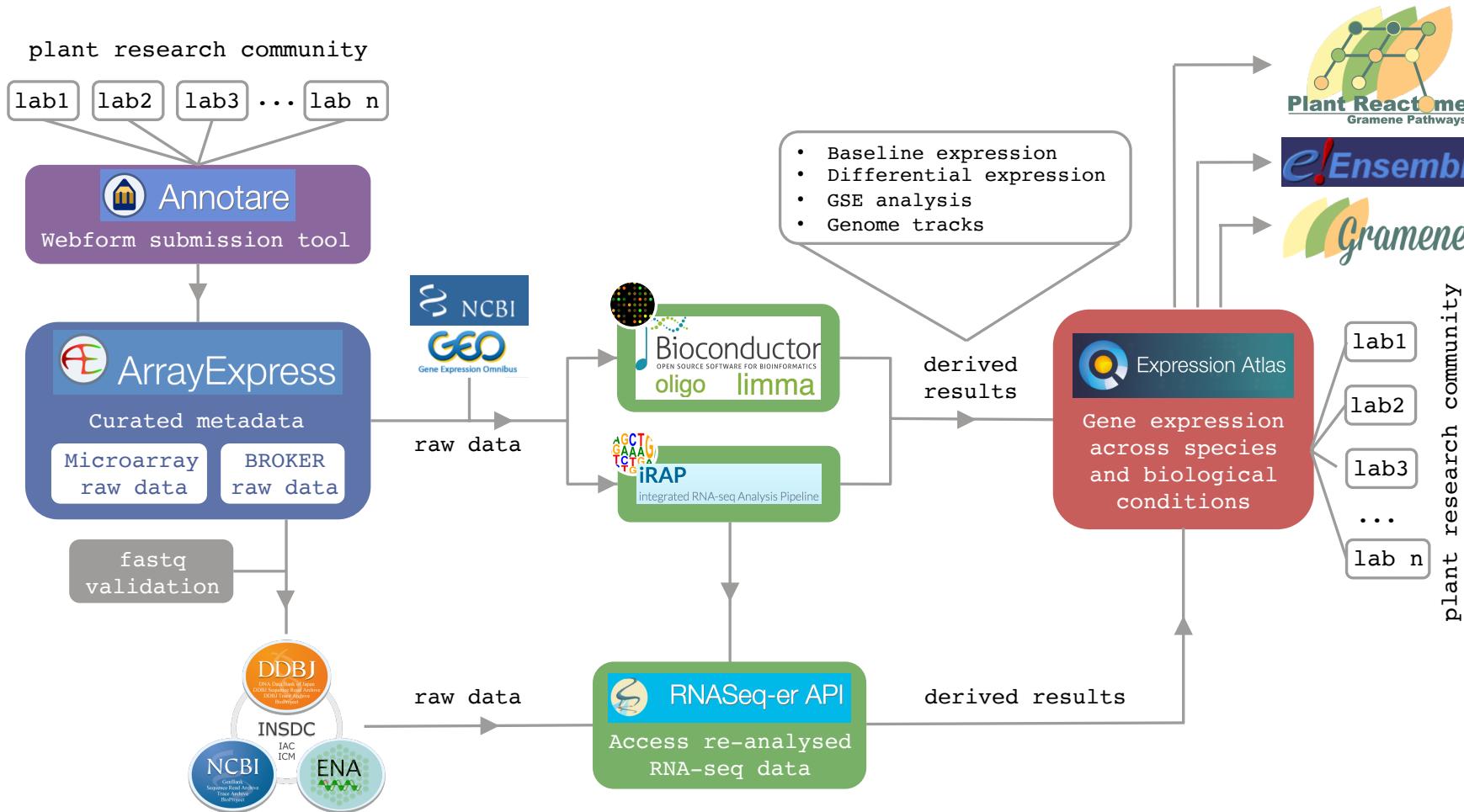
Laura Huerta, PhD

Senior Scientific Curator

lauhuema@ebi.ac.uk

10th SPPS PhD Student Conference

Plant expression data at EMBL-EBI





What is Plant Expression Atlas?



Exploring gene expression results across species under different biological conditions

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others. Expression Atlas aims to help answering questions such as 'where is a certain gene expressed?' or 'how does its expression change in a disease?'.

[Read more about Expression Atlas](#)

Search Gene set enrichment

Gene / Gene properties

Enter gene query...
Examples: REG1B, zinc finger, Q14777 (UniProt), GO:0010468 (regulation of gene expression)

Organism

Any

Biological conditions

Enter condition query...
Examples: lung, leaf, valproic acid, cancer

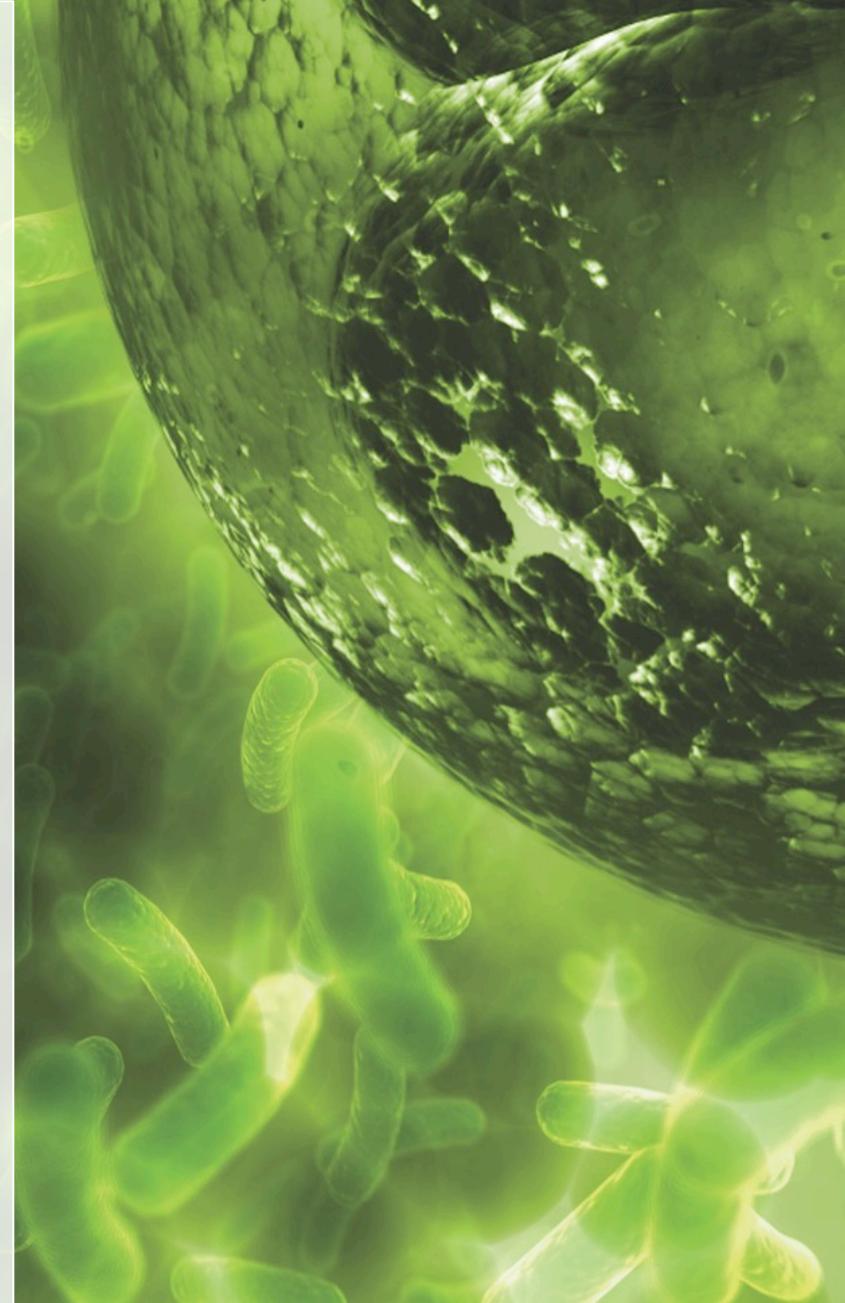
[Search](#) [Clear](#)

Open science resource for gene/protein expression across plant species and biological conditions

www.ebi.ac.uk/gxa/

OUR MISSION

To provide to the plant scientific community **freely available** information on the **abundance and localisation of RNA** (and proteins) across plant species and biological conditions



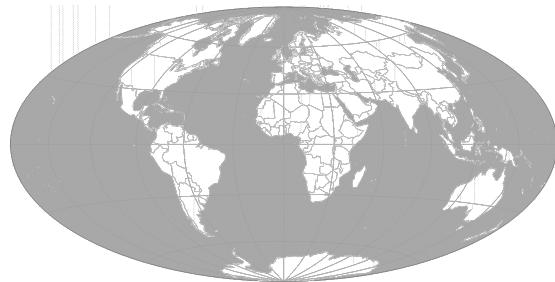


What is Plant Expression Atlas?

To achieve this objective, Expression Atlas involves **data curation**, **data analysis** and the development of a web application to **access** and **visualize** expression analysis results.



Plant Expression Atlas contains hundreds of selected and curated datasets



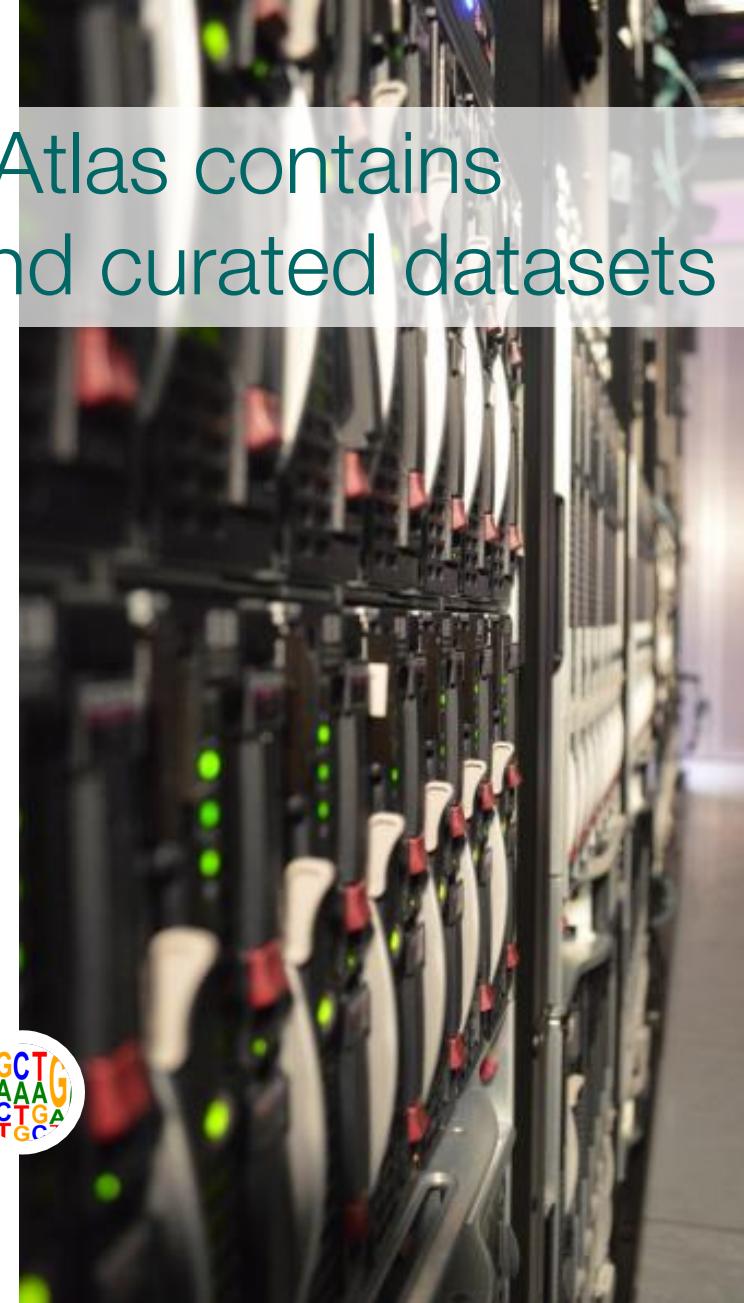
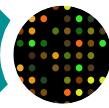
791 plant datasets



60 baseline expression



731 differential expression



... more than 25 plant species ...



www.gramene.org/

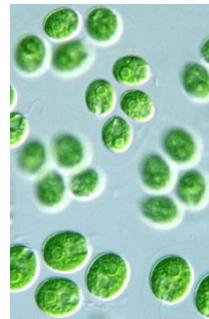
Data curation – new plant species in 2018



Beta vulgaris
(sugar beet)



Musa acuminata
Cavendish banana



Chlamydomonas
reinhardtii



Brassica napus
(rapeseed)



Setaria italica
(foxtail millet)



Triticum urartu
(AA, $2n = 2x = 14$)



Trifolium pratense
(red clover)



Prunus persica
(peach)

Data curation – new plant species in 2018

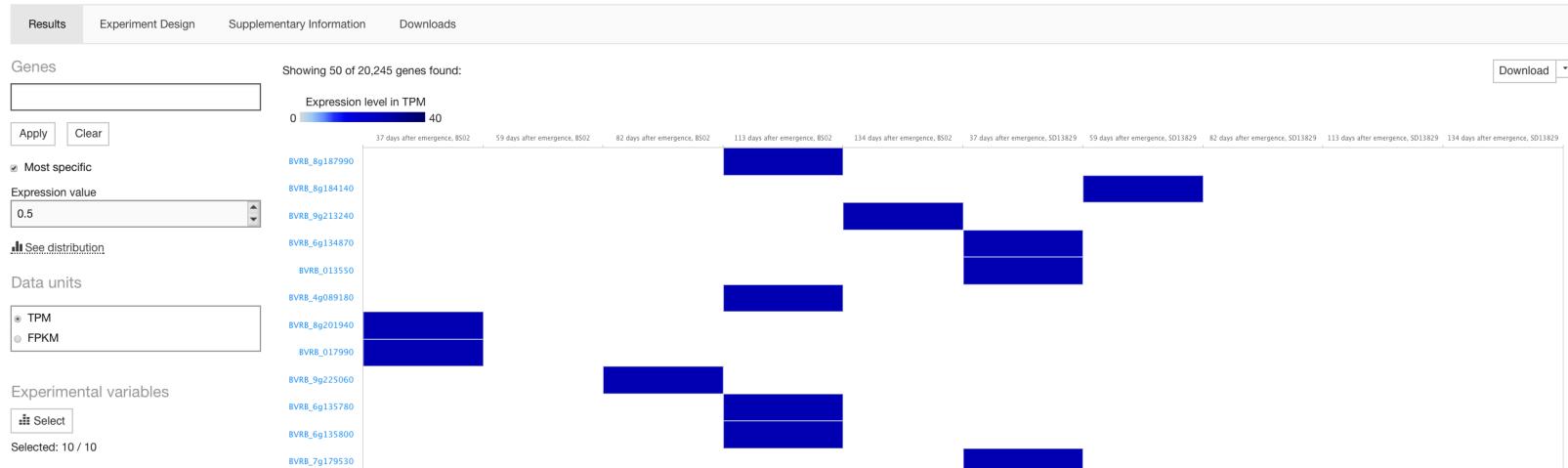


Beta vulgaris (sugar beet)

Transcriptomic analysis of the high taproot yield cultivar SD13829 and the high sucrose content cultivar BS02 at five developmental stages to study the regulation of these two **agronomic traits**

RNA-seq of a E-type and a Z-type cultivar of sugar beet at five different developmental stages

RNA-Seq mRNA baseline
Organism: *Beta vulgaris* subsp. *vulgaris*
Reference(s): 28406933 ([Filter by genes in paper](#))



www.ebi.ac.uk/gxa/experiments/E-ENAD-2/Results/

Data curation – new plant species in 2018



Brassica napus (canola)

Transcriptomic analysis of seed development in canola to identify potential target genes for the genetic improvement of the **yield and composition of canola oil**

RNA-seq of *Brassica napus* seeds at different developmental stages

RNA-Seq mRNA baseline
Organism: *Brassica napus*



www.ebi.ac.uk/gxa/experiments/E-GEOD-77637/Results/

Data curation – new plant species in 2018



Prunus persica (peach)

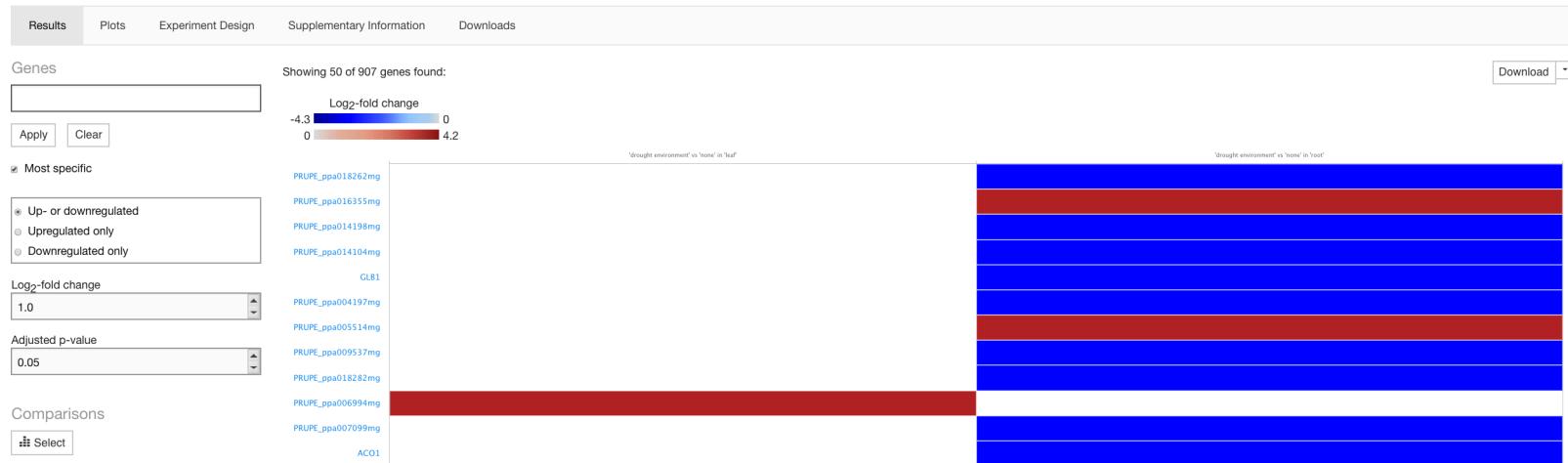
Transcriptomic analysis of peach root tissues (GF677 rootstock) and leaf tissues (graft, var. Catherina) after 16 days of drought stress to study the **response to drought**

RNA-seq of *Prunus persica* roots and leaves under drought stress

RNA-Seq mRNA differential

Organism: *Prunus persica*

Reference(s): 27933070 [Filter by genes in paper]



www.ebi.ac.uk/gxa/experiments/E-ENAD-3/Results/

Data curation – new plant species in 2018



Musa acuminata

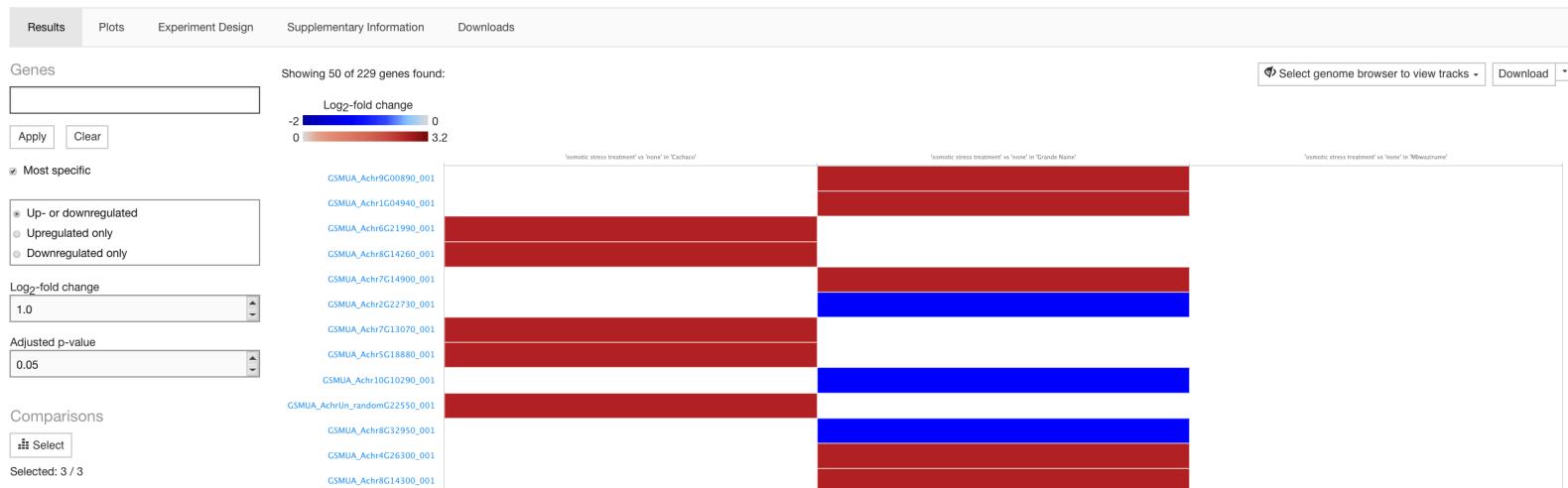
Transcriptomic analysis of three triploid banana genotypes grown under 5% PEG and control conditions to study the response to **mild osmotic stress in roots**

RNA-seq of banana in vitro roots under mild osmotic stress

RNA-Seq mRNA differential

Organism: *Musa acuminata*

Reference(s): 26935041 (Filter by genes in paper)



www.ebi.ac.uk/gxa/experiments/E-ENAD-8/Results/

Data curation – new anatomo^{grams}

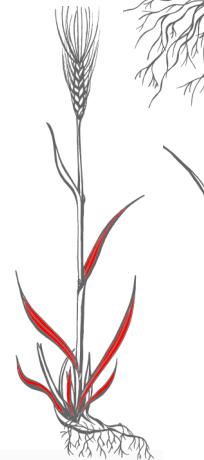
- ✓ *Arabidopsis lyrata*
- ✓ *Beta vulgaris*
- ✓ *Brassica napus*
- ✓ *Brassica oleracea*
- ✓ *Medicago truncatula*
- ✓ *Chlamydomonas reinhardtii*
- ✓ *Brassica rapa*
- ✓ *Glycine max*



Illustrations by Jana Eliasova

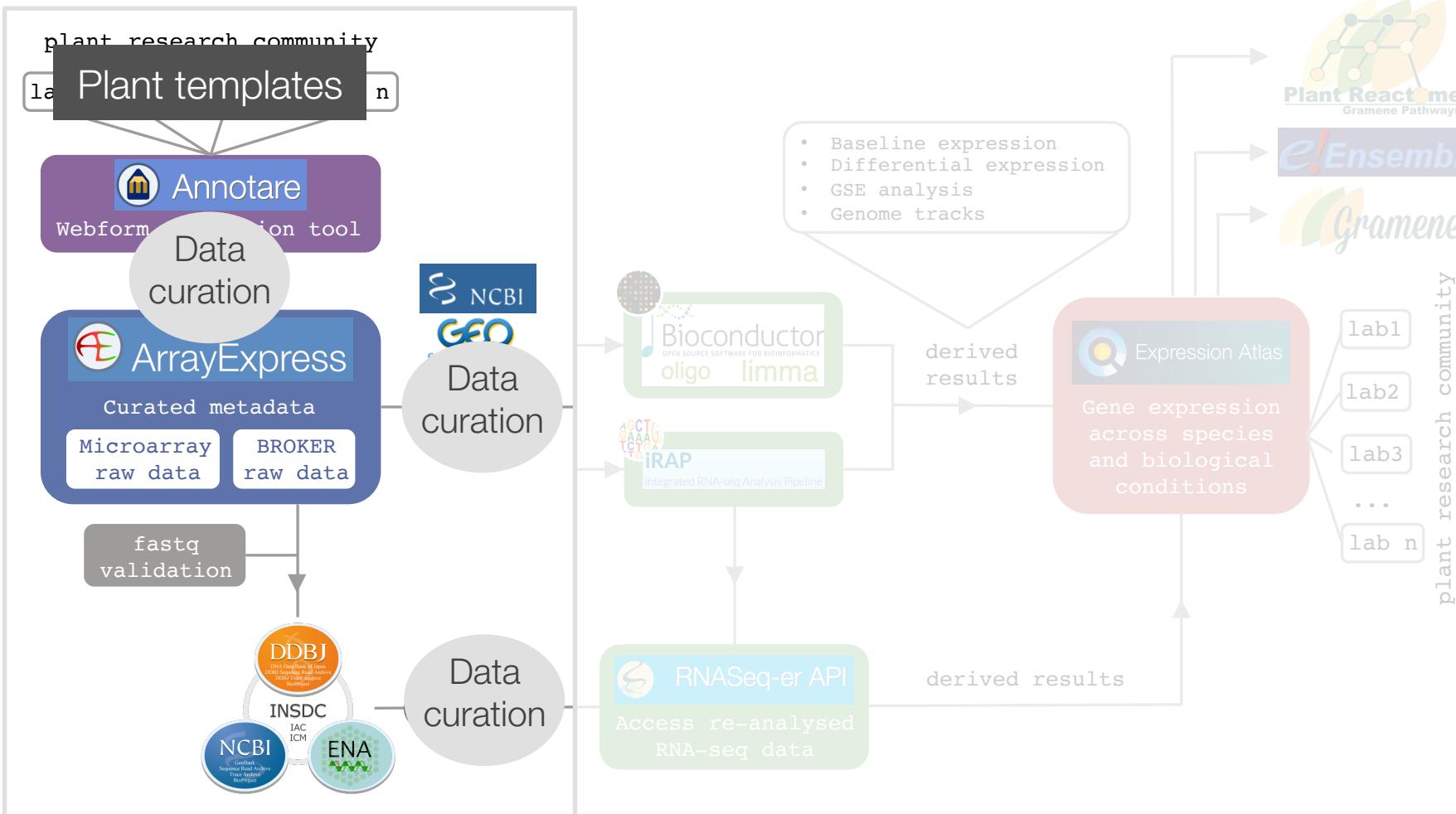
Data curation – new anatomo-grams

- ✓ *Musa acuminata*
- ✓ *Physcomitrella patens*
- ✓ *Prunus persica*
- ✓ *Setaria italica*
- ✓ *Solanum tuberosum*
- ✓ *Theobroma cacao*
- ✓ *Trifolium pratense*
- ✓ *Triticum urartu*
- ✓ *Vitis vinifera*



Illustrations by Jana Eliasova

Data curation – plant expression data





Plant data submissions via Annotare

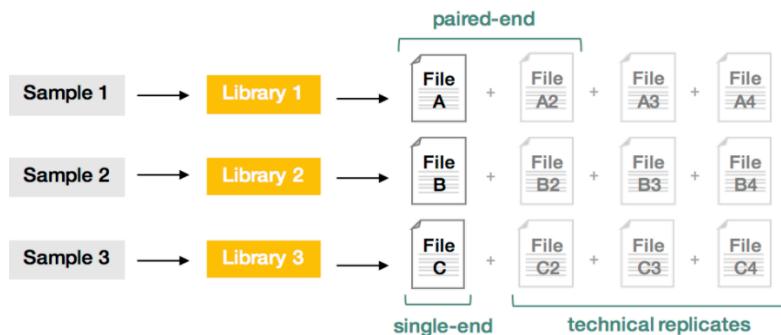
Plant templates

Select Submission Template Plant - High-throughput sequencing

⚠ Please ensure you have selected the correct template as this cannot be changed.

A sequencing experiment produces raw sequence data generated by next-generation sequencing instruments such as Illumina Analyzer/HiSeq or Ion Torrent. An example is [E-MTAB-4151](#).

Annotare expects the following experimental workflow (but is flexible with the number of files per sample, e.g. for paired-end layout or technical replicates):



This template is designed for plant species to reflect plant-specific metadata

www.ebi.ac.uk/fg/annotare/login/



Plant data submissions via Annotare

New Experiment Submission

Select Experiment Design(s) i

- binding site identification design
- case control design
- cell type comparison design
- clinical history design
- compound treatment design
- development or differentiation design
- disease state design
- dose response design
- genetic modification design
- genotype design
- growth condition design
- injury design
- organism part comparison design
- sex design
- species design
- stimulus or stress design
- strain or line design
- time series design

Provide Experiment Details

How many samples does your experiment have? i

6

Experiment design	Experiment variable
genotype design	genotype
compound treatment design	compound dose
pathogenicity design	infect
stimulus or stress design	environmental stress
time series design	time

Prev



Describe your experiment

UNACCESSIONED

General Information
Contacts
Publications
Create samples, add attributes and experimental variables
Assign ENA library information
Describe protocols
Assign data files

Plant - High-throughput sequencing

Title *

Description *
(at least 50 characters)

Experiment Type *

Experimental Designs

Date of Experiment

Date of Public Release *

Hide my identity from reviewers

Related Accession Number

Informative title highlighting its intention

Background biology and aim of the experiment

Keep your experiment private until you want

Cross referencing datasets

Save as Spreadsheets Validate Submit to ArrayExpress

Name	Date	Status	File Size (Bytes)
------	------	--------	-------------------

Upload Files

Drag-and-drop files here to start upload
or press the "Upload Files" button.

⚠ File names can only contain letters [A-Z,a-z], numbers [0-9], underscores [_], and dots [.]

File Upload



Upload files and assign samples

UNACCESSIONED

General Information

Contacts

Publications

Create samples, add attributes and experimental variables

Assign ENA library information

Describe protocols

Assign data files

Plant - High-throughput sequencing

Title * Transcription profiling by high throughput sequencing of maize primary root differentiation zone of rum1 mutant plants

Description * The maize (*Zea mays L.*) Aux/IAA protein RUM1 (ROOTLESS WITH UNDETECTABLE MERISTEMS 1) controls seminal and lateral root initiation. To identify RUM1-dependent gene expression patterns, RNA-Seq of the differentiation zone of primary roots of rum1 mutants and the wild type was performed in four biological replicates.
(at least 50 characters)

Experiment Type * RNA-seq of coding RNA

Experimental Designs genotype design

Date of Experiment YYYY-MM-DD

Date of Public Release * 2018-03-19

Release * Hide my identity from reviewers

Related Accession Number Enter the accession number of related experiments in ArrayExpress or PRIDE e.g. E-MTAB-4688, PXD123456

Upload multiple files in one go

Upload Files... FTP Upload... Delete Files

Name	Date	Status	File Size (Bytes)
FTP upload Drag-and-drop files here to start upload or press the "Upload Files" button.			
<p>⚠ File names can only contain letters [A-Z,a-z], numbers [0-9], underscores [_], and dots [.]</p> <p>Start uploading files from the beginning to save time</p>			



Annotate your samples

Mandatory plant
metadata

Sample Attributes and Experimental Variables

Organism	<	Cell Line
Cultivar	>	Cell Type
Age	↑	Clinical History
Developmental Stage	↓	Compound (Experimental Variable)
Organism Part	=	Description
Material Type		Diet
Genotype (Experimental Variable)		Disease

Sample Attribute Experimental Variable

Name: Genotype

EFO Term: genotype

Measurement Unit:

- Cell Line
- Cell Type
- Clinical History
- Compound (Experimental Variable)
- Description
- Diet
- Disease
- Disease Staging
- Dose (Experimental Variable)
- Ecotype
- Environmental History
- Environmental stress
- Fraction
- Genetic Modification
- Growth Condition
- Immunoprecipitate (Experimental Variable)
- Individual
- Infect
- Injury
- Irradiate (Experimental Variable)
- Karyotype
- Phenotype
- Provider
- RNA interference (Experimental Variable)
- Replicate
- Response to Treatment
- Sampling site
- Sex
- Single Cell Well Quantity
- [new attribute](#)

Select more attributes
to annotate your sample
in more detail

Experimental variable is automatically
identified based on the pre-selected
experiment design

Add your own attributes
if it is not listed

[Cancel](#) [OK](#)



Annotate your samples

UNACCESSIONED

Help

Contact Us

Save as Spreadsheets

Validate

Submit to ArrayExpress

General Information
Contacts
Publications
Create samples, add attributes and experimental variables
Assign ENA library information
Describe protocols
Assign data files

Add Sample Attributes and Variables *								
		Add Samples		Delete Samples		Fill Down Value		
						Paste Into Column		
<input type="checkbox"/>	Name	Organism	Cultivar	Age (no units)	Developmental Stage	Organism Part	Material Type	Genotype
	Sample 1	Zea mays	B73 inbred line	4 day	seedling development stage	root	organism part	wild type
	Sample 2	Zea mays	B73 inbred line	4 day	seedling development stage	root	organism part	wild type
	Sample 3	Zea mays	B73 inbred line	4 day	seedling development stage	root	organism part	wild type
	Sample 4	Zea mays	B73 inbred line	4 day	seedling development stage	root	organism part	rum1 mutant
	Sample 5	Zea mays	B73 inbred line	4 day	seedling development stage	root	organism part	rum1 mutant
	Sample 6	Zea mays	B73 inbred line	4 day	seedling development stage	root	organism part	rum1 mutant

Upload Files... FTP Upload... Delete File File Upload

Upload Files

Drag-and-drop files here to start upload
or press the "Upload Files" button.

File names can only contain letters [A-Z,a-z], numbers [0-9], underscores [_], and dots [.]

Experimental variable is automatically identified based on the pre-selected experiment design

Plant - High-throughput sequencing



Describe your protocols

UNACCESSIONED

General Information
Contacts
Publications
Create samples, add attributes and experimental variables
Assign ENA library information
Describe protocols
Assign data files

Plant - High-throughput sequencing

Add Protocol *

New Protocol

Minimum Protocols Required

sample collection protocol

Please enter the description for the sample collection protocol

nucleic acid extraction protocol

Please enter the description for the nucleic acid extraction protocol

nucleic acid library construction protocol

Please enter the description for the nucleic acid library construction protocol

nucleic acid sequencing protocol

Please enter the description for the nucleic acid sequencing protocol

growth protocol

Please enter the description for the growth protocol

Additional Protocols

treatment protocol

Please enter the description for the treatment protocol

normalization data transformation protocol

Please enter the description for the normalization data transformation protocol

conversion protocol

Please enter the description for the conversion protocol

ets

✓ Validate

Submit to ArrayExpress

FTP Upload...

Delete Files

Date

Status

File Size (Bytes)

Upload Files

Drop-and-drag files here to start upload
or press the "Upload Files" button.

names can only contain letters [A-Z,a-z],
numbers [0-9], underscores [_], and dots [.]



Assign files to samples

Experiment Details Experiment Description Preview Sample and Data Preview

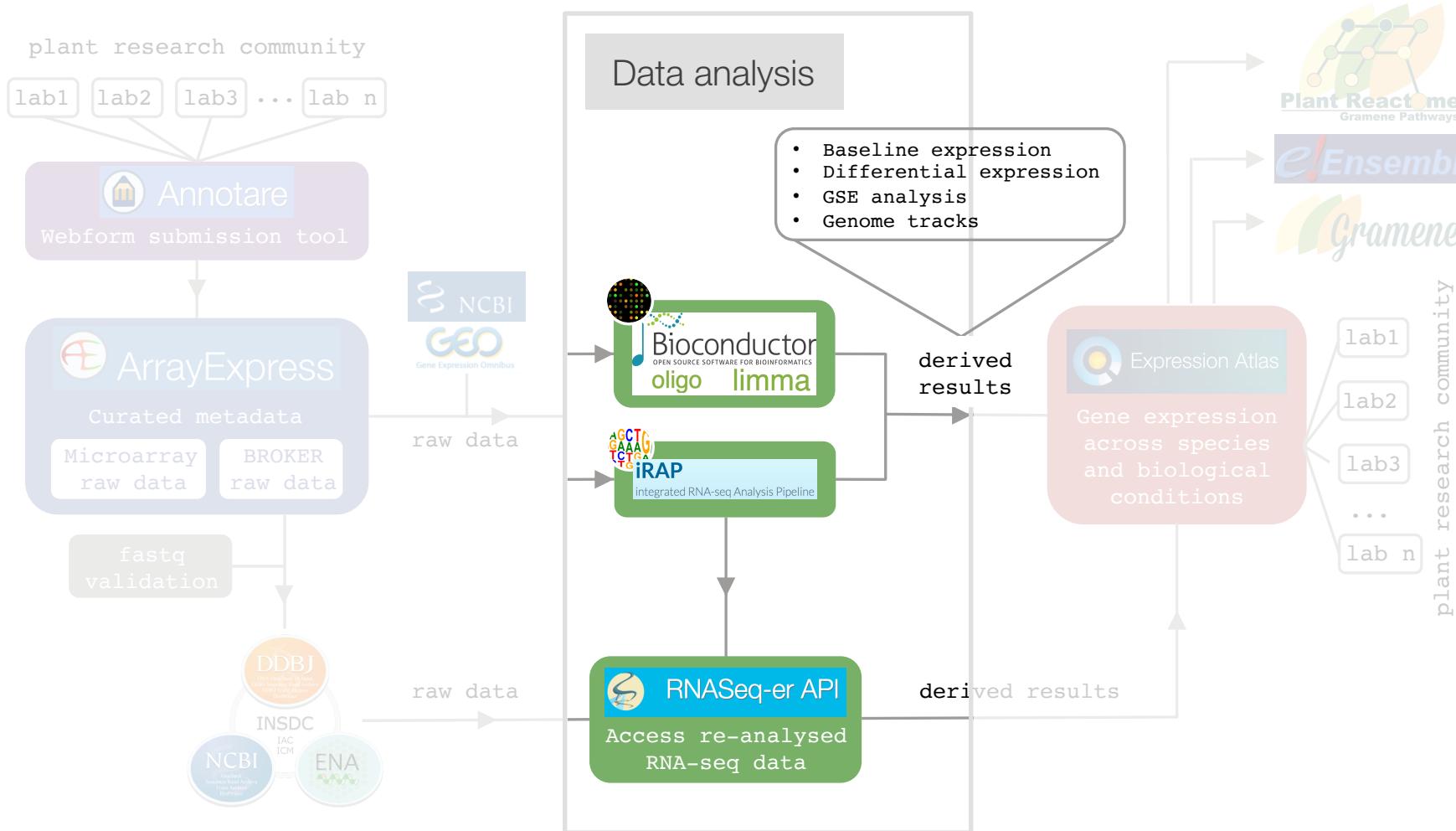
General Information	Add File Assignment Column *	Delete Column(s)	Fill Down Value	Paste Into Column
Contacts				
Publications				
Create samples, add attributes and experimental variables				
Assign ENA library information				
Describe protocols				
Assign data files				
High-throughput sequencing				

Name **Raw Data File** **Raw Data File (2)**

Leaf-control-untreated-replicate-1	Leaf_Control_set2_rep1_F.fastq.gz	Leaf_Control_set2_rep1_R.fastq.gz
Leaf-control-untreated-replicate-2	Leaf_Control_set2_rep2_F.fastq.gz	Leaf_Control_set2_rep2_R.fastq.gz
Leaf-control-untreated-replicate-3	Leaf_Control_set2_rep3_F.fastq.gz	Leaf_Control_set2_rep3_R.fastq.gz
Leaf-short-term-drought-replicate-1	Leaf_short-term_drought_rep1_F.fastq.gz	Leaf_short-term_drought_rep1_R.fastq.gz
Leaf-short-term-drought-replicate-2	Leaf_short-term_drought_rep2_F.fastq.gz	Leaf_short-term_drought_rep2_R.fastq.gz
Leaf-short-term-drought-replicate-3	Leaf_short-term_drought_rep3_F.fastq.gz	Leaf_short-term_drought_rep3_R.fastq.gz
Leaf-prolonged-drought-replicate-1	Leaf_prolonged_drought_rep1_F.fastq.gz	Leaf_prolonged_drought_rep1_R.fastq.gz
Leaf-prolonged-drought-replicate-2	Leaf_prolonged_drought_rep2_F.fastq.gz	Leaf_prolonged_drought_rep2_R.fastq.gz
Leaf-prolonged-drought-replicate-3	Leaf_prolonged_drought_rep3_F.fastq.gz	Leaf_prolonged_drought_rep3_R.fastq.gz
Leaf-short-term-heat-replicate-1	Leaf_short-term_heat_rep1_F.fastq.gz	Leaf_short-term_heat_rep1_R.fastq.gz
Leaf-short-term-heat-replicate-2	Leaf_short-term_heat_rep2_F.fastq.gz	Leaf_short-term_heat_rep2_R.fastq.gz
Leaf-short-term-heat-replicate-3	Leaf_short-term_heat_rep3_F.fastq.gz	Leaf_short-term_heat_rep3_R.fastq.gz
Leaf-prolonged-heat-replicate-1	Leaf_prolonged_h	
Leaf-prolonged-heat-replicate-2	Leaf_prolonged_h	
Leaf-prolonged-heat-replicate-3	Leaf_prolonged_h	

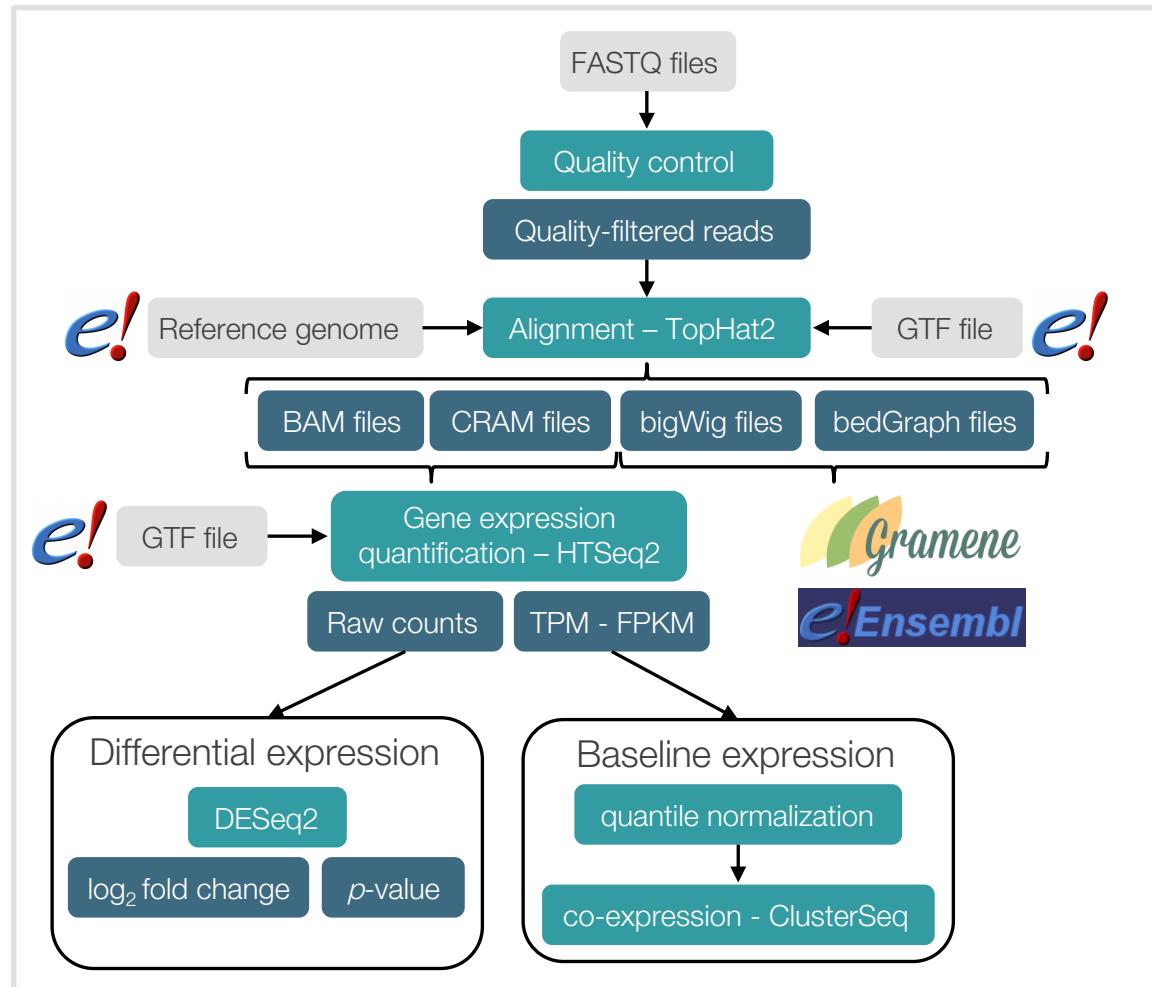
For paired-end RNA-sequencing libraries two data files are assigned to each sample

Plant expression data at EMBL-EBI





Data analysis – RNA-seq analysis pipeline



RNA-seq data analysis

<https://www.biorxiv.org/content/early/2014/06/06/005991>

iRAP - an integrated RNA-seq Analysis Pipeline

Nuno A. Fonseca, Robert Petryszak, John Marioni, Alvis Brazma

doi: <https://doi.org/10.1101/005991>

This article is a preprint and has not been peer-reviewed [what does this mean?].

Abstract

Info/History Metrics

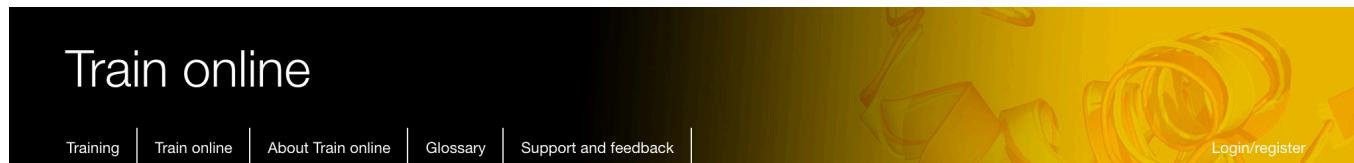
Preview PDF

Abstract

RNA-sequencing (RNA-Seq) has become the technology of choice for whole-transcriptome profiling. However, processing the millions of sequence reads generated requires considerable bioinformatics skills and computational resources. At each step of the processing pipeline many tools are available, each with specific advantages and disadvantages. While using a specific combination of tools might be desirable, integrating the different tools can be time consuming, often due to specificities in the formats of input/output files required by the different programs. Here we present iRAP, an integrated RNA-seq analysis pipeline that allows the user to select and apply their preferred combination of existing tools for mapping reads, quantifying expression, testing for differential expression. iRAP also includes multiple tools for gene set enrichment analysis and generates web browsable reports of the results obtained in the different stages of the pipeline. Depending upon the application, iRAP can be used to quantify expression at the gene, exon or transcript level. iRAP is aimed at a broad group of users with basic bioinformatics training and requires little experience with the command line. Despite this, it also provides more advanced users with the ability to customise the options used by their chosen tools.

<https://github.com/nunofonseca/irap>

EBI training: RNA-sequencing



Train online

Training | Train online | About Train online | Glossary | Support and feedback | Login/register

Functional genomics (II):
Common technologies and
data analysis methods

- Introduction
- Real-time PCR
- Microarrays
- Next Generation Sequencing (NGS)
- RNA sequencing
- Biological interpretation of gene expression data
- Genotyping, epigenetic and DNA/RNA-protein interaction methods
- Summary
- Quiz: Check your learning
- Your feedback
- Learn more
- References
- Contributors

Functional genomics (II): Common technologies and data analysis methods

This is the second part of our functional genomics course. This course will introduce you to common technologies in functional genomics studies, including microarrays and next generation sequencing (NGS), with a special focus on RNA-sequencing (RNA-seq). We will look at important considerations when designing your experiments, data analysis methods, and discuss when to use one technology over another.

We recommend that you look at [Functional genomics \(I\): Introduction and designing experiments](#) before starting this course.

About this course

Author(s): Laura Huerta, Melissa Burke



Beginner

2 hours

[Start the course](#)

Learning objectives:

- Evaluate different functional genomics technologies and know when to use them
- List some of the considerations when designing your own functional genomics experiments using these technologies
- Describe the steps involved in data analysis of microarray and RNA-seq experiments

<https://www.ebi.ac.uk/training/online>

EBI training: RNA-sequencing

Functional genomics (II):
Common technologies and
data analysis methods

- Introduction
- Real-time PCR
- Microarrays
- Next Generation Sequencing (NGS)
- RNA sequencing
 - Advantages of RNA-seq over hybridisation-based approaches
 - Applications of RNA-seq
 - Performing a RNA-seq experiment**
 - Design considerations
 - Number of replicates
 - Laboratory performance
 - Data analysis
 - Quality control
 - Read mapping or alignment
 - Quantification
 - Differential gene expression analysis
 - Biological interpretation of gene expression data
 - Genotyping, epigenetic and DNA/RNA-protein interaction methods
- Summary
- Quiz: Check your learning
- Your feedback
- Learn more
- References
- Contributors



All course materials in Train online
are free cultural works licensed
under a Creative Commons
Attribution-ShareAlike 4.0
International license

Performing a RNA-seq experiment

There is no common pipeline for all the different types of RNA-seq. Both the experimental design and the analysis procedures will be different depending on the organism being studied and the research goals. As the primary objective of many biological studies is a comparison of gene expression between samples, we outline a possible workflow for detecting differential expression using RNA-seq (Figure 7).

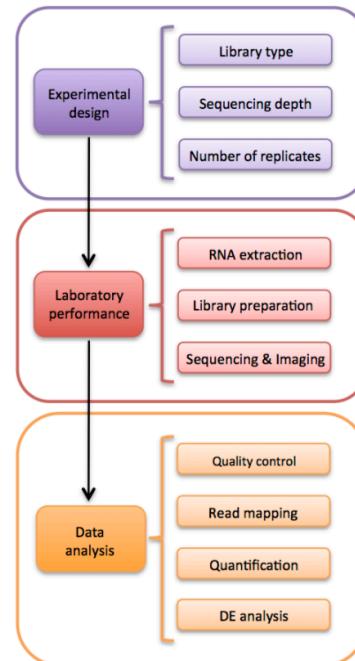


Figure 7 Overview of a RNA-seq experiment for detecting differential expression (DE).

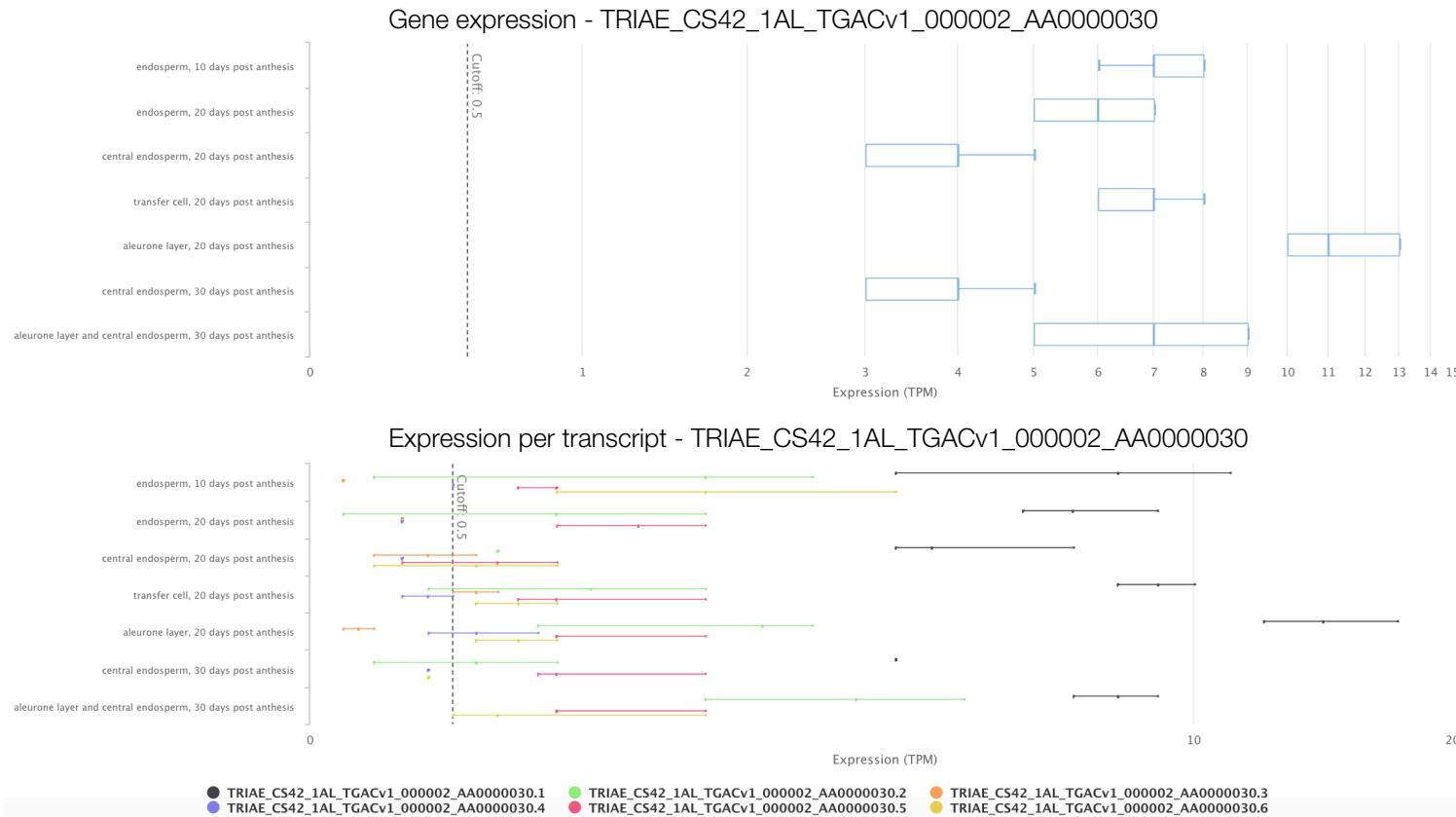
« Applications of RNA-seq

Next: Design considerations »

<https://www.ebi.ac.uk/training/online>



Data analysis – plant transcript quantification



www.ebi.ac.uk/gxa/experiments/E-MTAB-2137/Results/



On-the-fly gene set overlap analysis

Check if your set of genes is enriched in other comparisons

Search Gene set enrichment

Provide a set of Ensembl gene identifiers to test enrichment against differentially expressed genes by comparison

AT1G48030 AT1G53240 AT2G17130 AT2G20420 AT2G44350 AT2G47510 AT3G09810 AT3G15020 AT3G17240 AT3G27380 AT3G55410 AT3G60100 AT4G26910 AT4G35260 AT4G35650 AT4G35830 AT5G03290 AT5G08300 AT5G23250

Example: AT1G48030 AT1G53240 AT2G17130 AT2G20420 AT2G44350 AT2G47510 AT3G09810 AT3G15020 AT3G17240 AT3G27380 AT3G55410 AT3G60100 AT4G26910 AT4G35260 AT4G35650 AT4G35830 AT5G03290 AT5G08300 AT5G23250

Search Gene Set Enrichment API documentation

Gene set enrichment results

Species: *Arabidopsis thaliana*

Genes:

AT1G48030 AT1G53240 AT2G17130 AT2G20420 AT2G44350 AT2G47510 AT3G09810 AT3G15020 AT3G17240 AT3G27380 AT3G55410 AT3G60100 AT4G26910 AT4G35260 AT4G35650 AT4G35830 AT5G03290 AT5G08300 AT5G23250 (show fewer)

[Download results](#)

Show 10 entries

Search:

Experiment	Comparison	P-Value	Observed	Expected	Adjusted P-Value	Observed/Expected
Transcription profiling by array of <i>Arabidopsis</i> root cells after treatment with KNO ₃	'2hr KNO ₃ then incubation in protoplast-generating solution minus enzymes with KNO ₃ ' vs '2hr KCl then incubation in protoplast-generating solution minus enzymes' in 'whole root'	1.639x10 ⁻⁷	8	0.7	8.531x10 ⁻⁵	11.43
Transcription profiling by high throughput sequencing of <i>Arabidopsis</i> plants in response to a combination of salt and heat stress	'salt stress' vs 'none'	1.471x10 ⁻⁵	7	0.89	1.436x10 ⁻³	7.87
RNA-seq of coding RNA in frg1-2 mutant and wild-type strains of <i>Arabidopsis</i>	'frg1/2 mutant' vs 'wild type genotype'	4.612x10 ⁻⁶	8	1.08	6.003x10 ⁻⁴	7.41
CPIB Lateral Root timecourse	'30 hours' vs '27 hours'	1.209x10 ⁻⁴	6	0.85	4.329x10 ⁻³	7.06
Transcription profiling by array of <i>Arabidopsis</i> after infection with <i>Botrytis cinerea</i>	' <i>Botrytis cinerea</i> ' vs 'none' at '48 hour'	2.144x10 ⁻⁸	12	1.92	3.350x10 ⁻⁵	6.25
Transcription profiling by array of <i>Arabidopsis</i> mutant for npr1 after treatment with 3,5-dichloroanthranilic acid or 2,6-dichloroisonicotinic acid	'3,5-dichloroanthranilic acid' vs 'none' in 'wild type' at '2 day'	5.977x10 ⁻⁵	7	1.12	2.865x10 ⁻³	6.25

www.ebi.ac.uk/gxa/

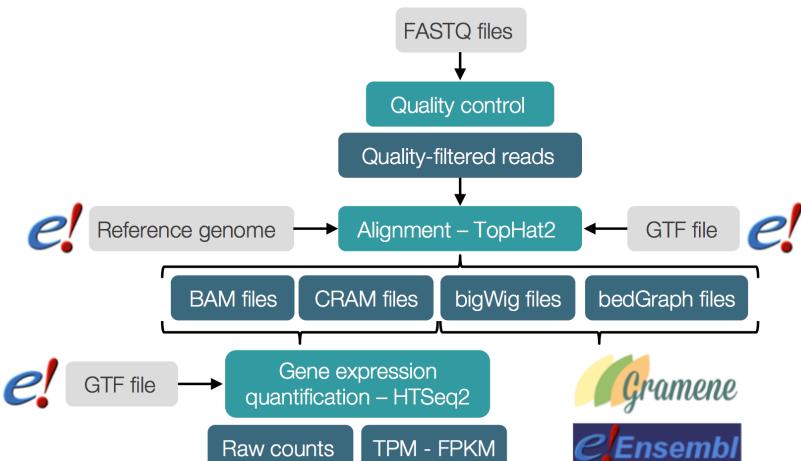


Data analysis - RNASeq-er API

The screenshot shows the RNASeq-er API homepage. At the top left is the logo, followed by the text "RNASeq-er API". Below the logo is a navigation bar with "Home" and "Documentation" links, and a "Feedback" link at the far right. The main content area features a large blue header with the text "Welcome to RNASeq-er API - a gateway to systematically updated analysis of public RNA-Seq data".

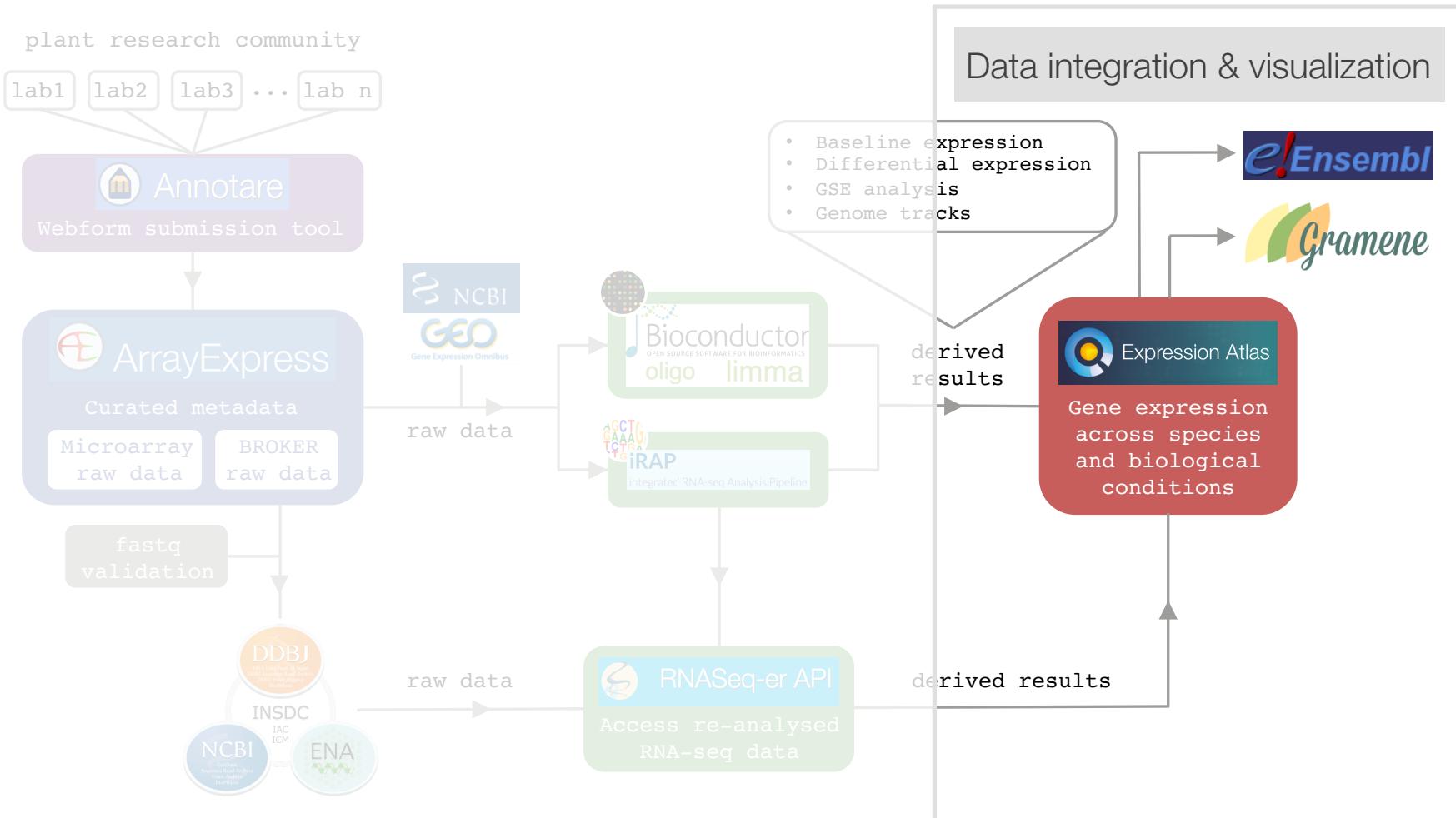
The RNASeq-er REST API provides easy access to the results of the systematically updated and continually growing analysis of public RNA-seq data in [European Nucleotide Archive \(ENA\)](#). The analysis of each sequencing run is performed by the EMBL-EBI's Gene Expression Team using the [iRAP pipeline](#).

- [1. Get Started](#)
- [2. What does the RNASeq-er pipeline do?](#)
- [3. How is the RNASeq-er performed?](#)
- [4. How to use the RNASeq-er API?](#)
- [5. What are the main classes of API calls?](#)
 - [5.1. Analysis results Per Run
 - \[5.1.1. Making per-run API calls\]\(#\)
 - \[5.1.2. Results of per-run API calls\]\(#\)](#)
 - [5.2. Analysis results Per Study
 - \[5.2.1. Making per-study API calls\]\(#\)
 - \[5.2.2. Results of per-study API calls\]\(#\)](#)
 - [5.3. Sample Attributes Per Run
 - \[5.3.1. Making sample attributes per-run API calls\]\(#\)
 - \[5.3.2. Results of sample attributes per-run API calls\]\(#\)](#)
 - [5.4. Baseline expression By Gene
 - \[5.4.1. Making baseline expression per-gene API calls\]\(#\)
 - \[5.4.2. Results of baseline expression per-gene API calls\]\(#\)](#)
 - [5.5. Mapping Quality Statistics For All Organisms
 - \[5.5.1. Retrieving the mean and standard deviation of mapping quality for all organisms\]\(#\)](#)



www.ebi.ac.uk/fg/rnaseq/api/

Plant expression data at EMBL-EBI



Data visualization – Plant Expression Atlas

The screenshot shows the Expression Atlas homepage. At the top left is the logo and the text "Expression Atlas Gene expression across species and biological conditions". Below the logo is a navigation bar with links: Home, Download, Release notes, FAQ, Help, Licence, About, and Feedback. To the right is a teal sidebar with the text "Discover and interpret gene expression analysis results quickly and easily". The main content area features a large image of a heatmap representing gene expression data across different species and conditions. Overlaid on this image is a callout box containing the question "In which conditions is my favourite gene expressed?". Below the heatmap are search and filter fields: "Search" and "Gene set enrichment" tabs; a "Gene / Gene properties" field containing "floury endosperm1 x"; an "Organism" dropdown set to "Zea mays"; and a "Biological conditions" field with a placeholder "Enter condition query...".

Exploring gene expression results across species under different biological conditions

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others. Expression Atlas aims to help answering questions such as 'where is a certain gene expressed?' or 'how does its expression change in a disease?'.

Expression Atlas provides gene expression results on more than 3,000 experiments (microarray and RNA-seq) from 40 different organisms, including metazoans and plants. Expression profiles of tissues from Human Protein Atlas, GTEx and F richer queries and re-analyses can be explored in Expression Atlas. All data are manually curated, annotated to ontology terms allowing for much

In which conditions is my favourite gene expressed?

Search Gene set enrichment

Gene / Gene properties

floury endosperm1 x

Examples: REG1B, zinc finger, O14777 (UniProt), GO:0010468 (regulation of gene expression)

Organism

Zea mays

Biological conditions

Enter condition query...

Examples: lung, leaf, valproic acid, cancer

Search Clear

www.ebi.ac.uk/gxa/



Plant Baseline Expression results

zm00001d003398 (floury endosperm1) Zea mays floury endosperm1

Baseline expression Differential expression floury endosperm1 information

Show anatomograms

Filter your results

Zea mays

Organism part

Cell type

Cultivar

Developmental stage

Sampling time point

Organism part

Showing 3 experiments:

6 Wang 5 x 4 Chettoor et al

ear embryo embryo sac endosperm immature st... plant ovary plant ovule pollen root shoot style tassel whole organ...

tissues

High Medium Low Below cutoff No data available

experiments

ear embryo embryo sac endosperm immature st... plant ovary plant ovule pollen root shoot style tassel whole organ...

6 Wang 5 x 4 Chettoor et al

High Medium Low Below cutoff No data available

Gene expression across maize tissues from different experiments is integrated and visualised in one heatmap



Plant Baseline Expression results

zm00001d003398 (floury endosperm1) Zea mays floury endosperm1

Sort results by gene expression rank to display higher values on the top left corner





Plant Baseline Expression results

zm00001d003398 (floury endosperm1) Zea mays floury endosperm1

Baseline expression

Differential expression

floury endosperm1 information

Show anatomograms

Filter your results

Zea mays

Organism part

Cell type

Cultivar

Developmental stage

Sampling time point

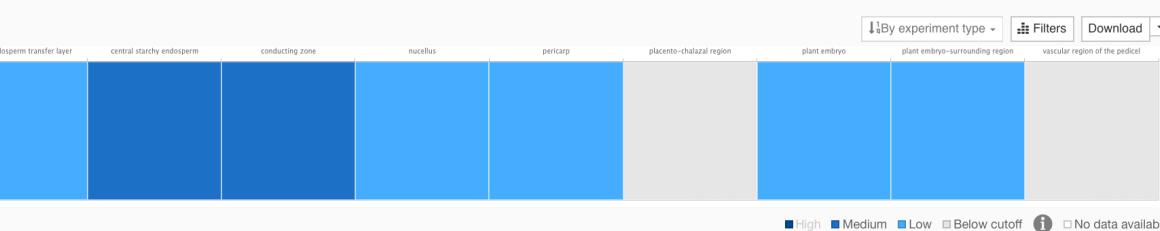
Organism part

Showing 3 experiments:



Cell type

Showing 1 experiment:



Cultivar

Showing 2 experiments:

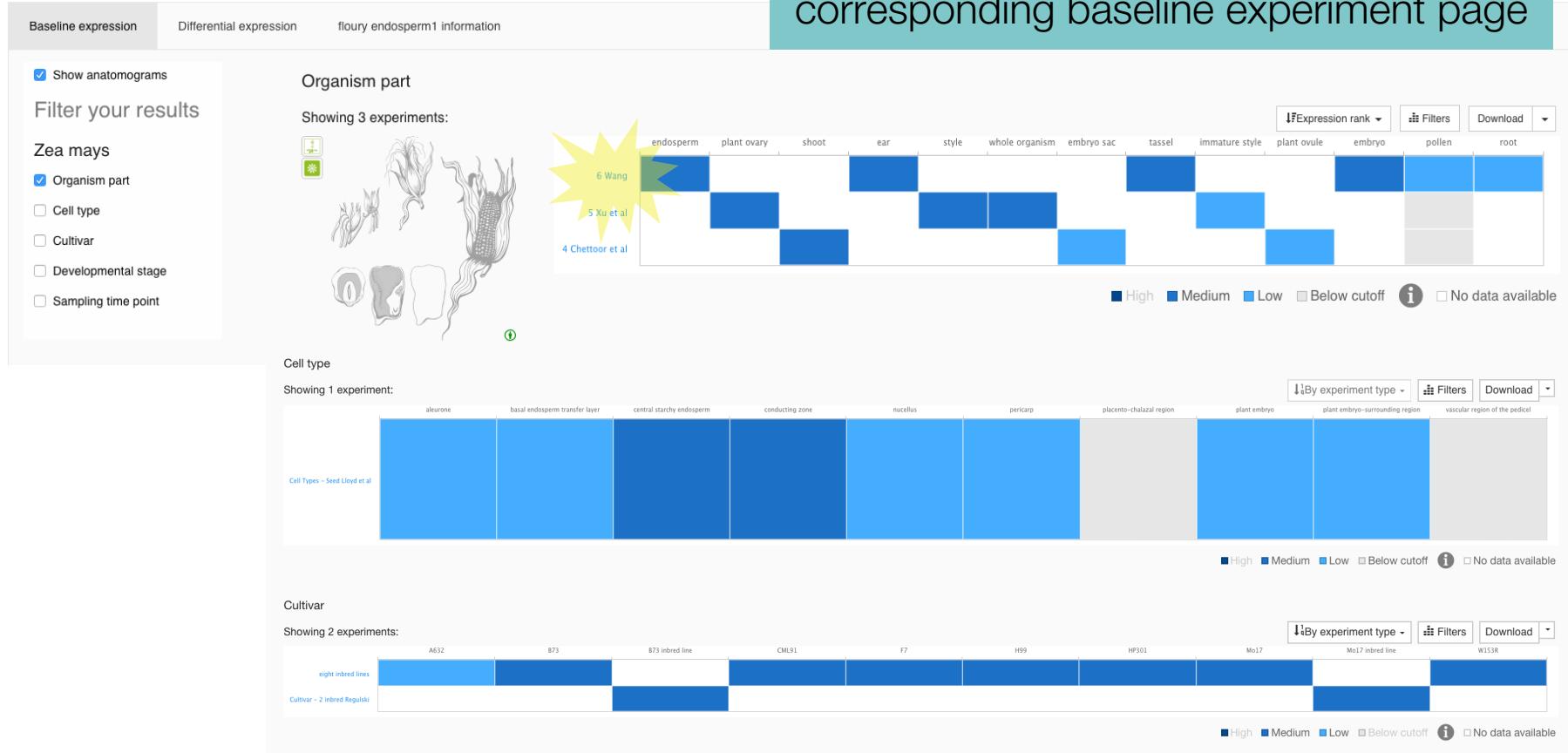


Visualise gene expression results for other experimental conditions using the filters



Plant Baseline Expression results

zm00001d003398 (floury endosperm1) Zea mays floury endosperm1



Click on a dataset name to see the corresponding baseline experiment page



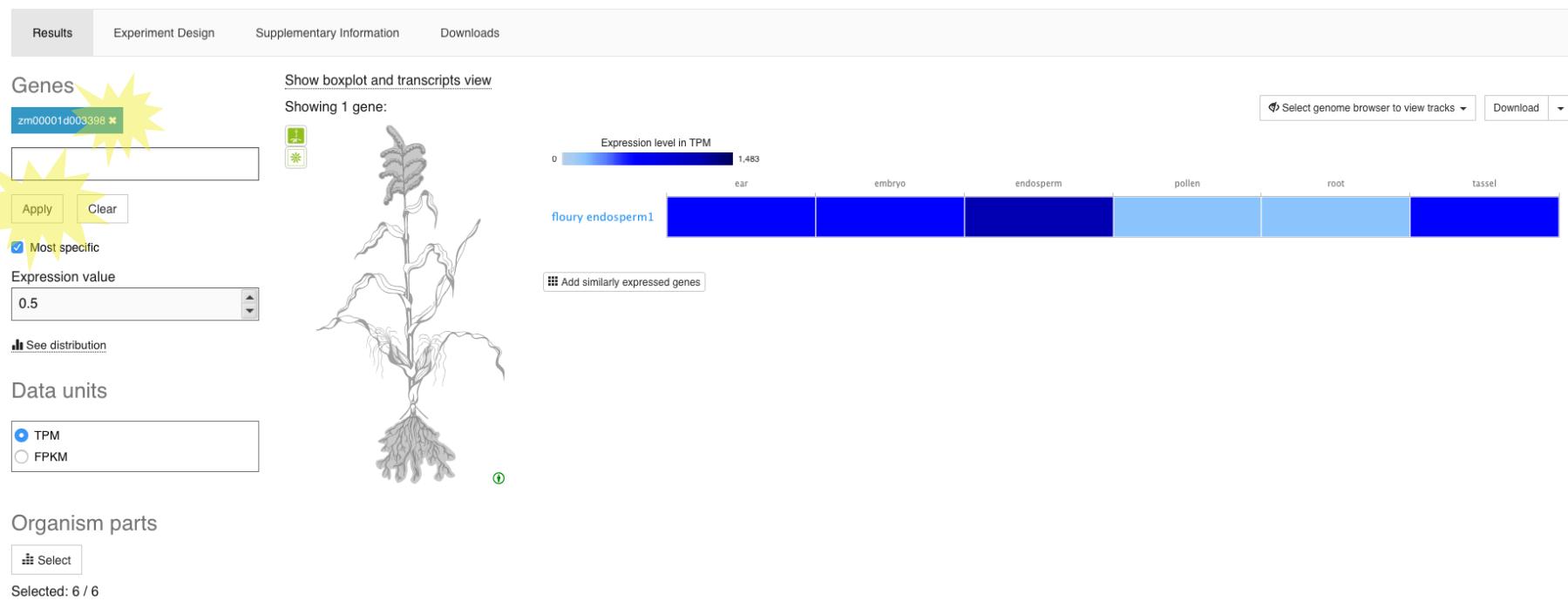
Plant Baseline Experiment page

RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, endosperm, pollen, root and tassel) from B73 strain

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)





Plant Baseline Experiment page

General information

RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, endosperm, pollen, root and tassel) from B73 strain

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): [27339440](#) (Filter by genes in paper)





Plant Baseline Experiment page

RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, endosperm,

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)

Results Experiment Design Supplementary Information Downloads

Genes

Showing 50 of 29,053 genes found:

Select genome browser to view tracks ▾ Download ▾

Specific checked

Most specific

Expression value

0.5

See distribution

Data units

TPM

FPKM

Organism parts

Select

Selected: 6 / 6



Expression level in TPM
0 606

Zm0001400473
Zm0001400410
Zm0001400481
Zm0001400511
Zm0001402732
Zm0001403203
Zm00014026163
guaiacol peroxidase2
Zm00014011419
Zm00014054042
Zm00014019042
Zm00014048419
Zm00014023148
Zm000140525151
endochitinase precursor4
Zm00014029387
Zm00014049168
Zm00014002004
Zm00014035756
Zm00014014405
Zm00014012416
Zm00014023116

barley mlo defense gene homolog8
Zm00014035779
Zm00014020739
Zm00014041618
Ustilago maydis induced11
Zm00014005105
benzoazepinone synthetase14
Zm00014024960
Zm00014051259
Zm00014045122
TIP6
Zm000140546138

HEATMAP: top 50 genes (rows)
most specifically expressed
across all tissues (columns)





Plant Baseline Experiment page

RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, endosperm,

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): [27339440](#) (Filter by genes in paper)

Results Experiment Design Supplementary Information Downloads

HEATMAP: top 50 genes (rows)
most highly expressed across all
tissues studied (columns)

Genes

Showing 50 of 29,053 genes found:

Select genome browser to view tracks ▾ Download ▾

Specific unchecked

Most specific

Expression value

0.5

See distribution

Data units

TPM

FPKM

Organism parts

Select

Selected: 6 / 6



Expression level in TPM

79.800

Zm00001d049476

Zm00001d019155

Zm00001d048850

Zm00001d048852

floury2

27-kDa zein protein

Zm00001d019239

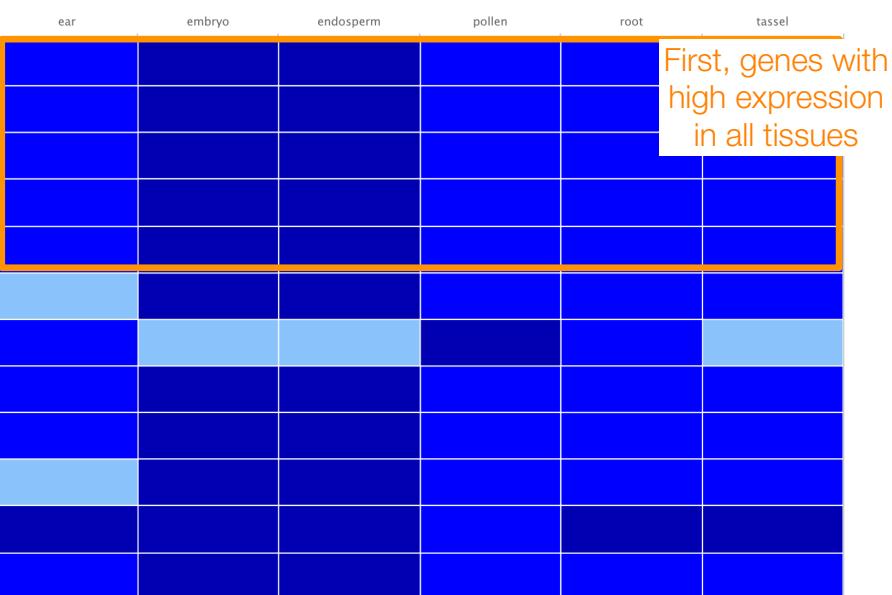
Zm00001d048848

50kD gamma zein

mucronate1

glycine-rich protein1

floury4





Plant Baseline Experiment page

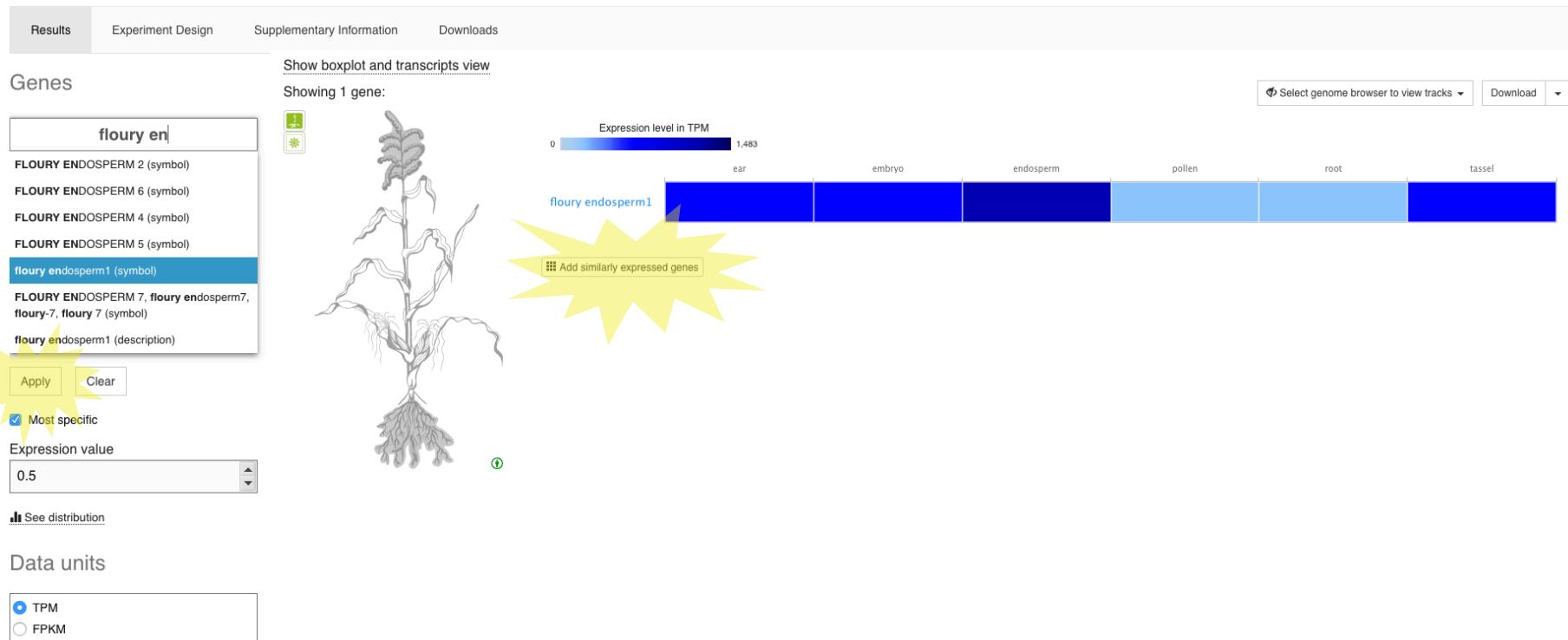
RNA-seq of coding RNA from 6 different maize tissues (ear, embryo...

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): [27339440](#) (Filter by genes in paper)

Search for a particular gene, e.g. *FL1* and find genes with similar expression





Plant Baseline Experiment page

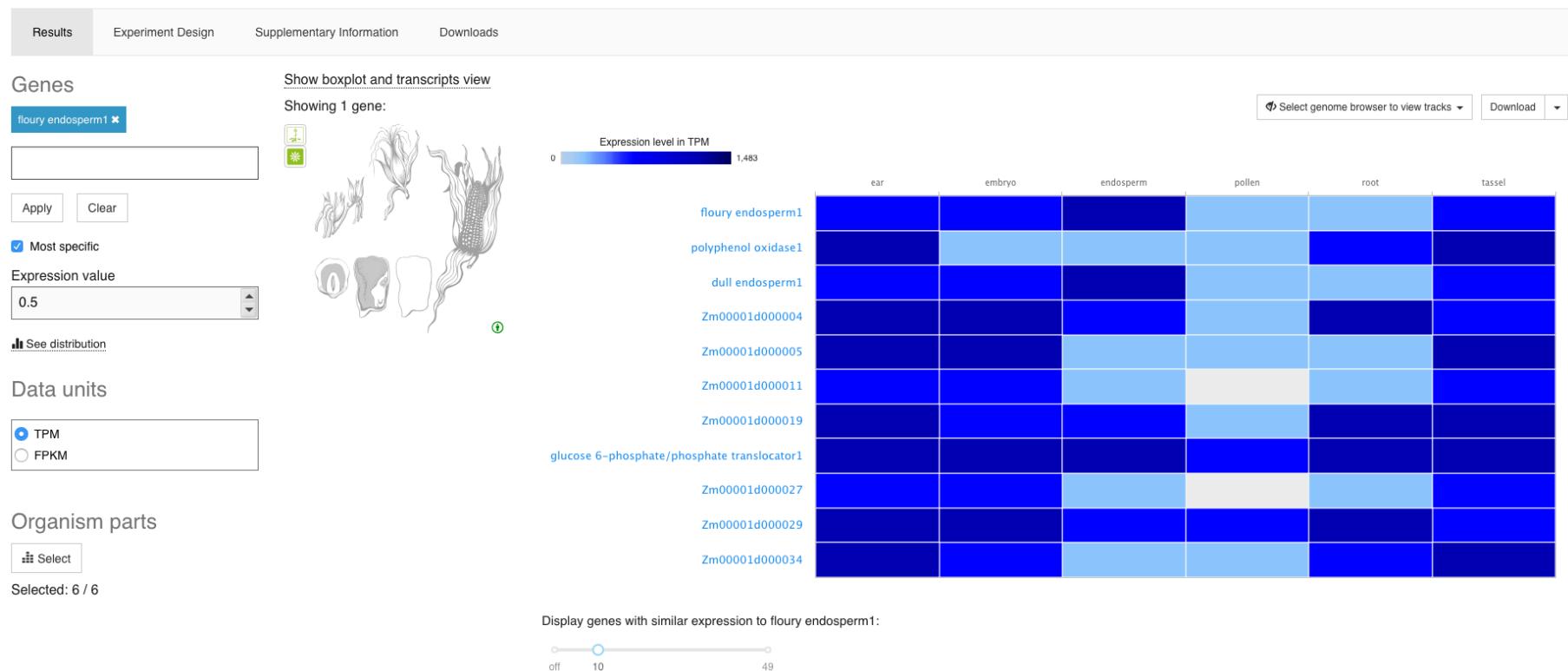
RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, endosperm, pollen, root, tassel)

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)

Visualise expression pattern of genes similarly expressed to *FL1*





Plant Baseline Experiment page

RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, ...)

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)

Results Experiment Design Supplementary Information

Genes

floury endosperm1 *

Show boxplot and transcript

Showing 1 gene:

floury endosperm1

Apply Clear

Most specific

Expression value

0.5

See distribution

Data units

TPM
 FPKM

Organism parts

Select

Selected: 6 / 6

Visualise expression pattern of genes similarly expressed to *FL1*

Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Home » Bioconductor 3.7 » Software Packages » clusterSeq

clusterSeq

platforms all downloads 1091 / 1561 posts 0 In Bioc 1.5 years
build warnings updated before release

DOI: [10.18129/B9.bioc.clusterSeq](https://doi.org/10.18129/B9.bioc.clusterSeq) [f](#) [t](#)

Clustering of high-throughput sequencing data by identifying co-expression patterns

Bioconductor version: Release (3.7)

Identification of clusters of co-expressed genes based on their expression across multiple (replicated) biological samples.

Author: Thomas J. Hardcastle & Irene Papathodorou

Maintainer: Thomas J. Hardcastle <tjh48 at cam.ac.uk>

Citation (from within R, enter `citation("clusterSeq")`):

Papatheodorou TJH&I (2016). *clusterSeq: Clustering of high-throughput sequencing data by identifying co-expression patterns*. R package version 1.4.0.



Plant Baseline Experiment page

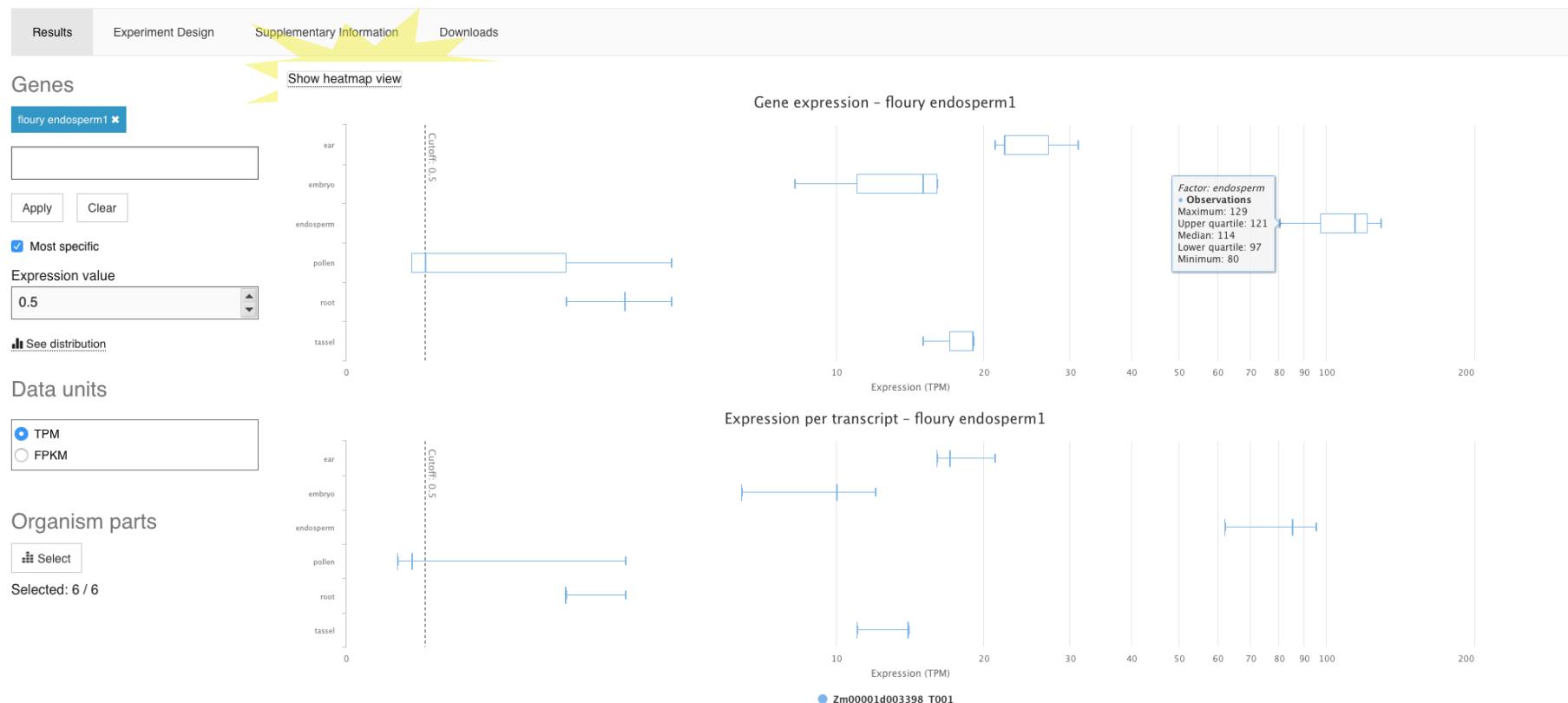
RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, endosperm, pollen)

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)

Visualise variation among biological replicates





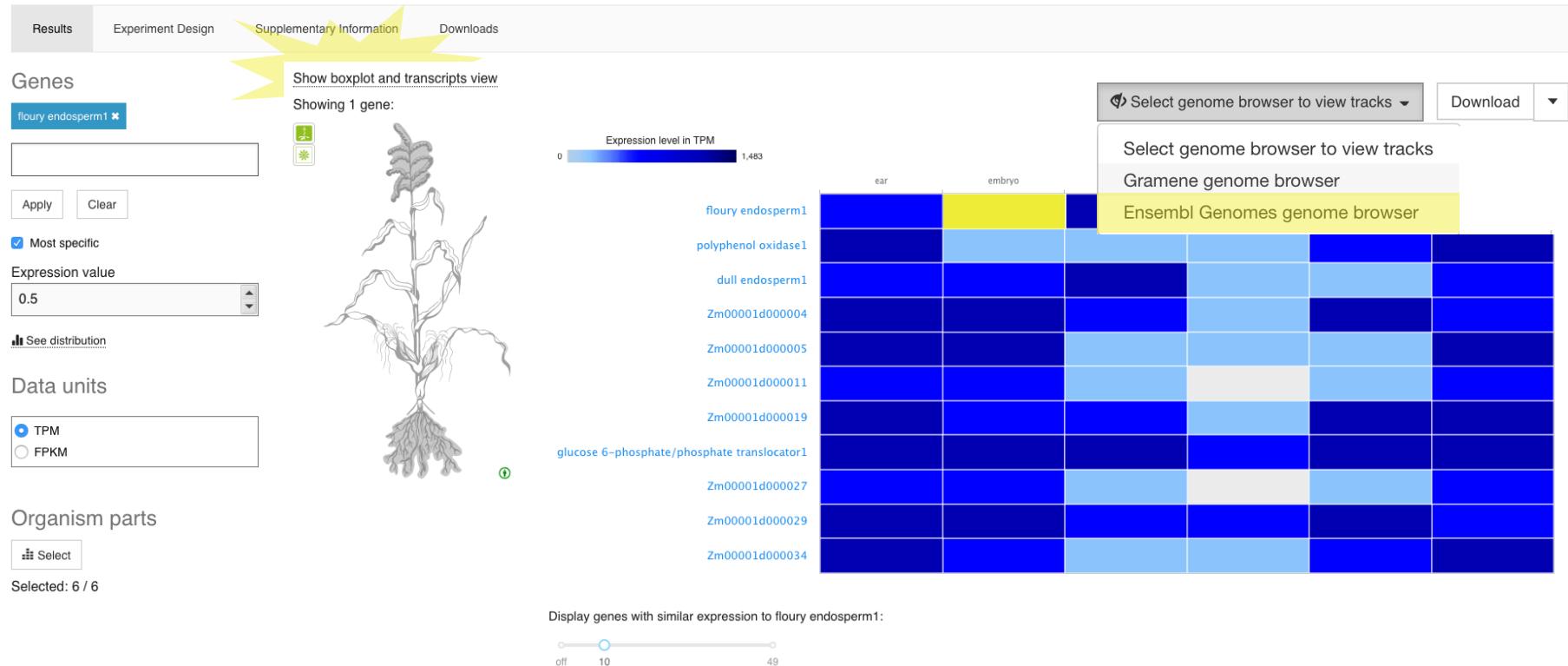
Plant Baseline Experiment page

RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, endosperm, ...)

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)





Plant Baseline Experiment page

e! Ensembl Plants ▾ HMMER | BLAST | BioMart | Tools | Downloads | Documentation | Website help

Location: 2:43,019,871-43,020,779

Gene: *floury endosperm* Trans: Zm00001d003398_T001

Search Ensembl Plants...

Location-based displays

- Whole genome
- Chromosome summary
- Region overview
- Region in detail

Comparative Genomics

- Synteny
- Alignments (image)
- Alignments (text)
- Region Comparison

Genetic Variation

- Variant table
- Resequencing
- Linkage Data
- Markers

Other genome browsers

- NCBI
- maizeGDB

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

Ensembl Plants is produced in collaboration with Gramene

Chromosome 2: 43,019,871-43,020,779

Region in detail

Contigs Genes ncRNA GrMZAGPv3 mapped Maize AGPv4 low confidence coding Gene Legend

Location: 2:43019871-43020779 Gene: Go

Drag>Select: Forward strand

8.00 E-MTAB-3826_g2.gene 0.00 HC_IncRNAs Sequence Contigs Sequence Genes GrMZAGPv3 mapped

No features from HC_IncRNAs in this location on this strand

B73V4_ctg1129 >

< Zm00001d003398_T001 protein coding

< GRMZM2G094532_T01 protein coding

909 bp

This screenshot shows the Ensembl Plants 'Region in detail' view for Chromosome 2. The top panel displays a genomic track for B73V4_ctg1129, spanning from 42.80 Mb to 43.25 Mb. It highlights several genes: Zm00001d003398 (floury endosperm), Zm00001d003400 (general regulatory factor), Zm00001d000803, Zm00001d000802, GRMZM2G168968, GRMZM2G29901, GRMZM2G023356, GRMZM2G325131, GRMZM2G394321, GRMZM2G102489, GRMZM2G094532, and GRMZM2G103116. The middle panel shows a zoomed-in view of the Zm00001d003398 gene, with its 909 bp length indicated. The bottom panel shows experimental data from E-MTAB-3826_g2.gene, with a brown bar representing the signal. A legend indicates that red bars represent protein coding and purple bars represent RNA genes.



Plant Baseline Experiment page

RNA-seq of coding RNA from 6 different maize tissues (ear, e

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)

Visualise sample metadata and download
metadata annotated to ontologies

Results	Experiment Design	Supplementary Information	Downloads			
Analysed	Run	cultivar	Sample Characteristics		Experimental Variables	
			organism	organism part		organism part
Yes	ERR986118	B73	Zea mays	tassel		tassel
Yes	ERR986117	B73	Zea mays	tassel		tassel
Yes	ERR986106	B73	Zea mays	tassel		tassel
Yes	ERR986105	B73	Zea mays	tassel		tassel
Yes	ERR986094	B73	Zea mays	tassel		tassel
Yes	ERR986093	B73	Zea mays	tassel		tassel
Yes	ERR986116	B73	Zea mays	root		root
Yes	ERR986115	B73	Zea mays	root		root
Yes	ERR986104	B73	Zea mays	root		root
Yes	ERR986103	B73	Zea mays	root		root
Yes	ERR986092	B73	Zea mays	root		root
Yes	ERR986091	B73	Zea mays	root		root
Yes	ERR986114	B73	Zea mays	pollen		pollen



Plant Baseline Experiment page

RNA-seq of coding RNA from 6 different maize tissues (ear, embryo, endosp

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)

Learn more about our RNA-seq analysis pipeline

Results Experiment Design Supplementary Information Downloads

Analysis Methods

Pipeline version	iRAP 0.8.4d1
Analyzed Libraries	See experiment design
Filtering Step 1	Discard reads below minimum quality threshold
Filtering Step 2	Check of bacterial contamination; discard offending reads
Filtering Step 3	Discard reads with common uncalled characters (e.g. N)
Filtering Step 4	Remove reads from pair-end libraries that were orphaned by filtering steps 1-3
Read Mapping	Against genome reference (Ensembl Genomes release: 34) tophat2 version: 2.1.1
Gene Quantification	htseq2 version: 0.6.1p1
Transcript Quantification	kallisto version: 0.42.4
Normalized Counts per Gene	(TPMs/FPKMs) are calculated from the raw counts by iRAP. These are averaged for each set of technical replicates, and then quantile normalized within each set of biological replicates using limma. Finally, they are averaged for all biological replicates (if any)

Resources

 ArrayExpress: experiment E-MTAB-3826



Plant Baseline Experiment page

RNA-seq of coding RNA from 6 different maize tissues (ear, embr

RNA-Seq mRNA baseline

Organism: *Zea mays*

Reference(s): 27339440 (Filter by genes in paper)

Results Experiment Design Supplementary Information **Downloads**

Expression values across all genes (TPM)

Expression values across all genes (FPKM)

Summary of the expression results for this experiment ready to view in R

Heatmap of aggregated expression data

Experiment Design (tsv)

Download gene quantification matrices
(TPM and FPKM) and R object

Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Home » [Bioconductor 3.7](#) » [Software Packages](#) » [ExpressionAtlas](#)

ExpressionAtlas

platforms [all](#) downloads [700 / 1561](#) posts [0](#) in Bioc [2.5 years](#)
build [ok](#) updated [before release](#)

DOI: [10.18129/B9.bioc.ExpressionAtlas](https://doi.org/10.18129/B9.bioc.ExpressionAtlas) [f](#) [t](#)

Download datasets from EMBL-EBI Expression Atlas

Bioconductor version: Release (3.7)

This package is for searching for datasets in EMBL-EBI Expression Atlas, and downloading them into R for further analysis. Each Expression Atlas dataset is represented as a SimpleList object with one element per platform. Sequencing data is contained in a SummarizedExperiment object, while microarray data is contained in an ExpressionSet or MAList object.

Author: Maria Keays

Maintainer: Maria Keays <maria.keays@gmail.com>

Citation (from within R, enter `citation("ExpressionAtlas")`):

Keays M (2018). *ExpressionAtlas: Download datasets from EMBL-EBI Expression Atlas*. R package version 1.8.1.



Let's try Expression Atlas

Baseline expression

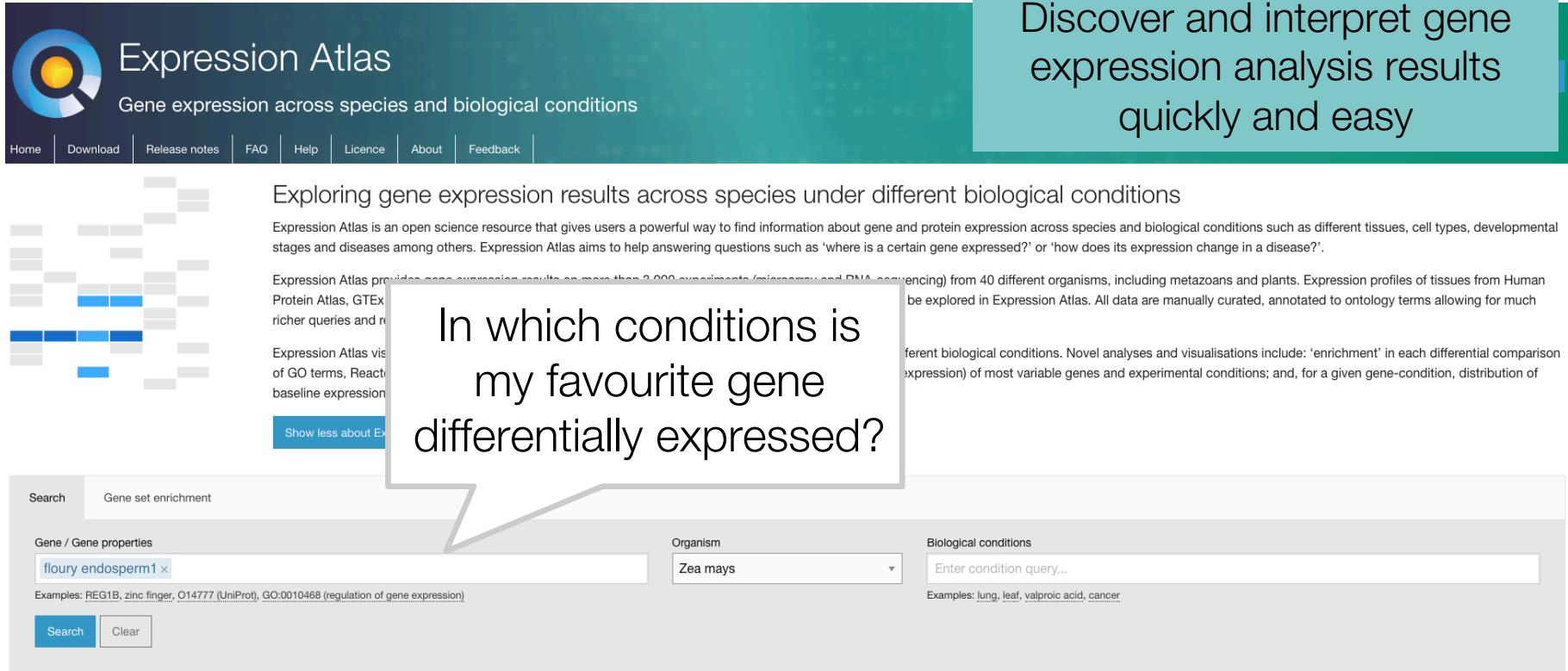
Hands-on activity

Look up expression of members of the
PIN gene family in maize



In pairs

Data visualization – Plant Expression Atlas



The screenshot shows the Expression Atlas homepage. At the top left is the logo and the text "Expression Atlas Gene expression across species and biological conditions". Below the logo is a navigation bar with links: Home, Download, Release notes, FAQ, Help, Licence, About, and Feedback. To the right is a teal box containing the text: "Discover and interpret gene expression analysis results quickly and easily". The main content area features a large image of a heatmap representing gene expression data. Overlaid on this image is a large, semi-transparent text box containing the question: "In which conditions is my favourite gene differentially expressed?". Below this text box are search and filter fields. The "Search" field contains "floury endosperm1 x". The "Organism" dropdown is set to "Zea mays". The "Biological conditions" field has "Enter condition query..." and examples like "lung, leaf, valproic acid, cancer".

Exploring gene expression results across species under different biological conditions

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others. Expression Atlas aims to help answering questions such as 'where is a certain gene expressed?' or 'how does its expression change in a disease?'.

Expression Atlas provides gene expression results on more than 2,000 experiments (microarray and RNA sequencing) from 40 different organisms, including metazoans and plants. Expression profiles of tissues from Human Protein Atlas, GTEx richer queries and results

Expression Atlas visualizations of GO terms, Reactome, baseline expression

Show less about Expression Atlas

In which conditions is my favourite gene differentially expressed?

Search Gene set enrichment

Gene / Gene properties

floury endosperm1 x

Examples: REG1B, zinc finger, O14777 (UniProt), GO:0010468 (regulation of gene expression)

Organism

Zea mays

Biological conditions

Enter condition query...

Examples: lung, leaf, valproic acid, cancer

Search Clear

www.ebi.ac.uk/gxa/



Plant Differential Expression results

Zm00001d003398 (FL1) Zea mays floury endosperm1

Baseline expression Differential expression FL1 information

Filter your results

Display log₂-fold change Download results

Kingdom

Plants

Species

Zea mays

Experiment type

Microarray 1-colour mRNA differential

RNA-seq mRNA differential

Experimental variables

Infect

Genotype

Compound

Organism part

Time

Growth condition

Regulation

Down

Up

Log₂-fold change

Species

Gene name

Comparison

Experimental variables

Experiment name

Adjusted p-value	Log ₂ -fold change	Property	Test value (N=3)	Reference value (N=3)
1.0313 × 10 ⁻⁹	-3.4	genotype	nkd mutant	wild type genotype
		organism part	aleurone layer	aleurone layer
		age	15 day	15 day
		cultivar	B73 inbred line	B73 inbred line
		growth condition	field study	field study
		initial time point	pollination	pollination
		organism	Zea mays	Zea mays

Sample attributes are annotated to ontologies

FL1

FL1

Curated comparisons in which *FL1* gene is differentially expressed

Transcription profiling by high throughput sequencing of aleurone and starchy endosperm cells were isolated by laser-capture microdissection in maize nkd mutants

Transcription profiling by high throughput sequencing of aleurone and starchy endosperm cells were isolated by laser-capture microdissection in maize nkd mutants

Transcription profiling by high throughput sequencing of maize after treatment with 12-oxo-phytodienoic acid or 10-oxo-11-phytodienoic acid

Transcription profiling by array of maize root cells after infestation with nematode *Meloidogyne incognita* for 14 or 21 days against uninfected controls

Transcription profiling by array of maize plants infected with *Ustilago maydis* mutant for cluster 19A and individual subdeletions for genes tin1, tin3, tin4 and tin5

Transcription profiling by high throughput sequencing of maize after treatment with 12-oxo-phytodienoic acid or 10-oxo-11-phytodienoic acid

Transcription profiling by array of maize plants infected with *Ustilago maydis* mutant for cluster 19A and individual subdeletions for genes tin1, tin3, tin4 and tin5

Expression profile of contrasting maize genotypes grown in acid and control soil (leaves)

Transcription profiling by array of maize root cells after infestation with nematode *Meloidogyne incognita* for 14 or 21 days against uninfected controls



Plant Differential Experiment page

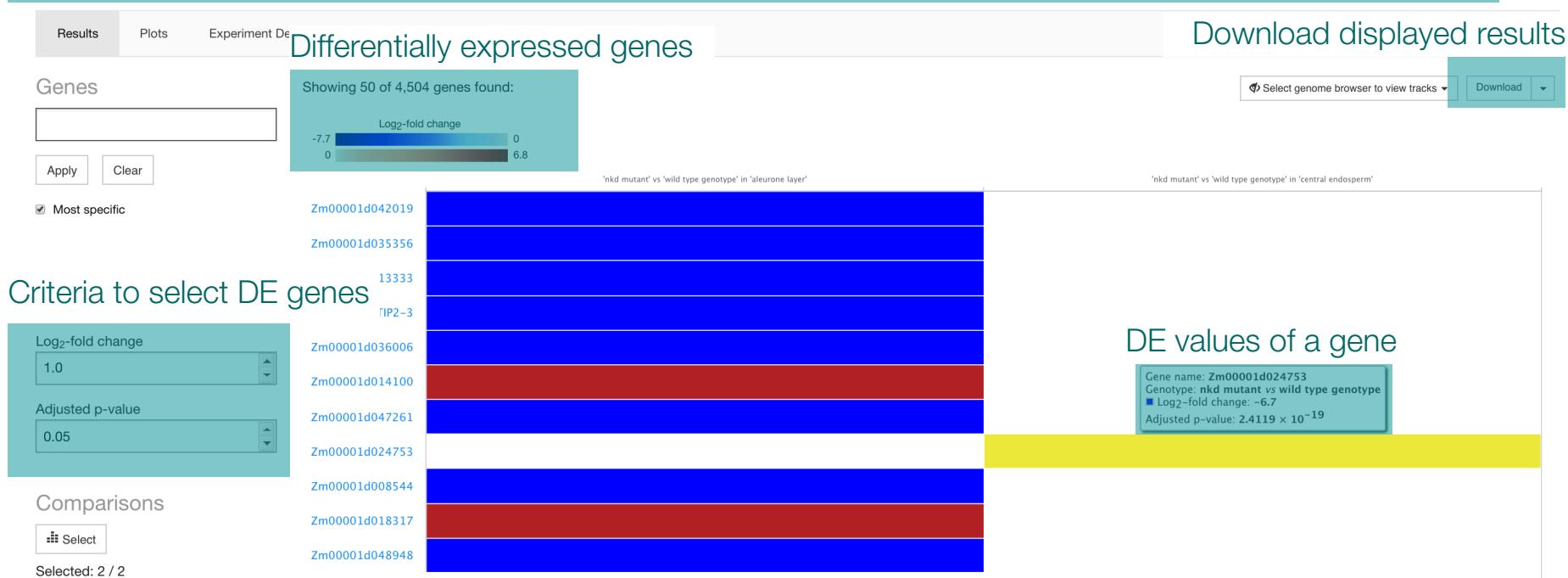
General information

Transcription profiling by high throughput sequencing of aleurone and starchy endosperm cells were isolated by laser-capture microdissection in maize nkd mutants

RNA-Seq mRNA differential

Organism: *Zea mays*

Reference(s): 25552497 (Filter by genes in paper)





Plant Differential Experiment page

Transcription profiling by high throughput sequencing of aleurone and starchy endosperm cells were isolated by laser-capture microdissection in maize nkd mutants

RNA-Seq mRNA differential

Organism: *Zea mays*

Reference(s): 25552497 (Filter by genes in paper)





Plant Differential Experiment page

Transcription profiling by high throughput sequencing of aleurone and starchy endosperm cells were isolated by laser-capture microdissection in maize nkd mutants

RNA-Seq mRNA differential

Organism: *Zea mays*

Reference(s): 25552497 (Filter by genes in paper)

Results Plots Experiment Design Supplementary Information Downloads

'nkd mutant' vs 'wild type genotype' in 'aleurone layer'

GSE GO GO terms enrichment analysis

GSE IPro Interpro domains enrichment analysis

MA MA plot for the contrast across all genes

'nkd mutant' vs 'wild type genotype' in 'central endosperm'

GSE GO GO terms enrichment analysis

GSE IPro Interpro domains enrichment analysis

MA MA plot for the contrast across all genes

The screenshot shows the Bioconductor piano package interface. At the top, there's a logo for Bioconductor with the text "OPEN SOURCE SOFTWARE FOR BIOINFORMATICS". Below it, the word "piano" is displayed in a large green font. A horizontal line separates this from the text "Platform for integrative analysis of omics data". Underneath, it says "Bioconductor version: Release (3.5)". The main text area contains information about the Piano package, mentioning gene set analysis using various statistical methods and combining results of multiple runs. It lists the author as Leif Varemo and the maintainer as Leif Varemo. A citation section is present, referencing a 2013 paper by Varemo, Nielsen, and Nookaew published in Nucleic Acids Research, volume 41(8), pp. 4378-4391. The doi is given as 10.1093/nar/gkt111.

Is there any GO, InterPro or Reactome term statistically over-represented in the differentially expressed set of genes?



Plant Differential Experiment page

Transcription profiling by high throughput sequencing of microdissection in maize nkd mutants

RNA-Seq mRNA differential

Organism: Zea mays

Reference(s): 25552497 (Filter by genes in paper)

Results Plots Experiment Design Supplementary Information

'nkd mutant' vs 'wild type genotype' in 'aleurone layer'

GSE GO GO terms enrichment analysis

GSE IPro Interpro domains enrichment analysis

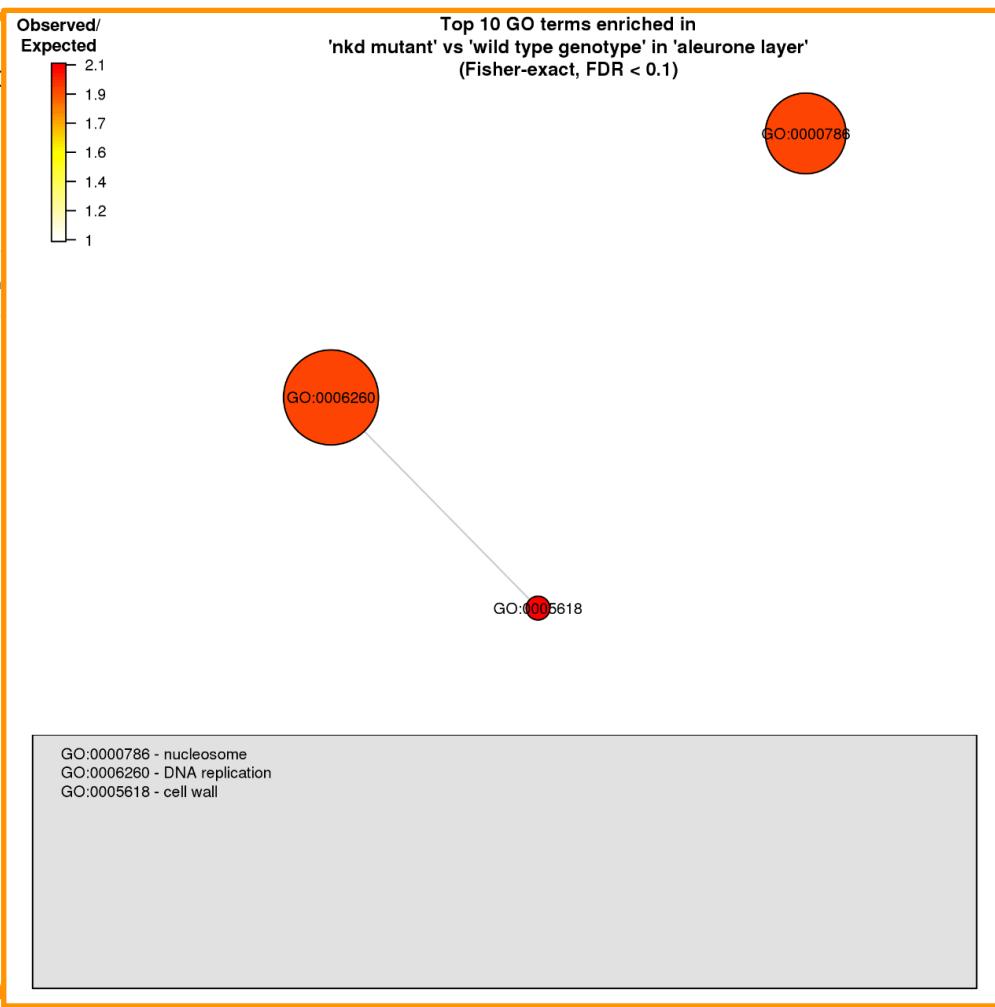
MA MA MA plot for the contrast across all genes

'nkd mutant' vs 'wild type genotype' in 'central endosperm'

GSE GO GO terms enrichment analysis

GSE IPro Interpro domains enrichment analysis

MA MA MA plot for the contrast across all genes





Plant Differential Experiment page

Transcription profiling by high throughput sequencing of aleuron microdissection in maize nkd mutants

RNA-Seq mRNA differential

Organism: *Zea mays*

Reference(s): 25552497 (Filter by genes in paper)

Visualise sample metadata and download metadata annotated to ontologies

Results	Plots	Experiment Design	Supplementary Information	Downloads
---------	-------	-------------------	---------------------------	-----------

Download

Pivoted				Sample Characteristics						Experimental Variables	
Comparison	Reference/Test	Run	age	cultivar	genotype	growth condition	initial time point	organism	organism part	genotype	organism part
▼ 'nkd mutant' vs 'wild	Reference, Test	6 runs	15 day	B73 inbred line	2 genotypes	field study	pollination	<i>Zea mays</i>	aleurone layer	2 genotypes	aleurone layer
	Reference	SRR1567893	15 day	B73 inbred line	wild type genotype	field study	pollination	<i>Zea mays</i>	aleurone layer	wild type genotype	aleurone layer
	Reference	SRR1567894	15 day	B73 inbred line	wild type genotype	field study	pollination	<i>Zea mays</i>	aleurone layer	wild type genotype	aleurone layer
	Reference	SRR1567895	15 day	B73 inbred line	wild type genotype	field study	pollination	<i>Zea mays</i>	aleurone layer	wild type genotype	aleurone layer
	Test	SRR1567896	15 day	B73 inbred line	nkd mutant	field study	pollination	<i>Zea mays</i>	aleurone layer	nkd mutant	aleurone layer
	Test	SRR1567897	15 day	B73 inbred line	nkd mutant	field study	pollination	<i>Zea mays</i>	aleurone layer	nkd mutant	aleurone layer
	Test	SRR1567898	15 day	B73 inbred line	nkd mutant	field study	pollination	<i>Zea mays</i>	aleurone layer	nkd mutant	aleurone layer
▼ 'nkd mutant' vs 'wild	Reference, Test	6 runs	15 day	B73 inbred line	2 genotypes	field study	pollination	<i>Zea mays</i>	central endosperm	2 genotypes	central endosperm
	Reference	SRR1567899	15 day	B73 inbred line	wild type genotype	field study	pollination	<i>Zea mays</i>	central endosperm	wild type genotype	central endosperm
	Reference	SRR1567900	15 day	B73 inbred line	wild type genotype	field study	pollination	<i>Zea mays</i>	central endosperm	wild type genotype	central endosperm
	Reference	SRR1567901	15 day	B73 inbred line	wild type genotype	field study	pollination	<i>Zea mays</i>	central endosperm	wild type genotype	central endosperm
	Test	SRR1567902	15 day	B73 inbred line	nkd mutant	field study	pollination	<i>Zea mays</i>	central endosperm	nkd mutant	central endosperm
	Test	SRR1567903	15 day	B73 inbred line	nkd mutant	field study	pollination	<i>Zea mays</i>	central endosperm	nkd mutant	central endosperm
	Test	SRR1567904	15 day	B73 inbred line	nkd mutant	field study	pollination	<i>Zea mays</i>	central endosperm	nkd mutant	central endosperm



Plant Differential Experiment page

Transcription profiling by high throughput sequencing of aleuron microdissection in maize nkd mutants

RNA-Seq mRNA differential

Organism: *Zea mays*

Reference(s): 25552497 ([Filter by genes in paper](#))

Learn more about our RNA-seq analysis pipeline for differential expression

Results Plots Experiment Design [Supplementary Information](#) Downloads

Analysis Methods

Pipeline version	iRAP 0.8.4d1
Analyzed Libraries	See experiment design
Filtering Step 1	Discard reads below minimum quality threshold
Filtering Step 2	Check of bacterial contamination; discard offending reads
Filtering Step 3	Discard reads with common uncalled characters (e.g. N)
Filtering Step 4	Remove reads from pair-end libraries that were orphaned by filtering steps 1-3
Read Mapping	Against genome reference (Ensembl Genomes release: 34) tophat2 version: 2.1.1
Quantification	htseq2 version: 0.6.1p1
Differential Expression	DESeq2 version: 1.10.1
Gene Set Overlap	Fisher (non-directional), FDR < 0.1 using piano version 1.14.5

Resources

Quality check report (generated by iRAP)

ArrayExpress: experiment E-GEO-61057



Plant Differential Experiment page

Transcription profiling by high throughput sequencing of aleuron microdissection in maize nkd mutants

RNA-Seq mRNA differential

Organism: *Zea mays*

Reference(s): 25552497 (Filter by genes in paper)

Results Plots Experiment Design Supplementary Information Download

All expression results in the experiment

All analytics for this experiment

All raw counts for the experiment

Summary of the expression results for this experiment ready to view in R

Experiment Design (tsv)

Download gene quantification matrices
(TPM and FPKM) and R object



[Home](#) » [Bioconductor 3.7](#) » [Software Packages](#) » ExpressionAtlas

ExpressionAtlas

platforms all downloads 700 / 1561 posts 0 in Bioc 2.5 years
build ok updated before release

DOI: [10.18129/B9.bioc.ExpressionAtlas](https://doi.org/10.18129/B9.bioc.ExpressionAtlas) [f](#) [t](#)

Download datasets from EMBL-EBI Expression Atlas

Bioconductor version: Release (3.7)

This package is for searching for datasets in EMBL-EBI Expression Atlas, and downloading them into R for further analysis. Each Expression Atlas dataset is represented as a SimpleList object with one element per platform. Sequencing data is contained in a SummarizedExperiment object, while microarray data is contained in an ExpressionSet or MAlist object.

Author: Maria Keays

Maintainer: Maria Keays <maria.keays@gmail.com>

Citation (from within R, enter `citation("ExpressionAtlas")`):

Keays M (2018). *ExpressionAtlas: Download datasets from EMBL-EBI Expression Atlas*. R package version 1.8.1.

Efficient search via ontology-expansion



Expression Atlas
Gene expression across species and biological conditions

Home | Download | Release notes | FAQ | Help | Licence | About | Feedback

EMBL-EBI Experimental Factor Ontology (EFO) includes now terms from Plant Ontology, Plant Environment Ontology and Plant Trait Ontology

Exploring gene expression results across species under different biological conditions

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others. Expression Atlas aims to help answering questions such as 'where is a certain gene expressed?' or 'how does its expression change in a disease?'.

Expression Atlas provides gene expression results on more than 3,000 experiments (microarray and RNA-sequencing) from 40 different organisms, including metazoans and plants. Expression profiles of tissues from Human Protein Atlas, GTEx and FANTOM5, and of cancer cell lines from ENCODE, CCLE and Genentech projects can be explored in Expression Atlas. All data are manually curated, annotated to ontology terms allowing for much richer queries and re-analysed using standardised methods.

Expression Atlas visualises gene expression results using heatmaps showing gene expression levels across different biological conditions. Novel analyses and visualisations include: 'enrichment' in each differential comparison of GO terms, Reactome, Plant Reactome pathways and InterPro domains; hierarchical clustering (by baseline expression) of most variable genes and experimental conditions; and, for a given gene-condition, distribution of baseline expression across biological replicates.

Show less about Expression Atlas

Search Gene set enrichment

Gene / Gene properties

Enter gene query...
Examples: REG1B, zinc finger, O14777 (UniProt), GO:0010468 (regulation of gene expression)

Organism Any

Biological conditions

seed ×
seed
endosperm
aleurone layer
basal endosperm transfer layer
central endosperm
seed coat



Let's try Expression Atlas

Differential expression

Hands-on activity

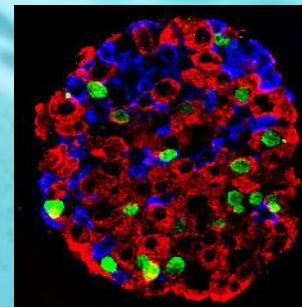
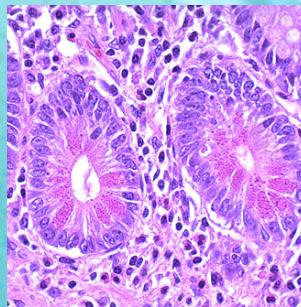
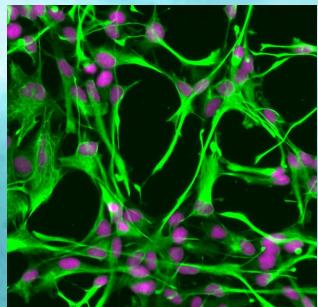
Find differentially expressed genes in response to powdery mildew in wheat



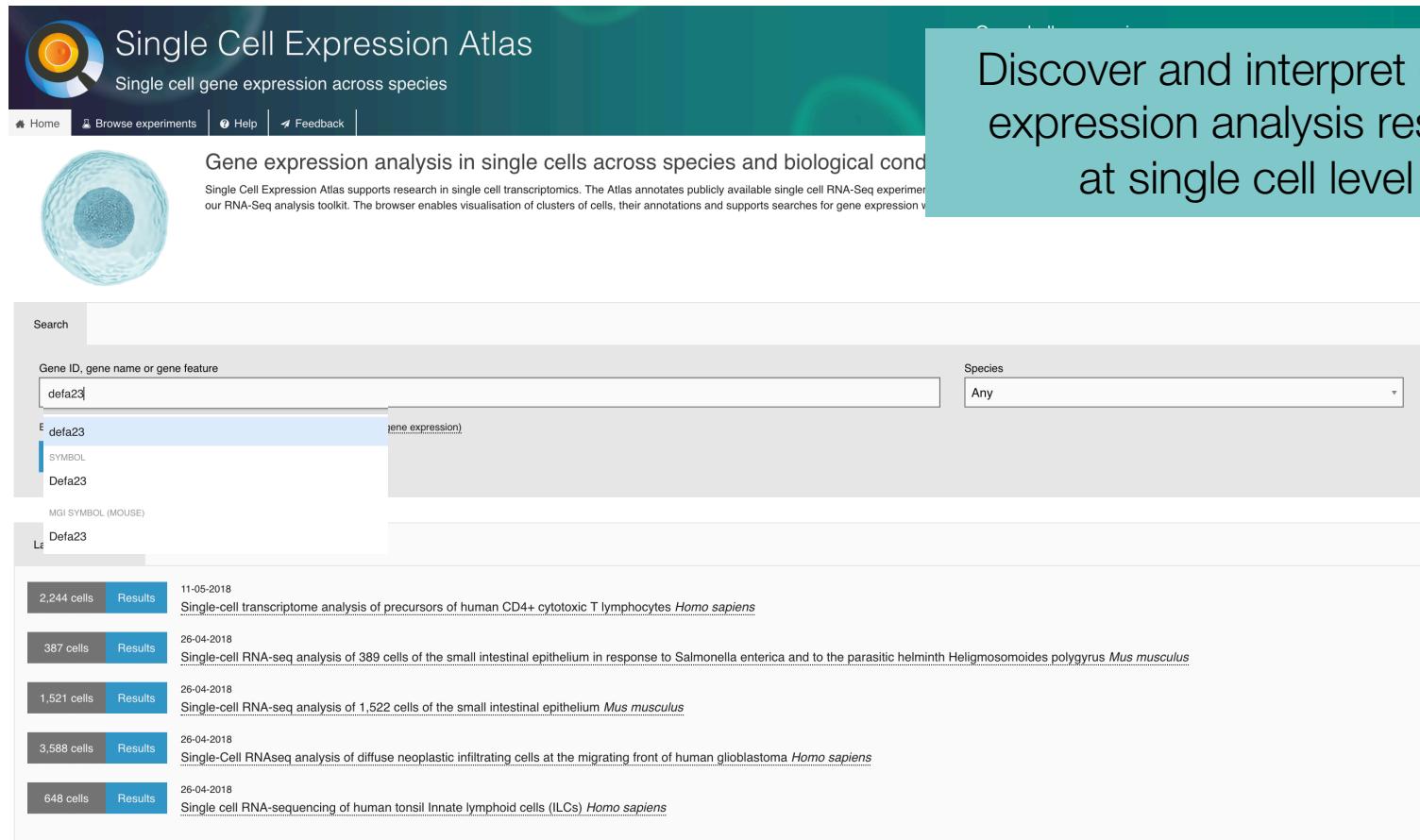
In pairs

Single Cell Expression Atlas

- ✓ Launch in May 218
- ✓ Human, mouse, zebrafish and common fruit fly data
- ✓ Analysis results for ~48,000 cells
- ✓ Including ~200 cell types



New resource - Single Cell Expression Atlas



The screenshot shows the homepage of the Single Cell Expression Atlas. At the top, there's a navigation bar with links for Home, Browse experiments, Help, and Feedback. Below the navigation is a large search bar with placeholder text "Gene ID, gene name or gene feature". A user has typed "defa23" into the search field. To the right of the search bar is a dropdown menu labeled "Species" set to "Any". Below the search bar, there's a section for "MGD SYMBOL (MOUSE)" which shows "Defa23". Underneath this, a list of experimental results is displayed, each with a "Cells" count, a "Results" button, and a timestamp. The results are:

- 2,244 cells | Results | 11-05-2018 | Single-cell transcriptome analysis of precursors of human CD4+ cytotoxic T lymphocytes *Homo sapiens*
- 387 cells | Results | 26-04-2018 | Single-cell RNA-seq analysis of 389 cells of the small intestinal epithelium in response to *Salmonella enterica* and to the parasitic helminth *Heligmosomoides polygyrus* *Mus musculus*
- 1,521 cells | Results | 26-04-2018 | Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium *Mus musculus*
- 3,588 cells | Results | 26-04-2018 | Single-Cell RNAseq analysis of diffuse neoplastic infiltrating cells at the migrating front of human glioblastoma *Homo sapiens*
- 648 cells | Results | 26-04-2018 | Single cell RNA-sequencing of human tonsil Innate lymphoid cells (ILCs) *Homo sapiens*

On the right side of the page, there's a teal sidebar with the text: "Discover and interpret gene expression analysis results at single cell level".

www.ebi.ac.uk/gxa/sc/

New resource - Single Cell Expression Atlas

Search results

Gene ID or gene symbol

defa23

Search

Marker genes

Experiments with marker genes

Species

Mus musculus

Inferred cell type

Select...

Organism part

Select...

defa23 is expressed in:

26-05-2018 | *Mus musculus*

E-GEOID-99235: Single cell RNA-seq of mouse lung vascular transcriptomes

About the experiment: By analyzing about 1500 vascular single cell transcriptomes from adult mouse lung, we provide molecular definitions of the principal vascular (endothelial and mural cells), and vascular-associated cell classes. We used different transgenic mouse reporter lines with fluorescent expression in different populations of vascular cells, Cldn5(BAC)eGFP mice that strongly labels lung endothelial cells, mice positive for Pdgfrb(BAC)eGFP and NG2dsRed. We chose to focus our efforts to adult mice 10 to 19 w... [\(Read more\)](#)

25-05-2018 | *Mus musculus*

E-GEOID-103334: Single cell RNA-seq of microglia cells isolated from the hippocampus of a mouse model of severe neurodegeneration with AD-like phenotypes

About the experiment: In this study, we used single-cell RNA-sequencing to gain unprecedented insight into the phenotypic heterogeneity and the transcriptional dynamics of microglia cells during the progression of neurodegeneration. Briefly, by using a severe neurodegeneration mouse model with Alzheimer's-like pathology and phenotypes (CK-p25 model), we surveyed microglia activation by RNA sequencing longitudinally at fine temporal- and single-cell resolution. In summary, our work identified previously unobserved het... [\(Read more\)](#)

25-05-2018 | *Mus musculus*

E-ENAD-13: Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium

- [View marker gene in clusters 15 for k = 16](#)
- [View marker gene in clusters 6 for k = 6](#)
- [View marker gene in clusters 7 for k = 7](#)
- [View marker gene in clusters 2 for k = 8](#)

Defa23 gene expression across mouse cells from different single cell experiments

New resource - Single Cell Expression Atlas

Search results

Gene ID or gene symbol

defa23

Search

Marker genes

Experiments with marker genes

Species

Mus musculus

Inferred cell type

Select...

Organism part

Select...

Single cell mouse experiments in which *defa23* is a marker gene

defa23 is expressed in:

25-05-2018 | *Mus musculus*

E-ENAD-13: Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium

- [View marker gene in clusters 15 for k = 16](#)
- [View marker gene in clusters 6 for k = 6](#)
- [View marker gene in clusters 7 for k = 7](#)
- [View marker gene in clusters 2 for k = 8](#)
- [View marker gene in clusters 9 for k = 9](#)
- [View marker gene in clusters 10 for k = 10](#)
- [View marker gene in clusters 7 for k = 11](#)
- [View marker gene in clusters 7 for k = 12](#)
- [View marker gene in clusters 13 for k = 13](#)
- [View marker gene in clusters 14 for k = 14](#)
- [View marker gene in clusters 15 for k = 15](#)

About the experiment: By analyzing about 1500 vascular single cell transcriptomes from adult mouse lung, we provide molecular definitions of the principal vascular (endothelial and mural cells), and vascular-associated cell classes. We used different transgenic mouse reporter lines with fluorescent expression in different populations of vascular cells, Cldn5(BAC)eGFP mice that strongly labels lung endothelial cells, mice positive for Pdgfrb(BAC)eGFP and NG2dsRed. We chose to focus our efforts to adult mice 10 to 19 w... [\(Read more\)](#)

New resource - Single Cell Expression Atlas

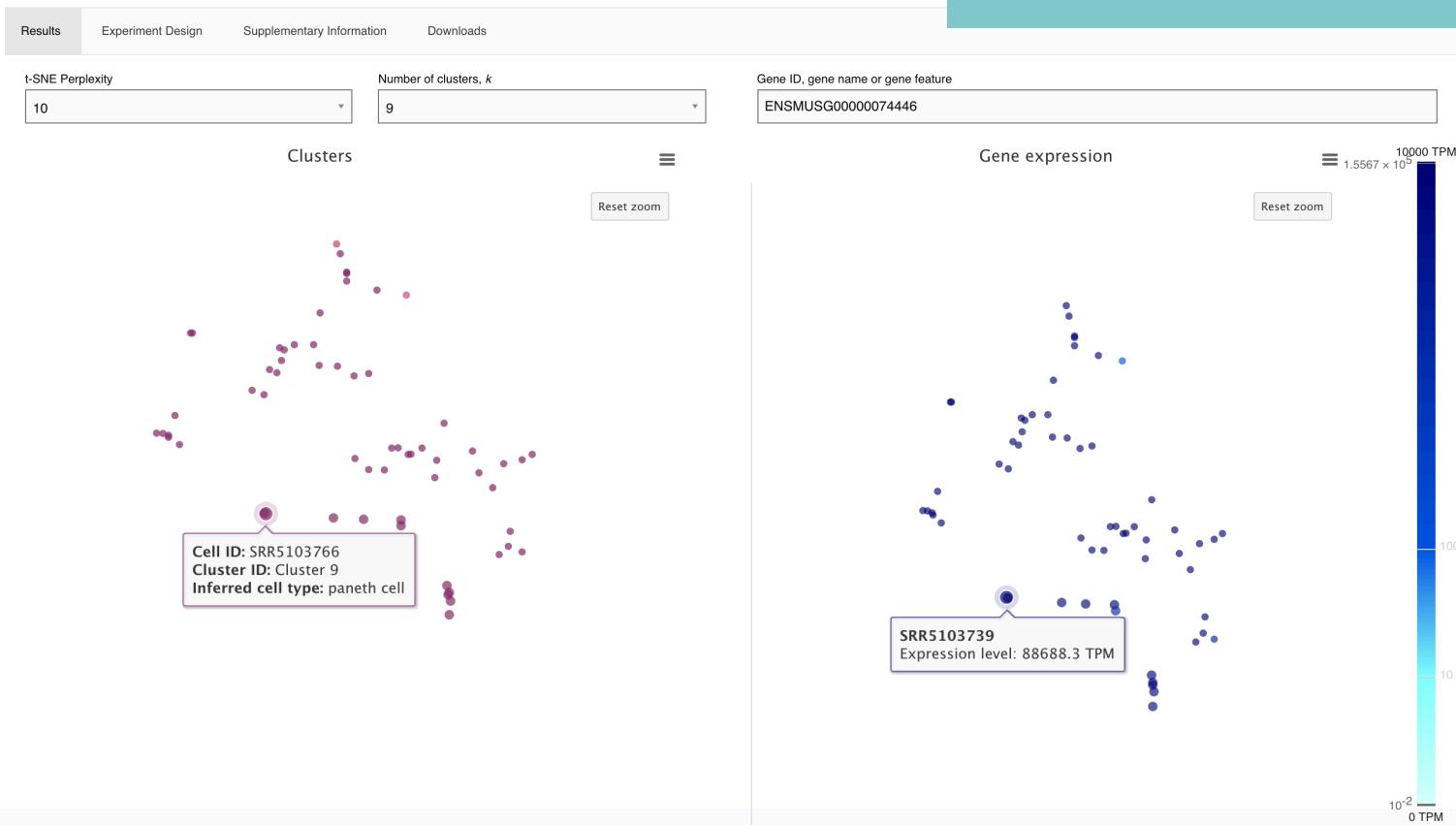
Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium

Single-cell RNA-Seq mRNA baseline

Organism: *Mus musculus*

Publication:

- Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K et al. (2017) *A single-cell survey of the small intestinal epithelium.*



New resource - Single Cell Expression Atlas

Results Experiment Design Supplementary Information Downloads

Metadata is manually curated
and annotated to ontologies

Cell	Sample Characteristics										Experimental Variables
	age	cell type	facs marker	genotype	individual	inferred cell type	organism	organism part	strain	single cell identifier	
SRR5102341	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	late enterocyte progen...	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 1	
SRR5102342	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	enterocyte	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 2	
SRR5102343	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 3	
SRR5102344	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 4	
SRR5102345	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 5	
SRR5102346	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 6	
SRR5102347	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 7	
SRR5102348	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 8	
SRR5102349	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	goblet cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 9	
SRR5102350	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 10	
SRR5102351	7 to 10 week	epithelial cell	LGR5-High	Lgr5-GFP	Mouse 7	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 11	
SRR5102352	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	late enterocyte progen...	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 12	
SRR5102353	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	goblet cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 13	
SRR5102354	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 14	
SRR5102355	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 15	
SRR5102356	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 16	

New resource - Single Cell Expression Atlas

Details on the methods followed to re-analyse each experiment

Results	Experiment Design	Supplementary Information	Downloads
---------	-------------------	---------------------------	-----------

Analysis Methods

Analysis	Software	Version	Citation
QC	FASTQC	0.11.7	http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Filtering and trimming	FASTX	0.0.13	http://hannonlab.cshl.edu/fastx_toolkit/
Pipeline	IRAP	1.0.0b	Nuno A. Fonseca, Robert Petryszak, John Marioni, Alvis Brazma. iRAP - an integrated RNA-seq Analysis Pipeline. bioRxiv 2014. doi: http://dx.doi.org/10.1101/005991
Clustering/Gene markers	SC3	1.4.2	Kiselev, Vladimir Yu, Kristina Kirschner, Michael T. Schaub, Tallulah Andrews, Andrew Yiu, Tamir Chandra, Kedar N. Natarajan et al. SC3: consensus clustering of single-cell RNA-seq data. Nature methods 14, no. 5 (2017): 483.
Reads alignment	bowtie2	2.3.2	Langmead, Ben, and Steven L. Salzberg. Fast gapped-read alignment with Bowtie 2. Nature methods 9, no. 4 (2012): 357-359.
QC	fastq_utils	0.14.7	https://github.com/nunofonseca/fastq_utils
Gene/transcript quantification	kallisto	0.44.0	Nicolas L Bray, Harold Pimentel, Pál Melsted, Lior Pachter. Near-optimal probabilistic RNA-seq quantification. Nature Biotechnology (2016).

New resource - Single Cell Expression Atlas

Download curated metadata
and analysis results

Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium

Single-cell RNA-Seq mRNA baseline

Organism: *Mus musculus*

Publication:

- Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K et al. (2017) [*A single-cell survey of the small intestinal epithelium.*](#)

Results

Experiment Design

Supplementary Information

Downloads

Metadata files

 SDRF file (.txt)

 IDF file (.txt)

 Experiment design file (.tsv)

Result files

 Clustering file (.tsv)

 Filtered quantification files (MatrixMarket)

Getting help and sending feedback

The screenshot shows the Expression Atlas homepage. At the top left is a blue circular logo with a yellow and white 'Q'. To its right is the text 'Expression Atlas' in large white letters, followed by 'Gene expression across species and biological conditions' in smaller white text. Below this is a horizontal navigation bar with six items: 'Home', 'Download', 'Release notes', 'FAQ' (which is highlighted in green), 'Help', 'Licence', and 'About' (which is also highlighted in green). The main content area has a teal background. A large teal box contains the heading 'About Expression Atlas'. Below it is a sub-section titled 'What is Expression Atlas?'. The text explains that Expression Atlas is an open science resource for gene and protein expression data across various species and conditions. It aims to answer questions like 'where is my favourite gene expressed?' and 'how its expression changes in a disease?'. To achieve this, the project involves data curation, analysis, search, and visualisation. On the right side of the slide, there is a diagram showing four teal boxes with icons: 'Data curation' (document with checkmark), 'Data analysis' (bar chart), 'Data search' (sunburst chart), and 'Data visualisation' (heat map). Below these are four descriptive labels: 'High quality data', 'High quality results', 'Accessible results', and '... easy to interpret'.

Expression Atlas

Gene expression across species and biological conditions

Home | Download | Release notes | **FAQ** | Help | Licence | **About**

About Expression Atlas

What is Expression Atlas?

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression. Our mission is to provide the scientific community with freely available information on the abundance and localisation of RNA (and proteins) across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others.

Expression Atlas aims to help answering questions such as 'where is my favourite gene expressed?' or 'how its expression changes in a disease?'

To achieve this objective, our project involves data curation, data analysis and the development of a web application to access and visualise publicly available data.

Data curation
High quality data

Data analysis
High quality results

Data search
Accessible results

Data visualisation
... easy to interpret

atlas-feedback@ebi.ac.uk

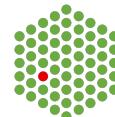


@ExpressionAtlas



Acknowledgements

- ✓ Gramene collaborators
- ✓ EBI teams:
 - SPOT
 - Reactome
 - ENA
 - Ensembl
 - Ensembl Genomes
 - PRIDE
- ✓ Chan Zuckerberg Initiative
- ✓ Human Cell Atlas
- ✓ Wellcome Trust





Expression Atlas: Who we are?



Irene
Papatheodorou
Team Leader

Data curation



Laura Huerta

Data analysis



Pablo Moreno

User Interface



Alfonso Fuentes



Monica Jianu



Nancy George



Anja Fullgrabe



Jon Manning



Suhaib Mohammed



Haider Iqbal

Please fill in the survey before you go!

<https://www.surveymonkey.co.uk/r/KobUni2018>



EMBL-EBI Workshop: Exploring plant gene expression in Expression Atlas, Copenhagen 2018

Workshop Evaluation Form

EMBL-EBI collects feedback from every course and workshop we run. The survey is a way for you to inform us about the workshop you have participated in, what you enjoyed, what you found useful and how we can make improvements. This information is also used to inform the development of new courses and workshops.

Next >>

EMBL-EBI workshop: Exploring plant gene expression in Expression Atlas

Expression Atlas: integrating and displaying plant gene expression

Laura Huerta, PhD

Senior Scientific Curator

lauhuema@ebi.ac.uk

10th SPPS PhD Student Conference