Mode of Action of FNRL in Arabidopsis roots using an integrated omics strategy

Commissioned by: **Dr Ghazanfar Khan** (PI)
Analysed by: **Dr Delphine Vincent** (Data scientist ML)

Completed on 03/03/2026

Google Scholar ORCID LinkedIn GitHub Tableau Public Website email

Table of Contents

1/ Context

Functional analysis of unannotated genes

Approach: Functional analysis of unannotated genes uses a combination of targeted mutations and multi-omics data to reveal biological roles that cannot be predicted by sequence alone.

Generation of Mutants: Mutants serve as the primary tool to observe the phenotypic change created by removing a gene's function.

- **Loss-of-Function (LoF):** Techniques like T-DNA insertion create knockouts to observe essentiality or developmental defects.

Multi-Omic Profiling of Mutants: By comparing the mutant to the wild type across multiple molecular layers, researchers can map the gene's influence.

- **Transcriptomics (RNA-seq):** Identifies gene networks that are up- or down-regulated in response to the mutation, placing the unknown gene within a transcriptional circuit.
- **Proteomics:** Validates if changes in RNA translate to protein levels and identifies protein-protein interaction partners.
- **Metabolomics:** Reveals the biochemical consequences of the gene's activity, such as identifying the specific metabolites that accumulate when a metabolic gene is disrupted.

Integrated Functional Annotation: Computational frameworks synthesize these data layers into a functional model.

- **Co-expression Networks:** Unannotated genes acting as "hubs" in these networks are often critical for stress responses or biofilm formation.
- **Enrichment Analysis:** Tools like Metascape or GSEA link the observed omic changes to established Gene Ontology (GO) terms or KEGG pathways.
- **Network Integration:** Frameworks like Mashup combine heterogeneous data (e.g., PPI and co-expression) to predict biological process annotations with high accuracy.

2/ Project goal

Functional analysis of root-specific FNRL gene in *Arabidopsis thaliana*

Rationale

Arabidopsis thaliana FNRL (Ferredoxin-NADP+ oxidoreductase-like) gene is involved in root architecture, mitochondrial signaling and tolerance to the isoxaben herbicide (1, 2). FNRL is highly conserved across all plant lineages suggesting the preservation of its function (2). FNRL-mediated isoxaben tolerance has potential for biotechnological applications, particularly in developing isoxaben tolerance in crop plants (2).
Another agricultural area of interest is crop nutrition via root nitrate assimilation. Nitrate (NO3 −) is the primary source of nitrogen, a key element in the biosynthesis of amino acids, proteins, and nucleic acids. Nitrate also serves as a signaling molecule regulating plant growth, development, and stress responses (3). Therefore, a deep analysis of FNRL mode of action could reveal its association with nitrate assimilation.

Strategy

To reveal the biochemical pathways linked FNRL (FNR-like; At1g15140) gene in *A. thaliana* roots by comparing wild type with two loss-of-function T-DNA insertion mutants (fnrl-1, fnrl-2) and GFP-based genetic complementation (fnrl-GFP) and using systems biology multi-omics (transcriptomics, proteomics and metabolomics) methods with phenotypic measurments.

Objective

To provide a thorough and robust analysis of the data provided to further the knowledge on FNRL function using a multi-omics integration, machine learning (ML) and bioinformatics strategy (4).

References:
1- Koskela MM, et al. Arabidopsis FNRL protein is an NADPH-dependent chloroplast oxidoreductase resembling bacterial ferredoxin-NADP+ reductases. Physiol Plant. 2018 Feb;162(2):177-190. doi: 10.1111/ppl.12621.
2- Broad RC, et al. The fnr-like mutants confer isoxaben tolerance by initiating mitochondrial retrograde signalling. Plant Biotechnol J. 2024 Nov;22(11):3000-3011. doi: 10.1111/pbi.14421.
3- Mao J, et al. The crosstalk between nitrate signaling and other signaling molecules in Arabidopsis thaliana. Front Plant Sci. 2025 Mar 10;16:1546011. doi: 10.3389/fpls.2025.1546011.
4- Kundu & Tanti, Decoding plant physiology through systems biology: Integrative multi-omics and computational perspectives for next-generation crop design, Plant Communications (2025), doi: 10.1016/j.xplc.2025.101668

3/ Wet lab study overview

Experimental design (wet lab - Ghaz)

Species & Genotypes:

- **Species**: *Arabidopsis thaliana*
- **Genotypes**: one wild type (Col-0), two independent mutants (*fnrl-1*, *fnrl-2*), and one complemented line (FNRL-GFP)

Materials:

- **Tissue**: Roots (primary tissue of FNRL expression)

Phenotyping:

- Nitrate content (ug/g) in roots and shoots
- Root length (cm) measured under varying external nitrate concentrations (0.1, 1, 10, 60 mM)

Omics analyses:

- **Transcriptomics**: RNA-seq
- **Proteomics**: LC-MS/MS
- **Metabolomics**: GC-MS

Replicates:

- **Biological replicates**: 3 for transcripts & proteins, 4 for metabolites (no technical replicates)

Samples:

- **Transcripts & proteins**: 4 genotypes (Col-0, *fnrl-1*, *fnrl-2*, GFP) x 3 replicates = 12
- **Metabolites**: 3 genotypes (Col-0, *fnrl-1*, *fnrl-2*; no complementation) x 4 replicates = 12

4/ Dry lab strategy

What EDA, statistical analyses and ML will bring to this study (Delphine)

Rationale

- Mutated *fnrl* gene impacts other genes whose expression will change and can be detected in the omics data
- Data analysis will extract responsive features (transcript, protein and metabolite biomarkers)
- Bioinformatics will highlight affected biochemical pathways

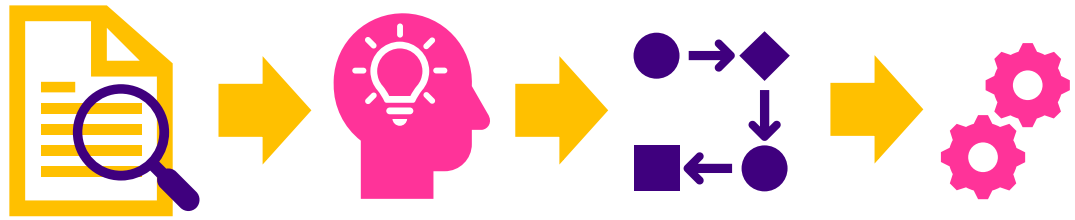
Analysis workflow

- Data handling
- Exploratory data analysis (EDA)
- Statistical tests (stats)
- Biomarker discovery
- Biomarker profiling
- Feature engineering
- Machine Learning (ML) modelling
- Data mining

Tools

- Python
- Online bioinformatics tools

Data files



4.1/ Dry lab steps

Data handling

Clean data for downstream analyses

- Handle missing values
- Normalise abundances

EDA:

Assess data quality and identify general patterns

- Bar charts of phenotypic data
- PCA of samples
- Loading plots of omics features

Statistical test:

Detect features significantly responding to mutation

- Pairwise differential analysis (DE contrast)

Biomarker discovery:

Use engineered variables to isolate FNRL-regulated omics features

- Define rules
- Apply rules to label responsive features
- Filter out unaffected features

Biomarker profiling:

Group FNRL-responsive features by expression profile

- Unsupervised clustering (HCA)
- Semi-Supervised clustering (k-means)
- Supervised clustering (LDA)

Feature engineering

Capture EDA and test as new variables for downstream analyses

- categorical
- binary
- numerical

ML modelling:

Does FNRL-dependent molecular variation explain phenotypic outcomes?

- Model 1: Elastic Net / Lasso for linear effects
- Model 2: Random Forest for non-linear relationships

Data mining and bioinformatics

Annotation and pathway analyses

- **UniprotKB** (https://www.uniprot.org/)
- **PlantReactome** (https://plantreactome.gramene.org/PathwayBrowser/#/SPECIES=170905&DTAB=AN&ANALYSIS=MjAyNjAxMTkwMTA3MThfMTQ%253D)
- **PMN** (https://pmn.plantcyc.org/)
- **PMN AraCyc** (Pathway Tools) (https://pmn.plantcyc.org/organism-summary?object=ARA)
- **KEGG** (https://www.kegg.jp/kegg/)
- **GO** (https://geneontology.org/)
- **NCBI Analyze** (https://www.ncbi.nlm.nih.gov/home/analyze/)
- **Cytoscape** (https://cytoscape.org/)
- **STRING** (https://string-db.org/)
- **PeptideAtlas** (https://peptideatlas.org/)
- **PaintOmics** (https://paintomics.uv.es/)
- **MetaboAnalyst** (https://www.metaboanalyst.ca/)
- **AtMAD** (http://119.3.41.228/atmad/index.php)

4.2/ Computational Pipeline

Objective
- Identify FNRL-regulated genes and by-products (biomarkers)

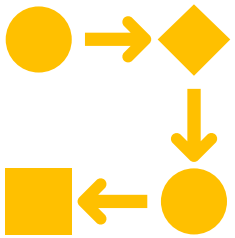
Data files
- Sample metadata file
- Phenotypic measurements
- Omics quantitative datasets
- TAIR annotations

Data handling
- Handle missing values
- Normalise abundances

Data analysis
- Exploratory data analyses (EDA)
- Statistical analyses
- Biomarker discovery and exploration

Data modelling
- Phenotype ML modelling based on biomarker profiles
- Model predictions

Bioinformatics
- Biomarker mining using bioinformatics tools
- Pathway analysis
- Interaction networks



4.3/ Tools

Data handling


python


Excel


python pandas


python NumPy

IP[y]: python iPython

/(reg) /ex/ python RegEx

Data analysis


python scikit-learn


python SciPy


python statsmodels


python seaborn


python matplotlib


python pandas

Data mining


PMN

Reporting & Publishing


JupyterLab


Anaconda Navigator


PowerPoint


GitHub

Debugging


ChatGPT

4.4/ Data files

List and description of files:

design

- **Sample_metadata.csv**: experimental design across all samples and omics streams.

phenotyping

- **Sample_Nitrate_Content_2025-12-17.csv**: nitrate content of shoot and root samples (3 reps) (original file name: "Nitrate Content 111022 from John 2.xlsx"). I only use the roots results that I've incorporated into the "Sample_metadata.csv" file under "nitrate_content" header. Only 3 reps were measured. As metabolomics samples comprise 4 reps, I averaged the 3 reps as a proxy for the 4th rep.
- **Sample_Root_length_2025-12-17.csv**: root lengths under various nitrate concentrations (14-29 reps) (original file name: "Data-FNRL-GFP lines on different nitrate.xlsx"). I added rep 1-3 to divide each lot of replicates into 3 sets to match the omics rep numbers and averaged across those subsets 1-3. I've incorporated the averaged into the "Sample_metadata.csv" file under "root_len_xxmM" header. As metabolomics samples comprise 4 reps, I averaged the 3 reps as a proxy for the 4th rep.

omics

- **Transcripts_fnrl_2025-11-12.csv**: transcript abundances for 4 genotypes x 3 reps (original file name: "normalized_counts_new.csv"). No NAs but lots of 0, therefore I think NAs were replaced by 0.
- **Proteins_fnrl_2025-11-12.csv**: protein abundances for 4 genotypes x 3 reps (original file name: "ME449_FNRL_gel fractioned_PROTEINS_by_rep.xlsx“, keeping only abundances columns AQ-BB). lots of NAs.
- **Metabolites_fnrl_2025-11-12.csv**: metabolite abundances for 3 genotypes x 4 reps (original file name: "normalized_counts_new.csv"). No NAs but lots of 0, therefore I think NAs were replaced by 0. Data harmonisation with TX and PR: 1 rep / genotype with most 0 dropped + 3 GFP reps introduced and filled with 0.

annotations

- **Proteins_IDs_2025-11-12.csv**: protein description (original file name: "ME449_FNRL_gel fractioned_PROTEINS_by_rep.xlsx“, keeping only description columns E-S). I might not use this.
- **TAIR10_functional_descriptions.csv**: gene functions that I retrieved from TAIR10 (https://www.arabidopsis.org/download/list?dir=Genes%2FTAIR10_genome_release).
- **ATH_GO_GOSLIM.txt**: gene ontology annotations that I retrieved from TAIR10 (https://www.arabidopsis.org/download/list?dir=GO_and_PO_Annotations%2FGene_Ontology_Annotations).
- **Arabidopsis_thaliana.TAIR10.pep.all.fa**: protein AA sequences fasta file from Gramene (plant ensembl https://ensembl.gramene.org/Arabidopsis_thaliana/Info/Index)

Colour code:
Files provided by Ghaz
Files generated or retrieved by Delphine

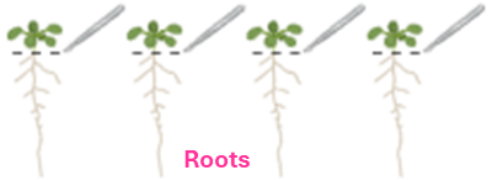
4.5/ Study workflow

1. Wet-lab experiment

Phenotyping : Nitrate content + Root length

Genotypes: WT Col-0 | Mutant 1 *fnrl-1* | Mutant2 *fnrl-2* | Complement FNRL-GFP

Sample collection:



Tissue: Roots

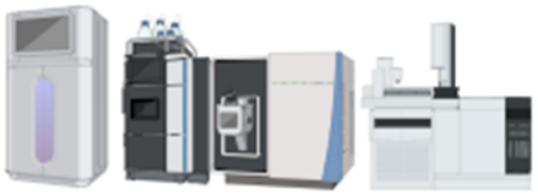
Replicates: 3 Biological replicates

Sample preparation:

Analytes: Transcripts | Proteins | Metabolites

Sample processing:

Omics technologies:



RNAseq | LC-MS/MS | GC-MS

Data processing:

Data export

2. Dry-lab experiment

Data files:
- Phenotypic measurements
- Transcript abundances
- Protein abundances
- Metabolite abundances

Data handling:
- Data harmonisation
- NA imputation
- Normalisation

EDA:
- Summary statistics
- PCA

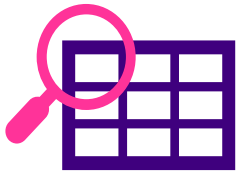
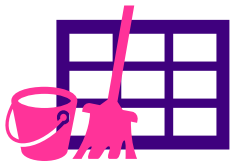
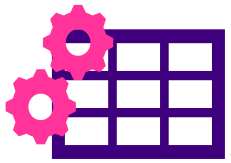
Statistical analyses:
- DE pairwise contrast
- Volcano plots

Biomarker discovery:
- FNRL-responsive features

Biomarker profiling:
- HCA
- K-means
- LDA
- Feature engineering

Phenotype ML modelling:
- Elastic Net/Lasso
- RF

Data mining Bioinformatics:
- **Feature annotation retrieval**
  TAIR10, UniProtKN, PeptideAtlas, MetaboAnalyst
- **Functional annotation**
  GO resource, AmiGO
- **Pathway mapping**
  KEGG, PlantReactome
- **Interaction network**
  STRING, Cytoscape
- **Multi-omics integration**
  PaintOmics 4, PMN AraCyc

5/ Data Handling

Overview:

1. Produce sample metadata file
- Needed for all data analysis

2. Format phenotyping data
- To be used as sample descriptors in analyses

3. Clean omics data
- To allow robust data analyses

4. Summary statistics and distribution
- To evaluate data transformation

5. Check FNRL expression
- To assess mutation effect on target gene

Colour code:
Files provided by Ghaz
Files generated or retrieved by Delphine

5.1/ Sample metadata

Sample description reflects study experimental design

The table below was produced to **link dataset headers (i.e. sample names) to genotypes, omics layers, replicates and phenotypic measurements**.
This sample stratification will allow **meaningful and relevant analyses** (EDA, stats, ML).

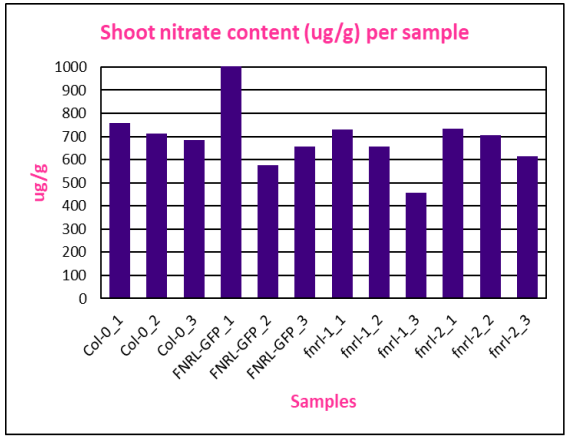
index	sample_id	genotype	genotype_group	genotype_type	replicate	omics_layer	root_nitrate_content	shoot_nitrate_content	root_len_0.1mM	root_len_1mM	root_len_10mM	root_len_60mM
1	Col-0_1	Col-0	WT	control	1	transcriptomics	611.18	759.13	2.13	2.50	2.48	1.59
2	Col-0_2	Col-0	WT	control	2	transcriptomics	830.30	712.90	2.15	2.34	2.37	1.64
3	Col-0_3	Col-0	WT	control	3	transcriptomics	765.93	684.29	2.25	2.47	2.45	1.60
4	fnrl-1_1	fnrl-1	mutant	mutant	1	transcriptomics	439.50	730.51	1.47	1.91	1.81	1.41
5	fnrl-1_2	fnrl-1	mutant	mutant	2	transcriptomics	504.23	657.87	1.45	1.94	1.80	1.49
6	fnrl-1_3	fnrl-1	mutant	mutant	3	transcriptomics	397.47	457.57	1.32	1.86	1.83	1.37
7	fnrl-2_1	fnrl-2	mutant	mutant	1	transcriptomics	589.24	732.71	1.36	1.48	1.81	1.20
8	fnrl-2_2	fnrl-2	mutant	mutant	2	transcriptomics	508.09	706.30	1.36	1.66	1.60	1.08
9	fnrl-2_3	fnrl-2	mutant	mutant	3	transcriptomics	508.03	613.85	1.22	1.57	1.79	1.19
10	FNRL-GFP_1	FNRL-GFP	complement	control	1	transcriptomics	736.33	1005.66	2.04	2.61	2.66	1.75
11	FNRL-GFP_2	FNRL-GFP	complement	control	2	transcriptomics	825.49	574.23	1.96	2.53	2.56	1.82
12	FNRL-GFP_3	FNRL-GFP	complement	control	3	transcriptomics	868.64	657.87	2.18	2.50	2.56	1.86
13	Col-0_1	Col-0	WT	control	1	proteomics	611.18	759.13	2.13	2.50	2.48	1.59
14	Col-0_2	Col-0	WT	control	2	proteomics	830.30	712.90	2.15	2.34	2.37	1.64
15	Col-0_3	Col-0	WT	control	3	proteomics	765.93	684.29	2.25	2.47	2.45	1.60
16	fnrl-1_1	fnrl-1	mutant	mutant	1	proteomics	439.50	730.51	1.47	1.91	1.81	1.41
17	fnrl-1_2	fnrl-1	mutant	mutant	2	proteomics	504.23	657.87	1.45	1.94	1.80	1.49
18	fnrl-1_3	fnrl-1	mutant	mutant	3	proteomics	397.47	457.57	1.32	1.86	1.83	1.37
19	fnrl-2_1	fnrl-2	mutant	mutant	1	proteomics	589.24	732.71	1.36	1.48	1.81	1.20
20	fnrl-2_2	fnrl-2	mutant	mutant	2	proteomics	508.09	706.30	1.36	1.66	1.60	1.08
21	fnrl-2_3	fnrl-2	mutant	mutant	3	proteomics	508.03	613.85	1.22	1.57	1.79	1.19
22	FNRL-GFP_1	FNRL-GFP	complement	control	1	proteomics	736.33	1005.66	2.04	2.61	2.66	1.75
23	FNRL-GFP_2	FNRL-GFP	complement	control	2	proteomics	825.49	574.23	1.96	2.53	2.56	1.82
24	FNRL-GFP_3	FNRL-GFP	complement	control	3	proteomics	868.64	657.87	2.18	2.50	2.56	1.86
25	Col-0_1	Col-0	WT	control	1	metabolomics	611.18	759.13	2.13	2.50	2.48	1.59
26	Col-0_2	Col-0	WT	control	2	metabolomics	830.30	712.90	2.15	2.34	2.37	1.64
27	Col-0_3	Col-0	WT	control	3	metabolomics	765.93	684.29	2.25	2.47	2.45	1.60
28	fnrl-1_1	fnrl-1	mutant	mutant	1	metabolomics	439.50	730.51	1.47	1.91	1.81	1.41
29	fnrl-1_2	fnrl-1	mutant	mutant	2	metabolomics	504.23	657.87	1.45	1.94	1.80	1.49
30	fnrl-1_3	fnrl-1	mutant	mutant	3	metabolomics	397.47	457.57	1.32	1.86	1.83	1.37
31	fnrl-2_1	fnrl-2	mutant	mutant	1	metabolomics	589.24	732.71	1.36	1.48	1.81	1.20
32	fnrl-2_2	fnrl-2	mutant	mutant	2	metabolomics	508.09	706.30	1.36	1.66	1.60	1.08
33	fnrl-2_3	fnrl-2	mutant	mutant	3	metabolomics	508.03	613.85	1.22	1.57	1.79	1.19
34	FNRL-GFP_1	FNRL-GFP	complement	control	1	metabolomics	736.33	1005.66	2.04	2.61	2.66	1.75
35	FNRL-GFP_2	FNRL-GFP	complement	control	2	metabolomics	825.49	574.23	1.96	2.53	2.56	1.82
36	FNRL-GFP_3	FNRL-GFP	complement	control	3	metabolomics	868.64	657.87	2.18	2.50	2.56	1.86

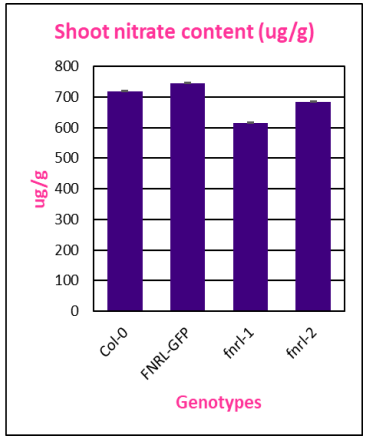
NOTE: The original metabolomics data was wrangled so that it could be aligned with transcript and protein samples. Read file description on "**4.1 Data files**" slide for more details on data harmonisation.

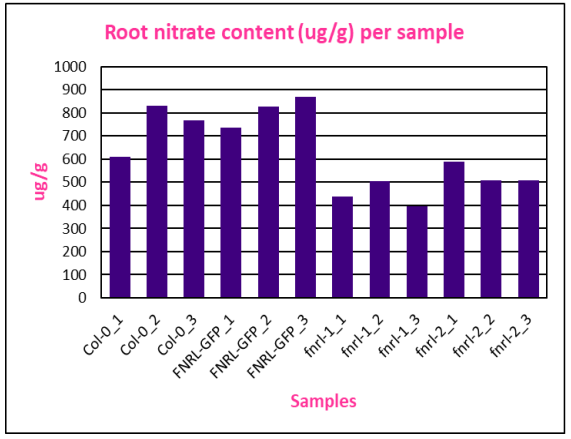
5.2/ Phenotyping

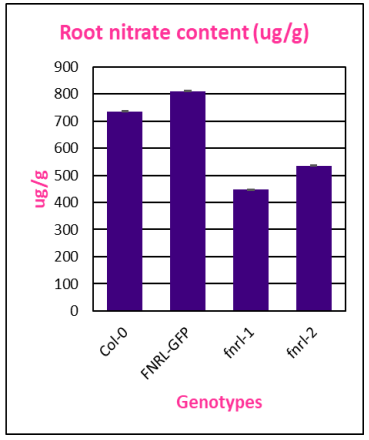
Nitrate assimilation impacted by *fnrl* mutation

- Nitrate content is **reduced in both fnrl mutants** in roots and shoots, with a **stronger effect observed in fnrl-1 than fnrl-2**.
- **Roots are more strongly affected than shoots**, consistent with the root-specific expression of FNRL.
- GFP complementation leads to a **partial recovery of nitrate content**, supporting a direct role of FNRL in nitrate regulation.



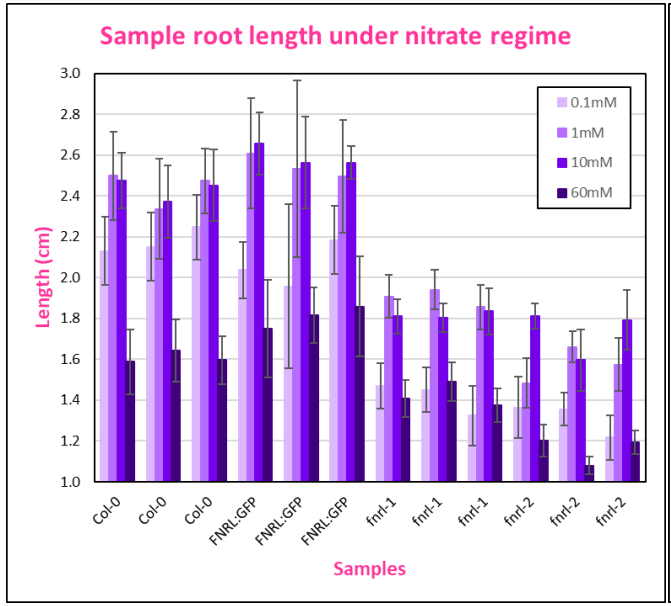


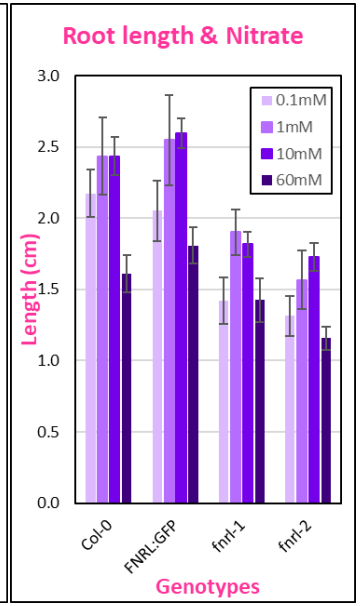


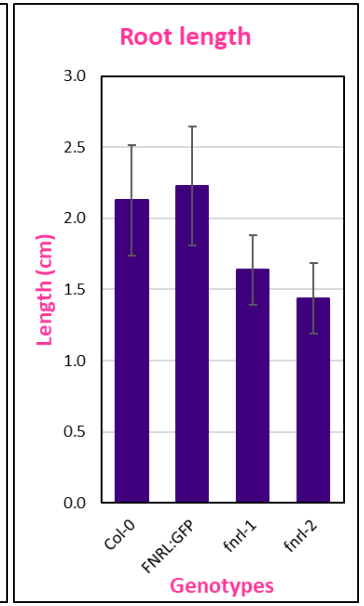


Root growth affected by *fnrl* mutation and nitrate supply

- Root length is **maximised under intermediate nitrate concentrations (1–10 mM)** across all genotypes.
- Both **low nitrate (0.1 mM) and high nitrate (60 mM) significantly reduce root growth**.
- **fnrl mutants exhibit reduced root length**, with **fnrl-2** showing a stronger growth penalty.
- GFP complementation results in a **modest but consistent improvement in root development** relative to wild type.







NOTE: The original data was wrangled so that it could be aligned with omics samples. Read file description on "**4.1 Data files**" slide for more details.

5.3/ Omics data cleaning steps

1. Missing value handling

- Zero values were treated as missing and converted to NA
- Features containing only NA values were removed
- Missing values were imputed using **genotype-specific means** when at least one replicate was present
- Features entirely missing within a genotype were conservatively set to zero
- **Genotype-wise imputation was chosen to preserve biologically meaningful differences between mutant and complemented lines**.

2. Data transformation

- Transcriptomic and proteomic data exhibited strong right skew and wide dynamic ranges
- A $\log_2(1+x)$ transformation was applied to stabilise variance and approximate normality
- Metabolomics data were already approximately normally distributed and were not transformed

3. Scaling

- To enable **cross-omics comparability**, all datasets were Min–Max scaled to a common range.
- **Min–Max scaling to a shared 0–20 range was applied to enable cross-omics comparison while preserving relative feature distributions. The upper bound was chosen to reflect the observed dynamic range of log-transformed intensities across transcriptomic, proteomic, and metabolomic layers**.

5.4/ Summary statistics

Summary statistics show how data conversion corrects skew and range

Raw transcriptomic and proteomic data exhibited **extreme right skewness and kurtosis**, reflecting wide dynamic ranges and sparse expression.

$Log_2(1+x)$ transformation substantially reduced both skewness and tail heaviness, yielding approximately symmetric distributions.

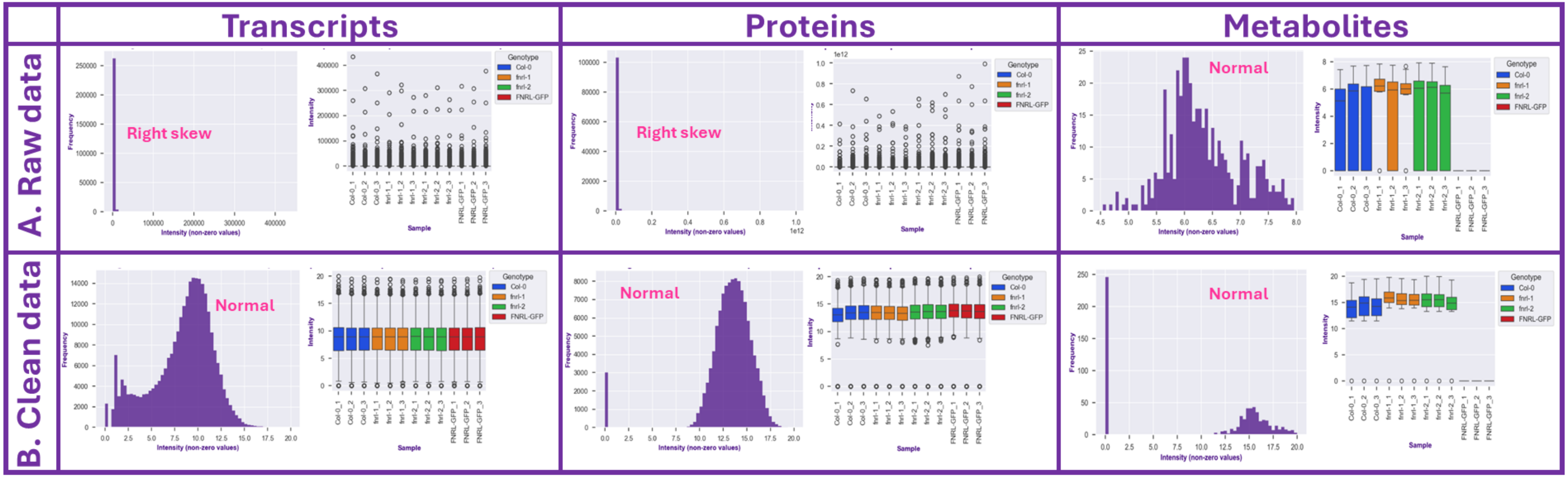
Min–Max scaling to a common range preserved distributional shape while enabling cross-omics comparability.

omics	data	cleaning methiod	total_values	missing_values	zeros	mean	median	std	min	max	skewness	kurtosis
TX	raw	none	276792	0	9030	1041	341	3963	0	433569	40	2737
PR	raw	none	120540	15110	0	2.E+09	2.E+08	1.E+10	3.E+04	1.E+12	31	1411
MT	raw	none	648	0	231	4	6	3	0	8	0	0
TX	clean1	NA imputation --> log2(x+1)	276708	0	2316	7.77	8.42	3.08	0	18.73	-0.64	-0.18
PR	clean1	NA imputation --> log2(x+1)	115320	0	3012	26.46	27.01	5.47	0	39.85	-2.80	11.74
MT	clean1	NA imputation	648	0	108	5.19	6.01	2.39	0	7.96	-1.54	0.79
TX	clean 2	NA imputation --> log2(x+1) --> min-max_0-20	276708	0	2316	8.30	8.99	3.29	0	20	-0.64	-0.18
PR	clean 2	NA imputation --> log2(x+1) --> min-max_0-20	115320	0	3012	13.28	13.56	2.74	0	20	-2.80	11.74
MT	clean 2	NA imputation --> min-max_0-20	648	0	108	13.03	15.09	6.02	0	20	-1.54	0.79

5.5/ Data distribution

Data cleaning achieves gaussian distribution

Histograms and box plots confirm that **clean data** are **unskewed, normally distributed and uniformly scaled**.

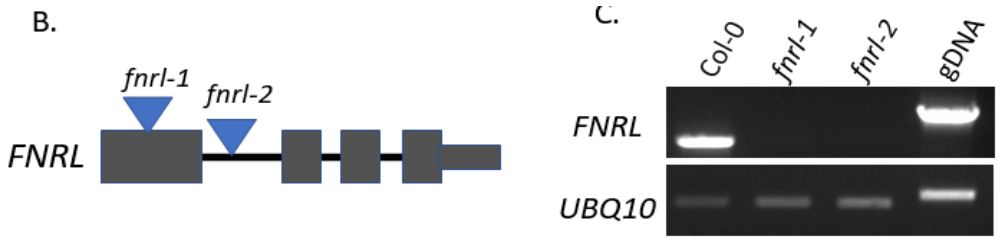


5.6/ FNRL (AT1G15140) expression profiles

FNRL is post-transcriptionally regulated

Literature summary:

Published qRT-PCR analysis demonstrated **complete loss of functional FNRL expression in fnrl-1 and fnrl-2 mutants** (1) (Figure below).



Reference:

1- Broad RC, et al. The fnr-like mutants confer isoxaben tolerance by initiating mitochondrial retrograde signalling. Plant Biotechnol J. 2024 Nov;22(11):3000-3011. doi: 10.1111/pbi.14421.

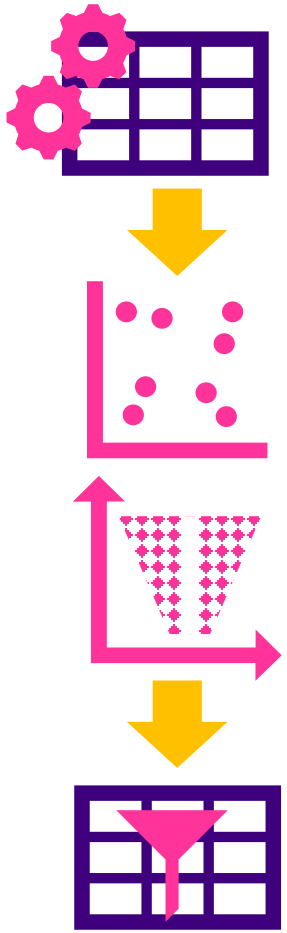
This study:

In the current study, **RNA-seq detects residual transcript signal for AT1G15140 in fnrl mutants**, likely reflecting partial or non-functional transcripts captured by sequencing.

In contrast, **proteomics data confirm loss of FNRL protein in fnrl-1 and strong knockdown in fnrl-2**, consistent with published knockout validation PCR test.

GFP complementation strongly elevates transcript abundance and partially restores protein accumulation.

Raw intensities illustrate absolute expression differences.
Cleaned intensities reflect relative expression after cross-omics normalisation.



6/ EDA & Stats

Overview:

1. Principal Component Analysis (PCA) on all data for each omics layer

- Assess sample distribution and reproducibility. Identify global patterns across genotypes.

2. Differential Analysis (DE) Contrasts on all data for each omics layer

- Robust statistical method to isolate FNRL-responsive features.
- Pairwise comparisons:
    - WT vs fnrl-1
    - WT vs fnrl-2
    - WT vs fnrl-GFP
- Scatterplots of all 3 pairwise DE comparisons by highlighting FDR significant ($q < 0.05$) features of high fold-change ($|FC| > 2$).

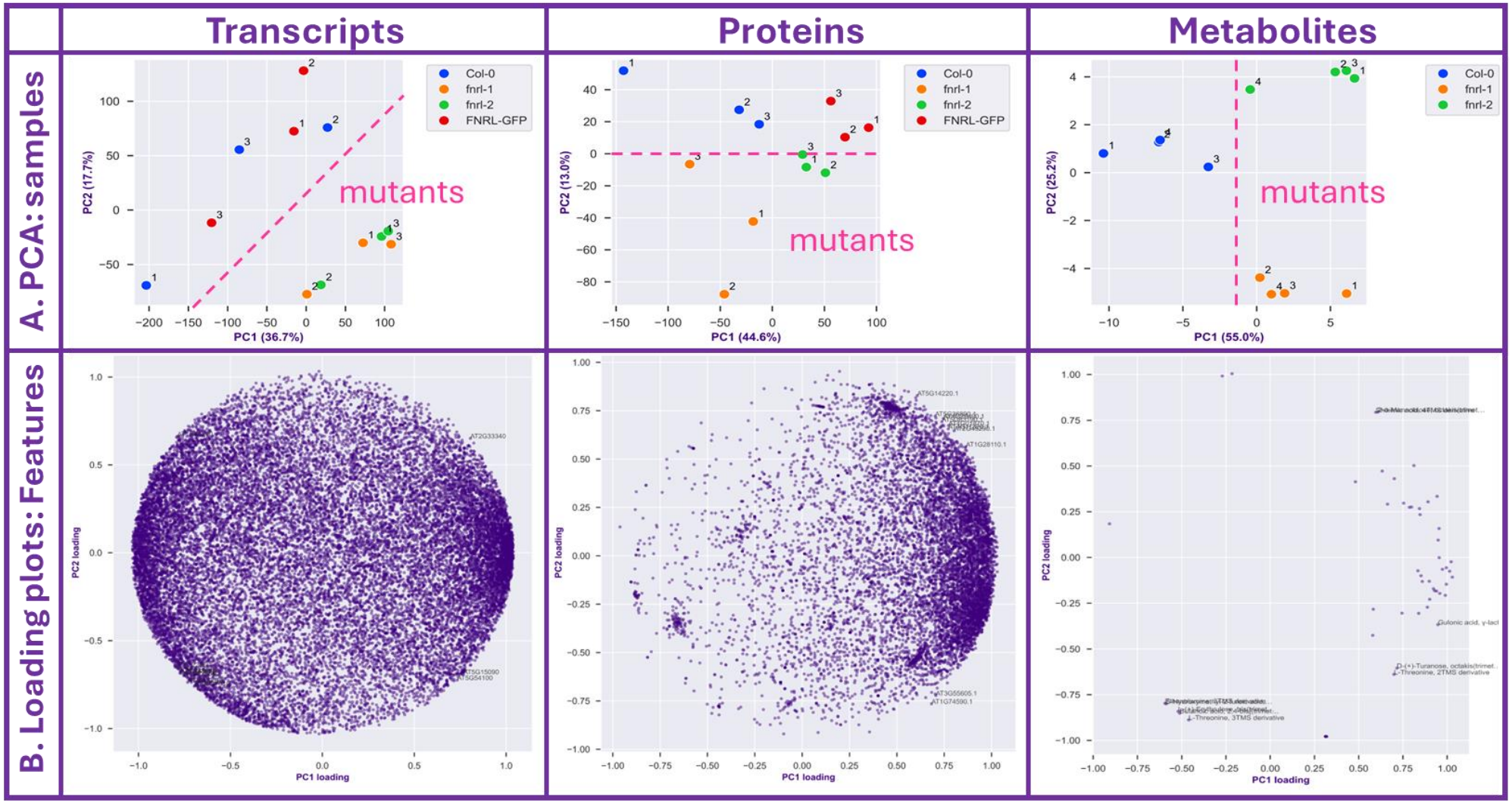
3. Biomarker discovery

- Identify FNRL-regulated omics features

6.1/ PCA on all features

Abundances reflect genotypic differences

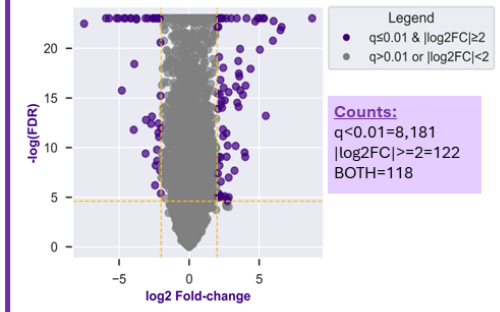
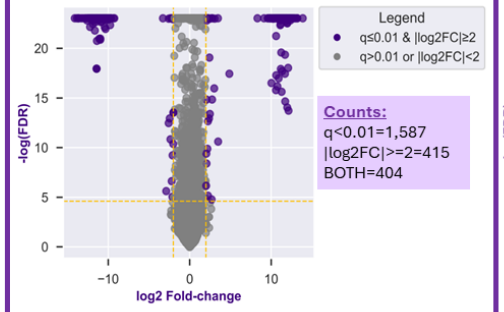
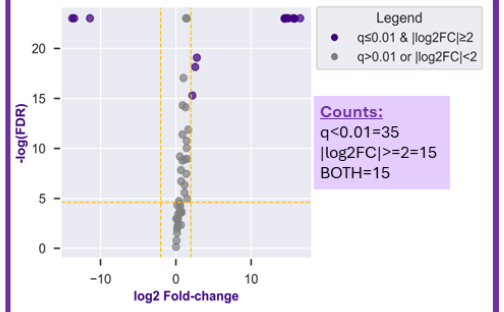
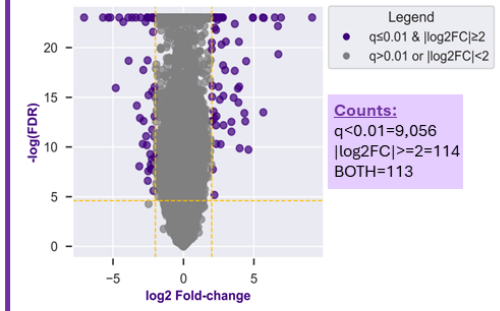
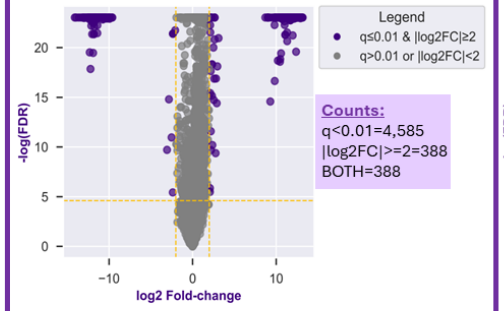
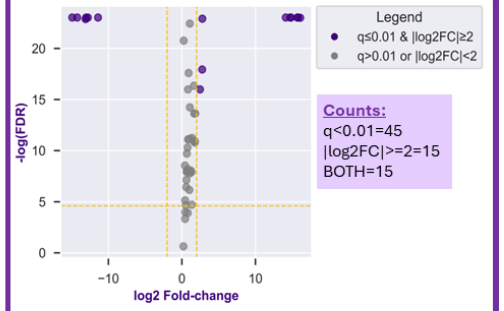
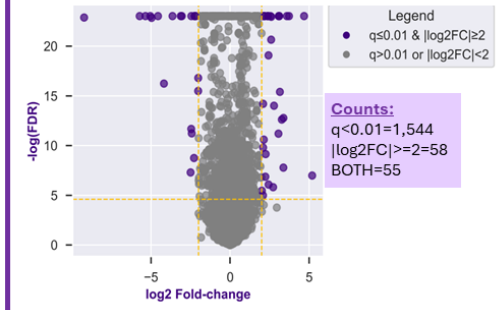
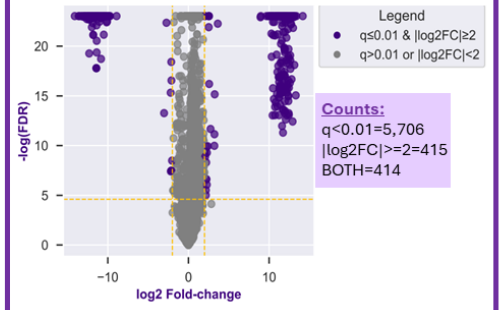
PCA reveals strong **genotype-driven structure across all omics layers**, with increasing signal sparsity from transcripts to metabolites due to reduction in dataset sizes. **Acceptable reproducibility** (fair grouping of replicates, more so in PR and MT than TX).



6.2/ DE contrasts & Volcano plots

Univariate analyses isolate responsive features

Differential expression (DE) analysis using pairwise contrast tests reveals **strong, allele-consistent FNRL-dependent regulation across all omics layers**, with **partial and incomplete rescue** in the GFP-complemented line.

	Transcripts	Proteins	Metabolites
A. WT vs fnrl-1			
B. WT vs fnrl-2			
C. WT vs GFP			No metabolomics performed on FNRL-GFP samples

6.3/ Biomarker discovery

Principle on activation/repression by FNRL gene

Activation:

- If features are **more abundant** in **mutants**, they are **FNRL-repressed**.

Repression:

- If features are **less abundant** in **mutants**, they are **FNRL-activated**.

Rules:

Experimental design implies **three conditions**:

- fnrl-1 vs WT → loss of FNRL (allele 1)
- fnrl-2 vs WT → loss of FNRL (allele 2)
- FNRL-GFP vs WT → functional rescue

A bona fide FNRL-regulated must fulfil the following rules:

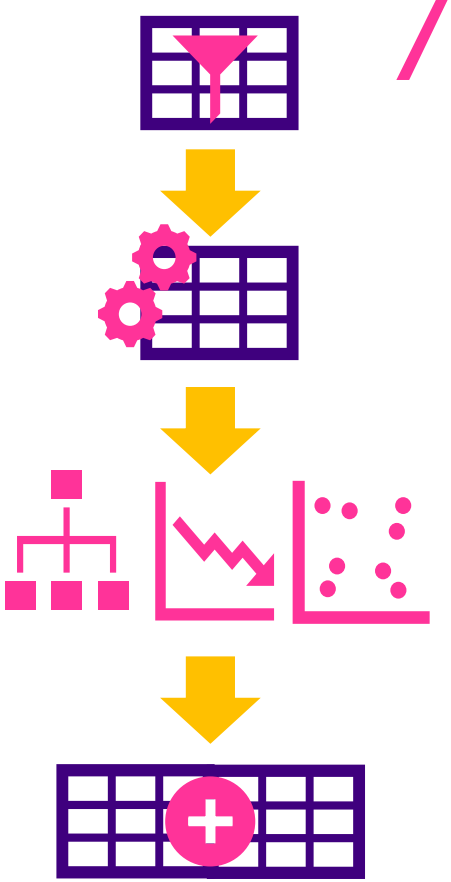
- **Rule A: Reproducible perturbation**
  - Significant contrast DE in both fnrl-1 and fnrl-2
  - Same direction of change in both mutants
- **Rule B: Rescue**
  - Effect size in FNRL-GFP vs WT is reduced in magnitude compared to both mutant contrasts, and/or directionally reversed

Results:

	Transcripts	Proteins	Metabolites
Summary	sig fnrl1: 870 sig fnrl2: 945 GFP FC non-missing: 23059 Consistent across mutant: 395 Core FNRL targets: **341** rescued_by_GFP (True): 341 rescued_by_GFP (NaN): 0	sig fnrl1: 694 sig fnrl2: 825 GFP FC non-missing: 9610 Consistent across mutant : 325 Core FNRL targets : **148** rescued_by_GFP (True): 148 rescued_by_GFP (NaN): 0	sig fnrl1: 31 sig fnrl2: 24 GFP FC non-missing: 0 Consistent across mutant : 15 Core FNRL targets: **15** GFP rescue not assessed

There are in total **504 biomarkers** (341 TX + 148 PR + 15MT).
The number of FNRL-regulated features decreases from transcripts to proteins to metabolites, reflecting increasing **regulatory constraint and buffering across molecular layers, but also dataset size**.
Applying stringent genetic and rescue-based criteria identifies a high-confidence set of 189 FNRL-regulated biomarkers, revealing a robust transcriptional regulon with partial propagation to the proteome and limited downstream metabolic effects.

7/ Biomarker profiling



Overview:

1. Filtering in FNRL-regulated features → biomarkers

- Isolating DE-significant features with consistent fold-change direction across mutants and rescue in the complemented genotype.

2. Hierarchical Clustering Analysis (HCA) on biomarkers

- Clustering and ordering of biomarkers exhibiting similar expression profiles visualised as a heatmap.

3. k-means clustering on biomarkers

- Clustering biomarkers exhibiting similar expression profiles, visualised with line plots per cluster.

4. Linear Discriminant Analysis (LDA) using k-means cluster as a factor on biomarkers

- Project k-means-defined clusters into a supervised space to maximise between-cluster separation and further structure biomarkers with similar expression profiles.

5. Data engineering

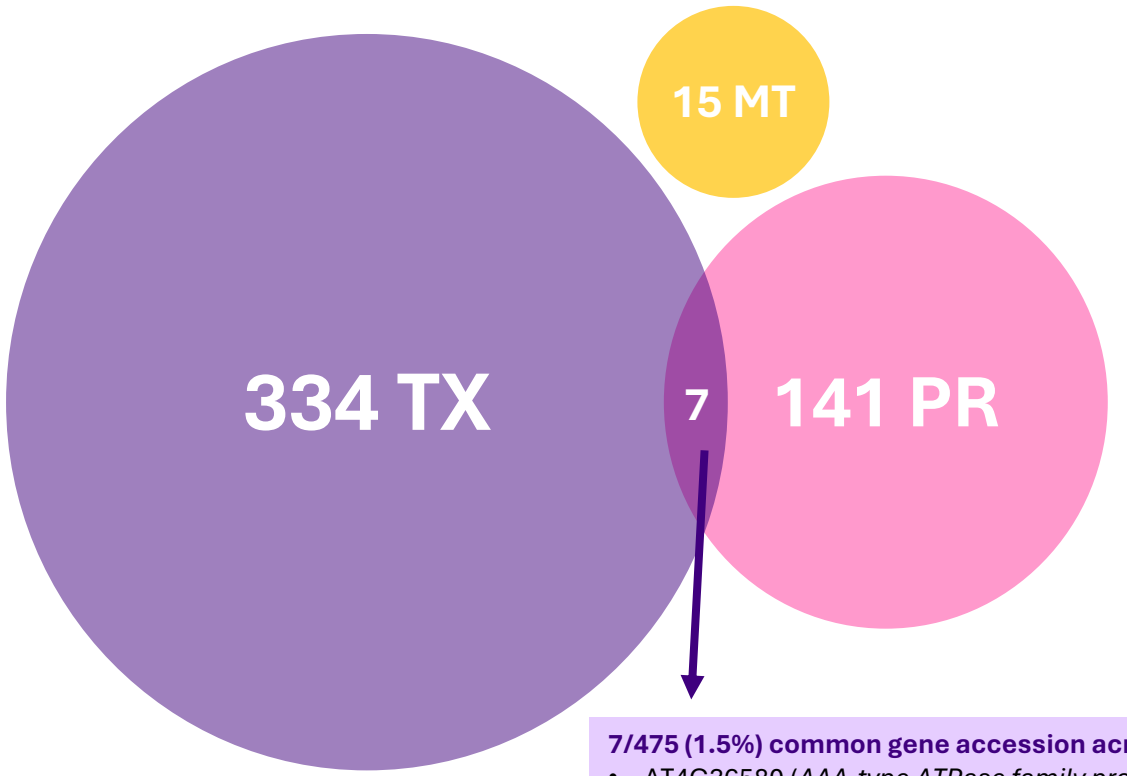
- All analysis results are captured as additional columns in the quantitative datasets for downstream modelling alongside phenotypic measurements.

NOTE: MT omitted from analysis.

7.1/ Filtering of biomarkers & Venn diagram

Filtering from 32723 features to 504 (1.5%) biomarkers

- **TX**: 341/23059 (1.48%) biomarkers → **further EDA & ML**
- **PR**: 148/9610 (1.54%) biomarkers → **further EDA & ML**
- **Intersection TX/PR**: Venn diagram below shows only 2 common AT IDs across transcriptomics and proteomics. **Highly complementary**.
- **MT**: 1/54 (27.78%) biomarker → **further EDA & ML**



	All features	Biomarkers	%
TX	23059	341	1.5
PR	9610	148	1.5
MT	54	15	27.8
All	32723	504	1.5

7/475 (1.5%) common gene accession across TX and PR:

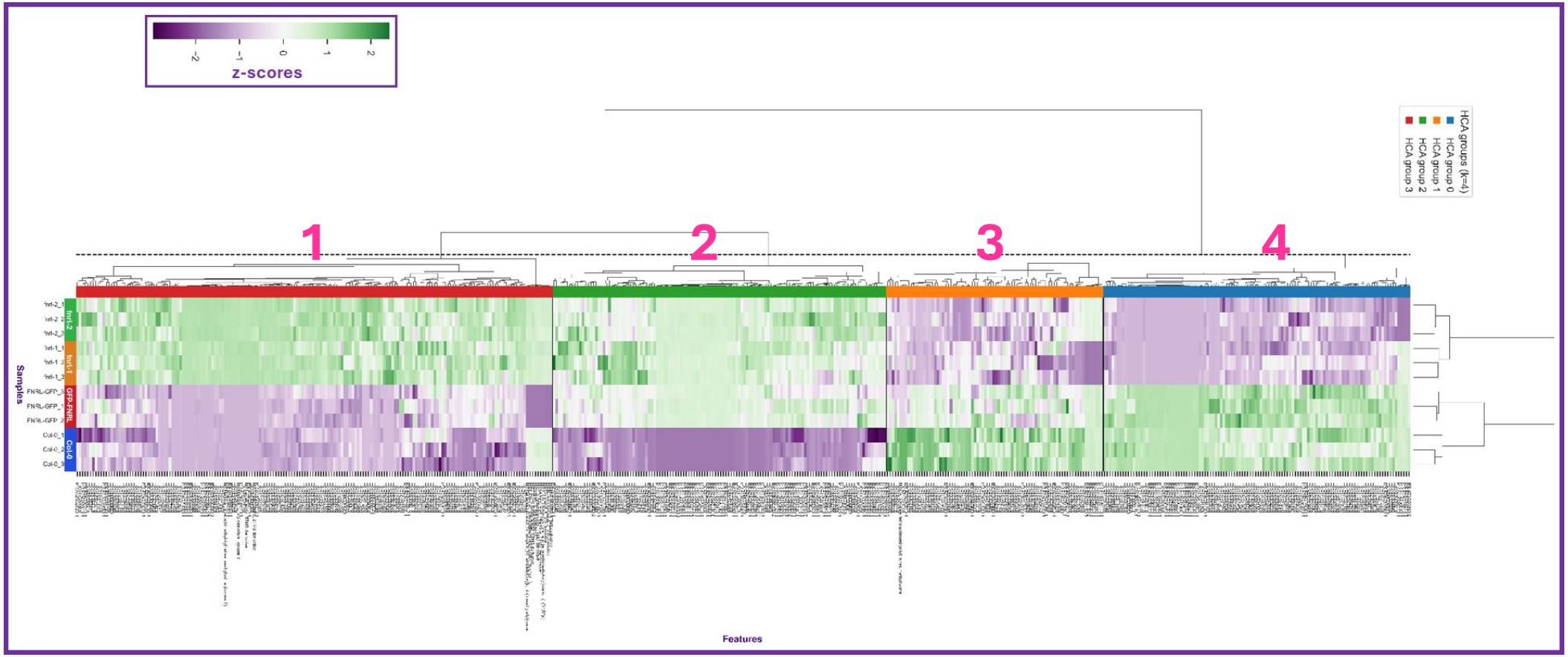
- AT4G36580 (*AAA-type ATPase family protein*)
- AT1G72330 (*alanine aminotransferase 2*)
- AT5G51440 (*HSP20-like chaperones superfamily protein*)
- AT1G21120 (*O-methyltransferase family protein*)
- AT2G18193 (*P-loop containing nucleoside triphosphate hydrolases superfamily protein*)
- AT4G39740 (*Thioredoxin superfamily protein*)
- AT5G13210 (*Uncharacterised conserved protein UCP015417, vWA*)

7.2/ HCA (unsupervised clustering)

HCA reveals 4 main expression profiles

Hierarchical clustering of 188 FNRL biomarkers exhibits **4 coherent expression modules** corresponding to FNRL-activated and FNRL-repressed programs, with strong allelic consistency and GFP-mediated rescue:

- **Profile 1:** Higher abundance in mutants; reduced in WT and GFP → These features are **repressed by FNRL in WT**. Loss of FNRL releases repression.
- **Profile 2:** Reduced abundance in mutants; partially restored by GFP → Another **activation-like pattern**, with **stronger dynamic range or delayed regulation**.
- **Profile 3:** Reduced abundance in WT; partially restored by GFP → Another **repression-like pattern**, with **stronger dynamic range or delayed regulation.**
- **Profile 4:** Higher abundance in WT and GFP; reduced in mutants → These features are **positively regulated (activated) by FNRL**.

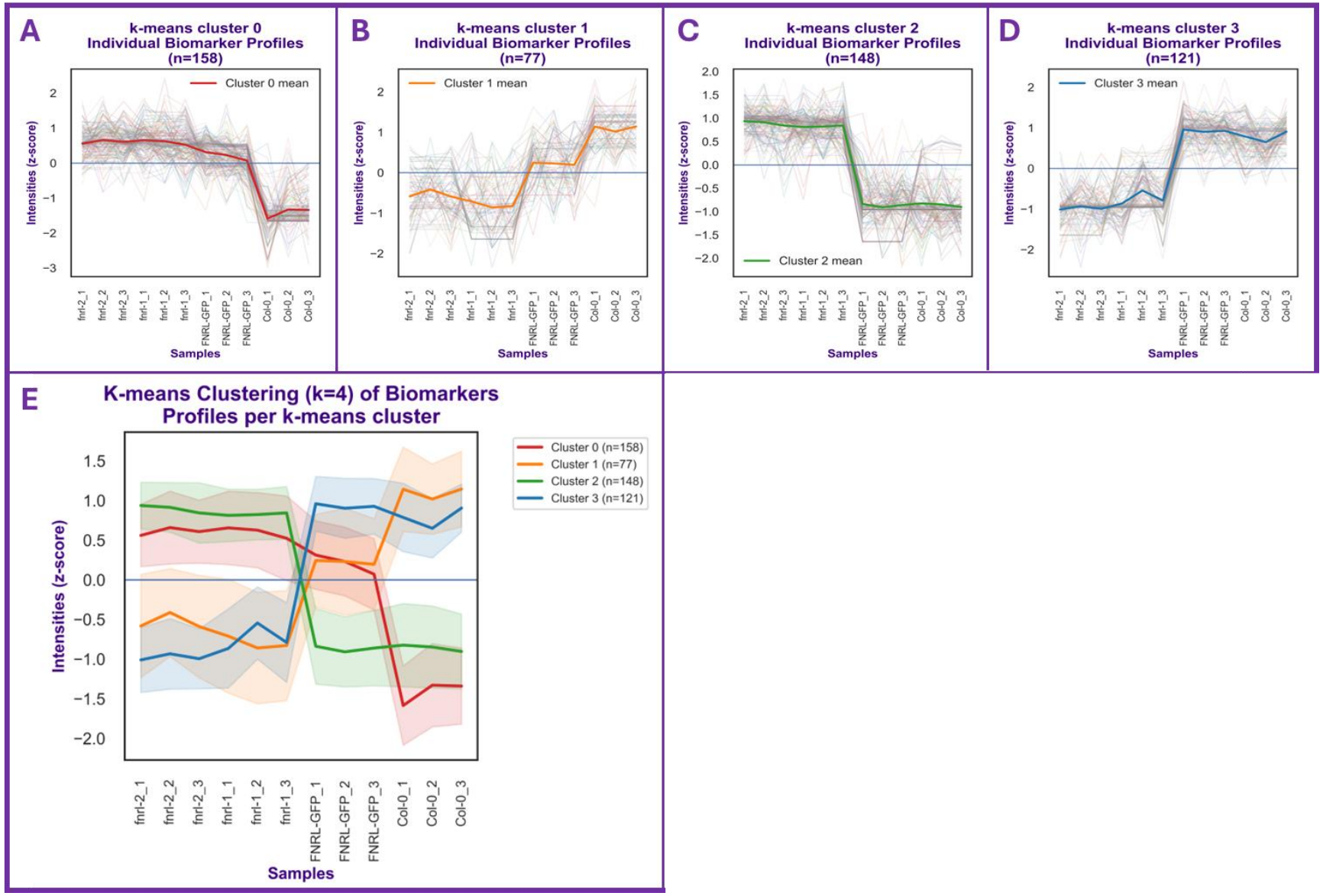


7.3/ k-means clustering with k=4

K-means highlights regulation directionality and rescue strength

Based on hierarchical clustering, four dominant expression programs were identified.
k-means clustering using k = 4 was therefore applied to formalise these patterns.
K-means clustering resolves FNRL core targets into **four regulatory programs** distinguished by **directionality** (**activated vs repressed**) and **degree of GFP-mediated rescue**:

- **Cluster 0: FNRL-represeed** targets with **partial GFP rescue** (158 features – 32%)
- **Cluster 1: FNRL-activated** targets with **partial GFP rescue** (77 features – 15%)
- **Cluster 2: FNRL-repressed** targets with **strong GFP rescue** (148 features – 29%)
- **Cluster 3:FNRL-activated** targets with **strong GFP rescue** (121features – 24%)



7.4/ LDA using k-means cluster

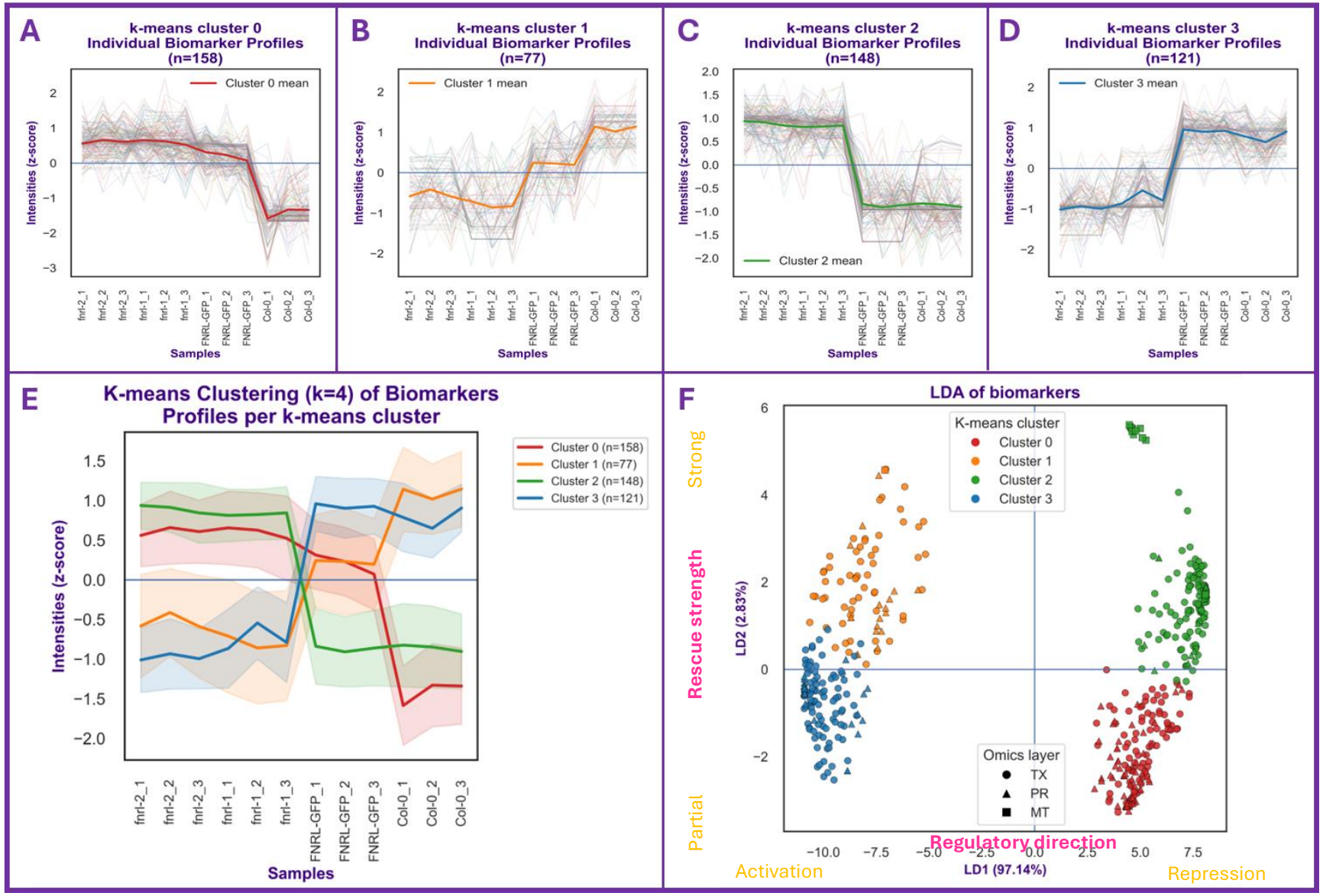
LDA further demonstrates regulation directionality and rescue strength

K-means cluster number is used as a discriminant factor in LDA.

LDA confirms clear separation of the four k-means clusters.

LD1 accounts for 97.14% of the discriminative separation and captures the primary **regulatory directionality**, separating FNRL-repressed from FNRL-activated targets.

LD2 accounts for 2.83% of the discriminative separation and highlights differences in the **extent of GFP-mediated rescue**, distinguishing clusters with strong versus partial functional complementation.



7.5/ Conceptual model of FNRL-mediated regulation

Core principle

FNRL functions as a bidirectional regulator controlling distinct target classes, with differential sensitivity to genetic rescue.

Biological implication

FNRL is **not a simple on/off regulator.**
It acts through **at least two mechanistic layers**:

- **Directionality of regulation** (activation vs repression)
- **Strength and completeness of regulatory control**

Partial rescue suggests:

- dosage sensitivity
- temporal mis-expression
- loss of regulatory fine-tuning
- requirement for native genomic context

Two orthogonal axes of regulation

Axis 1 — Regulatory direction (LD1): Does FNRL activate or repress the target?

- **FNRL-activated targets**
  ↓ expression in fnrl mutants, require FNRL for activation
- **FNRL-repressed targets**
  ↑ expression in fnrl mutants, repressed by FNRL in WT

This axis explains the **dominant separation** of targets (≈97% of discriminative structure).

Axis 2 — Rescue strength (LD2): How completely is regulation restored by FNRL-GFP?

- **Strong rescue:** GFP restores expression close to WT
  → strong functional complementation
- **Partial rescue:** GFP produces intermediate expression
  → incomplete recovery of regulatory control

This axis captures **qualitative differences in FNRL dependency** (~3%), not noise.

Summary

FNRL regulates gene expression along two orthogonal dimensions: regulatory direction and rescue sensitivity, revealing multiple functional classes of downstream targets rather than a single mode of action.

		Regulatory direction	
		activated	**repressed**
Rescue strength	**strong**	Canonical direct targets requiring FNRL	Robust repression mediated by FNRL
	partial	Activation requires precise regulation, stoichiometry, or chromatin context	Repression sensitive to expression level, timing, or protein context

		Regulatory direction	
		Activated (198)	**Repressed (306)**
Rescue strength	**Strong (269)**		
	Partial (235)		

7.6/ Feature engineering

Feature augmentation enables modelling and filtering

Each data analysis (PCA, DE, HCA, k-means, LDA) was captured as additional columns to each respective omics dataset.
These new variables can be used in down stream analyses for filtering and modelling purposes.

Initial omics dataset

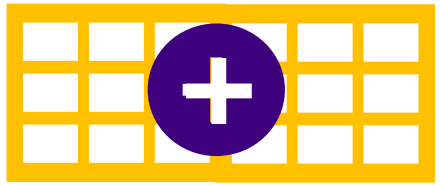
13 columns:

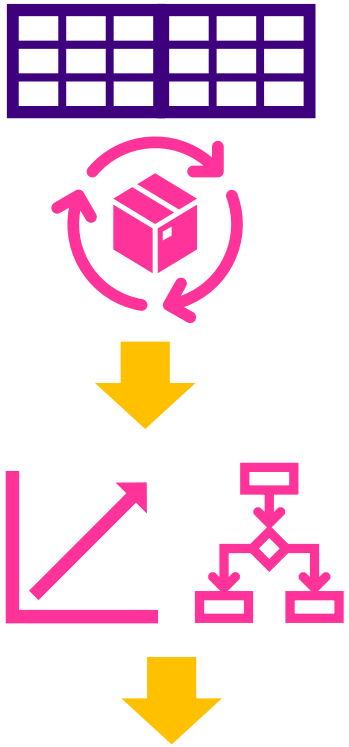
- 1 feature ID
- 12 sample clean abundances

Augmented omics dataset

58 columns:

- 1 feature ID
- 12 sample clean abundances
- 2 PCA (PC1, PC2)
- 12 DE (FC, p-values, FDR, biomarkers) for each pairwise test
- 12 volcano plots (log10FC, FDR, scores)
- 6 biomarker rules (consistent, rescue, biomarker, score, strength, integrated_score)
- 4 HCA (order, rank, cluster, biomarker)
- 5 k-means (cluster, size, distance, biomarker, strength)
- 4 LDA (LD1, LD2, radius, score)





8/ ML modelling

Overview:

1. Data needed
- Phenotypic measurements
- Augmented datasets

2. Model overview
- Predictors
- Targets
- Algorithms

3. Prediction results
- Metrics
- Top predictors

8.1/ Data required

Phenotypic measurements

- genotypes
- root_nitrate_content
- shoot_nitrate_content
- root_len_0.1mM
- root_len_1mM
- root_len_10mM
- root_len_60mM

sample_id	genotype	root_nitrate_content	shoot_nitrate_content	root_len_0.1mM	root_len_1mM	root_len_10mM	root_len_60mM
Col-0_1	Col-0	611.18	759.13	2.13	2.50	2.48	1.59
Col-0_2	Col-0	830.30	712.90	2.15	2.34	2.37	1.64
Col-0_3	Col-0	765.93	684.29	2.25	2.47	2.45	1.60
fnrl-1_1	fnrl-1	439.50	730.51	1.47	1.91	1.81	1.41
fnrl-1_2	fnrl-1	504.23	657.87	1.45	1.94	1.80	1.49
fnrl-1_3	fnrl-1	397.47	457.57	1.32	1.86	1.83	1.37
fnrl-2_1	fnrl-2	589.24	732.71	1.36	1.48	1.81	1.20
fnrl-2_2	fnrl-2	508.09	706.30	1.36	1.66	1.60	1.08
fnrl-2_3	fnrl-2	508.03	613.85	1.22	1.57	1.79	1.19
FNRL-GFP_1	FNRL-GFP	736.33	1005.66	2.04	2.61	2.66	1.75
FNRL-GFP_2	FNRL-GFP	825.49	574.23	1.96	2.53	2.56	1.82
FNRL-GFP_3	FNRL-GFP	868.64	657.87	2.18	2.50	2.56	1.86

Augmented datasets

- biomarkers only
- kmeans_cluster
- clean abundances

omics	ID1	clean_Col-0_1	clean_Col-0_2	clean_Col-0_3	clean_fnrl-1_1	clean_fnrl-1_2	clean_fnrl-1_3	clean_fnrl-2_1	clean_fnrl-2_2	clean_fnrl-2_3	clean_FNRL-G	clean_FNRL-G	clean_FNRL-G	kmeans_cluster
TX:	AT2G18193	9.3177518	9.1529504	9.198313	11.530097	11.654997	11.649193	11.799517	11.750896	11.801992	9.370797	9.2320883	9.5095819	1
TX:	AT1G68090	1.5049872	1.5049872	1.5049872	0	0	0	0	0	0	1.6093447	1.6093447	1.6093447	0
PR:	AT4G33110.1	10.831863	10.831863	10.831863	0	0	0	0	0	0	12.215978	12.126421	12.024202	0
PR:	AT1G23080.1	10.429873	10.429873	10.429873	0	0	0	0	0	0	11.033816	11.033816	11.033816	0
PR:	AT5G64070.1	11.374511	11.374511	11.374511	0	0	0	0	0	0	11.990769	11.990769	11.990769	0
PR:	AT1G06970.1	10.009706	10.009706	10.009706	0	0	0	0	0	0	10.429146	10.429146	10.429146	0
PR:	AT2G01070.1	9.6709403	9.6709403	9.6709403	0	0	0	0	0	0	11.24452	11.008941	10.657492	0
PR:	AT1G70070.1	11.674457	12.080291	11.905446	0	0	0	0	0	0	12.368973	12.368973	12.368973	0
PR:	AT3G23700.1	11.520949	11.816774	11.593774	0	0	0	0	0	0	11.902761	12.22577	12.29869	0
TX:	AT4G39340	0	0	0	1.1110531	1.1110531	1.1110531	1.0620504	1.0620504	1.0620504	1.0306074	1.0306074	1.0306074	2
TX:	AT3G58150	2.1900506	2.2494578	2.2200405	7.5597773	7.4596144	7.6799169	7.9490938	7.7114054	8.1229365	2.0416345	2.380661	2.54668	1
PR:	AT5G06240.1	12.255831	12.18197	12.322849	0	0	0	0	0	0	12.554626	12.5488	12.542926	0
PR:	AT2G31880.1	10.575718	10.575718	10.575718	0	0	0	0	0	0	10.844526	10.844526	10.844526	0
PR:	AT4G22000.1	0	0	0	11.749817	11.749817	11.749817	12.297549	12.297549	12.297549	0	0	0	1
PR:	AT5G49950.1	10.453783	10.453783	10.453783	0	0	0	0	0	0	10.6839	10.6839	10.6839	0
PR:	AT2G32350.1	0	0	0	11.854283	11.854283	11.854283	12.248417	12.248417	12.248417	0	0	0	1
PR:	AT1G61960.1	0	0	0	10.946793	10.946793	10.946793	11.741411	11.741411	11.741411	0	0	0	1
PR:	AT1G72120.1	10.986712	11.071255	10.890976	0	0	0	0	0	0	11.462732	11.272043	11.012437	0

8.2/ Modelling approach

Hypotheses

- Does FNRL-dependent molecular variation explain phenotypic outcomes?
- **Specifically, can variation in FNRL core targets quantitatively predict nitrate accumulation and root growth?**

Predictors (X)

Build a compact, biologically structured feature matrix at the sample level by aggregating FNRL core-target intensities by **k-means cluster (i.e. main expression patterns)**:

- Transcripts: TX_km0 … TX_km3
- Proteins: PR_km0 … PR_km3
- Metabolites: MT_km0 … MT_km3

This yields 12 predictors total, which is appropriate for the limited number of biological samples.
Each predictor represents the mean activity of a coherent FNRL-regulated module, ensuring stability and interpretability while reducing overfitting.

Targets (y)

- Root nitrate content, shoot nitrate content
- Root length at 0.1, 1, 10 and 60 mM nitrate

Inputs

Only FNRL core target features because the biological unit of regulation is the FNRL-controlled module, not individual genes or proteins.
Start with:

- transcripts (TX) only
- proteins (PR) only
- metabolites (MT)
- combined TX + PR + MT

Algorithms

- **Elastic Net / Lasso**: feature selection and interpretability (linear effects)
- **Random Forest Regressor**: captures non-linear genotype–phenotype relationships

Benefits

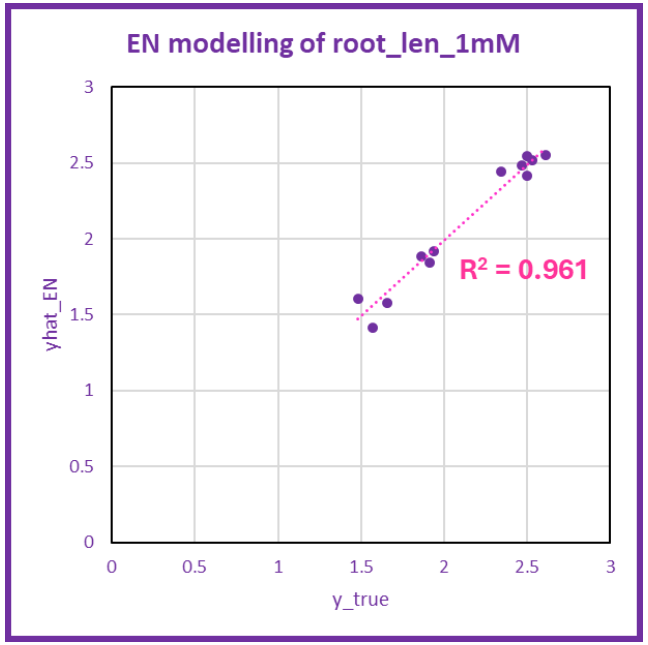
- Direct genotype → mechanism → phenotype link
- Models are both predictive and mechanistically interpretable

8.3/ Prediction results

Linear relationship between root length and main abundance profiles

Elastic Net linear model applied on biomarkers aggregated by main profiles (k-means clusters) robustly predict root growth phenotypes across nitrate regimes ($R^2 > 0.9$), indicating that **root developmental responses are largely encoded by coordinated transcriptional programs under FNRL control**.
In contrast, root nitrate accumulation is only partially predictable (maximum $R^2 = 0.66$), with Random Forest models outperforming Elastic Net. This suggests substantial **non-linear regulation, with protein-level clusters contributing most strongly, consistent with post-transcriptional control of nitrate handling**.
Nitrate content in shoots is not predictable from biomarker main profiles ($R^2 \leq 0$), supporting tissue-specific regulation and validating the root-focused experimental design. The scatterplot illustrates the high predictive accuracy of the biomarker-derived model for root length at 1 mM nitrate.
Together, these results support a model in which FNRL orchestrates discrete transcriptional modules driving root growth, while nitrate accumulation is modulated downstream through protein-level regulatory mechanisms.

target(y)	ElasticNet_R2	ElasticNet_RMSE	RF_R2	RF_RMSE
shoot_nitrate_content	-0.29	140.72	-0.20	135.64
root_nitrate_content	0.52	110.60	0.66	93.08
root_len_0.1mM	0.89	0.13	0.92	0.11
root_len_60mM	0.91	0.07	0.85	0.10
root_len_10mM	0.94	0.09	0.88	0.13
root_len_1mM	0.96	0.08	0.85	0.15



NOTE: The original data was wrangled so that it could be aligned with omics samples.
This is acceptable for nitrate content data in roots but probably not for nitrate content in shoots.
For the **root length** measurements, we should **limit to the nitrate regime that was applied to the 12 omics samples** used in this study (Ghaz, which nitrate concentration was used?).

9/ Data mining

Annotation and pathway mapping using bioinformatics:

1/ feature annotation retrieval

- **TAIR10** (https://www.arabidopsis.org/) **& biomarkers descriptions**
- **UniprotKB** (https://www.uniprot.org/)
- **PeptideAtlas** (https://peptideatlas.org/)
- **MetaboAnalyst** (https://www.metaboanalyst.ca/)

2/ Functional Annotation and Gene Ontology Enrichment

- **GO** (https://geneontology.org/) & **AMiGO**

3/ Pathway Mapping and Metabolic Contextualisation

- **KEGG** (https://www.kegg.jp/kegg/)
- **PlantReactome** (https://plantreactome.gramene.org/PathwayBrowser/#/SPECIES=170905&DTAB=AN&ANALYSIS=MjAyNjAxMTkwMTA3MThfMTQ%253D)
- **MetaboAnalyst** (https://www.metaboanalyst.ca/)

4/ Protein–Protein Interaction Network Construction

- **STRING** (https://string-db.org/) & **Cytoscape** (https://cytoscape.org/)

5/ Multi-Omics Integration and Biochemistry

- **PMN AraCyc** (Pathway Tools) (https://pmn.plantcyc.org/organism-summary?object=ARA)
- **PaintOmics** (https://paintomics.uv.es/)

NOTE: FNRL gene product (PR:AT1G15140.1) is one of the biomarkers, but not at the transcript level (TX:AT1G15140).

9.1.1/ TX+PR Biomarker TAIR10 annotations (top feat.)

Mapping of TX + PR biomarkers (using ID1)

om	ID1	FNRL_inte	kmea	direction_	Short_description
PR:	AT1G45616.1	50767.49	3	act_full	receptor like protein 6
PR:	AT4G11980.1	42687.88	3	act_full	nudix hydrolase homolog 14
PR:	AT5G64070.1	36710.95	3	act_full	phosphatidylinositol 4-OH kinase beta1
PR:	AT3G16030.1	34559.04	3	act_full	lectin protein kinase family protein
PR:	AT2G31880.1	32700.89	3	act_full	Leucine-rich repeat protein kinase family protein
PR:	AT5G49950.1	32062.80	3	act_full	alpha/beta-Hydrolases superfamily protein
PR:	AT4G33110.1	31005.75	3	act_full	S-adenosyl-L-methionine-dependent methyltransferases superf
PR:	AT1G23080.1	30744.96	3	act_full	Auxin efflux carrier family protein
PR:	AT2G36570.1	29901.67	3	act_full	Leucine-rich repeat protein kinase family protein
PR:	AT1G80640.1	29621.27	3	act_full	Protein kinase superfamily protein
PR:	AT1G06970.1	28798.73	3	act_full	cation/hydrogen exchanger 14
PR:	AT2G01070.1	24288.27	3	act_full	Lung seven transmembrane receptor family protein
PR:	AT3G54270.1	13585.43	3	act_full	sucrose-6F-phosphate phosphohydrolase family protein
PR:	AT3G55380.2	11067.52	3	act_full	ubiquitin-conjugating enzyme 14
TX:	AT1G62190	6575.41	3	act_full	Kua-ubiquitin conjugating enzyme hybrid localisation
PR:	AT5G06240.1	2595.39	3	act_full	embryo defective 2735
PR:	AT5G47550.1	2240.97	3	act_full	Cystatin/monellin superfamily protein
PR:	AT1G72120.1	1941.19	3	act_full	Major facilitator superfamily protein
PR:	AT3G23700.1	1929.07	3	act_full	Nucleic acid-binding proteins superfamily
PR:	AT1G70070.1	1896.93	3	act_full	DEAD/DEAH box helicase, putative
PR:	AT1G71500.1	1741.01	3	act_full	Rieske (2Fe-2S) domain-containing protein
PR:	AT4G30340.1	1487.80	3	act_full	diacylglycerol kinase 7

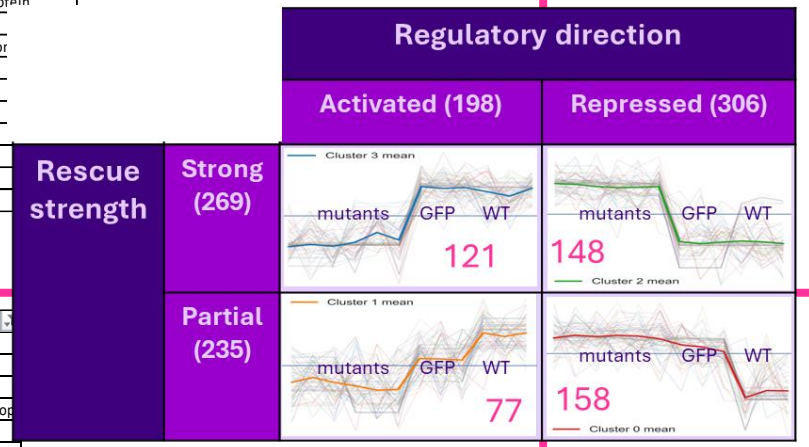
activation

100/121 annotated
21/121 (17%) unannotated

repression

100/148 annotated
48/148 (32%) un.

om	ID1	FNRL_inte	kmea	direction_	Short_description
PR:	AT2G32350.1	43570.51	2	rep_full	Ubiquitin-like superfamily protein
PR:	AT1G61960.1	38606.59	2	rep_full	Mitochondrial transcription termination factor family protein
PR:	AT1G16650.1	36769.26	2	rep_full	S-adenosyl-L-methionine-dependent methyltransferases superf
PR:	AT3G45720.1	30418.32	2	rep_full	Major facilitator superfamily protein
PR:	AT4G26910.1	26511.77	2	rep_full	Dihydrolipoamide succinyltransferase
PR:	AT4G14370.1	26033.25	2	rep_full	Disease resistance protein (TIR-NBS-LRR class) family
PR:	AT1G05700.1	21757.17	2	rep_full	Leucine-rich repeat transmembrane protein kinase protein
PR:	AT1G63250.1	20603.08	2	rep_full	DEA(D/H)-box RNA helicase family protein
PR:	AT2G07750.1	20603.08	2	rep_full	DEA(D/H)-box RNA helicase family protein
PR:	AT4G36580.1	2223.30	2	rep_full	AAA-type ATPase family protein
PR:	AT5G04440.1	1794.37	2	rep_full	Protein of unknown function (DUF1997)
PR:	AT3G52970.2	1631.94	2	rep_full	cytochrome P450, family 76, subfamily G, polypeptide 1
TX:	AT5G09570	1533.34	2	rep_full	Cox19-like CHCH family protein
TX:	AT1G76190	1156.52	2	rep_full	SAUR-like auxin-responsive protein family
TX:	AT2G42930	1135.97	2	rep_full	Carbohydrate-binding X8 domain superfamily protein
TX:	AT5G67120	1122.16	2	rep_full	RING/U-box superfamily protein
TX:	AT3G44840	632.87	2	rep_full	S-adenosyl-L-methionine-dependent methyltransferases superf
TX:	AT5G52940	497.48	2	rep_full	Protein of unknown function (DUF295)
TX:	AT1G19320	424.83	2	rep_full	Pathogenesis-related thaumatin superfamily protein



om	ID1	FNRL_inte	kmea	direction_	Short_description
PR:	AT1G68150.1	13281.62	1	act_part	WRKY DNA-binding protein 9
PR:	AT1G37150.2	11899.75	1	act_part	holocarboxylase synthetase 2
PR:	AT1G09140.1	10210.21	1	act_part	SERINE-ARGININE PROTEIN 30
PR:	AT2G46300.1	8194.21	1	act_part	Late embryogenesis abundant (LEA) hydroxyproline-rich glycop
TX:	AT5G60470	822.43	1	act_part	C2H2 and C2HC zinc fingers superfamily protein
PR:	AT1G15140.1	767.03	1	act_part	FAD/NAD(P)-binding oxidoreductase
PR:	AT2G34710.1	659.80	1	act_part	Homeobox-leucine zipper family protein / lipid-binding START do
TX:	AT1G09665	518.43	1	act_part	Toll-Interleukin-Resistance (TIR) domain family protein
TX:	AT5G58170	403.55	1	act_part	SHV3-like 5
TX:	AT5G36240	272.62	1	act_part	zinc knuckle (CCHC-type) family protein
TX:	AT5G09640	236.13	1	act_part	serine carboxypeptidase-like 19
TX:	AT5G46795	202.74	1	act_part	microspore-specific promoter 2
TX:	AT5G38430	160.01	1	act_part	Ribulose bisphosphate carboxylase (small chain) family protein
TX:	AT3G10510	141.30	1	act_part	Galactose oxidase/kelch repeat superfamily protein
TX:	AT5G46660	103.45	1	act_part	protein kinase C-like zinc finger protein
TX:	AT5G37760	79.10	1	act_part	Chaperone DnaJ-domain superfamily protein
PR:	AT5G10810.1	22.77	1	act_part	enhancer of rudimentary protein, putative
TX:	AT4G00180	15.56	1	act_part	Plant-specific transcription factor YABBY family protein
PR:	AT1G77020.1	10.89	1	act_part	DNAJ heat shock N-terminal domain-containing protein
PR:	AT5G25220.1	10.40	1	act_part	KNOTTED1-like homeobox gene 3
PR:	AT5G11060.1	10.40	1	act_part	KNOTTED1-like homeobox gene 4
TX:	AT5G53660	10.33	1	act_part	growth-regulating factor 7

FNRL protein

59/77 annotated
18/77 (23%) unannotated

135/158 annotated
23/158 (15%) un.

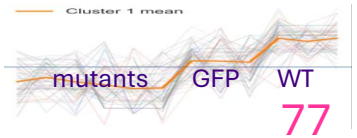
om	ID1	FNRL_inte	kmea	direction_	Short_description
PR:	AT1G65040.2	3666.89	0	rep_part	RING/U-box superfamily protein
PR:	AT4G32208.1	2958.64	0	rep_part	heat shock protein 70 (Hsp 70) family protein
PR:	AT3G19960.2	2694.64	0	rep_part	myosin 1
PR:	AT3G10500.1	2587.36	0	rep_part	NAC domain containing protein 53
PR:	AT1G02816.1	2431.50	0	rep_part	Protein of unknown function, DUF538
PR:	AT5G52540.1	1320.34	0	rep_part	Protein of unknown function (DUF819)
PR:	AT3G47110.1	965.11	0	rep_part	Leucine-rich repeat protein kinase family protein
PR:	AT3G02580.1	845.14	0	rep_part	sterol 1
PR:	AT1G14390.1	843.99	0	rep_part	Leucine-rich repeat protein kinase family protein
PR:	AT5G53100.1	479.25	0	rep_part	NAD(P)-binding Rossmann-fold superfamily protein
TX:	ATMG00860	343.32	0	rep_part	DNA/RNA polymerases superfamily protein
PR:	AT1G61475.1	307.61	0	rep_part	ATP binding;protein kinases
TX:	AT5G58050	306.20	0	rep_part	SHV3-like 4
TX:	AT1G14800	302.42	0	rep_part	Nucleic acid-binding, OB-fold-like protein
TX:	AT1G24010	232.98	0	rep_part	Polyketide cyclase/dehydrase and lipid transport superfamily p
TX:	AT4G38560	198.53	0	rep_part	Arabidopsis phospholipase-like protein (PEARLI 4) family
TX:	AT5G53742	184.17	0	rep_part	Protein of unknown function (DUF1278)
TX:	AT2G17860	178.08	0	rep_part	Pathogenesis-related thaumatin superfamily protein
TX:	AT5G19360	177.26	0	rep_part	calcium-dependent protein kinase 34

394/504 (78%) annotated biomarkers

9.1.2/ TX+PR Biomarkers – all 394 annotated features

Activation

121



77

om	ID1	FNRL_inte	kmea	direction	Short_description
PR:	AT1G45616.1	50767.49	3	act_full	receptor like protein 6
PR:	AT4G11980.1	42687.88	3	act_full	nudix hydrolase homolog 14
PR:	AT5G64070.1	36710.95	3	act_full	phosphatidylinositol 4-OH kinase beta1
PR:	AT3G16030.1	34559.04	3	act_full	lectin protein kinase family protein
PR:	AT2G31880.1	32700.89	3	act_full	Leucine-rich repeat protein kinase family protein
PR:	AT5G49950.1	32062.80	3	act_full	alpha/beta-Hydrolases superfamily protein
PR:	AT4G33110.1	31005.75	3	act_full	S-adenosyl-L-methionine-dependent methyltransferases superf
PR:	AT1G23080.1	30744.96	3	act_full	Auxin efflux carrier family protein
PR:	AT2G36570.1	29901.67	3	act_full	Leucine-rich repeat protein kinase family protein
PR:	AT1G80640.1	29621.27	3	act_full	Protein kinase superfamily protein
PR:	AT1G06970.1	28798.73	3	act_full	cation/hydrogen exchanger 14
PR:	AT2G01070.1	24288.27	3	act_full	Lung seven transmembrane receptor family protein
PR:	AT3G54270.1	13585.43	3	act_full	sucrose-6F-phosphate phosphohydrolase family protein
PR:	AT3G55380.2	11067.52	3	act_full	ubiquitin-conjugating enzyme 14
TX:	AT1G62190	6575.41	3	act_full	Kua-ubiquitin conjugating enzyme hybrid localisation domain
PR:	AT5G06240.1	2595.39	3	act_full	embryo defective 2735
PR:	AT5G47550.1	2240.97	3	act_full	Cystatin/monellin superfamily protein
PR:	AT1G72120.1	1941.19	3	act_full	Major facilitator superfamily protein
PR:	AT3G23700.1	1929.07	3	act_full	Nucleic acid-binding proteins superfamily
PR:	AT1G70070.1	1896.93	3	act_full	DEAD/DEAH box helicase, putative
PR:	AT1G71500.1	1741.01	3	act_full	Rieske (2Fe-2S) domain-containing protein
PR:	AT4G30340.1	1487.80	3	act_full	diacylglycerol kinase 7
TX:	AT1G68090	632.38	3	act_full	annexin 5
PR:	AT1G55600	591.34	3	act_full	WRKY DNA-binding protein 10
PR:	AT1G18370.1	589.81	3	act_full	ATP binding microtubule motor family protein
TX:	AT1G37000	572.83	3	act_full	Beta-galactosidase related protein
PR:	AT3G06630.1	517.54	3	act_full	protein kinase family protein
PR:	AT1G21840.1	404.44	3	act_full	urease accessory protein F
TX:	AT5G12020	265.23	3	act_full	17.6 kDa class II heat shock protein
TX:	AT1G18010	143.43	3	act_full	Major facilitator superfamily protein
TX:	AT2G14690	137.27	3	act_full	Glycosyl hydrolase superfamily protein
TX:	ATMG00290	50.24	3	act_full	mitochondrial ribosomal protein S4
PR:	AT4G19180.1	43.50	3	act_full	GDA1/CD39 nucleoside phosphatase family protein
TX:	AT2G44220	34.82	3	act_full	Protein of Unknown Function (DUF239)
TX:	AT2G41940	29.21	3	act_full	zinc finger protein 8
TX:	AT2G47770	24.06	3	act_full	TSPO(outer membrane tryptophan-rich sensory protein)-related
TX:	AT1G19610	23.85	3	act_full	Arabidopsis defensin-like protein
TX:	AT3G46190	21.28	3	act_full	TRAF-like family protein
TX:	AT4G26560	20.92	3	act_full	calcineurin B-like protein 7
TX:	AT4G10530	20.17	3	act_full	Subtilase family protein
TX:	AT5G47450	19.19	3	act_full	tonoplast intrinsic protein 2;3
TX:	AT5G54700	18.81	3	act_full	Ankyrin repeat family protein
TX:	AT2G19900	17.08	3	act_full	NADP-malic enzyme 1
TX:	AT3G51420	15.80	3	act_full	strictosidine synthase-like 4
TX:	AT1G04445	15.64	3	act_full	C2H2-like zinc finger protein
TX:	AT2G22990	12.55	3	act_full	sinapoylglucose 1
TX:	AT5G23660	12.25	3	act_full	homolog of Medicago truncatula MTN3
TX:	AT2G05400	11.05	3	act_full	Ubiquitin-specific protease family C19-related protein
TX:	AT3G47040	10.83	3	act_full	Glycosyl hydrolase family protein
TX:	AT4G28040	10.82	3	act_full	nodulin MtN21 /EamA-like transporter family protein
TX:	AT2G22860	10.77	3	act_full	phytosulfokine 2 precursor
TX:	AT2G45360	10.47	3	act_full	Protein of unknown function (DUF1442)
TX:	AT5G07500	9.97	3	act_full	Zinc finger C-x8-C-x5-C-x3-H type family protein
TX:	AT3G46090	9.53	3	act_full	C2H2 and C2HC zinc fingers superfamily protein
TX:	AT3G48740	9.23	3	act_full	Nodulin MtN3 family protein
TX:	AT4G09110	8.84	3	act_full	RING/U-box superfamily protein
TX:	AT2G37180	8.67	3	act_full	Aquaporin-like superfamily protein
TX:	AT5G24220	8.63	3	act_full	Lipase class 3-related protein
TX:	AT3G10450	8.61	3	act_full	serine carboxypeptidase-like 7
TX:	AT4G12550	8.38	3	act_full	Auxin-Induced in Root cultures 1
PR:	AT4G15360.1	8.10	3	act_full	cytochrome P450, family 705, subfamily A, polypeptide 3
TX:	AT5G60060	8.00	3	act_full	Protein of unknown function (DUF295)
TX:	AT4G15400	7.67	3	act_full	HXXXD-type acyl-transferase family protein
TX:	AT4G20370	7.65	3	act_full	PEBP (phosphatidylethanolamine-binding protein) family protei
TX:	AT5G25250	7.53	3	act_full	SPFH/Band 7/PHB domain-containing membrane-associated pro
TX:	AT3G14225	7.46	3	act_full	GDSL-motif lipase 4
TX:	AT5G07650	7.23	3	act_full	Actin-binding FH2 protein
TX:	AT3G47050	6.87	3	act_full	Glycosyl hydrolase family protein
TX:	AT2G36110	6.67	3	act_full	Polynucleotidyl transferase, ribonuclease H-like superfamily pro
TX:	AT2G20880	6.27	3	act_full	Integrase-type DNA-binding superfamily protein
TX:	AT2G14210	6.22	3	act_full	AGAMOUS-like 44
TX:	AT1G64710	6.12	3	act_full	GroES-like zinc-binding dehydrogenase family protein
TX:	AT1G07550	5.83	3	act_full	Leucine-rich repeat protein kinase family protein
TX:	AT2G02250	5.61	3	act_full	phloem protein 2-B2
TX:	AT2G05070	5.50	3	act_full	photosystem II light harvesting complex gene 2.2

om	ID1	FNRL_inte	kmea	direction	Short_description
PR:	AT1G68150.1	13281.62	1	act_part	WRKY DNA-binding protein 9
PR:	AT1G37150.2	11899.75	1	act_part	holocarboxylase synthetase 2
PR:	AT1G09140.1	10210.21	1	act_part	SERINE-ARGININE PROTEIN 30
PR:	AT2G46300.1	8194.21	1	act_part	Late embryogenesis abundant (LEA) hydroxyproline-rich glycopr
TX:	AT5G60470	822.43	1	act_part	C2H2 and C2HC zinc fingers superfamily protein
PR:	AT1G15140.1	767.03	1	act_part	FAD/NAD(P)-binding oxidoreductase
PR:	AT2G34710.1	659.80	1	act_part	Homeobox-leucine zipper family protein / lipid-binding START do
TX:	AT1G09665	518.43	1	act_part	Toll-Interleukin-Resistance (TIR) domain family protein
TX:	AT5G58170	403.55	1	act_part	SHV3-like 5
TX:	AT5G36240	272.62	1	act_part	zinc knuckle (CCHC-type) family protein
TX:	AT5G09640	236.13	1	act_part	serine carboxypeptidase-like 19
TX:	AT5G46795	202.74	1	act_part	microspore-specific promoter 2
TX:	AT5G38430	160.01	1	act_part	Ribulose bisphosphate carboxylase (small chain) family protein
TX:	AT3G10510	141.30	1	act_part	Galactose oxidase/kelch repeat superfamily protein
TX:	AT5G46660	103.45	1	act_part	protein kinase C-like zinc finger protein
TX:	AT5G37760	79.10	1	act_part	Chaperone DnaJ-domain superfamily protein
PR:	AT5G10810.1	22.77	1	act_part	enhancer of rudimentary protein, putative
TX:	AT4G00180	15.56	1	act_part	Plant-specific transcription factor YABBY family protein
PR:	AT1G77020.1	10.89	1	act_part	DNAJ heat shock N-terminal domain-containing protein
PR:	AT5G25220.1	10.40	1	act_part	KNOTTED1-like homeobox gene 3
PR:	AT5G11060.1	10.40	1	act_part	KNOTTED1-like homeobox gene 4
TX:	AT5G53660	10.33	1	act_part	growth-regulating factor 7
TX:	AT1G28450	9.52	1	act_part	AGAMOUS-like 58
TX:	AT3G28040	9.40	1	act_part	RING/U-box superfamily protein
PR:	AT4G11810.1	8.44	1	act_part	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) doma
PR:	AT4G22990.2	8.44	1	act_part	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) doma
TX:	AT2G35130	8.25	1	act_part	Tetratricopeptide repeat (TPR)-like superfamily protein
TX:	AT3G02620	7.36	1	act_part	Plant stearoyl-acyl-carrier-protein desaturase family protein
TX:	AT2G44920	7.26	1	act_part	Tetratricopeptide repeat (TPR)-like superfamily protein
TX:	AT1G78450	7.04	1	act_part	SOUL heme-binding family protein
TX:	AT5G45240	6.49	1	act_part	Disease resistance protein (TIR-NBS-LRR class)
TX:	AT2G05540	6.28	1	act_part	Glycine-rich protein family
TX:	AT5G47280	6.23	1	act_part	ADR1-like 3
TX:	AT5G07400.1	4.95	1	act_part	forkhead-associated domain-containing protein / FHA domain-c
PR:	AT2G47490.1	4.88	1	act_part	NAD+ transporter 1
TX:	AT4G30880	4.25	1	act_part	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S alb
PR:	AT3G05790.1	4.24	1	act_part	Ion protease 4
TX:	AT1G78720	3.83	1	act_part	SecY protein transport family protein
TX:	AT1G58330	3.64	1	act_part	transcription factor-related
PR:	AT3G25810.1	3.61	1	act_part	Terpenoid cyclases/Protein prenyltransferases superfamily pro
TX:	AT5G24080	3.19	1	act_part	Protein kinase superfamily protein
TX:	AT2G18050	3.04	1	act_part	histone H1-3
TX:	AT3G06100	2.68	1	act_part	NOD26-like intrinsic protein 7;1
TX:	AT5G45120	2.60	1	act_part	Eukaryotic aspartyl protease family protein
TX:	AT5G28237	2.45	1	act_part	Pyridoxal-5'-phosphate-dependent enzyme family protein
TX:	AT4G14368	2.35	1	act_part	Regulator of chromosome condensation (RCC1) family protein
TX:	AT3G13710	2.28	1	act_part	prenylated RAB acceptor 1.F4
TX:	AT1G14686	2.17	1	act_part	ENTH/ANTH/VHS superfamily protein
TX:	AT1G00580	2.05	1	act_part	photosystem II reaction center protein E
TX:	AT5G09930	1.88	1	act_part	ABC transporter family protein
TX:	AT4G15680	1.79	1	act_part	Thioredoxin superfamily protein
TX:	AT1G50760	1.77	1	act_part	Aminotransferase-like, plant mobile domain family protein
TX:	AT5G57810	1.05	1	act_part	tetraspanin15
TX:	AT1G64290	0.65	1	act_part	F-box protein-related
TX:	AT1G52650	0.48	1	act_part	F-box/RNI-like superfamily protein
PR:	AT1G65700.3	0.47	1	act_part	Small nuclear ribonucleoprotein family protein
TX:	AT1G12130	0.28	1	act_part	Flavin-binding monooxygenase family protein
TX:	AT3G61230	0.11	1	act_part	GATA type zinc finger transcription factor family protein
TX:	AT4G18220	0.09	1	act_part	Drug/metabolite transporter superfamily protein

Repression

148

158

om	ID1	FNRL_inte	kmea	direction	Short_description
PR:	AT2G32350.1	43570.51	2	rep_full	Ubiquitin-like superfamily protein
PR:	AT1G61960.1	38606.59	2	rep_full	Mitochondrial transcription termination factor family protein
PR:	AT1G16650.1	36769.26	2	rep_full	S-adenosyl-L-methionine-dependent methyltransferases superf
PR:	AT3G45720.1	30418.32	2	rep_full	Major facilitator superfamily protein
PR:	AT4G26910.1	26511.77	2	rep_full	Dihydrolipoamide succinyltransferase
PR:	AT4G14370.1	26033.25	2	rep_full	Disease resistance protein (TIR-NBS-LRR class) family
PR:	AT1G05700.1	21757.17	2	rep_full	Leucine-rich repeat transmembrane protein kinase protein
PR:	AT1G63250.1	20603.08	2	rep_full	DEA(D/H)-box RNA helicase family protein
PR:	AT2G07750.1	20603.08	2	rep_full	DEA(D/H)-box RNA helicase family protein
PR:	AT4G36580.1	2223.30	2	rep_full	AAA-type ATPase family protein
PR:	AT5G04440.1	1794.37	2	rep_full	Protein of unknown function (DUF1997)
PR:	AT3G52970.2	1631.94	2	rep_full	cytochrome P450, family 76, subfamily G, polypeptide 1
TX:	AT5G09570	1533.34	2	rep_full	Cox19-like CHCH family protein
TX:	AT1G76190	1156.52	2	rep_full	SAUR-like auxin-responsive protein family
TX:	AT2G42930	1135.97	2	rep_full	Carbohydrate-binding X8 domain superfamily protein
TX:	AT5G67120	1122.16	2	rep_full	RING/U-box superfamily protein
TX:	AT3G44840	632.87	2	rep_full	S-adenosyl-L-methionine-dependent methyltransferases superf
TX:	AT5G52940	497.48	2	rep_full	Protein of unknown function (DUF295)
TX:	AT1G19320	424.83	2	rep_full	Pathogenesis-related thaumatin superfamily protein
TX:	AT3G58150	387.03	2	rep_full	Optic atrophy 3 protein (OPA3)
TX:	AT1G21250	386.76	2	rep_full	cell wall-associated kinase
TX:	AT2G20800	365.95	2	rep_full	NAD(P)H dehydrogenase B4
TX:	AT3G04290	363.76	2	rep_full	Li-tolerant lipase 1
TX:	AT5G52930	308.31	2	rep_full	Protein of unknown function (DUF295)
TX:	AT1G09080	276.52	2	rep_full	Heat shock protein 70 (Hsp 70) family protein
TX:	AT5G09940	217.20	2	rep_full	Protein of unknown function (DUF1635)
TX:	AT2G36800	130.40	2	rep_full	don-glucosyltransferase 1
TX:	AT2G04050	123.16	2	rep_full	MATE efflux family protein
TX:	AT1G53480	118.32	2	rep_full	mto 1 responding down 1
TX:	AT2G18193	91.71	2	rep_full	P-loop containing nucleoside triphosphate hydrolases superfam
TX:	AT2G47520	89.18	2	rep_full	Integrase-type DNA-binding superfamily protein
TX:	AT5G54560	88.91	2	rep_full	Protein of unknown function (DUF295)
TX:	AT5G54450	88.44	2	rep_full	Protein of unknown function (DUF295)
TX:	AT5G54550	87.01	2	rep_full	Protein of unknown function (DUF295)
TX:	AT1G07180	71.22	2	rep_full	alternative NAD(P)H dehydrogenase 1
TX:	AT4G25930	68.69	2	rep_full	Protein of unknown function (DUF295)
TX:	AT1G18830	60.41	2	rep_full	Transducin/WD40 repeat-like superfamily protein
TX:	AT4G21490	55.74	2	rep_full	NAD(P)H dehydrogenase B3
TX:	AT5G54100	55.38	2	rep_full	SPFH/Band 7/PHB domain-containing membrane-associated pro
TX:	AT3G01600	50.74	2	rep_full	NAC domain containing protein 44
TX:	AT4G39740	33.55	2	rep_full	Thioredoxin superfamily protein
TX:	AT5G13210	27.15	2	rep_full	Uncharacterised conserved protein UCP015417, vWA
TX:	AT4G25010	24.47	2	rep_full	Nodulin MtN3 family protein
TX:	AT3G08970	23.26	2	rep_full	DNAJ heat shock N-terminal domain-containing protein
TX:	AT5G51440	21.20	2	rep_full	HSP20-like chaperones superfamily protein
TX:	AT5G14490	20.14	2	rep_full	NAC domain containing protein 85
TX:	AT2G03290	18.10	2	rep_full	emp24/gp25L/p24 family/GOLD family protein
PR:	AT2G18193.1	18.04	2	rep_full	P-loop containing nucleoside triphosphate hydrolases superfam
TX:	AT4G27585	17.20	2	rep_full	SPFH/Band 7/PHB domain-containing membrane-associated pro
TX:	AT3G60540	16.71	2	rep_full	Preprotein translocase Sec, Sec61-beta subunit protein
TX:	AT5G66150	16.04	2	rep_full	Glycosyl hydrolase family 38 protein
TX:	AT5G64060	15.81	2	rep_full	NAC domain containing protein 103
TX:	ATMG01010	15.42	2	rep_full	Unknown conserved protein
TX:	AT1G32870	12.22	2	rep_full	NAC domain protein 13
TX:	AT4G10440	12.17	2	rep_full	S-adenosyl-L-methionine-dependent methyltransferases superf
PR:	AT2G18190.1	11.84	2	rep_full	P-loop containing nucleoside triphosphate hydrolases superfam
TX:	AT1G72330	11.06	2	rep_full	alanine aminotransferase 2
TX:	AT1G05680	10.78	2	rep_full	Uridine diphosphate glycosyltransferase 74E2
TX:	AT1G24095	10.65	2	rep_full	Putative thiol-disulphide oxidoreductase DCC
PR:	AT1G72330.3	10.12	2	rep_full	alanine aminotransferase 2
TX:	AT3G59820	10.06	2	rep_full	LETM1-like protein
TX:	AT5G55150	9.95	2	rep_full	Protein of unknown function (DUF295)
TX:	AT2G07675	9.40	2	rep_full	Ribosomal protein S12/S23 family protein
TX:	AT1MG00980	9.40	2	rep_full	Ribosomal protein S12/S23 family protein
TX:	AT5G06250	9.33	2	rep_full	AP2/B3-like transcriptional factor family protein
PR:	AT5G51440.1	9.21	2	rep_full	HSP20-like chaperones superfamily protein
TX:	AT3G58160	8.72	2	rep_full	P-loop containing nucleoside triphosphate hydrolases superfam
TX:	AT1G02520	8.39	2	rep_full	P-glycoprotein 11
TX:	AT4G19770	8.06	2	rep_full	Glycosyl hydrolase family protein with chitinase insertion doma
TX:	AT3G54800	7.92	2	rep_full	Pleckstrin homology (PH) and lipid-binding START domains-conta
TX:	AT1G17180	7.70	2	rep_full	glutathione S-transferase TAU 25
TX:	AT4G13260	7.60	2	rep_full	Flavin-binding monooxygenase family protein
TX:	AT2G03760	7.52	2	rep_full	sulphotransferase 12
TX:	AT5G37490	7.31	2	rep_full	ARM repeat superfamily protein
TX:	AT3G28500	7.16	2	rep_full	60S acidic ribosomal protein family

om	ID1	FNRL_inte	kmea	direction	Short_description
PR:	AT1G65040.2	3666.89	0	rep_part	RING/U-box superfamily protein
PR:	AT4G32208.1	2958.64	0	rep_part	heat shock protein 70 (Hsp 70) family protein
PR:	AT3G19960.2	2694.64	0	rep_part	myosin 1
PR:	AT3G10500.1	2587.36	0	rep_part	NAC domain containing protein 53
PR:	AT1G02816.1	2431.50	0	rep_part	Protein of unknown function, DUF538
PR:	AT5G52540.1	1320.34	0	rep_part	Protein of unknown function (DUF819)
PR:	AT3G47110.1	965.11	0	rep_part	Leucine-rich repeat protein kinase family protein
PR:	AT3G02580.1	845.14	0	rep_part	sterol 1
PR:	AT1G14390.1	843.99	0	rep_part	Leucine-rich repeat protein kinase family protein
PR:	AT5G53100.1	479.25	0	rep_part	NAD(P)-binding Rossmann-fold superfamily protein
TX:	ATMG00860	343.32	0	rep_part	DNA/RNA polymerases superfamily protein
PR:	AT1G61475.1	307.61	0	rep_part	ATP binding;protein kinases
TX:	AT5G58050	306.20	0	rep_part	SHV3-like 4
TX:	AT1G14800	302.42	0	rep_part	Nucleic acid-binding, OB-fold-like protein
TX:	AT1G24010	232.98	0	rep_part	Polyketide cyclase/dehydrase and lipid transport superfamily p
TX:	AT4G38560	198.53	0	rep_part	Arabidopsis phospholipase-like protein (PEARLI 4) family
TX:	AT5G53742	184.17	0	rep_part	Protein of unknown function (DUF1278)
TX:	AT2G17860	178.08	0	rep_part	Pathogenesis-related thaumatin superfamily protein
TX:	AT5G19360	177.26	0	rep_part	calcium-dependent protein kinase 34
TX:	AT1G08440	159.89	0	rep_part	Aluminium activated malate transporter family protein
TX:	AT3G30200	141.60	0	rep_part	Plant transposase (Ptta/En/Spm family)
PR:	ATCG00740.1	127.18	0	rep_part	RNA polymerase subunit alpha
PR:	AT1G01650.1	118.41	0	rep_part	SIGNAL PEPTIDE PEPTIDASE-LIKE 4
PR:	AT4G32360.1	115.08	0	rep_part	Pyridine nucleotide-disulphide oxidoreductase family protein
TX:	ATMG00370	108.90	0	rep_part	Ycf1 protein
TX:	AT3G55780	105.60	0	rep_part	Glycosyl hydrolase superfamily protein
PR:	AT3G12900.1	96.76	0	rep_part	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfam
TX:	AT2G27380	95.56	0	rep_part	extensin proline-rich 1
TX:	AT1G22240	81.99	0	rep_part	pumilio 8
TX:	AT5G20740	74.83	0	rep_part	Plant invertase/pectin methylesterase inhibitor superfamily pro
PR:	AT4G19190.1	69.42	0	rep_part	zinc knuckle (CCHC-type) family protein
TX:	AT2G37435	62.21	0	rep_part	Cystatin/monellin superfamily protein
PR:	AT2G45690.1	62.01	0	rep_part	shrunken seed protein (SSE1)
TX:	AT4G11890	60.66	0	rep_part	Protein kinase superfamily protein
PR:	AT1G34200.1	60.36	0	rep_part	Glyceraldehyde-3-phosphate dehydrogenase-like family protein
PR:	AT1G59710.1	48.68	0	rep_part	Protein of unknown function (DUF569)
TX:	AT3G16600	46.25	0	rep_part	SNF2 domain-containing protein / helicase domain-containing p
PR:	AT5G13210.1	44.49	0	rep_part	Uncharacterised conserved protein UCP015417, vWA
PR:	AT4G12340.1	44.15	0	rep_part	copper ion binding
PR:	AT3G47980.1	41.09	0	rep_part	Integral membrane HPP family protein
TX:	AT3G63095	36.29	0	rep_part	Tetratricopeptide repeat (TPR)-like superfamily protein
TX:	AT4G20040.1	35.70	0	rep_part	Pectin lyase-like superfamily protein
TX:	AT4G29305	32.80	0	rep_part	low-molecular-weight cysteine-rich 25
PR:	AT1G52510.1	31.54	0	rep_part	alpha/beta-Hydrolases superfamily protein
PR:	AT5G14740.1	29.90	0	rep_part	carbonic anhydrase 2
TX:	AT1G20790	23.55	0	rep_part	F-box family protein
PR:	AT4G36730.1	20.78	0	rep_part	G-box binding factor 1
PR:	AT2G35320.1	19.30	0	rep_part	EYES ABSENT homolog
TX:	AT4G39340	18.24	0	rep_part	Protein of unknown function (DUF1278)
TX:	AT1G20350	16.67	0	rep_part	translocase inner membrane subunit 17-1
PR:	AT1G61980.1	15.57	0	rep_part	Mitochondrial transcription termination factor family protein
PR:	AT3G05520.2	15.20	0	rep_part	Subunits of heterodimeric actin filament capping protein Capz s
TX:	AT3G21860	14.16	0	rep_part	SKP1-like 10
PR:	AT5G56580.1	14.06	0	rep_part	MAP kinase kinase 6
PR:	AT1G10810.1	12.37	0	rep_part	NAD(P)-linked oxidoreductase superfamily protein
TX:	AT4G36590	11.58	0	rep_part	MADS-box transcription factor family protein
PR:	AT4G00190.1	10.71	0	rep_part	pectin methylesterase 38
TX:	AT2G27220	9.74	0	rep_part	BEL1-like homeodomain 5
PR:	AT2G31955.1	9.57	0	rep_part	cofactor of nitrate reductase and xanthine dehydrogenase 2
TX:	AT5G22460	7.95	0	rep_part	alpha/beta-Hydrolases superfamily protein
PR:	AT3G14470.1	7.12	0	rep_part	NB-ARC domain-containing disease resistance protein
PR:	AT5G23070.1	7.09	0	rep_part	Thymidine kinase
TX:	AT3G44290	6.33	0	rep_part	NAC domain containing protein 60
TX:	ATMG00300	5.96	0	rep_part	Gag-Pol-related retrotransposon family protein
TX:	AT4G17710	5.95	0	rep_part	homeodomain GLABROUS 4
TX:	AT1G51460	5.66	0	rep_part	ABC-2 type transporter family protein
TX:	AT4G34440	5.37	0	rep_part	Protein kinase superfamily protein
TX:	AT3G28790	5.21	0	rep_part	Protein of unknown function (DUF1216)
PR:	AT5G12320.1	4.97	0	rep_part	ankyrin repeat family protein
PR:	AT1G78090.1	4.87	0	rep_part	trehalose-6-phosphate phosphatase
TX:	AT5G58610	4.62	0	rep_part	PHD finger transcription factor, putative
TX:	AT5G39100	4.53	0	rep_part	germin-like protein 6
TX:	AT5G55270	4.51	0	rep_part	Protein of unknown function (DUF295)
TX:	AT5G52350	4.38	0	rep_part	exocyst subunit exo70 family protein A3
PR:	AT4G39740.1	4.17	0	rep_part	Thioredoxin superfamily protein

9.1.3/ TX+PR Biomarkers – UniProt Retrieve/ID mapping

Mapping of TX + PR biomarkers (using ID2)

478/482 unique AT ID2 hits (99%)

Job type	Name	Created	Status
ID MAPPING 49ADD836a8	AT1G15140 +481 Araport → UniPr	2026-02-26 10:36	**Completed** (708 hits) Source database: Araport Target database: UniProtKB

NOTE: FNRL (PR:AT1G15140.1) mapped (**F4HXZ3, Q9XI55**).

478 IDs mapped to **708** results

4 IDs not mapped:

AT5G56555
AT1G04105
AT4G24615
AT5G41763

Linked DBs

- Pfam
- KEGG (EC no.)
- GO
- BioCyc
- UniPathway

From	Entry	Reviewed	Entry Name	Protein name	Gene Names	Organism	Length	Pfam	Organism (ID)	Sequence	EC number	Function [CC	Pathway	Gene Ontolo	Gene Ontolo	Gene Ontolo	Gene Ontolo	Gene Ontolo	BioCyc	PathwayCom	PlantReacton	Reactome	UniPathway
AT1G15140	F4HXZ3	unreviewed	F4HXZ3_ARAT	FAD/NAD(P)-b	FNRL At1g151	Arabidopsis th	271	PF00175;	3702	MSTLPFAPSV					chloroplast [G	copper ion bir	chloroplast [G	GO:0004324;					
AT1G15140	Q9XI55	unreviewed	Q9XI55_ARAT	F9L1.8 protein	FNRL At1g151	Arabidopsis th	295	PF00175;	3702	MSTLPFAPSVTHAHFSHSLSPMFILRHLPLTRHLRLSRNNRVASVVSAAAVRQ					chloroplast [G	copper ion bir	chloroplast [G	GO:0004324;	GO:0005507;	GO:0005576;	GO:0009507;	GO:0009570;	GO:0009579;
AT1G02520	Q9FWX7	reviewed	AB11B_ARATH	ABC transport	ABCB11 MDR	Arabidopsis th	1278	PF00664;PF0	3702	MNGDGAREG		FUNCTION: In		auxin export a	extracellular re	ABC-type tran	extracellular re	GO:0005524;	ARA:AT1G025				
AT1G04445	A0A1P8AWD1	unreviewed	A0A1P8AWD1	C2H2-like zinc	At1g04445	Arabidopsis th	174		3702	MIMPFSEPQE						zinc ion bindir	zinc ion bindir	GO:0008270					
AT1G04445	P93815	unreviewed	P93815_ARAT	C2H2-like zinc	F19P19.9 At1	Arabidopsis th	172		3702	MPFSEPQECA						zinc ion bindir	zinc ion bindir	GO:0008270					
AT1G05680	Q9SYK9	reviewed	U74E2_ARAT	UDP-glycosyl	UGT74E2 At1	Arabidopsis th	453	PF00201;	3702	MREGSHLIVL	2.4.1.-	FUNCTION: G		cellular hyper	cytoplasm [G	indole-3-buty	cytoplasm [G	GO:0005737;	ARA:AT1G056				
AT1G06330	F4IC29	reviewed	HIP28_ARATH	Heavy metal-a	HIPP28 At1g0	Arabidopsis th	159	PF00403;	3702	MNLTLNLIQL		FUNCTION: H				metal ion bind	metal ion bind	GO:0046872					
AT1G06330	A0A178WC78	unreviewed	A0A178WC78	Heavy metal t	At1g06330 T2	Arabidopsis th	150	PF00403;	3702	MTTIEMRVHM						metal ion bind	metal ion bind	GO:0046872					
AT1G07180	Q8GWA1	reviewed	NDA1_ARATH	Internal altern	NDA1 NDI1 At	Arabidopsis th	510	PF22366;PF0	3702	MLWIKNLARIS	1.6.5.9	FUNCTION: Al		cellular respo	mitochondrial	NADH dehydr	mitochondrial	GO:0003954;	ARA:AT1G071				
AT1G07550	C0LGD8	reviewed	Y1755_ARATH	Probable LRR	At1g07550 F2	Arabidopsis th	864	PF13855;PF1	3702	MDTCTRLLFA	2.7.11.1				membrane [G	ATP binding [G	membrane [G	GO:0004674;					
AT1G08090	O82811	reviewed	NRT21_ARATH	High-affinity n	NRT2.1 ACH1	Arabidopsis th	530	PF07690;	3702	MGDSTGEPGS		FUNCTION: In		cellular respo	membrane [G	nitrate transm	membrane [G	GO:0005886;					
AT1G08165	A0A654E9T7	unreviewed	A0A654E9T7	Uncharacteriz	At1g08165 At	Arabidopsis th	47		3702	MNYDRHLSPC													
AT1G08440	Q9SJE8	reviewed	ALMT2_ARATH	Aluminum-ac	ALMT2 At1g08	Arabidopsis th	501	PF11744;	3702	MEKVREIVRE		FUNCTION: M		malate transp	plant-type vac		plant-type vac	GO:0009705;					
AT1G09080	Q8H1B3	reviewed	BIP3_ARATH	Heat shock 70	BIP3 BIP-L HS	Arabidopsis th	675	PF00012;	3702	MIFIKENTAKM		FUNCTION: In		endoplasmic	cytoplasm [G	ATP binding [G	cytoplasm [G	GO:0000304;					
AT1G09080	A0A178WHW	unreviewed	A0A178WHW	(thale cress) h	At1g09080 A	Arabidopsis th	665	PF00012;	3702	MIFIKENTAKM				pollen tube gr	endoplasmic	ATP binding [G	endoplasmic	GO:0005524;					
AT1G09665	Q3EDF9	unreviewed	Q3EDF9_ARA	Toll-interleuki	At1g09665	Arabidopsis th	165	PF01582;	3702	MKRSSSNRKI				defense respo	nucleus [GO:0		nucleus [GO:0	GO:0005634;					
AT1G09720	F4I131	reviewed	NET2B_ARATH	Protein NETW	NET2B At1g09	Arabidopsis th	928	PF07765;PF2	3702	MLQRAASNAY		FUNCTION: Pl			membrane [G	actin binding [	membrane [G	GO:0003779;					
AT1G10875	A0A1P8AME6	unreviewed	A0A1P8AME6	Core-2/I-bran	At1g10875	Arabidopsis th	302		3702	MEVAKIGRKK					endoplasmic		endoplasmic	GO:0005789;					
AT1G12130	Q9FWW3	reviewed	GSXL6_ARATH	Flavin-contain	At1g12130 T2	Arabidopsis th	470	PF00743;	3702	MTPPPNSISS	1.8.-.-	FUNCTION: C		glucosinolate		flavin adenine	flavin adenine	GO:0004497;	ARA:AT1G121				
AT1G12211	B3H6A6	unreviewed	B3H6A6_ARA	Uncharacteriz	At1g12211	Arabidopsis th	82		3702	MGKNHHPLG													
AT1G13200	Q9SAF4	reviewed	FBK3_ARATH	Putative F-box	At1g13200 F3	Arabidopsis th	435	PF00646;PF0	3702	MKDAEKREVL								GO:0000149;					
AT1G14686	Q9LQW4	reviewed	CAP15_ARAT	Putative clathr	At1g14686 F1	Arabidopsis th	339	PF07651;	3702	MKLWKRAAVV				clathrin coat a	clathrin-coate	1-phosphatid	clathrin-coate	GO:0000149;					
AT1G14800	F4HXU8	unreviewed	F4HXU8_ARA	Nucleic acid-b	At1g14800 F1	Arabidopsis th	384	PF02721;PF2	3702	MAEVDVVAIP													
AT1G17180	Q9SHH7	reviewed	GSTUP_ARAT	Glutathione S	GSTU25 At1g	Arabidopsis th	221	PF13410;PF0	3702	MADEVILLDF	2.5.1.18	FUNCTION: M		2,4,6-trinitrot	cytoplasm [G	glutathione tra	cytoplasm [G	GO:0004364;	ARA:AT1G171				
AT1G18010	Q8LG53	reviewed	UN932_ARAT	UNC93-like pr	At1g18010 T1	Arabidopsis th	459	PF05978;	3702	MNVRDEGKTT					membrane [G		membrane [G	GO:0016020					
AT1G18320	A1XJK0	reviewed	TI224_ARATH	Mitochondrial	TIM22-4 At1g	Arabidopsis th	142	PF02466;	3702	MPTIEEIRAQE		FUNCTION: Es		protein inserti	TIM22 mitoch	mitochondrio	TIM22 mitoch	GO:0008320;					
AT1G18830	F4ICD9	reviewed	SC31A_ARATH	Protein transp	SEC31A At1g	Arabidopsis th	969	PF12931;PF0	3702	MDCIKSIGRSA		FUNCTION: Re		COPII-coated	COPII vesicle	structural mol	COPII vesicle	GO:0005198;					
AT1G18830	A0A1P8AP97	unreviewed	A0A1P8AP97	Transducin/W	SEC31A At1g	Arabidopsis th	963	PF12931;PF0	3702	MDCIKSIGRSA		FUNCTION: Re		endoplasmic	endoplasmic		endoplasmic	GO:0005783;					
AT1G18830	A0A5S9V1Q4	unreviewed	A0A5S9V1Q4	Sec16 Sec23-	At1g18830 C	Arabidopsis th	963	PF12931;PF0	3702	MDCIKSIGRSA		FUNCTION: Re		endoplasmic	endoplasmic		endoplasmic	GO:0005783;					
AT1G19320	A0A654EBB0	unreviewed	A0A654EBB0	Pathogenesis-	At1g19320 At	Arabidopsis th	247	PF00314;	3702	MAIFSTSHLLF													
AT1G19610	P82787	reviewed	DEF19_ARATH	Defensin-like	PDF1.4 LCR7	Arabidopsis th	78	PF00304;	3702	MASSYTLMLF		FUNCTION: Cl		defense respo	extracellular re		extracellular re	GO:0005576;					
AT1G20180	Q6DYE5	reviewed	U496K_ARAT	UPF0496 prot	At1g20180 T2	Arabidopsis th	390	PF05055;	3702	MLKVKNFLGS					membrane [G		membrane [G	GO:0016020					
AT1G20350	Q9LN27	reviewed	TI171_ARATH	Mitochondrial	TIM17-1 At1g	Arabidopsis th	218	PF02466;	3702	MGTPESSREF		FUNCTION: Es		protein import	TIM23 mitoch		TIM23 mitoch	GO:0005744;					
AT1G20790	Q9LM75	reviewed	FBK6_ARATH	Putative F-box	At1g20790 F2	Arabidopsis th	435	PF00646;PF0	3702	MKRLPLHLLD													
AT1G21120	Q9LPU7	reviewed	IGMT2_ARATH	Indole glucos	IGMT2 At1g21	Arabidopsis th	373	PF08100;PF0	3702	MGYLFEETLSS	2.1.1.-	FUNCTION: In	PATHWAY: Se	callus formati	cytoplasmic s	methyltransfe	cytoplasmic s	GO:0005829;	ARA:AT1G211				
AT1G21120	A0A1P8AP65	unreviewed	A0A1P8AP65	O-methyltrans	IGMT2 At1g21	Arabidopsis th	421	PF08100;PF0	3702	MHQKYTNSSS				methylation [G		O-methyltrans	O-methyltrans	GO:0008171;					
AT1G21250	Q39191	reviewed	WAK1_ARATH	Wall-associate	WAK1 PRO25	Arabidopsis th	735	PF07645;PF1	3702	MKVQEGLFLV	2.7.11.-	FUNCTION: Se		cell surface re	plant-type cel	ATP binding [G	plant-type cel	GO:0000325;					
AT1G22240	Q9LM20	reviewed	PUM8_ARATH	Putative pumi	APUM8 At1g2	Arabidopsis th	515	PF00806;	3702	MMRGEFGEAS		FUNCTION: Se		post-transcrip	chloroplast [G	mRNA binding	chloroplast [G	GO:0003729;					
AT1G23700	A0A1P8AMH8	unreviewed	A0A1P8AMH8	Protein kinase	At1g23700 F5	Arabidopsis th	395	PF00069;	3702	MTSSPETRFP						ATP binding [G	ATP binding [G	GO:0004672;					
AT1G23700	A0A1P8AML6	unreviewed	A0A1P8AML6	Protein kinase	At1g23700 F5	Arabidopsis th	483	PF00069;	3702	MTSSPETRFP						ATP binding [G	ATP binding [G	GO:0004672;					

9.1.4/ PR Biomarkers – PeptideAtlas – Get Proteins

Annotation retrieval of PR biomarkers (using ID1)

143/148 AT ID1 hits (97%); qualitative analysis

Good to retrieve annotations and genome coordinates (from column “Protein Description”).

NOTE: FNRL (PR:AT1G15140.1) **mapped.**

biosequence_name	presence_lev	represented_t	group_	protein_description	gene_chromosome	gene_locus	gene_direction	protein_length
AT1G15140.1	canonical	AT1G15140.1	709	FAD/NAD(P)-binding oxidoreductase (FNRL)	Chr1	5210403-5212137	REVERSE	295
AT1G02816.1	canonical	AT1G02816.1	86	pectinesterase (Protein of unknown function%2C DUF538)	Chr1	621637-622137	FORWARD	166
AT1G06970.1	canonical	AT1G06970.1	297	cation/hydrogen exchanger 14	Chr1	2138109-2140818	FORWARD	829
AT1G16650.1	canonical	AT1G16650.1	778	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	Chr1	5687994-5691101	FORWARD	521
AT1G72330.3	subsumed	AT1G17290.1	812	alanine aminotransferase 2	Chr1	27233637-27236571	FORWARD	553
AT1G21840.1	weak	AT1G21840.1	1010	urease accessory protein F	Chr1	7666859-7667581	FORWARD	240
AT2G47490.1	canonical	AT1G25380.1	1158	NAD+ transporter 1	Chr2	19487549-19489311	FORWARD	312
AT1G59710.1	canonical	AT1G27100.1	1214	actin cross-linking protein (DUF569)	Chr1	21938623-21939673	FORWARD	300
AT1G28490.1	canonical	AT1G28490.1	1288	syntaxin of plants 61	Chr1	10016433-10017842	FORWARD	245
AT1G30300.1	canonical	AT1G30300.1	1353	Metallo-hydrolase/oxidoreductase superfamily protein	Chr1	10673084-10675061	FORWARD	324
AT2G34710.1	canonical	AT1G30490.1	1359	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein	Chr2	14639548-14643993	REVERSE	852
AT1G48200.1	canonical	AT1G48200.1	1650	hypothetical protein	Chr1	17797070-17797852	REVERSE	118
AT5G42930.1	canonical	AT1G56630.1	2018	alpha/beta-Hydrolases superfamily protein	Chr5	17210738-17214152	REVERSE	467
AT4G22990.2	marginally dis	AT1G63010.1	2181	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein	Chr4	12048240-12050984	REVERSE	700
AT4G11810.1	subsumed	AT1G63010.1	2181	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein	Chr4	7105446-7108450	FORWARD	707
AT1G01650.1	canonical	AT1G63690.1	2207	SIGNAL PEPTIDE PEPTIDASE-LIKE 4	Chr1	233188-237647	REVERSE	540
AT1G65040.2	canonical	AT1G65040.2	2270	RING/U-box superfamily protein	Chr1	24160105-24163365	REVERSE	460
AT1G65700.3		AT1G65700.1	2295	Small nuclear ribonucleoprotein family protein	Chr1	24434463-24435870	REVERSE	141
AT1G69070.1	canonical	AT1G69070.1	2461	nucleolar-like protein	Chr1	25967421-25971389	REVERSE	901
AT1G70070.1	canonical	AT1G70070.1	2503	DEAD/DEAH box helicase	Chr1	26390016-26394148	REVERSE	1171
AT1G71500.1	canonical	AT1G71500.1	2561	Rieske (2Fe-2S) domain-containing protein	Chr1	26936084-26937331	FORWARD	287
AT1G72120.1	canonical	AT1G72120.1	2594	Major facilitator superfamily protein	Chr1	27132133-27133975	FORWARD	557
AT1G23080.1	canonical	AT1G73590.1	2671	Auxin efflux carrier family protein	Chr1	8180768-8183406	REVERSE	619
AT4G15960.1	canonical	AT1G74580.1	2728	alpha/beta-Hydrolases superfamily protein	Chr4	9045763-9047199	REVERSE	375
AT1G77020.1	canonical	AT1G77020.1	2836	DNAJ heat shock N-terminal domain-containing protein	Chr1	28945056-28946867	REVERSE	379
AT1G78090.1	canonical	AT1G78090.1	2893	trehalose-6-phosphate phosphatase	Chr1	29373955-29376295	FORWARD	374
AT1G35780.1	canonical	AT1G78150.1	2897	N-lysine methyltransferase	Chr1	13277778-13280113	REVERSE	286
AT1G80640.1	canonical	AT1G80640.1	3026	Protein kinase superfamily protein	Chr1	30311979-30314238	FORWARD	427
AT2G01070.1	canonical	AT2G01070.1	3050	Lung seven transmembrane receptor family protein	Chr2	75596-77625	FORWARD	496
AT5G18100.1	canonical	AT2G15620.1	3329	copper/zinc superoxide dismutase 3	Chr5	5987221-5988706	FORWARD	164
AT2G22795.1	canonical	AT2G22795.1	3649	hypothetical protein	Chr2	9697380-9699584	REVERSE	734
AT1G37150.2	subsumed	AT2G25710.1	3772	holocarboxylase synthetase 2	Chr1	14175029-14177098	REVERSE	329
AT2G27680.1	canonical	AT2G27680.1	3873	NAD(P)-linked oxidoreductase superfamily protein	Chr2	11803981-11805965	REVERSE	384
AT2G28800.1	canonical	AT2G28800.1	3929	63 kDa inner membrane family protein	Chr2	12356669-12359158	REVERSE	462
AT2G29220.1	canonical	AT2G29220.1	3942	Concanavalin A-like lectin protein kinase family protein	Chr2	12562781-12564664	REVERSE	627

9.1.5/ MT Biomarkers – MetaboAnalyst – Pathway Analysis

Annotation retrieval of MT biomarkers (using ID2)

14/15 metabolites ID1 hits (93%); qualitative analysis

Good to HMDB, PubChem, KEGG, SMILES annotations .

Query(ID2)	Match	HMDB	PubChem	KEGG	SMILES
1,5-Anhydroglucitol	1,5-Anhydrosorbitol	HMDB0002712	64960	C07326	OC[C@H]1OC[C@H](O)[C@@H](O)[C@@H]1O
2,3,4-Trihydroxybutyric acid	Threonic acid	HMDB0000943	5460407	C01620	OC[C@H](O)[C@@H](O)C(O)=O
Amphetamine	Amphetamine	HMDB0014328	3007	C07514	CC(N)CC1=CC=CC=C1
Butanedioic acid	Succinic acid	HMDB0000254	1110	C00042	OC(=O)CCC(O)=O
Citric acid	Citric acid	HMDB0000094	311	C00158	OC(=O)CC(O)(CC(O)=O)C(O)=O
D-(+)-Xylose	D-Xylose	HMDB0000098	135191	C00181	O[C@@H]1COC(O)[C@H](O)[C@H]1O
Gulonic acid, Î^3-lactone	L-Gulonolactone	HMDB0003466	439373	C01040	[H][C@@]1(OC(=O)[C@@H](O)[C@H]1O)[C@@H](O)CO
L-(-)-Arabitol	L-Arabitol	HMDB0001851	439255	C00532	OC[C@H](O)C(O)[C@@H](O)CO
Lactulose	Lactulose	HMDB0000740	11333	C07064	OC[C@H]1O[C@](O)(CO)[C@@H](O)[C@@H]1O[C@@H]1O[C@H](CO)[C@H](O)[C@H](O)[C@H]1O
Maltose	D-Maltose	HMDB0000163	10991489	C00208	OC[C@H]1O[C@H](O[C@H]2[C@H](O)[C@H](O)[C@@H](O)O[C@@H]2CO)[C@H](O)[C@@H](O)[C@@H]1O
Myo-Inositol	myo-Inositol	HMDB0000211	NA	C00137	O[C@H]1[C@H](O)[C@@H](O)[C@H](O)[C@H](O)[C@@H]1O
Pentanedioic acid	Glutaric acid	HMDB0000661	743	C00489	OC(=O)CCCC(O)=O
Serine	Serine	HMDB0000187	5951	C00065	N[C@@H](CO)C(O)=O
β-D-(+)-Talopyranose	NA	NA	NA	NA	NA

9.2.1/ GO & AmiGO 2

Biological Process analysis of TX and PR biomarkers (using ID2)

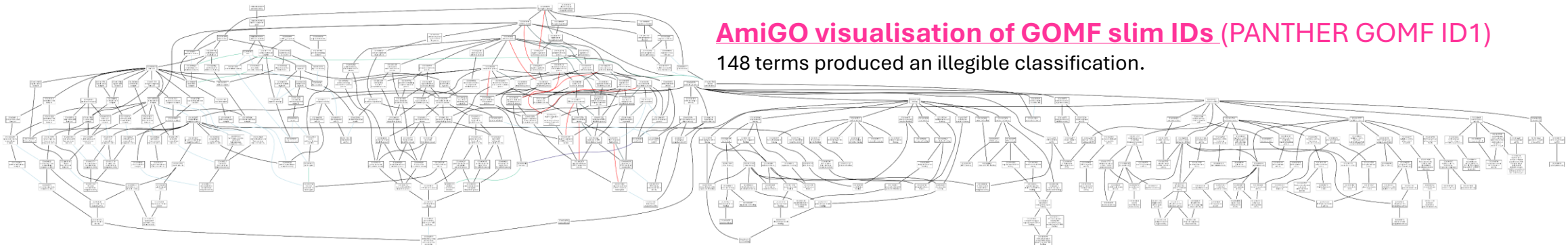
466/490 AT ID2 hits (95%) ; quantitative analysis.
GO to retrieve GO terms.
Root-related GO terms shown on the right
AmiGO to visualize tree hierarchical annotations.
Molecular Function (GOMF) classification shown below.

	Reference list	upload_1
Uniquely Mapped IDS:	27475 out of 27475	466 out of 468
Unmapped IDs:	0	16
Multiple mapping information:	0	2

NOTE: FNRL (PR:AT1G15140.1) mapped (**F9L1.8 PROTEIN**; **PTHR47215:SF1**) but no classification available (no Panther Protein Class).

GOBP classification zoomed in on root terms

	Arabidopsis thaliana (REF)	upload_1 (▽ Hierarchy NEW!)					
GO biological process complete	#	#	expected	Fold Enrichment	+/-	raw P value	FDR
root hair cell tip growth	31	1	.53	1.89	+	4.13E-01	1.00E00
↳cell tip growth	168	4	2.86	1.40	+	3.77E-01	1.00E00
↳unidimensional cell growth	289	6	4.92	1.22	+	6.42E-01	1.00E00
↳cell growth	383	6	6.52	.92	-	1.00E00	1.00E00
↳developmental growth involved in morphogenesis	337	6	5.74	1.05	+	8.31E-01	1.00E00
↳cell morphogenesis	375	7	6.39	1.10	+	6.89E-01	1.00E00
↳root hair elongation	81	1	1.38	.72	-	1.00E00	1.00E00
↳root hair cell development	111	1	1.89	.53	-	1.00E00	1.00E00
↳cell development	288	4	4.91	.82	-	1.00E00	1.00E00
↳root hair cell differentiation	138	1	2.35	.43	-	7.34E-01	1.00E00
↳trichoblast maturation	138	1	2.35	.43	-	7.34E-01	1.00E00
↳trichoblast differentiation	141	1	2.40	.42	-	7.37E-01	1.00E00
↳root epidermal cell differentiation	157	1	2.67	.37	-	5.29E-01	1.00E00
↳plant epidermal cell differentiation	212	1	3.61	.28	-	2.76E-01	1.00E00
↳root morphogenesis	339	5	5.77	.87	-	1.00E00	1.00E00
↳root development	614	8	10.46	.76	-	5.29E-01	1.00E00
↳root system development	616	8	10.49	.76	-	5.29E-01	1.00E00
↳plant epidermis development	299	2	5.09	.39	-	2.54E-01	1.00E00
↳cell maturation	138	1	2.35	.43	-	7.34E-01	1.00E00
↳anatomical structure maturation	148	1	2.52	.40	-	5.25E-01	1.00E00
↳developmental maturation	239	3	4.07	.74	-	8.02E-01	1.00E00
↳developmental cell growth	208	4	3.54	1.13	+	7.84E-01	1.00E00



AmiGO visualisation of GOMF slim IDs (PANTHER GOMF ID1)

148 terms produced an illegible classification.

9.2.2/ GO & AmiGO 2

Biological Process analysis of TX and PR biomarkers (using ID2)

466/489 AT ID2 hits (95%) ; quantitative analysis.
GO to retrieve GO terms.
GOBP sorted by number of hits (occurrences).
Not very informative...

GO biological process	Arath Ref	occurence	expected	over/unde	fold Enrichment	raw P-value	FDR
biological_process (GO:0008150)	22442	393	382.27	+	1.03	2.06E-01	1.00E+00
cellular process (GO:0009987)	14823	243	252.49	-	0.96	3.75E-01	1.00E+00
metabolic process (GO:0008152)	8963	129	152.67	-	0.84	1.93E-02	1.00E+00
response to stimulus (GO:0050896)	6925	122	117.96	+	1.03	6.68E-01	1.00E+00
biological regulation (GO:0065007)	6370	107	108.5	-	0.99	9.12E-01	1.00E+00
regulation of biological process (GO:0050789)	6106	100	104.01	-	0.96	6.95E-01	1.00E+00
primary metabolic process (GO:0044238)	7516	100	128.03	-	0.78	3.35E-03	1.00E+00
regulation of cellular process (GO:0050794)	5641	91	96.09	-	0.95	6.03E-01	1.00E+00
Unclassified (UNCLASSIFIED)	5033	75	85.73	-	0.87	2.06E-01	1.00E+00
response to stress (GO:0006950)	4262	73	72.6	+	1.01	9.49E-01	1.00E+00
cellular response to stimulus (GO:0051716)	3603	68	61.37	+	1.11	3.69E-01	1.00E+00
macromolecule metabolic process (GO:0043170)	5385	64	91.73	-	0.7	8.01E-04	1.00E+00
response to chemical (GO:0042221)	3233	63	55.07	+	1.14	2.47E-01	1.00E+00
localization (GO:0051179)	2869	56	48.87	+	1.15	2.85E-01	1.00E+00
transport (GO:0006810)	2618	55	44.59	+	1.23	1.12E-01	1.00E+00
establishment of localization (GO:0051234)	2729	55	46.48	+	1.18	1.85E-01	1.00E+00
biosynthetic process (GO:0009058)	4343	52	73.98	-	0.7	3.99E-03	1.00E+00
regulation of metabolic process (GO:0019222)	3496	48	59.55	-	0.81	1.23E-01	1.00E+00
developmental process (GO:0032502)	3164	47	53.89	-	0.87	3.43E-01	1.00E+00
signal transduction (GO:0007165)	2204	46	37.54	+	1.23	1.45E-01	1.00E+00
signaling (GO:0023052)	2242	46	38.19	+	1.2	2.01E-01	1.00E+00
cell communication (GO:0007154)	2273	46	38.72	+	1.19	2.35E-01	1.00E+00
anatomical structure development (GO:0048856)	3051	45	51.97	-	0.87	3.35E-01	1.00E+00
cellular component organization or biogenesis (GO:0071840)	3450	45	58.77	-	0.77	5.71E-02	1.00E+00
response to abiotic stimulus (GO:0009628)	2348	44	40	+	1.1	5.04E-01	1.00E+00
multicellular organismal process (GO:0032501)	2907	44	49.52	-	0.89	4.48E-01	1.00E+00
regulation of biosynthetic process (GO:0009889)	3162	44	53.86	-	0.82	1.65E-01	1.00E+00
regulation of macromolecule metabolic process (GO:0060255)	3212	44	54.71	-	0.8	1.28E-01	1.00E+00
regulation of primary metabolic process (GO:0080090)	2982	43	50.79	-	0.85	2.61E-01	1.00E+00
regulation of macromolecule biosynthetic process (GO:001055	3032	42	51.65	-	0.81	1.58E-01	1.00E+00
transmembrane transport (GO:0055085)	1414	41	24.09	+	1.7	9.63E-04	1.00E+00
regulation of gene expression (GO:0010468)	2993	41	50.98	-	0.8	1.55E-01	1.00E+00
cellular component organization (GO:0016043)	3068	41	52.26	-	0.78	1.03E-01	1.00E+00
response to oxygen-containing compound (GO:1901700)	1788	39	30.46	+	1.28	1.08E-01	1.00E+00
defense response (GO:0006952)	1845	39	31.43	+	1.24	1.62E-01	1.00E+00
regulation of nucleobase-containing compound metabolic proc	2607	39	44.41	-	0.88	4.27E-01	1.00E+00
protein metabolic process (GO:0019538)	2835	38	48.29	-	0.79	1.25E-01	1.00E+00
regulation of RNA metabolic process (GO:0051252)	2507	37	42.7	-	0.87	4.18E-01	1.00E+00
multicellular organism development (GO:0007275)	2610	37	44.46	-	0.83	2.65E-01	1.00E+00
regulation of RNA biosynthetic process (GO:2001141)	2381	36	40.56	-	0.89	5.07E-01	1.00E+00
regulation of DNA-templated transcription (GO:0006355)	2381	36	40.56	-	0.89	5.07E-01	1.00E+00
cellular response to chemical stimulus (GO:0070887)	1735	32	29.55	+	1.08	6.31E-01	1.00E+00
catabolic process (GO:0009056)	1879	31	32.01	-	0.97	9.26E-01	1.00E+00
response to external stimulus (GO:0009605)	1939	31	33.03	-	0.94	7.85E-01	1.00E+00

9.3.1/ KEGG – TX+PR+MT biomarkers

Pathway analysis of protein biomarkers

452/485 UniProt ID hits (93%) + 14/15 MT (93%) hits; qualitative analysis

Step 1: Conversion ID (UniProt → KEGG ath:xxxx)

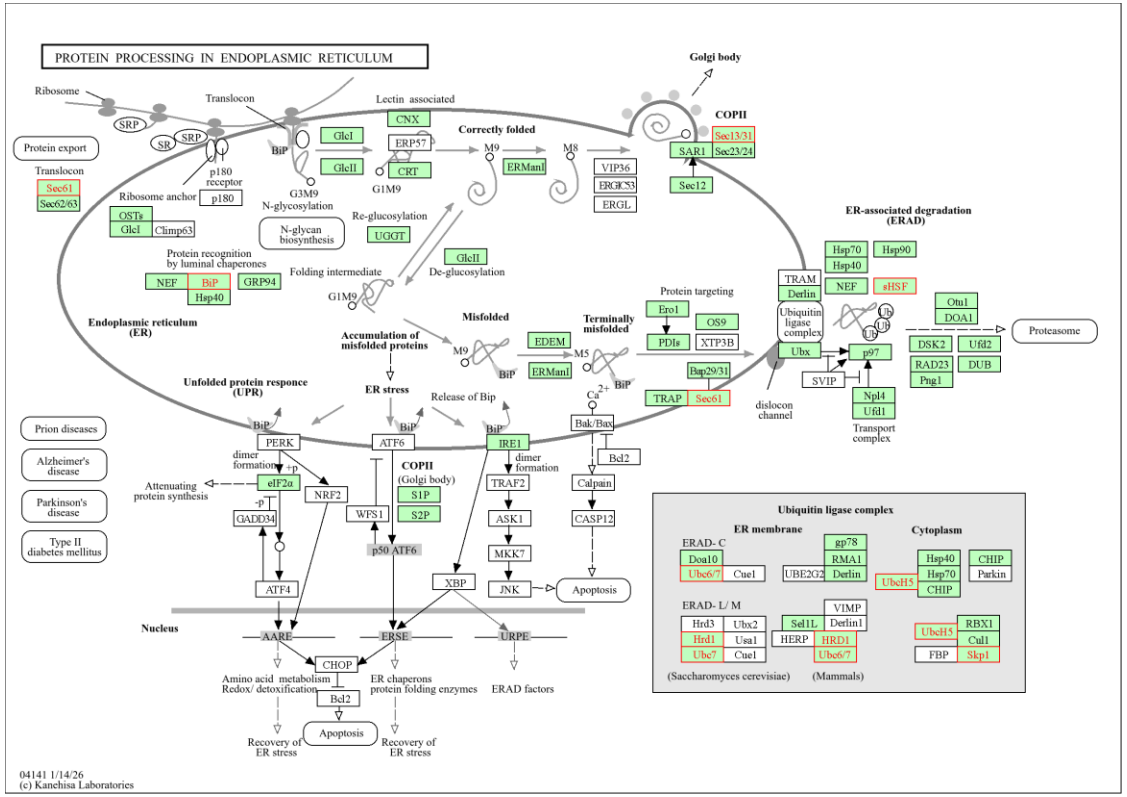
Step 2: Combine KEGG ath:xxxx with compounds Cxxxx IDs

Step 3: Mapper Search of combined IDs

Pathways with most hits:

- **Protein processing in endoplasmic reticulum**
- **Carbon metabolism**
- **ABC transporters**
- **Ubiquitin mediated proteolysis**

NOTE: FNRL (PR:AT1G15140.1) **unmapped.**



KEGG Mapper Search Result

Pathway (95) | Brite (30) | Module (27)

KEGG categ	KEGG ID	Category description	Occurrenc
pathway	ath01100	Metabolic pathways	46
pathway	ath01110	Biosynthesis of secondary metabolites	26
pathway	ath04141	Protein processing in endoplasmic reticulum	10
pathway	ath01200	Carbon metabolism	7
pathway	ath02010	ABC transporters	6
pathway	ath04120	Ubiquitin mediated proteolysis	6
pathway	ath00500	Starch and sucrose metabolism	5
pathway	ath00190	Oxidative phosphorylation	5
pathway	ath01240	Biosynthesis of cofactors	5
pathway	ath01230	Biosynthesis of amino acids	5
pathway	ath04075	Plant hormone signal transduction	5
pathway	ath00630	Glyoxylate and dicarboxylate metabolism	4
pathway	ath03083	Polycomb repressive complex	4
pathway	ath03010	Ribosome	4
pathway	ath04814	Motor proteins	3
pathway	ath01210	2-Oxocarboxylic acid metabolism	3
pathway	ath00020	Citrate cycle (TCA cycle)	3
pathway	ath00940	Phenylpropanoid biosynthesis	3
pathway	ath00710	Carbon fixation by Calvin cycle	3
pathway	ath00040	Pentose and glucuronate interconversions	3
pathway	ath04070	Phosphatidylinositol signaling system	3
pathway	ath00310	Lysine degradation	3
pathway	ath00620	Pyruvate metabolism	3
pathway	ath00350	Tyrosine metabolism	3
pathway	ath00910	Nitrogen metabolism	3
pathway	ath00564	Glycerophospholipid metabolism	3
pathway	ath00053	Ascorbate and aldarate metabolism	3
pathway	ath03060	Protein export	3
pathway	ath00250	Alanine, aspartate and glutamate metabolism	3
brite	ath00001	KEGG Orthology (KO)	161
brite	ath01000	Enzymes	59
brite	ath02000	Transporters	31
brite	ath03000	Transcription factors	16
brite	ath04131	Membrane trafficking	16
brite	ath03029	Mitochondrial biogenesis	13
brite	ath03036	Chromosome and associated proteins	9
brite	ath04121	Ubiquitin system	9
brite	ath04147	Exosome	7
brite	ath01002	Peptidases and inhibitors	6
brite	ath03019	Messenger RNA biogenesis	4
brite	ath01007	Amino acid related enzymes	4
brite	ath03110	Chaperones and folding catalysts	4
brite	ath04812	Cytoskeleton proteins	3
brite	ath02044	Secretion system	3
brite	ath01009	Protein phosphatases and associated proteins	3
brite	ath03041	Spliceosome	3
brite	ath01001	Protein kinases	3
module	ath_M00009	Citrate cycle (TCA cycle, Krebs cycle)	3

9.3.2/ PlantReactome

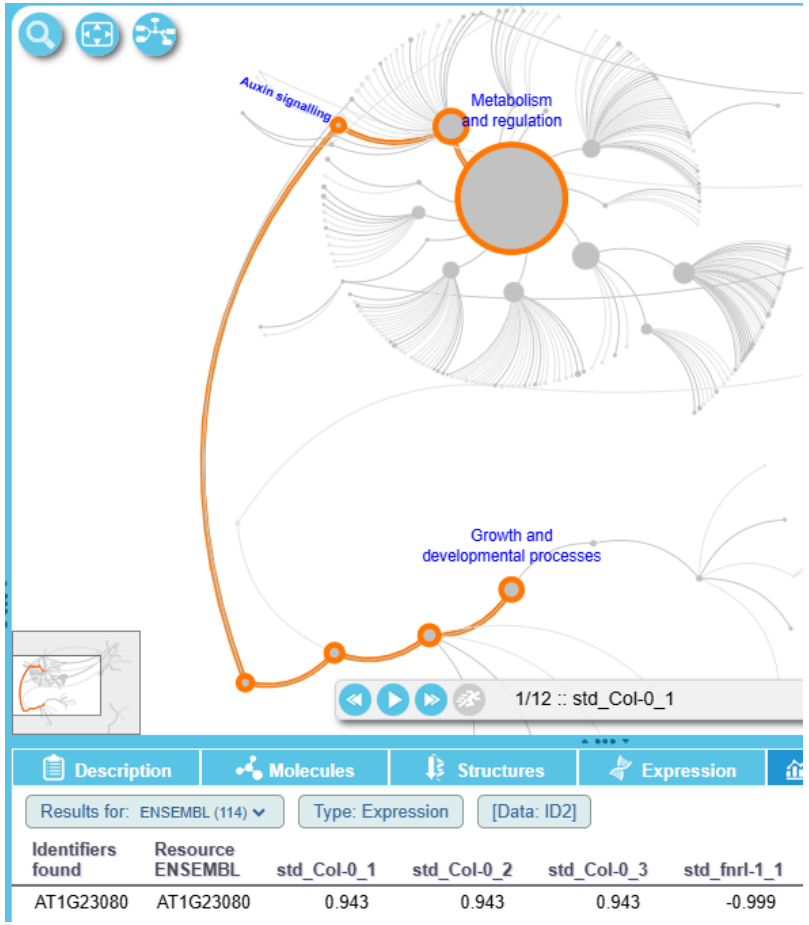
Pathway analysis of TX and PR unique biomarkers (using ID2)

63/482 unique AT ID hits (13%); quantitative analysis using standardized abundances

Pathways of interest:

- **Auxin signaling**
- **Seed development**
- **gravitropism**

NOTE: FNRL (PR:AT1G15140.1) **unmapped**.



Pathway identifie	Pathway name	#Entities f
R-ATH-2744345	Metabolism and regulation	13
R-ATH-2744343	Amino acid metabolism	3
R-ATH-2744344	Secondary metabolism	2
R-ATH-8932729	Responses to stimuli: abiotic stimuli and stresses	2
R-ATH-9031669	Reproductive structure development	2
R-ATH-5368291	Inorganic nutrients metabolism	2
R-ATH-2744341	Hormone signaling, transport, and metabolism	2
R-ATH-9030769	Growth and developmental processes	2
R-ATH-1119586	Cyanate degradation	2
R-ATH-2883407	Carbohydrate metabolism	2
R-ATH-5655124	Amino acid catabolism	2
R-ATH-5655122	Amino acid biosynthesis	2
R-ATH-1119506	tyrosine degradation I	1
R-ATH-4827054	Tetrapyrrole biosynthesis I	1
R-ATH-1119533	TCA cycle (plant)	1
R-ATH-1119465	Sucrose biosynthesis	1
R-ATH-9623902	Seed development	1
R-ATH-1119438	Secologanin and strictosidine biosynthesis	1
R-ATH-9618278	Response to heavy metals	1
R-ATH-9035605	Regulation of seed size	1
R-ATH-9826782	Regulation of seed germination and coleoptile growth under su	1
R-ATH-1119312	Photorespiration	1
R-ATH-1119402	Phospholipid biosynthesis I	1
R-ATH-1119389	Phenylalanine biosynthesis I	1
R-ATH-1119308	Momilactone biosynthesis	1
R-ATH-1119400	Methionine biosynthesis II	1
R-ATH-1119365	Lysine degradation II	1
R-ATH-1119419	Lysine biosynthesis VI	1
R-ATH-8934036	Long day regulated expression of florigens	1
R-ATH-8934088	Inflorescence development	1
R-ATH-1119486	IAA biosynthesis I	1
R-ATH-9826798	Gravitropism under normal or artificial gravity environments	1
R-ATH-2961031	Generation of precursor metabolites and energy	1
R-ATH-3906998	Fatty acid and lipid metabolism	1
R-ATH-1119348	Ent-kaurene biosynthesis	1
R-ATH-2867929	Cofactor biosyntheses	1
R-ATH-1119519	Calvin cycle	1
R-ATH-5608118	Auxin signalling	1
R-ATH-9618218	Arsenic uptake and detoxification	1

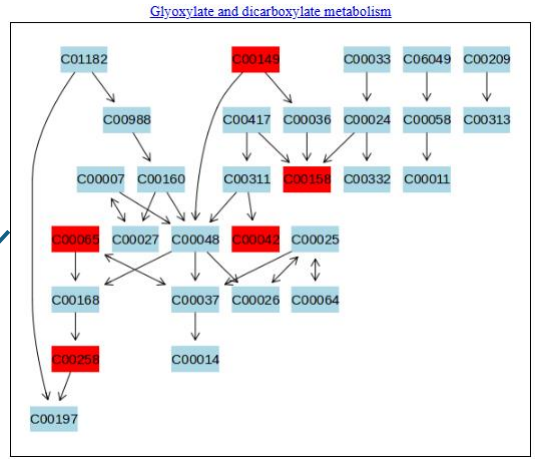
9.3.3/ MetaboAnalyst – all MT – Pathway Analysis

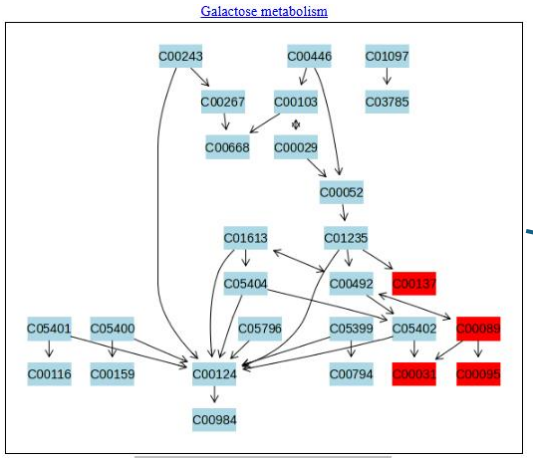
Pathway analysis of all quantified metabolites

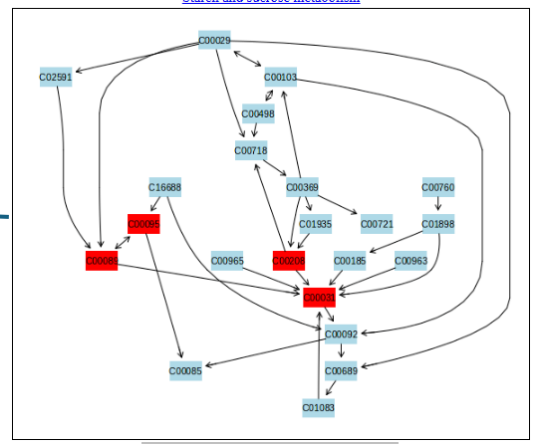
38/54 name hits (70%); qualitative analysis.

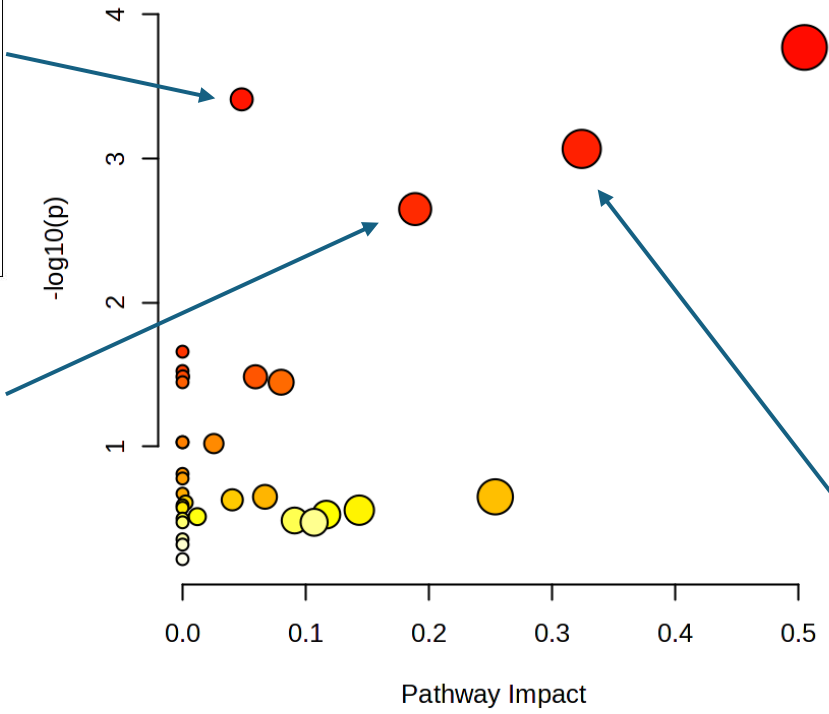
NOTE: compound names edited by removing derivatisation chemical functions (e.g "3TMS derivative").

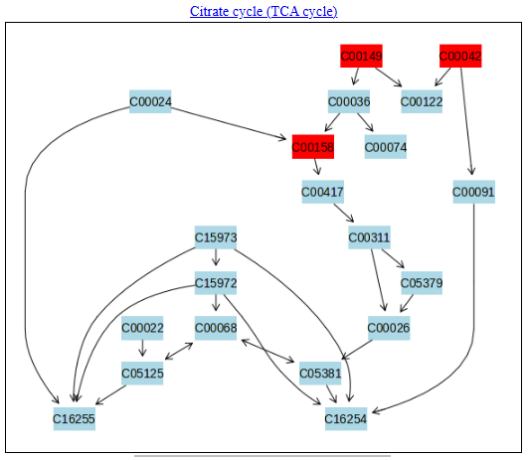
Few pathways flagged mostly from carbohydrate metabolisms.

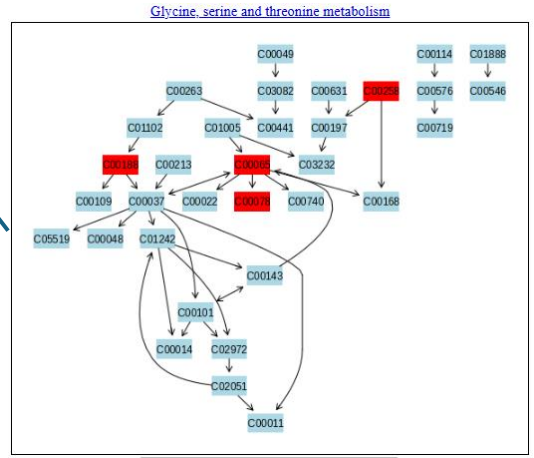












9.4.1/ STRING – Interaction Network

Interaction analysis of TX and PR biomarkers (ID2)

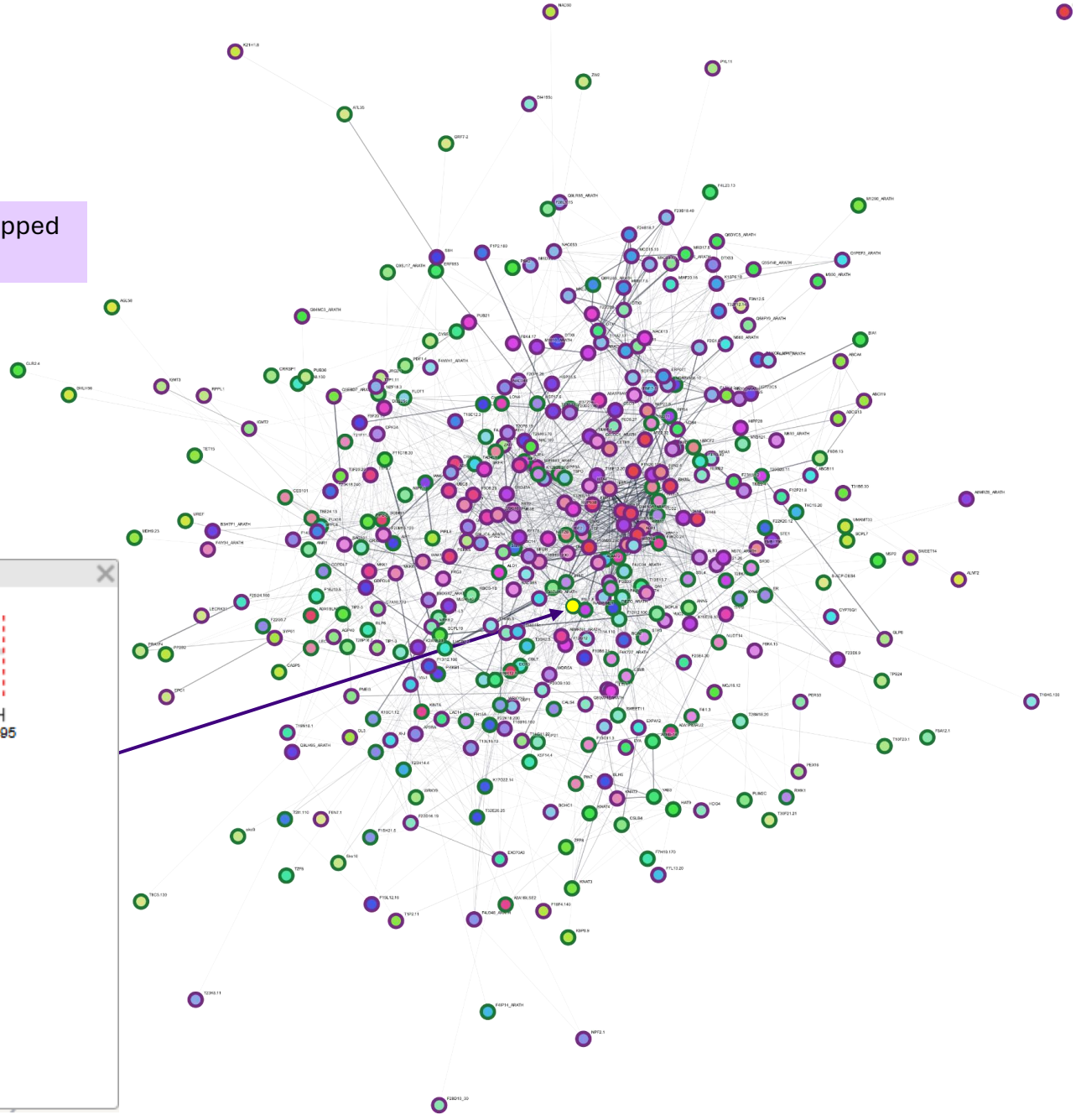
490/491 unique AT ID hits (99%); qualitative analysis

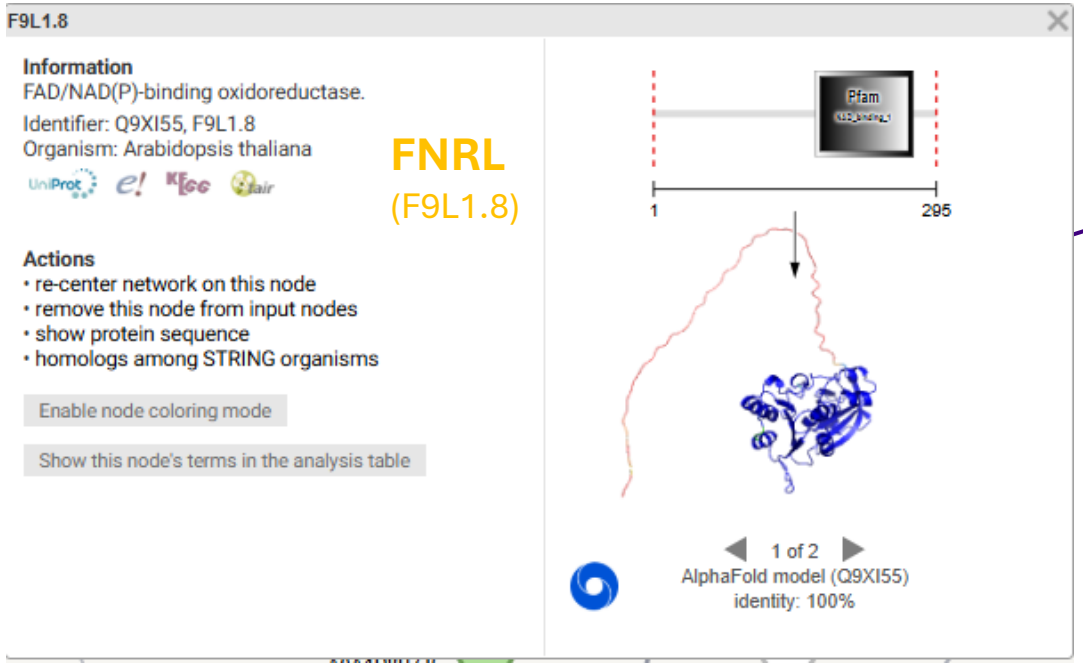
- number of nodes: 480
- number of edges: 358
- average node degree: 1.49
- avg. local clustering coefficient: 0.273
- expected number of edges: 275
- PPI enrichment p-value: 8.51e-07

Quick access to feature information.

Network exported to Cytoscape and merged with quantitative data and mapping annotations.

NOTE: FNRL (PR:AT1G15140.1) mapped (**3702.Q9XI55**).



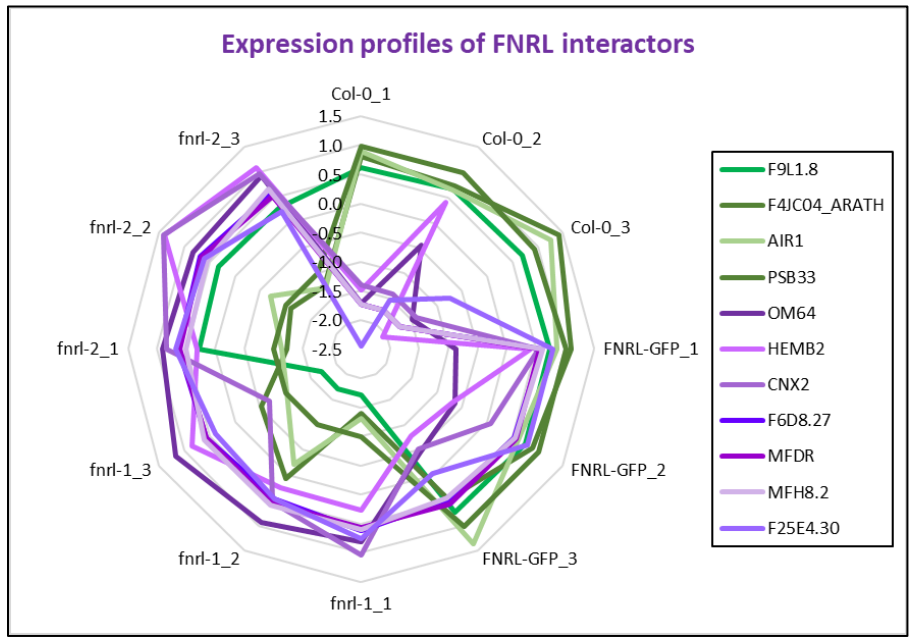


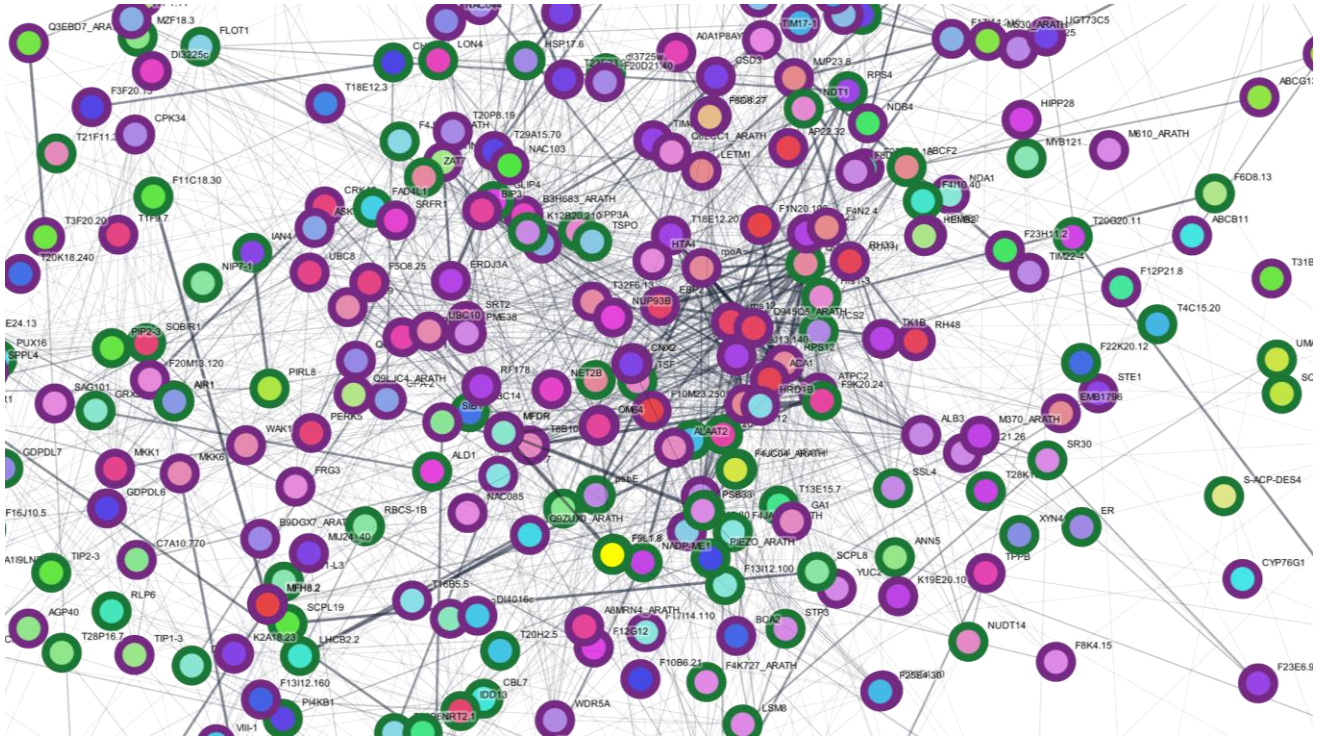
9.4.2/ Cytoscape

FNRL level 1 interactors

FNRL (F9L1.8) closest neighbours selected and isolated:

- **Activated** (like FNRL = F9L1.8): F4JC04_ARATH, AIR1, PSB33.
- **Repressed**: OM64, HEMB2, CNX2, F6D8.27, MFDR, MFH8.2, F25E4.30.





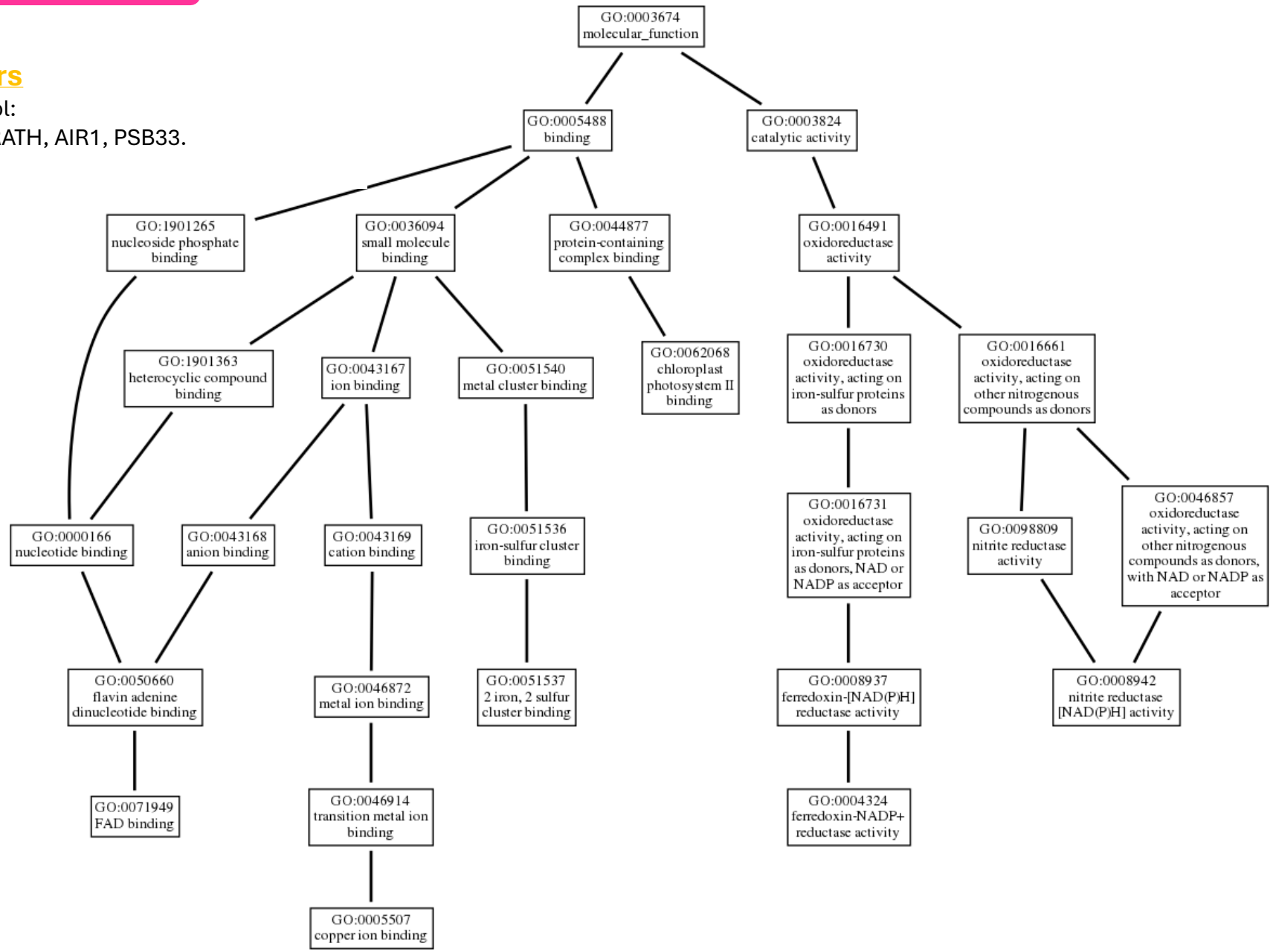
GeneName	std_Col-0_1	std_Col-0_2	std_Col-0_3	d_FNRL-GFP_1	d_FNRL-GFP_2	d_FNRL-GFP_3	std_fnrl-1_1	std_fnrl-1_2	std_fnrl-1_3	std_fnrl-2_1	std_fnrl-2_2	std_fnrl-2_3	TAIR_Short_description
F9L1.8	0.607	0.672	0.702	0.744	0.758	0.727	-1.708	-1.708	-1.708	0.275	0.332	0.306	FAD/NAD(P)-binding oxidoreductase
F4JC04_ARATH	0.806	0.725	1.432	1.107	0.898	0.467	-1.391	0.069	-0.518	-1.224	-1.117	-1.254	
AIR1	0.899	0.655	1.256	0.860	0.614	1.365	-1.299	-0.205	-1.027	-1.132	-0.700	-1.285	Auxin-Induced in Root cultures 1
PSB33	0.988	1.004	0.933	1.024	1.024	1.024	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	Rieske (2Fe-2S) domain-containing protein
OM64	-1.722	-0.444	-1.479	-0.864	-0.632	-0.454	0.819	0.938	1.170	0.902	0.834	0.933	translocon at the outer membrane of chloroplasts 64-V
HEMB2	-1.480	0.417	-2.071	0.410	-0.678	-0.781	0.274	0.258	0.842	0.315	1.406	1.088	Aldolase superfamily protein
CNX2	-1.404	-1.404	-1.404	0.548	0.075	-0.517	1.038	0.528	-0.694	0.832	1.424	0.977	cofactor of nitrate reductase and xanthine dehydrogenase 2
F6D8.27	-1.730	-1.730	-1.730	0.607	0.542	0.490	0.611	0.483	0.560	0.623	0.683	0.588	alpha/beta-Hydrolases superfamily protein
MFDR	-1.731	-1.731	-1.731	0.569	0.569	0.569	0.562	0.562	0.562	0.611	0.663	0.526	Pyridine nucleotide-disulphide oxidoreductase family protein
MFH8.2	-1.729	-1.729	-1.729	0.626	0.564	0.447	0.604	0.604	0.604	0.514	0.550	0.676	NAD(P)-binding Rossmann-fold superfamily protein
F25E4.30	-2.448	-1.525	-0.745	0.793	0.799	-0.036	0.756	0.478	0.394	0.688	0.618	0.227	NAD(P)-binding Rossmann-fold superfamily protein

9.4.3/ GOBP of activated interactors

Activated FNRL level 1 interactors

GO BP terms used in AmiGO visualization tool:

- **Activated** (like FNRL = F9L1.8): F4JC04_ARATH, AIR1, PSB33.



9.5.1/ PMN AraCyc – Cellular overview

Pathway analysis of transcript and protein biomarkers + all metabolites

482/482 unique AT ID hits (100%) + 14/15 MT name hits (93%); quantitative analysis using standardized abundances.
Cellular overview focuses on the pathways.

9.5.2/ PaintOmics 4 - Mapping

Pathway analysis of TX, PR, and PR biomarkers using full dataset as background

All data: 22720/22726 unique AT ID2 hits (99.97%) + 3
Biomarkers: 482/482 unique AT ID2 hits (100%) + 10/1
quantitative analysis using standardized abundances.
162 mapped pathways (18 MapMan + 144 KEGG), 5 sigr

Multiple databases used

The selected species had more than one database available. Your final analysis contains inforr

KEGG

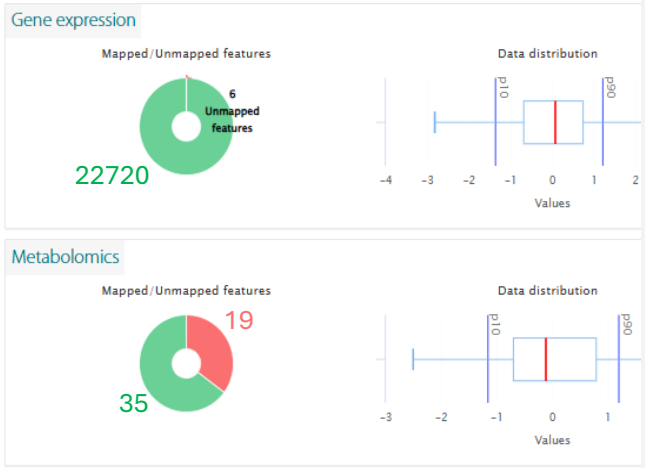
Kyoto Encyclopedia of Genes and Genomes is a database resource for understandir biological system, such as the cell, the organism and the ecosystem, from molecul molecular datasets generated by genome sequencing and other high-throughput exp

Omic	Matched
Gene expression	22720 (98%)
Metabolomics	undefined (NaN%)

MapMan

Oriented towards plant species, in combination with GoMapMan, it provides additio more consolidated annotation for the model species Arabidopsis, and several crop spe

Omic	Matched
Gene expression	23281 (100%)
Metabolomics	undefined (NaN%)

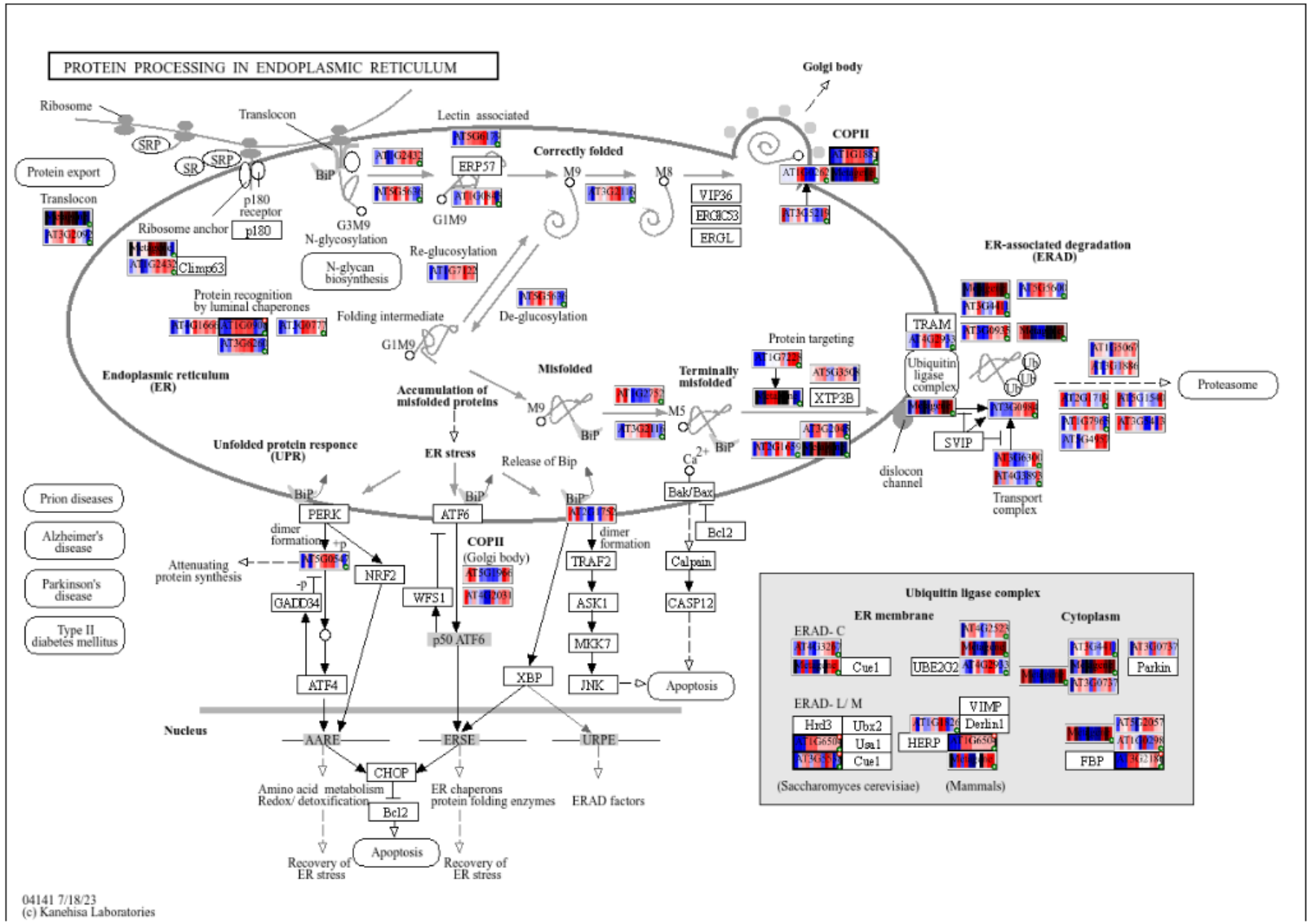


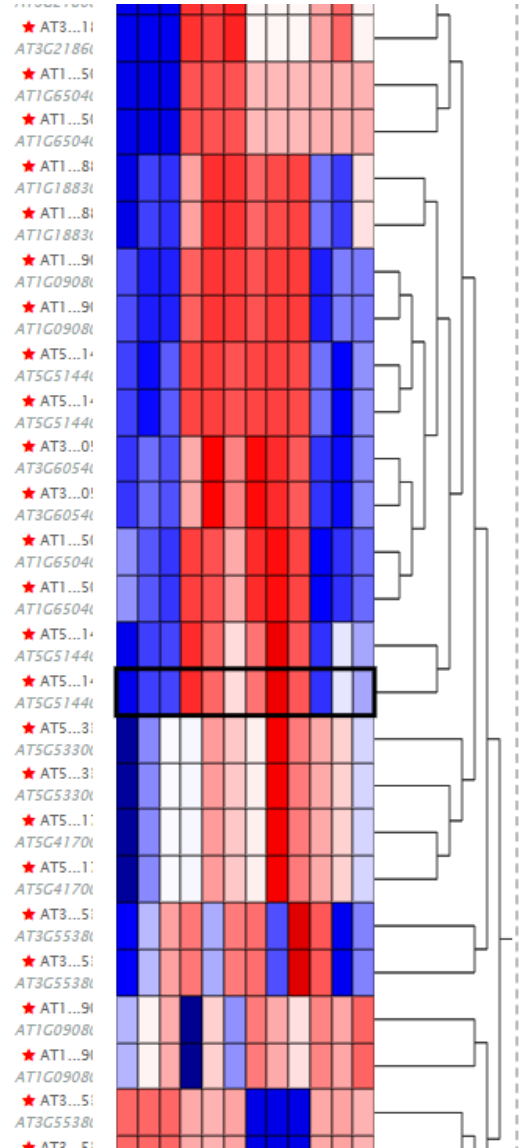
Paint		Pathway name	Features: Unique genes	Features: Unique metabol.	Significance tests: Gene expression	Gene expression (FDR BH)	Metabolomics	Metabolomics (FDR BH)	Combined pValue (Fisher)
☑	K	Protein processing in endoplasmic reticulum	205	0	5.6683e-4	0.09183	-	-	5.6683e-4
☑	K	Polycomb repressive complex	53	0	0.01165	0.94333	-	-	0.01165
☑	K	Protein export	61	0	0.01876	1.00000	-	-	0.01876
☑	K	Nitrogen metabolism	38	0	0.02533	1.00000	-	-	0.02533
☑	K	Ubiquitin mediated proteolysis	142	0	0.03143	1.00000	-	-	0.03143
☑	M	Shikimate Synthesis	53	0	0.09025	1.00000	-	-	0.09025
☑	M	GA Synthesis	5	0	0.09673	1.00000	-	-	0.09673
☑	K	Monoterpenoid biosynthesis	6	0	0.09675	1.00000	-	-	0.09675
☑	K	Cytoskeleton in muscle cells	8	0	0.12690	1.00000	-	-	0.12690
☑	K	Tyrosine metabolism	39	0	0.13901	1.00000	-	-	0.13901
☑	K	Motor proteins	96	0	0.21828	1.00000	-	-	0.21828
☑	K	Lysine biosynthesis	15	0	0.22479	1.00000	-	-	0.22479
☑	K	Biotin metabolism	16	0	0.23786	1.00000	-	-	0.23786
☑	K	Starch and sucrose metabolism	161	6	0.13317	1.00000	0.50877	1.00000	0.25014
☑	K	Phosphatidylinositol signaling system	75	2	0.36061	1.00000	0.20000	1.00000	0.26176
☑	K	Oxidative phosphorylation	150	0	0.27047	1.00000	-	-	0.27047
☑	K	Other glycan degradation	19	0	0.27578	1.00000	-	-	0.27578
☑	K	Lysine degradation	30	2	0.39951	1.00000	0.20000	1.00000	0.28181
☑	K	Isoquinoline alkaloid biosynthesis	20	0	0.28799	1.00000	-	-	0.28799
☑	K	Diterpenoid biosynthesis	21	0	0.30001	1.00000	-	-	0.30001

9.5.3/ PaintOmics 4 – Paint pathways

Enriched pathway: combined p value = 0.0006

Flagged pathways

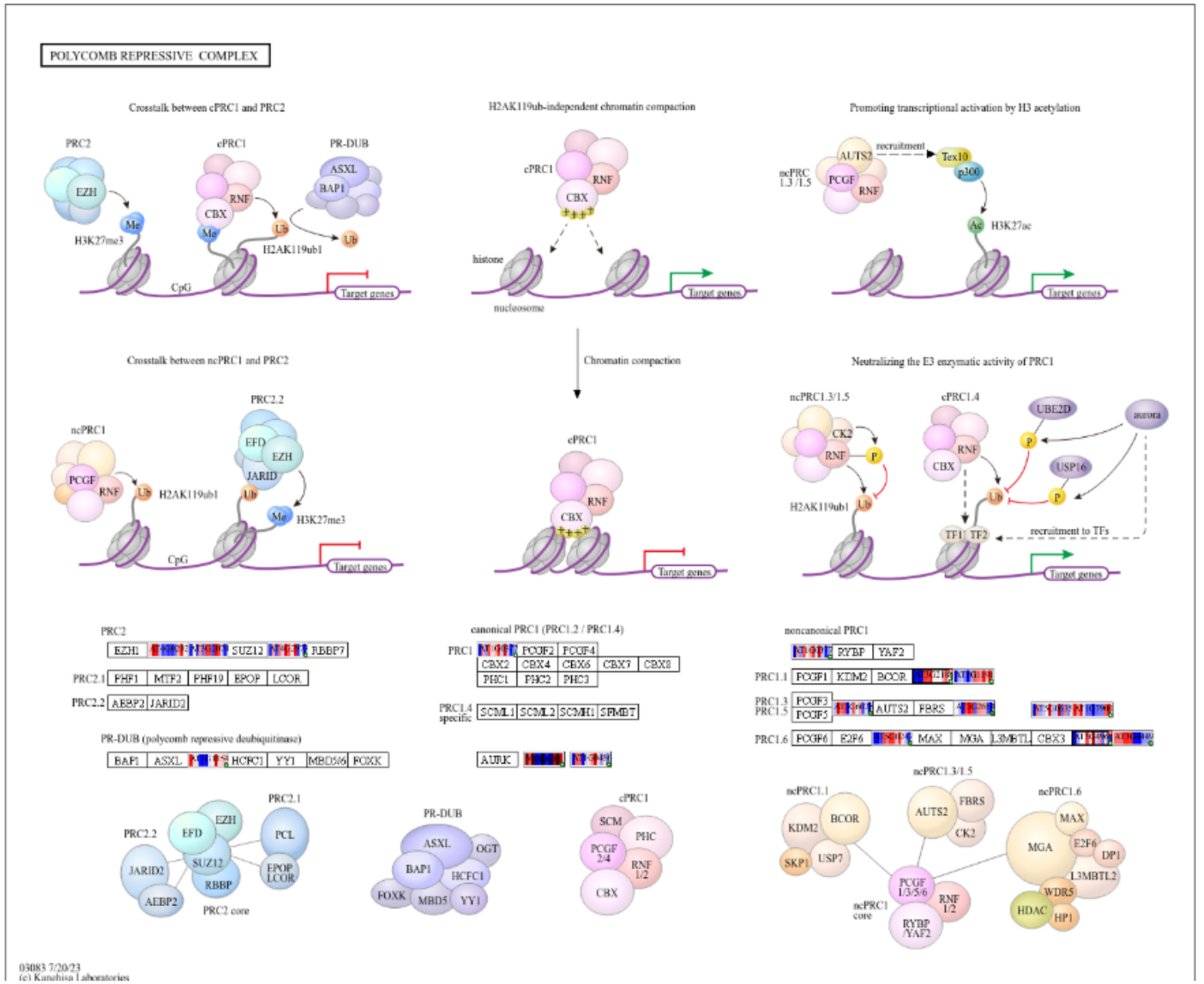




9.5.3/ PaintOmics 4 – Paint pathways

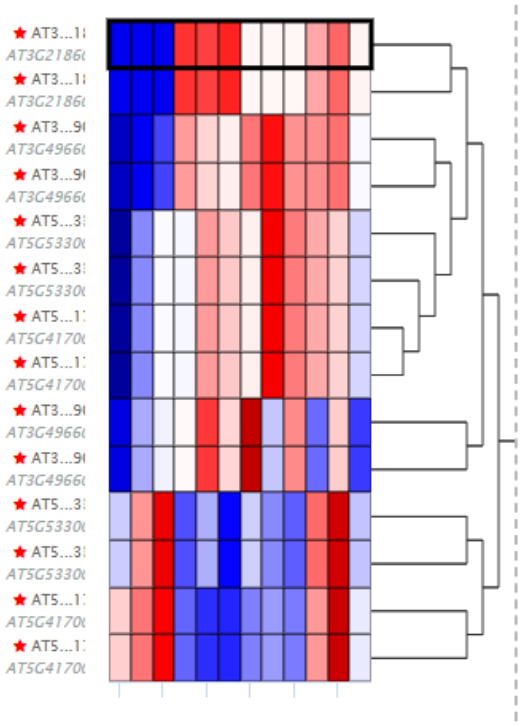
Enriched pathway: combined p value = 0.0118

Flagged pathways



This panel contains the heatmap for all the features involved on this pathway.
Choose the visible omics features will be visible using the Settings button.

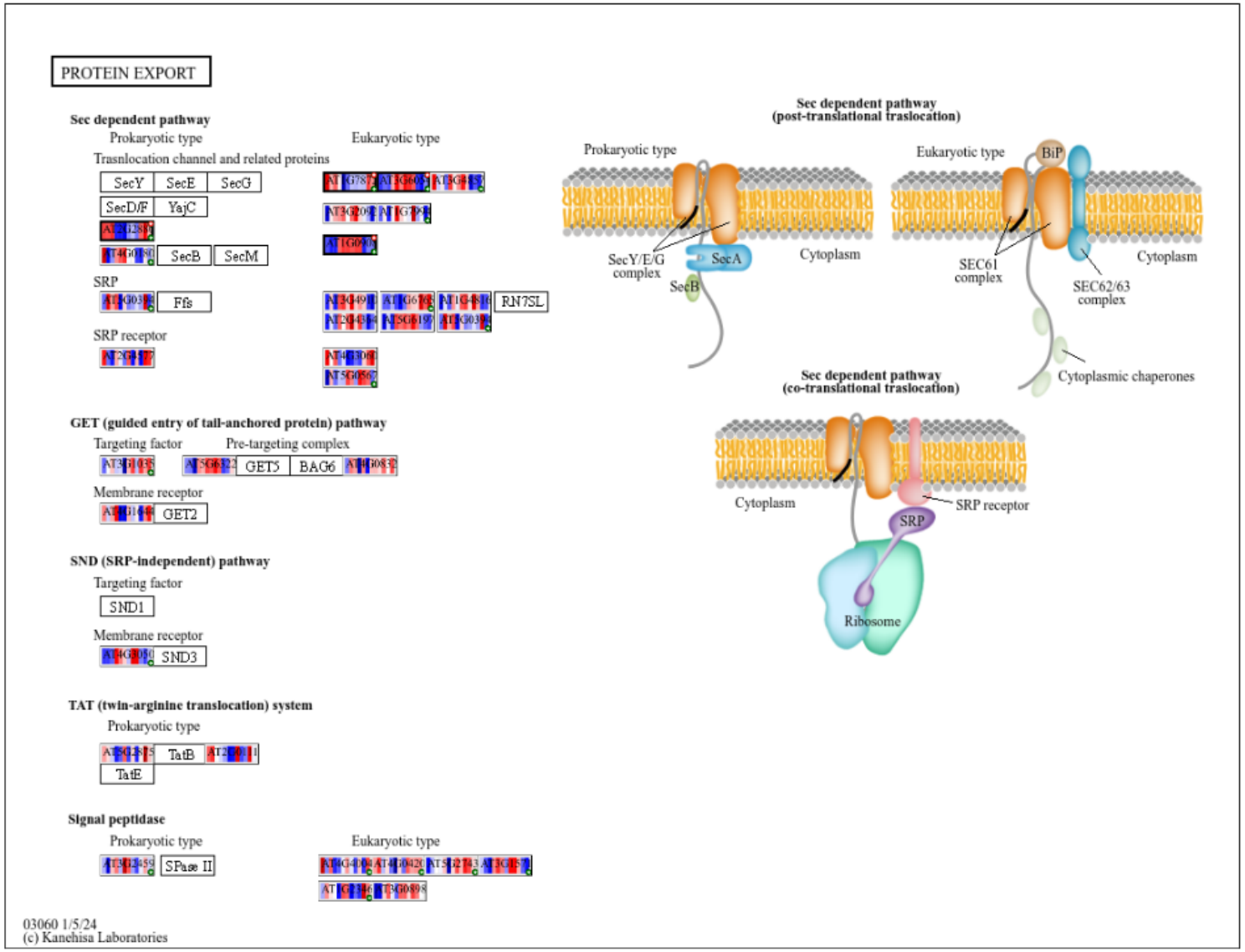
Gene expression



9.5.3/ PaintOmics 4 – Paint pathways

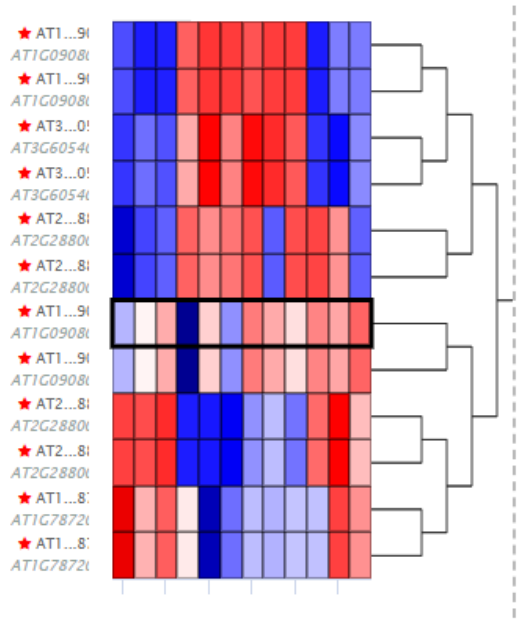
Enriched pathway: combined p value = 0.0187

Flagged pathways



This panel contains the heatmap for all the features involved on this pathway.
Choose the visible omics features will be visible using the Settings button.

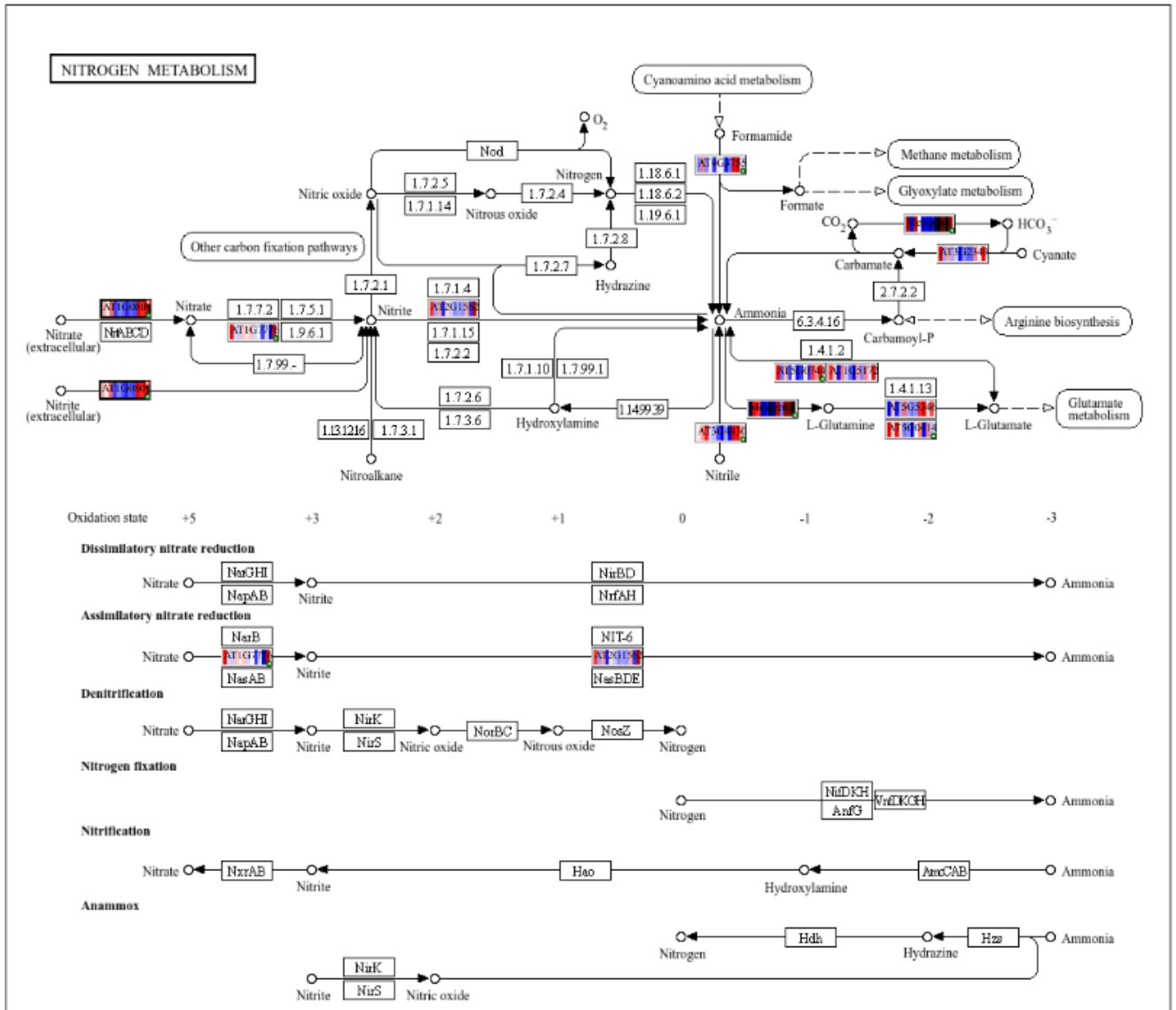
Gene expression



9.5.3/ PaintOmics 4 – Paint pathways

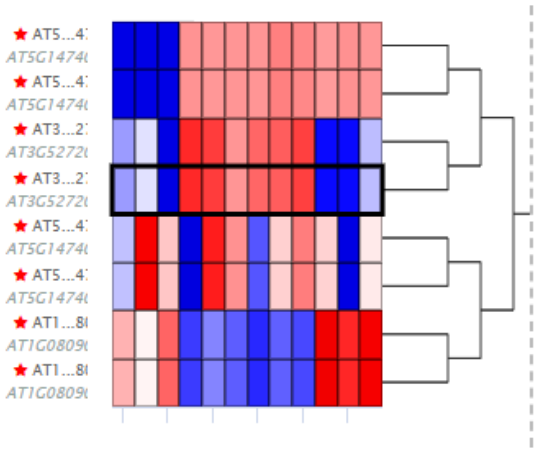
Enriched pathway: combined p value = 0.0253

Flagged pathways



This panel contains the heatmap for all the features involved on this pathway.
Choose the visible omics features will be visible using the Settings button.

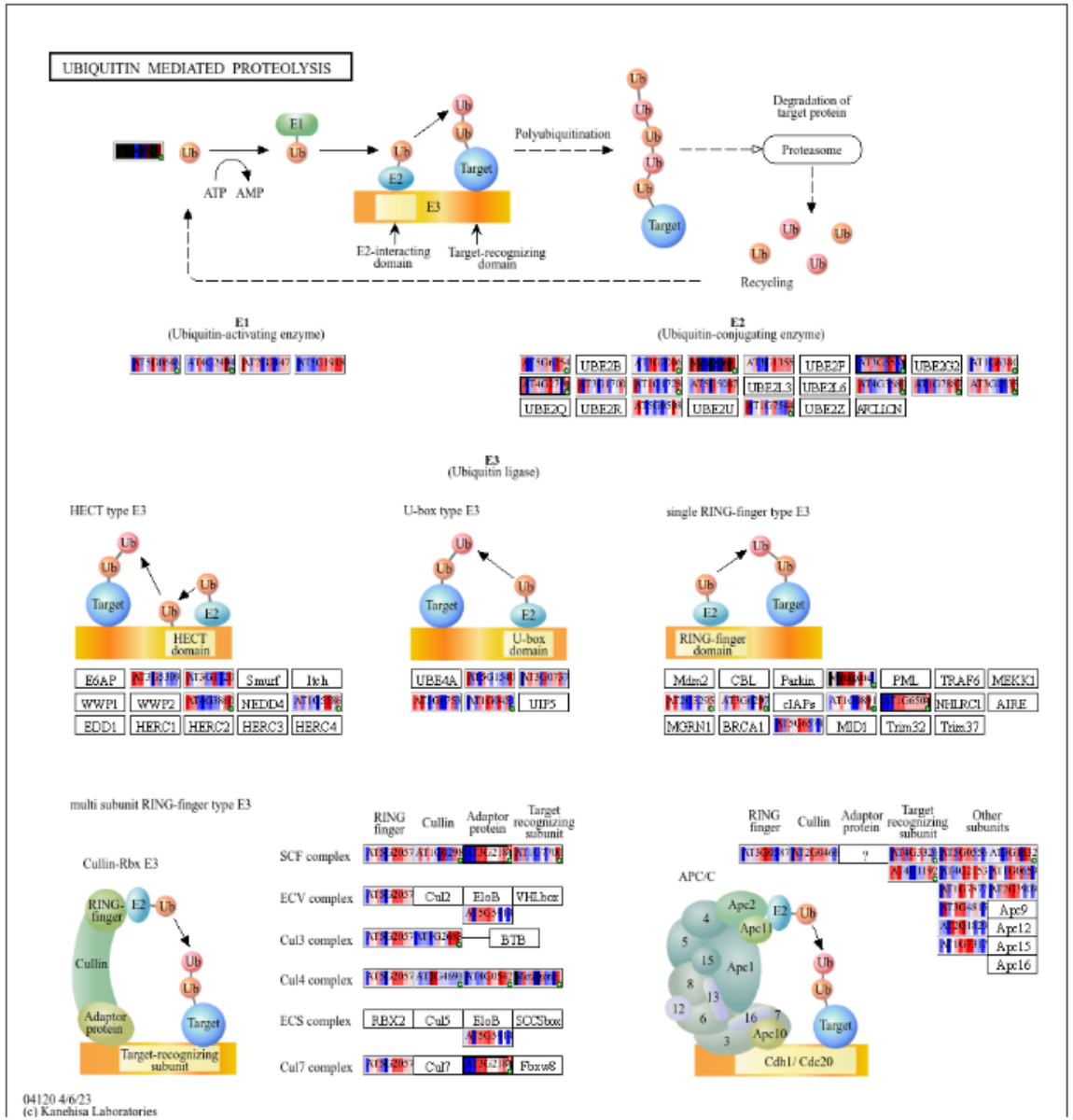
Gene expression

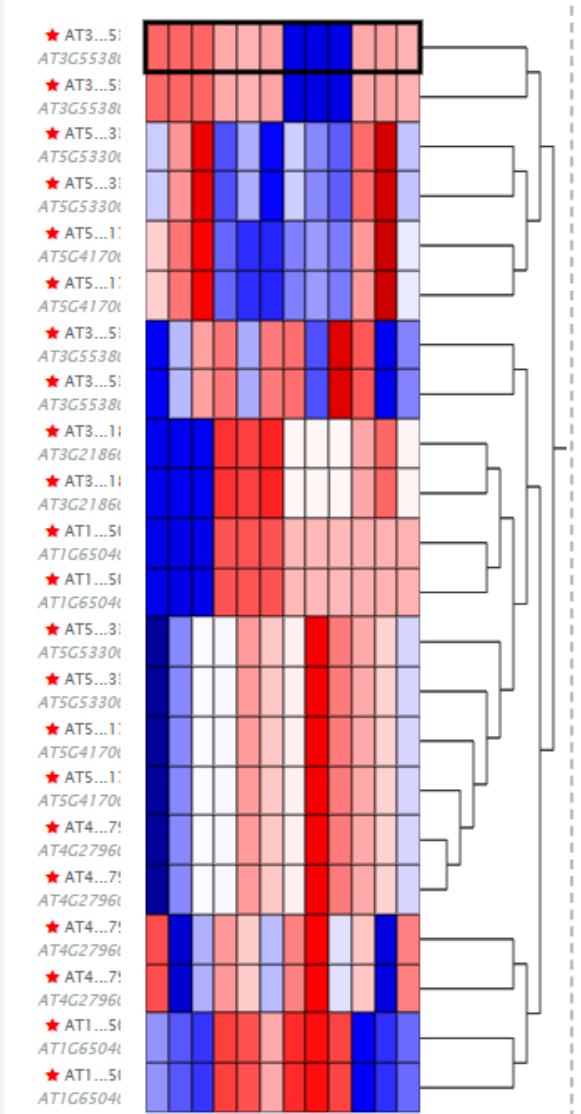


9.5.3/ PaintOmics 4 – Paint pathways

Enriched pathway: combined p value = 0.0314

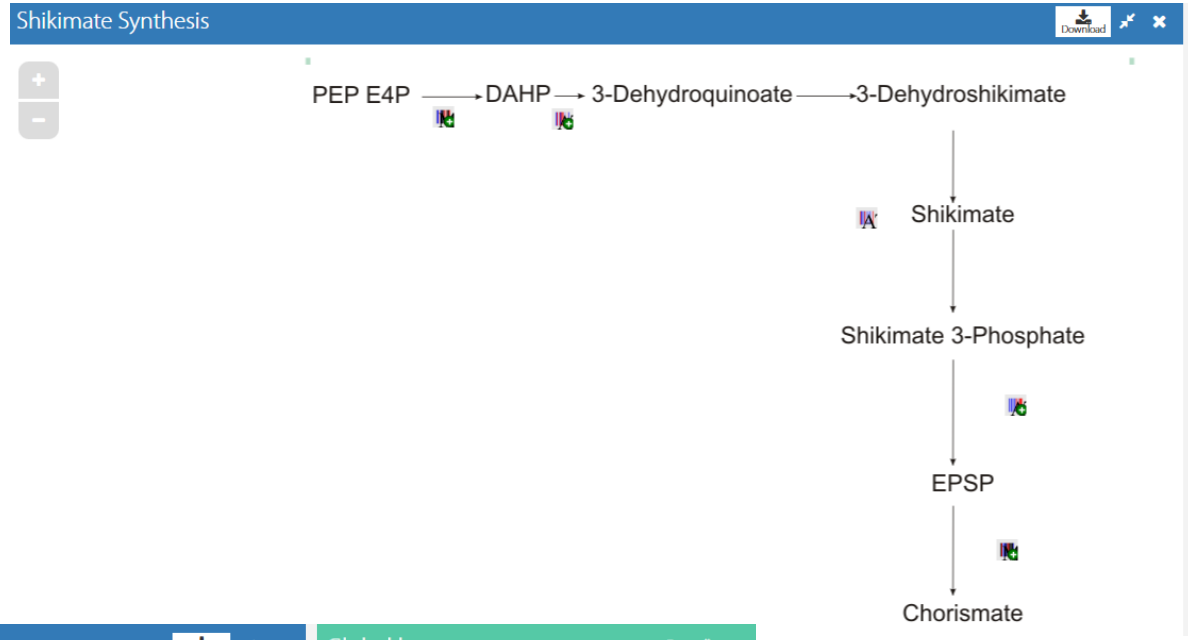
Flagged pathways

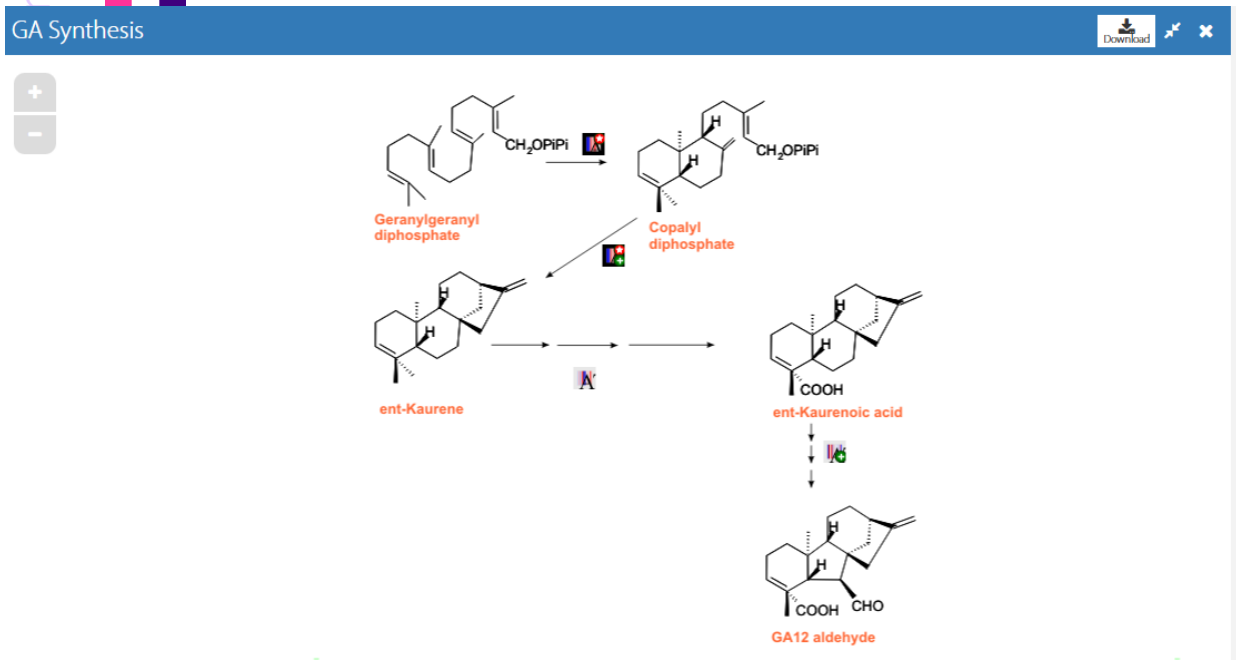


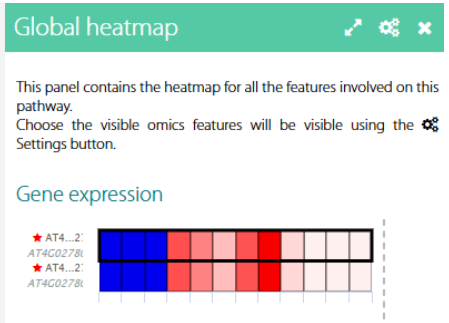


9.5.3/ PaintOmics 4 – Paint pathways

Flagged pathways

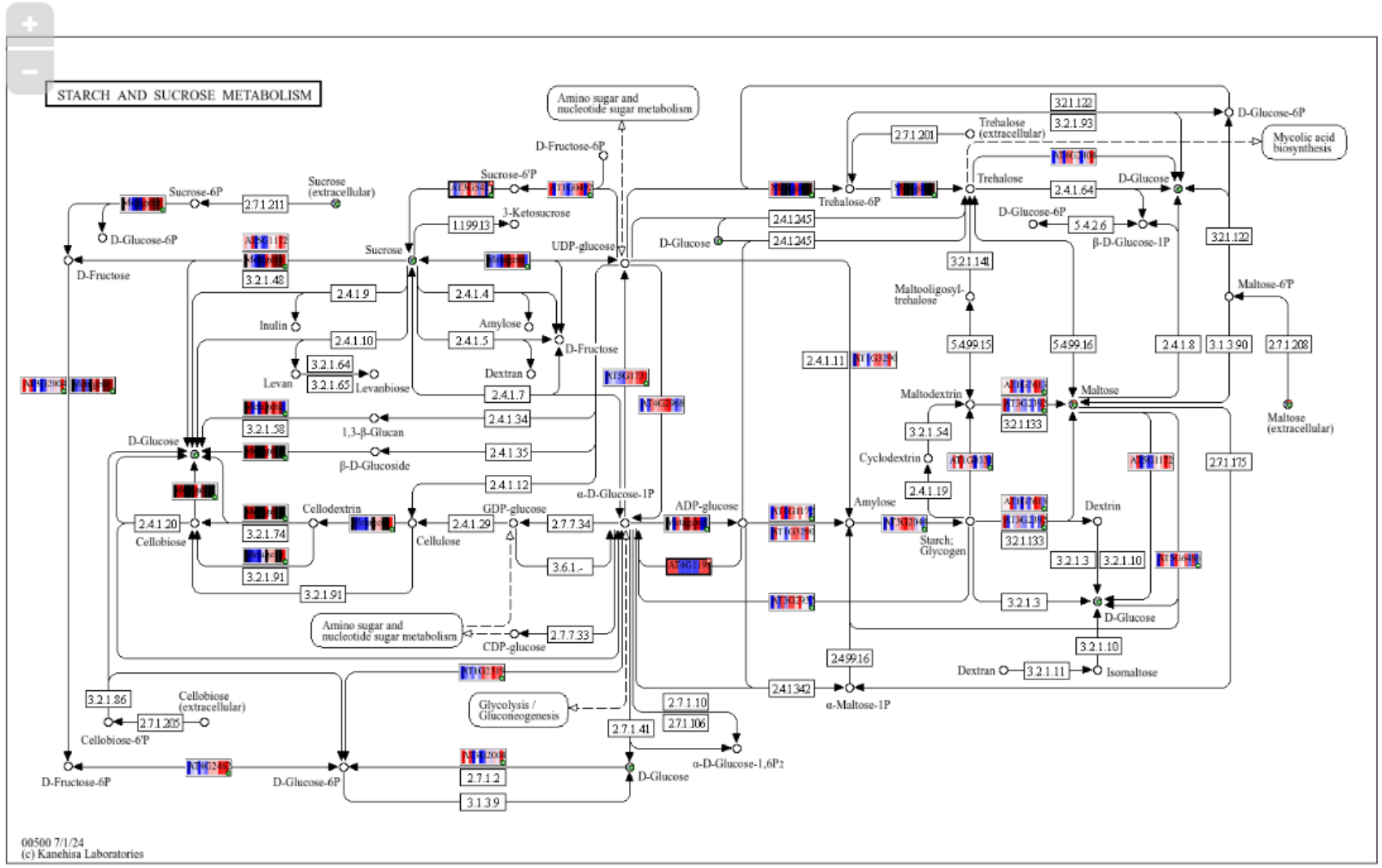


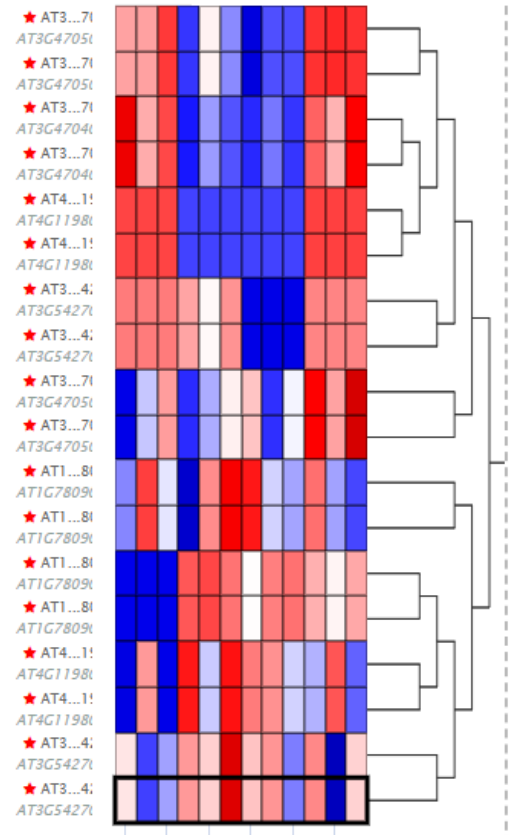




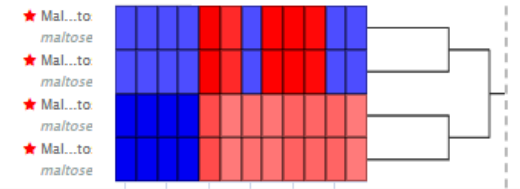
9.5.3/ PaintOmics 4 – Paint pathways

Flagged pathways



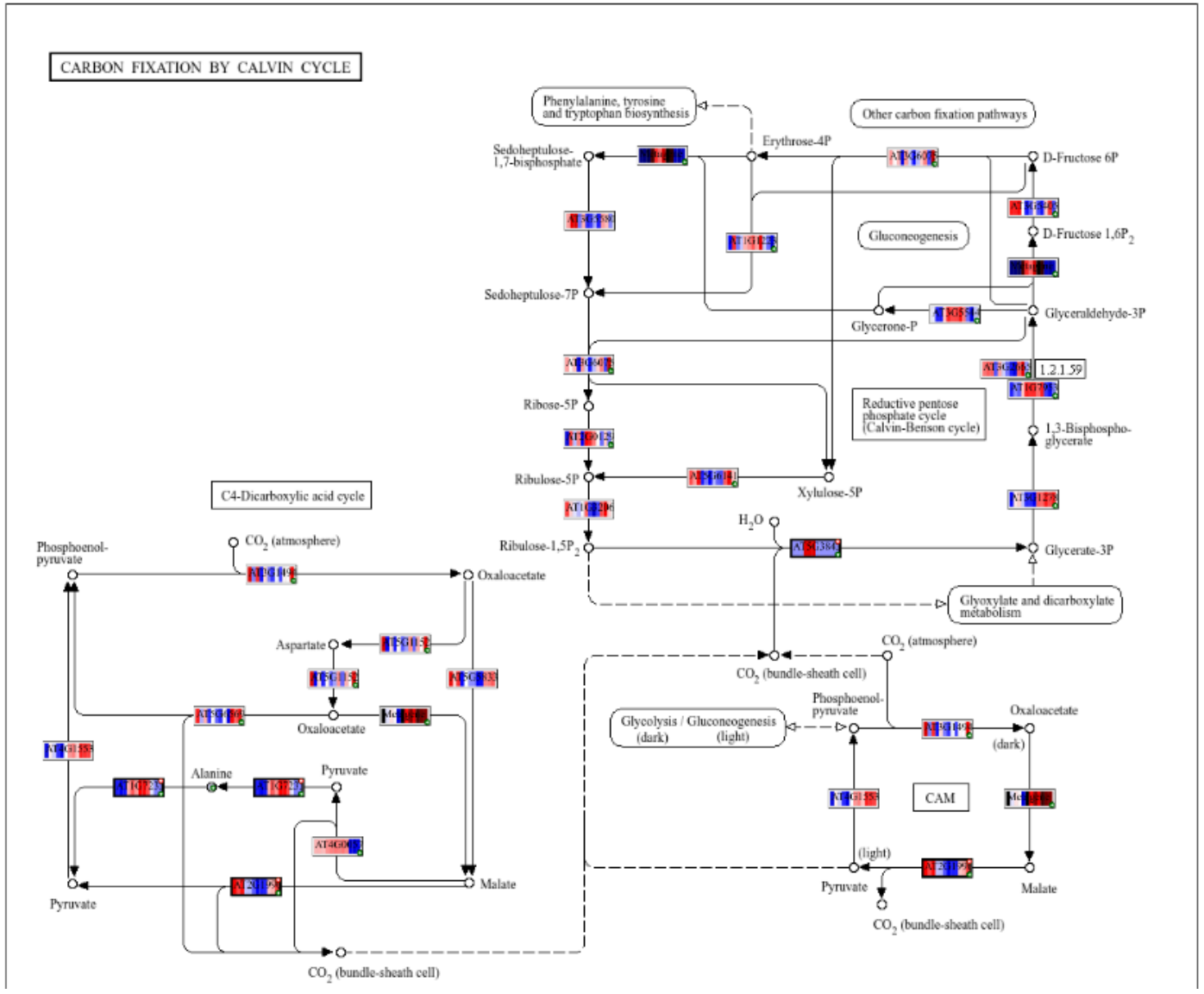


Metabolomics



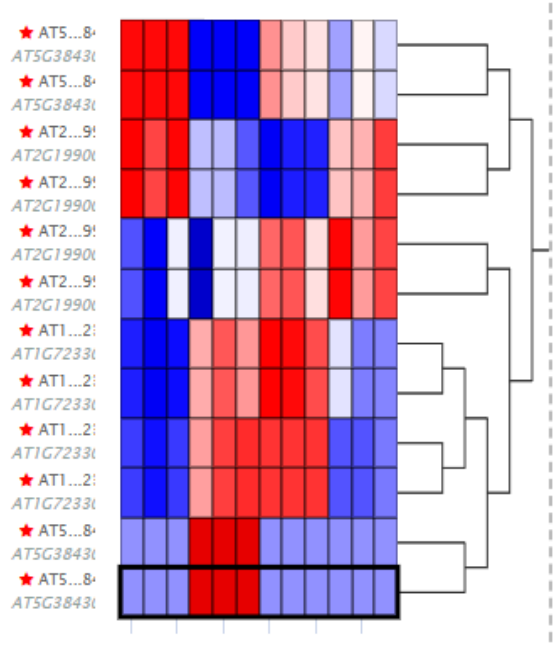
9.5.3/ PaintOmics 4 – Paint pathways

Flagged pathways



This panel contains the heatmap for all the features involvec pathway.
Choose the visible omics features will be visible using Settings button.

Gene expression

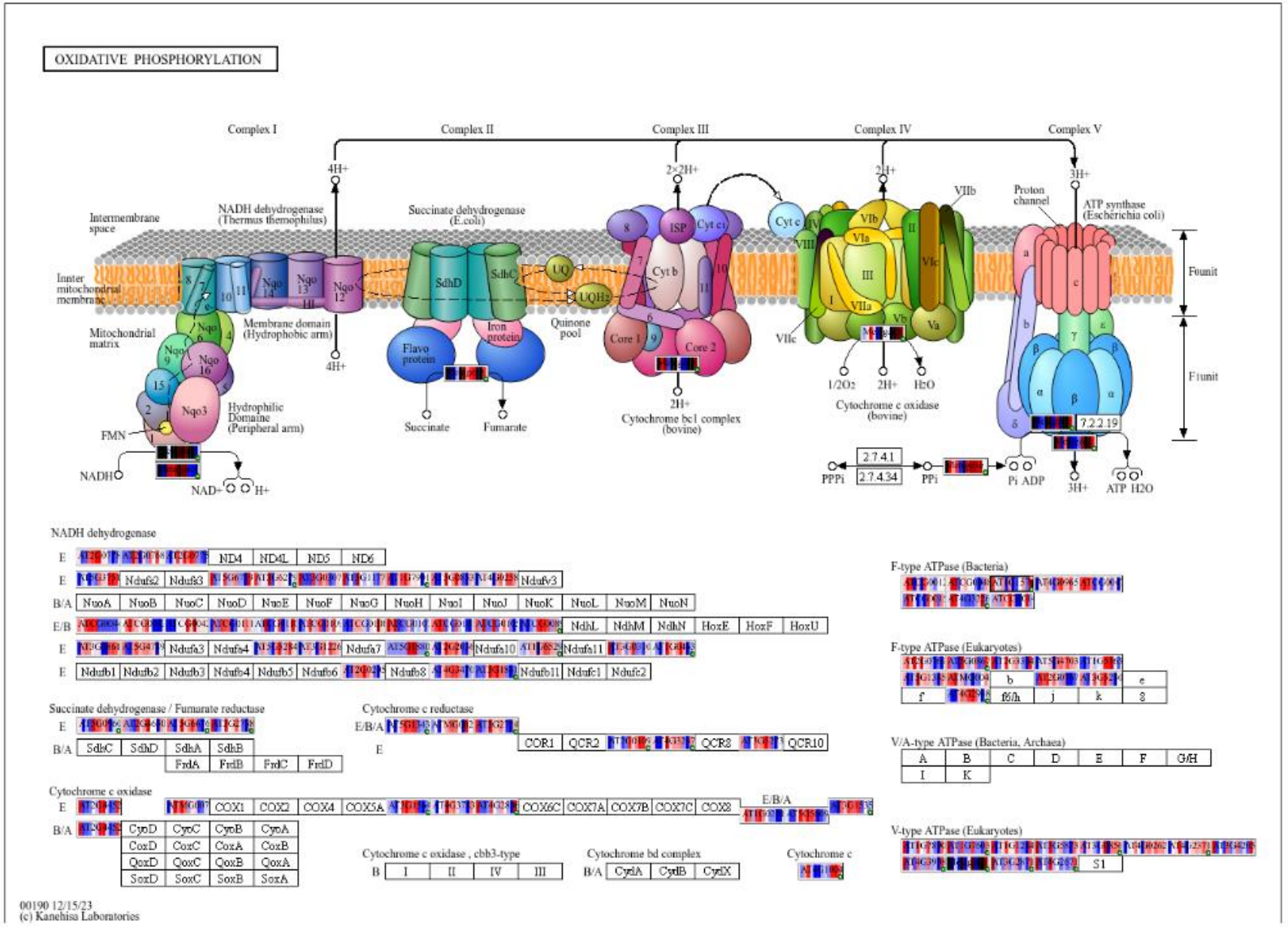


Metabolomics

No data

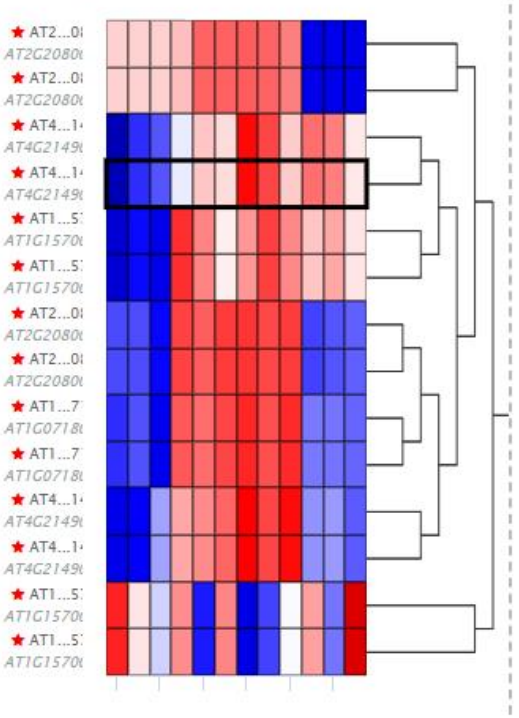
9.5.3/ PaintOmics 4 – Paint pathways

Flagged pathways



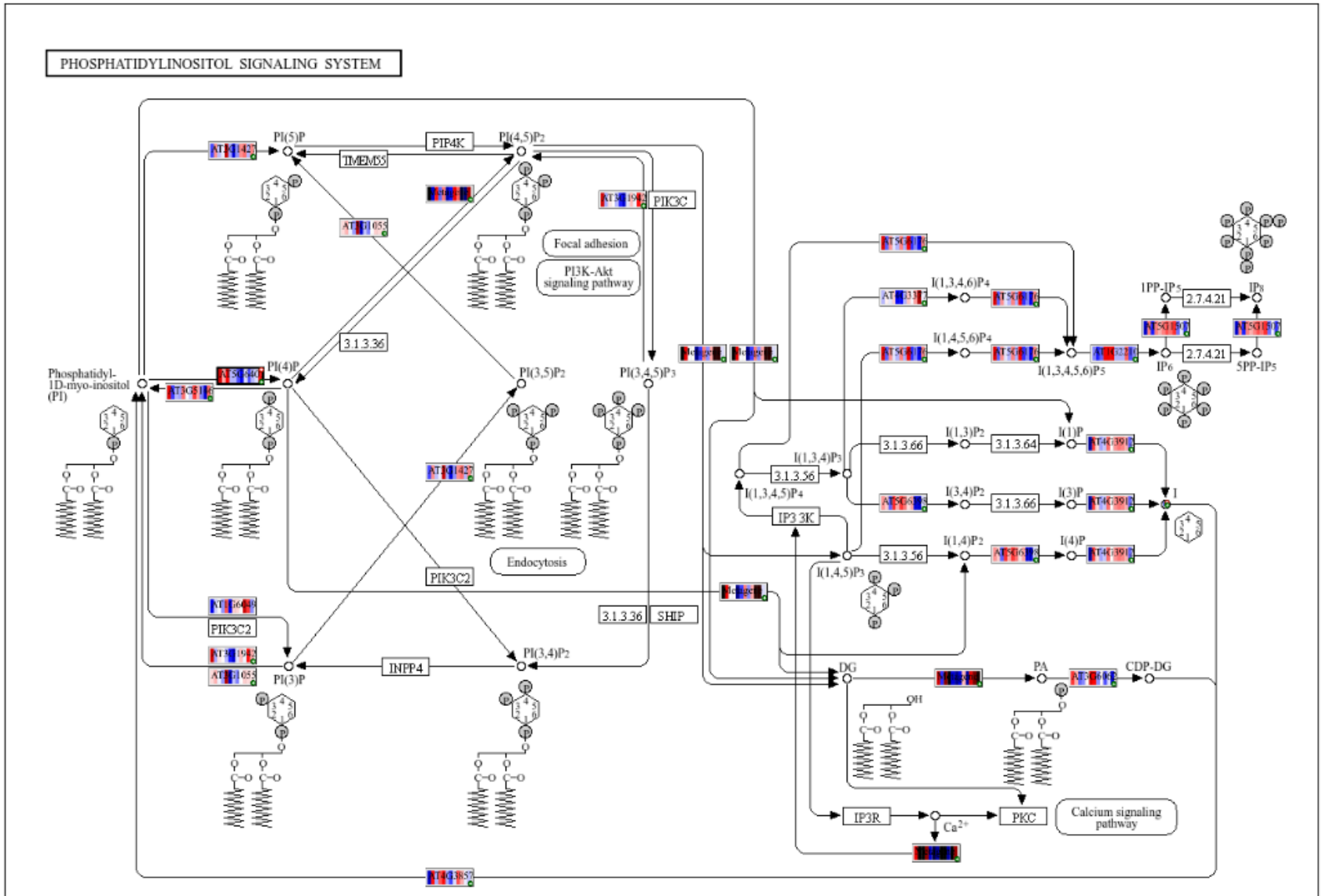
This panel contains the heatmap for all the features involved c
pathway.
Choose the visible omics features will be visible using tl
Settings button.

Gene expression



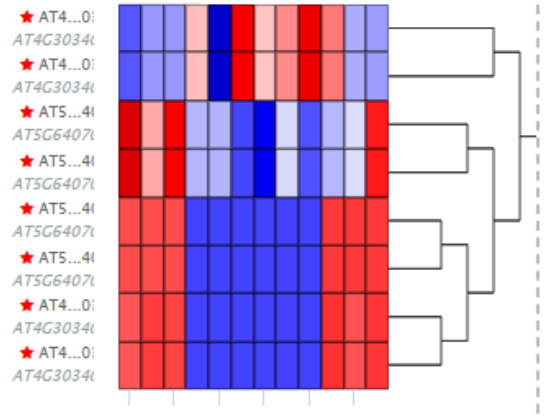
9.5.3/ PaintOmics 4 – Paint pathways

Flagged pathways

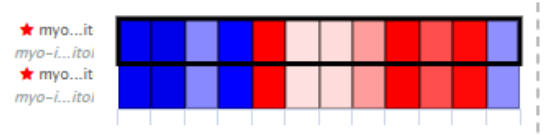


This panel contains the heatmap for all the features involved on this pathway.
Choose the visible omics features will be visible using the Settings button.

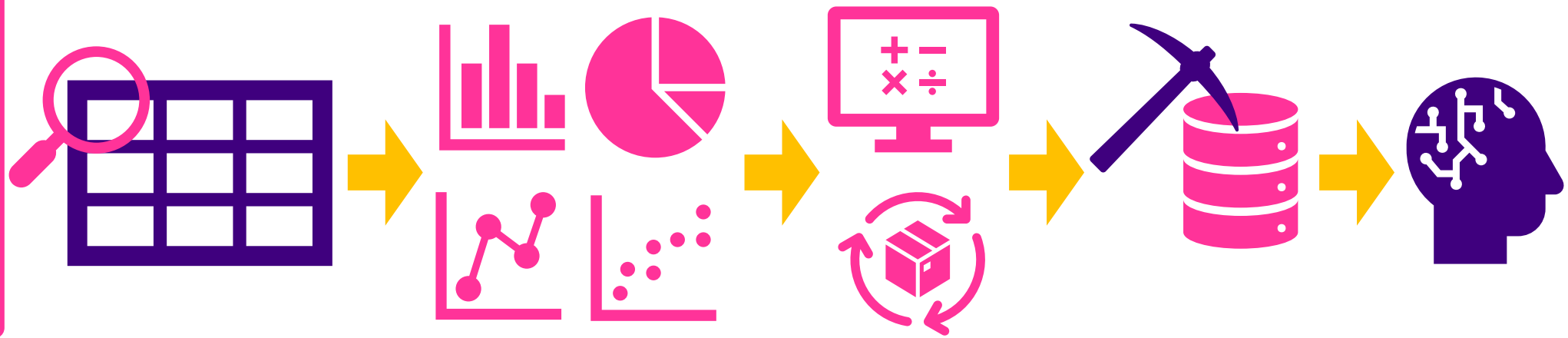
Gene expression



Metabolomics



10/ Conclusions



10.1/ Summary

Key findings:

- Statistical analyses (DE, volcano plot) → 341 TX + 148 PR + 15 MT= **504 biomarkers** (including FNRL (PR:AT1G15140.1)
- Biomarker EDA (HCA, k-means, LDA) → **4 main expression profiles**
- FNRL-regulation conceptual model→ **activation (down in mutant) or repression (up in mutant) with full or partial gene rescue**
- **Root length** phenotype accurately **predicted by linear regression model (Elastic Net)**
- Some data mining tools proved less useful (hit rates in **Table**):
  - PlantReactome (too few hits)
  - PeptideAtlas (no pathway annotations)
  - MetaboAnalyst (MT only)
  - AraCyc (too unstable, inaccessible for weeks at a time)
- Best data mining tools were:
  - PaintOmics (perfect for omics integration of A. thaliana)
  - STRING
  - GO/AmiGO
- Enriched biochemical **pathways**:
  - Protein processing in endoplasmic reticulum
  - Polycomb repressive complex
  - Protein export
  - Nitrogen metabolism
  - Ubiquitin mediated proteolysis
- Network analysis flagged 10 **associated features**:
  - 3 co-activated (AIR1, F4JC04_ARATH, PSB33)
  - 7 repressed (CNX2, HEMB2, F25E4.30, F6D8.27, MFDR, MFH8.2, OM64)

	All biomarker counts per omics layer				Mapped biomarker counts per omics layer			
Database	TX	PR	MT	TX+PR+MT	TX	PR	MT	TX+PR+MT (%)
TAIR	341	148	15	504	341	148	0	489 (97%)
UniProtKB	341	148	15	504	338	148	0	486 (96%)
PeptideAtlas	341	148	15	504	0	143	0	143 (28%)
MetaboAnalyst	341	148	15	504	0	0	14	14 (3%)
GO Resource	341	148	15	504	266	147	0	413 (82%)
KEGG2	341	148	15	504	105	70	14	189 (38%)
PlantReactome	341	148	15	504	10	6	0	16 (3%)
STRING	341	148	15	504	341	148	0	489 (97%)
PaintOmics 4	341	148	15	504	341	148	10	499 (99%)
PMN AraCyc	341	148	15	504	341	148	14	503 (99%)

FNRL Mode of Action:

- Well, I think I'll leave this to the expert (Ghaz) based on all my results! ☺

10.2/ Next steps (Delphine then Ghaz)

Draft manuscript:

- **Abstract:** Ghaz
- **Introduction**: Ghaz
- **Methods**:
  - Wet-lab study: Ghaz
  - Dry-lab study: Delphine (DONE)
- **Results:** Delphine (DONE)
- **Discussion**: Ghaz
- **Figures, Tables, Supplementary data**: Delphine (DONE)
  - 4 main Tables
  - 4 Suppl. Tables
  - 6 main Figures
  - 8 Suppl. Figures

Thank you!